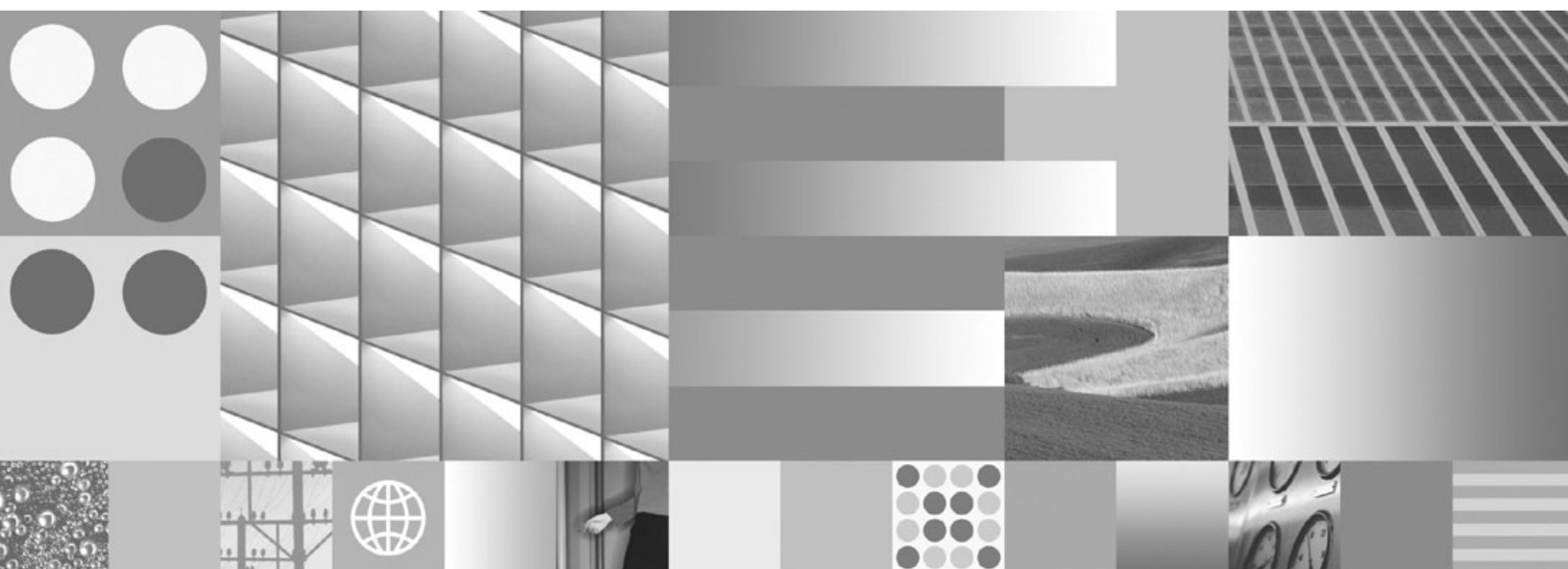


Administrowanie wyszukiwaniem korporacyjnym



Administrowanie wyszukiwaniem korporacyjnym

Uwaga

Przed rozpoczęciem korzystania z niniejszych informacji oraz z produktu, którego one dotyczą, należy zapoznać się z informacjami zawartymi w sekcji "Uwagi i znaki towarowe" na stronie 417.

Uwaga dotycząca wydania

To wydanie dotyczy wersji 8, wydania 5, modyfikacji 0 produktu IBM OmniFind Enterprise Edition (numer produktu 5724-C74) oraz wszystkich jego następnych wydań i modyfikacji, chyba że w nowych wydaniach podręcznika zostanie zaznaczone, że jest inaczej.

Wysłanie informacji do firmy IBM daje jej prawo do ich używania i dystrybucji w dowolny sposób, jaki uzna za właściwy, bez żadnych zobowiązań wobec ich nadawcy.

© Copyright International Business Machines Corporation 2004, 2008. Wszelkie prawa zastrzeżone.

Spis treści

Serwis ibm.com i zasoby pokrewne . . . vii

Jak przysłać swoje uwagi vii

Kontakt z firmą IBM. viii

Co to jest wyszukiwanie korporacyjne? 1

Typy źródeł danych obsługiwane w wyszukiwaniu korporacyjnym 2

Przegląd komponentów wyszukiwania korporacyjnego . . . 3

Przeszukiwacze wyszukiwania korporacyjnego 4

Analizatory składni wyszukiwania korporacyjnego . . . 5

Indeksy wyszukiwania korporacyjnego 7

Serwery wyszukiwania dla wyszukiwania korporacyjnego 8

Konsola administracyjna wyszukiwania korporacyjnego 9

Monitorowanie systemu wyszukiwania korporacyjnego 9

Pliki dzienników wyszukiwania korporacyjnego . . . 10

Dostosowywanie wyszukiwania korporacyjnego . . . 11

Przykładowa aplikacja wyszukująca dla wyszukiwania korporacyjnego 11

Przepływ danych w wyszukiwaniu korporacyjnym . . . 12

Administrowanie systemem wyszukiwania korporacyjnego. 15

Logowanie się do konsoli administracyjnej 19

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji jednoserwerowej. 20

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji wieloserwerowej. 21

Numery portów TCP używane w wyszukiwaniu korporacyjnym. 23

Zmianianie numeru portu dla systemu wyszukiwania korporacyjnego 24

Zmianianie nazw hosta lub adresów IP serwera wyszukiwania korporacyjnego 25

Konfigurowanie obsługi podwójnych adresów IP 27

Włączanie obsługi protokołu IPv6 28

Kolekcje wyszukiwania korporacyjnego 31

Tworzenie kolekcji przy użyciu Kreatora kolekcji 31

Tworzenie kolekcji przy użyciu widoku kolekcji. . . . 33

Edytowanie kolekcji 34

Usuwanie kolekcji. 35

Określanie identyfikatora kolekcji 36

Administrowanie przeszukiwaczem . . . 37

Tworzenie przeszukiwacza 39

Edycja właściwości przeszukiwacza 40

Edytowanie przestrzeni przeszukiwania 40

Usuwanie przeszukiwacza 41

Harmonogramy przeszukiwacza. 41

Przeszukiwacze Content Edition 42

Dostęp do repozytoriów Content Edition w trybie bezpośrednim 43

Dostęp do repozytoriów WebSphere II Content Edition w trybie serwera 44

Przeszukiwacze DB2 47

Konfigurowanie serwera przeszukiwania w systemie UNIX do przeszukiwania źródeł DB2 49

Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł DB2 50

Konfigurowanie programu WebSphere Information Integrator Event Publisher Edition dla przeszukiwaczy DB2 51

Konfigurowanie programu WebSphere MQ dla przeszukiwaczy DB2 54

Przeszukiwanie baz danych DB2 na klasycznym serwerze źródła danych 55

Przeszukiwacze DB2 Content Manager 56

Konfigurowanie serwera przeszukiwaczy w systemie UNIX do współpracy z przeszukiwaczami DB2 Content Manager. 57

Konfigurowanie serwera przeszukiwaczy w systemie Windows do współpracy z przeszukiwaczami DB2 Content Manager 59

Przeszukiwacze Domino Document Manager. 61

Przeszukiwacze Exchange Server 63

Przeszukiwacze baz danych JDBC 64

Odwzorowania związków dla baz danych JDBC 66

Przeszukiwanie wielu ustrukturyzowanych tabel bazy danych JDBC 68

Przeszukiwacze NNTP 72

Przeszukiwacze Notes 73

Wskazówki dotyczące przeszukiwania baz danych Lotus Domino 75

Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino 76

Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino . . . 78

Konfigurowanie serwerów korzystających z protokołu DIIOP 80

Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino 81

Przeszukiwacze QuickPlace 82

Przeszukiwacze listy adresów początkowych 86

Przeszukiwacze systemu plików UNIX 87

Przeszukiwacze sieci WWW. 88

Konfiguracja agenta użytkownika 89

Jak przeszukiwacz sieci WWW korzysta z protokołu Robots Exclusion 90

Obsługa języka JavaScript 92

Reguły ograniczania przestrzeni przeszukiwania w sieci WWW 93

Testowanie połączeń URL za pomocą przeszukiwacza sieci WWW. 98

Ustawienia przeszukiwacza sieci WWW określające odstęp czasu przed ponownym przeszukiwaniem. . . . 98

Opcje odwiedzania adresów URL przy użyciu przeszukiwacza sieci WWW. 99

Jak przeszukiwacz sieci WWW obsługuje strony błędów programowych	99
Przeszukiwanie zabezpieczonych serwisów WWW	101
Serwisy WWW obsługiwane przez serwery proxy	103
Administrowanie informacjami cookie	104
Konfiguracja globalnej przestrzeni przeszukiwania sieci WWW	105
Dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie	107
Ignorowanie dyrektyw blokujących wykorzystywanie odsyłaczy i indeksowanie na stronach WWW	108
Konfigurowanie daty używanej przez przeszukiwacze sieci WWW dla przeszukiwanych dokumentów	108
Przeszukiwacze Web Content Management	110
Przeszukiwacze WebSphere Portal	111
Kopiowanie adresu URL do przeszukania z portalu WebSphere Portal	113
Przeszukiwacze systemu plików Windows	114
Konfiguracja obsługi aplikacji nasłuchiwanie danych	116
Niestandardowe wtyczki przeszukiwacza	117
Obsługa przeszukiwania plików archiwum	119
Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego	120

Administrowanie analizatorem składni 131

Praca z kategoriami	132
Kategorie oparte na regułach	132
Drzewa kategorii	134
Wybór typu klasyfikacji	135
Konfiguracja kategorii	136
Praca z polami wyszukiwania XML	137
Pola wyszukiwania XML	137
Odwzorowanie elementów XML na pola wyszukiwania	138
Praca z polami wyszukiwania HTML	140
Pola wyszukiwania HTML	141
Odwzorowywanie elementów metadanych HTML na pola wyszukiwania	141
Niestandardowe przetwarzanie tekstu	143
Dodawanie mechanizmów analizy tekstu do systemu	144
Powiązanie mechanizmu analizy tekstu z kolekcją	145
Odwzorowywanie elementów XML na wspólną strukturę analizy	146
Odwzorowywanie wspólnej struktury analizy na indeks	147
Odwzorowywanie wspólnej struktury analizy na relacyjną bazę danych	148
Konfigurowanie wątków dla usługi analizatora składni	150
Włączanie zaawansowanej analizy terminów złożonych	150
Włączanie obsługi rodzimego wyszukiwania XML	151

Wykrywanie formatu dokumentów . . . 153

Domyślne typy obsługiwanych dokumentów	154
Typy dokumentów związane z analizatorami kolekcji i analizatorami Stellent	155
Powiązanie typów dokumentów z analizatorem składni kolekcji	155
Domyślne reguły usługi analizatora składni kolekcji	157
Analizowanie dokumentów nieznanymi typów	158

Zmiana reguł wymiany dla niektórych znaczników HTML	158
Domyślne reguły wymiany HTML	159
Powiązanie typów dokumentów z analizatorem składni Stellent	160
Domyślne reguły analizowania przez analizatory składni Stellent	163

Obsługa języków i stron kodowych 167

Automatyczne rozpoznawanie języka	168
Automatyczne wykrywanie strony kodowej	169
Analiza lingwistyczna dokumentów chińskich, japońskich i koreańskich	170
Podział na ciągi n-znakowe	170
Usuwanie znaków spacji, tabulacji lub nowego wiersza z tekstu	170

Administrowanie indeksem 173

Planowanie operacji budowania indeksu	174
Zmiana harmonogramu operacji na indeksie	175
Włączanie i wyłączanie harmonogramów operacji na indeksach	175
Konfiguracja współbieżnych operacji budowania indeksu	176
Budowanie indeksów tylko w przypadku wykrycia zmian	177
Zatrzymywanie budowania indeksu	179
Opcje wpływające na podlegający wyszukiwaniu widok indeksu	179
Indeksowane opcje wyszukiwania dokumentów	179
Wykrywanie duplikatów dokumentów	184
Znaki wieloznaczne w zapytaniach	184
Zasięgi	188
Konfiguracja zasięgów	189
Zwinięte identyfikatory URI	190
Zwijanie identyfikatorów URI w wynikach wyszukiwania	191
Usuwanie z indeksu identyfikatorów URI	191

Administrowanie serwerem wyszukiwania 193

Pamięci podręczne wyszukiwania	193
Konfiguracja pamięci podręcznej wyszukiwania	194
Niestandardowe słowniki synonimów	194
Dodawanie słowników synonimów do systemu	196
Powiązanie słownika synonimów z kolekcją	196
Niestandardowe słowniki słów ignorowanych	197
Dodawanie do systemu słowników słów ignorowanych	198
Powiązanie słownika słów ignorowanych z kolekcją	199
Ponowne wdrażanie słowników niestandardowych	199
Podsumowywanie dynamiczne	200
Dostosowywanie podsumowań dokumentów w konsoli administracyjnej	201
Dostosowywanie podsumowań dokumentów za pomocą edycji właściwości	201
Praca z szybkimi odsyłaczami	202
Szybkie odsyłacze	202
Konfiguracja szybkich odsyłaczy	203

Ranking dokumentów 205

Wartościowanie danych tekstowych	205
Ocena statyczna	206
Odtwarzanie domyślnych wartości w rankingu dokumentów statycznych	207
Niestandardowe słowniki kryteriów	209
Dodawanie do systemu słowników kryteriów	210
Powiązanie słownika kryteriów z kolekcją	211
Ocena dokumentów na podstawie wzorców URI	211
Wpływanie na wyniki dokumentów zgodnych z wzorcami URI	212
Ocena dokumentów na podstawie klas kryteriów	213
Odwzorowywanie pól na klasy kryteriów	214
Konfigurowanie współczynników kryteriów dla klas kryteriów	215
Domyślne wartości klas kryteriów.	216

Aplikacje wyszukujące do wyszukiwania korporacyjnego 219

Przypisywanie aplikacji wyszukujących do kolekcji	220
Funkcje przykładowej aplikacji wyszukującej	220
Właściwości aplikacji wyszukującej	222
Edycja właściwości przykładowej aplikacji wyszukującej	238
Konfigurowanie aplikacji wyszukujących	239
Klonowanie przykładowej aplikacji wyszukiwania	241
Analizowanie najczęstszych wyników	243
Dostęp do aplikacji wyszukujących	246
Konfigurowanie serwerów wyszukiwania w celu zaakceptowania tylko bezpiecznych żądań (SSL)	247
Konfigurowanie serwerów wyszukiwania w celu zaakceptowania żądań z serwerów proxy	248

Obsługa źródeł zewnętrznych 251

Dodawanie do systemu źródeł zewnętrznych	251
Powiązanie aplikacji wyszukujących ze źródłami zewnętrznymi.	254

Zabezpieczenia wyszukiwania korporacyjnego 255

Bezpieczeństwo instalacji	256
Uwierzytelnianie a kontrola dostępu	257
Role administracyjne	258
Konfiguracja użytkowników z uprawnieniami administracyjnymi	259
Zabezpieczenia na poziomie kolekcji	260
Analiza duplikatów dokumentów i bezpieczeństwo kolekcji	260
Identyfikatory aplikacji wyszukiwania	260
Zabezpieczenia na poziomie dokumentu	261
Wstępne i końcowe filtrowanie wyników wyszukiwania.	262
Sprawdzanie poprawności na podstawie zapisanych znaczników bezpieczeństwa	263
Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań	264
Analiza tekstu zakotwiczenia	271
Włączanie zabezpieczeń w wyszukiwaniu korporacyjnym	273
Konfigurowanie zabezpieczeń globalnych oraz rejestru użytkowników protokołu LDAP w serwerze aplikacji WebSphere.	274

Włączanie zabezpieczeń systemu wyszukiwania korporacyjnego z jednym serwerem	276
Włączanie zabezpieczeń systemu wyszukiwania korporacyjnego z wieloma serwerami.	277
Wymagania konfiguracyjne przeszukiwacza w zakresie bezpieczeństwa	278
Sprawdzanie dostępu do zabezpieczonych dokumentów serwera Exchange	281
Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino	282
Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows	285
Wyłączanie zabezpieczeń w wyszukiwaniu korporacyjnym	288
Wyłączanie zabezpieczenia dla aplikacji korporacyjnej na serwerze WebSphere Application Server	288
Wyłączanie zabezpieczeń na poziomie dokumentu	290
Wyłączanie zabezpieczeń w zwiniętych wynikach wyszukiwania.	291

Uruchamianie i zatrzymywanie systemu wyszukiwania korporacyjnego 293

Uruchamianie systemu wyszukiwania korporacyjnego	293
Zatrzymywanie systemu wyszukiwania korporacyjnego	295
Wybieranie komponentów do uruchamiania lub zatrzymywania	297
Administrowanie serwerami wyszukiwania pracującymi w trybie autonomicznym	298

Monitorowanie działań związanych z wyszukiwaniem korporacyjnym. 301

Szacowanie liczby dokumentów w kolekcji	301
Monitorowanie kolekcji.	302
Wyświetlanie szczegółów dotyczących identyfikatora URI.	302
Monitorowanie przeszukiwaczy	304
Wyświetlanie szczegółów na temat działalności przeszukiwacza sieci WWW	305
Szczegółowe informacje o wątkach przeszukiwacza sieci WWW	306
Aktywne serwisy przeszukiwacza sieci WWW	307
Szybkość przeszukiwania przeszukiwacza sieci WWW	307
Tworzenie raportów przeszukiwacza sieci WWW	308
Kody statusu HTTP zwracane do przeszukiwacza sieci WWW	310
Monitorowanie analizatora składni	313
Monitorowanie aktywności indeksu dla kolekcji	315
Monitorowanie kolejki indeksu wyszukiwania korporacyjnego	315
Monitorowanie serwerów wyszukiwania	316
Zmiana sposobu obliczania statystyk zapytań	317
Monitorowanie nasłuchiwanie danych	318
Śledzenie dokumentu	319
Konfigurowanie plików dziennika do śledzenia dokumentów	320
Wyświetlanie raportów o usuniętych dokumentach	320
Wyświetlanie plików dzienników dotyczących usuniętych dokumentów	321

Alerty i pliki dziennika 323

Alerty	323
Konfiguracja alertów na poziomie kolekcji	324
Konfiguracja alertów na poziomie systemu	325
Konfiguracja plików dziennika	326
Konfigurowanie informacji o serwerze SMTP	328
Odbieranie wiadomości poczty elektronicznej dotyczących zarejestrowanych komunikatów	328
Zmianianie wielkości dziennika zapytań	330
Wyświetlanie plików dziennika	331

Tworzenie kopii zapasowej i odtwarzanie systemu wyszukiwania korporacyjnego 333

Tworzenie kopii zapasowej wyszukiwania korporacyjnego	334
Odtwarzanie systemu wyszukiwania korporacyjnego	335
Eksportowanie i importowanie konfiguracji kolekcji	336

Integracja z produktem Lotus Notes wersja 8. 341

Tworzenie serwisu aktualizacyjnego wtyczki wyszukiwania korporacyjnego	341
Instalowanie wtyczki wyszukiwania korporacyjnego w kliencie Lotus Notes, wersja 8	341

Integracja z produktem WebSphere Portal. 343

Skrypty konfiguracyjne do integrowania wyszukiwania korporacyjnego z programem WebSphere Portal	344
Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 5.1	345
Przystosowywanie paska wyszukiwania portalu WebSphere w wersji 5.1 do wykorzystania wyszukiwania korporacyjnego	348
Usuwanie wyszukiwania korporacyjnego z portalu WebSphere w wersji 5.1	349
Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 6	350
Konfigurowanie centrum wyszukiwania portalu WebSphere w wersji 6 do współpracy z wyszukiwaniem korporacyjnym	353

Konfigurowanie paska wyszukiwania portalu WebSphere w wersji 6 do korzystania z wyszukiwania korporacyjnego	354
Konfigurowanie portletu wyszukiwania korporacyjnego dla produktu Lotus Quickr	356
Usuwanie wyszukiwania korporacyjnego z portalu WebSphere w wersji 6	357
Integracja wyszukiwania korporacyjnego z systemami klastrowymi WebSphere Portal	358
Konfigurowanie wyszukiwania korporacyjnego w systemie klastrowym WebSphere Portal	358
Usuwanie wyszukiwania korporacyjnego z systemu klastrowego WebSphere Portal.	362

Migracja z portalu WebSphere Portal do wyszukiwania korporacyjnego. 365

Migrowanie kolekcji z portalu WebSphere Portal	365
Ustawienia zmigrowanej kolekcji	367
Plik dziennika Kreatora migracji	368

Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego 369

Rozróżnianie wielkości liter w wyszukiwaniu korporacyjnym 399

Dokumentacja wyszukiwania korporacyjnego 401

Ułatwienia dostępu 403

Glosariusz terminów związanych z wyszukiwaniem korporacyjnym. 405

Uwagi i znaki towarowe 417

Uwagi	417
Znaki towarowe	419

Indeks 421

Serwis [ibm.com](http://www.ibm.com) i zasoby pokrewne

Pomoc techniczna i dokumentacja produktu są dostępne za pośrednictwem serwisu [ibm.com](http://www.ibm.com).

Wsparcie i asysta

Wsparcie dla produktu jest dostępne za pośrednictwem sieci WWW.

IBM OmniFind Enterprise Edition

<http://www.ibm.com/software/data/enterprise-search/omnifind-enterprise/support.html>

IBM OmniFind Discovery Edition

<http://www.ibm.com/software/data/enterprise-search/omnifind-discovery/support.html>

IBM OmniFind Yahoo! Edition

<http://www.ibm.com/software/data/enterprise-search/omnifind-yahoo/support.html>

Centrum informacyjne

Dokumentację produktu można przeglądać opartym na technologii Eclipse centrum informacyjnym, korzystając z przeglądarki WWW. Centrum informacyjne dostępne jest pod adresem <http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/>.

Publikacje w formacie PDF

Zawartość plików PDF można przeglądać w trybie z połączeniem za pomocą aplikacji Adobe Acrobat Reader dla danego systemu operacyjnego. Można ją pobrać z serwisu WWW firmy Adobe dostępnego pod adresem <http://www.adobe.com>.

Należy zapoznać się z następującymi serwisami WWW z publikacjami w formacie PDF:

Produkt	Adres serwisu WWW
OmniFind Enterprise Edition, wersja 8.5	http://www.ibm.com/support/docview.wss?rs=63&uid=swg27010938
OmniFind Discovery Edition, wersja 8.4	http://www.ibm.com/support/docview.wss?rs=3035&uid=swg27008552
OmniFind Yahoo! Edition, wersja 8.4	http://www.ibm.com/support/docview.wss?rs=3193&uid=swg27008932

Jak przesyłać swoje uwagi

Opinie użytkowników są dla nas ważne i pomagają nam dostarczać najdokładniejsze informacje o najwyższej jakości.

Można je do nas przesyłać za pomocą elektronicznego formularza na uwagi czytelników dostępnego pod adresem https://www14.software.ibm.com/webapp/iwm/web/signup.do?lang=en_US&source=swg-rcf.

Kontakt z firmą IBM

Aby skontaktować się z działem obsługi klientów firmy IBM w Stanach Zjednoczonych lub Kanadzie, należy zadzwonić pod numer 1-800-IBM-SERV (1-800-426-7378).

Aby uzyskać informacje o dostępnych opcjach serwisowych, należy zadzwonić pod jeden z poniższych numerów:

- W Stanach Zjednoczonych: 1-888-426-4343
- W Kanadzie: 1-800-465-9600

Więcej informacji na temat kontaktowania się z firmą IBM zawiera serwis WWW Contact IBM pod adresem <http://www.ibm.com/contact/us/>.

Co to jest wyszukiwanie korporacyjne?

System wyszukiwania korporacyjnego dostarcza wiele możliwości przeszukiwania dowolnej liczby źródeł danych o określonej strukturze lub bez określonej struktury przy użyciu pojedynczego zapytania. Krótkie czasy odpowiedzi na zapytania oraz skonsolidowane, ocenione zestawy wynikowe, które są oparte na rozległej analizie tekstu, umożliwiają nie tylko znajdowanie właściwych dokumentów, lecz również wydobywanie znaczeń z treści dokumentów.

Komponenty wyszukiwania korporacyjnego, które instalowane są razem z programem IBM OmniFind Enterprise Edition, gromadzą informacje z całego przedsiębiorstwa. Wprowadzając odpowiednie zapytanie w przeglądarce sieci WWW, można przeszukać jednocześnie lokalne i zdalne bazy danych, systemy wspomagające pracę grupową, systemy zarządzania treścią, a także wewnętrzne i zewnętrzne serwisy WWW.

System wyszukiwania korporacyjnego bezproblemowo integruje się z istniejącymi systemami i zapewnia możliwość gromadzenia danych z różnych źródeł oraz indeksowania ich w celu przyspieszenia pobierania. Dzięki zastosowaniu w wyszukiwaniu korporacyjnym analizy lingwistycznej i innych typów analizy, otrzymywane wyniki wyszukiwania są niezwykle trafne. Nie ma potrzeby uczenia się obsługi różnych interfejsów w celu wyszukania różnych typów repozytoriów.

Istnieje możliwość dodania obsługi wyszukiwania w źródłach danych, których uwzględnienie w indeksie wyszukiwania korporacyjnego jest niepożądane. Przy użyciu funkcji wyszukiwania stowarzyszonego dla wyszukiwania korporacyjnego w tych źródłach zewnętrznych można wyszukiwać dane w tym samym czasie, co w poindeksowanych źródłach danych.

Jakość wyszukiwania

Aby zapewnić jak największą skuteczność wyszukiwania przez użytkowników interesujących ich informacją, w produkcie OmniFind Enterprise Edition udostępniono obsługę architektury IBM Unstructured Information Management Architecture (UIMA). Architektura UIMA to otwarte środowisko definiujące wspólny, standardowy interfejs przeznaczony do analizy tekstu. Dzięki rozległej analizie semantycznej proces wyszukiwania korporacyjnego może identyfikować pojęcia, ukryte znaczenia, związki, fakty i inne pokrewne dane, często ukryte w nieustrukturyzowanym tekście. Informacje wyodrębnione podczas analizy mogą zostać użyte do poprawy jakości wyników wyszukiwania lub do poprawy jakości działania innych aplikacji, takich jak aplikacje inteligentnej analizy danych lub eksploracji danych.

Bezpieczeństwo

Bezpieczeństwo to integralny element systemu wyszukiwania korporacyjnego. Tylko użytkownicy posiadający odpowiednie uprawnienia administracyjne mogą administrować systemem. Za pomocą mechanizmów bezpieczeństwa dostępnych w programie IBM WebSphere Application Server możliwe jest skonfigurowanie ról administratora i uwierzytelnianie użytkowników z uprawnieniami administracyjnymi.

Można także określić opcje przypisania tokenów zabezpieczeń do danych podczas gromadzenia danych. Po włączeniu zabezpieczeń w aplikacjach wyszukujących tokeny te - przechowywane z dokumentami w indeksie - mogą służyć do wymuszania kontroli dostępu i zapewnienia, że tylko użytkownicy z odpowiednimi referencjami mogą wprowadzać zapytania o dane i przeglądać wyniki wyszukiwania.

Dla określonych typów źródeł danych można skonfigurować opcje sprawdzania poprawności referencji logowania użytkownika z bieżącymi prawami dostępu w czasie przetwarzania zapytania. Ta dodatkowa warstwa bezpieczeństwa zapewnia, że poprawność uprawnień użytkownika jest sprawdzana w czasie rzeczywistym z rodzimym źródłem danych. Ta opcja może chronić przed przypadkami, w których referencje użytkownika zmieniły się od czasu poindeksowania dokumentu i jego znaczników bezpieczeństwa.

Kurs produktu

Kurs dostępny jest pod adresem <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html>. Kurs zawiera opis czynności wykonywanych podczas instalacji i konfiguracji, pokazuje w jaki sposób należy przeszukiwać różne typy źródeł danych, a także przedstawia sposoby rozszerzenia możliwości wyszukiwania korporacyjnego za pomocą aplikacyjnych interfejsów programistycznych produktu. Kurs odwołuje się do starszej wersji produktu OmniFind Enterprise Edition, ale większość pojęć i procedur jest nadal aktualna.

Pojęcia pokrewne

“Przeгляд komponentów wyszukiwania korporacyjnego” na stronie 3

“Przepływ danych w wyszukiwaniu korporacyjnym” na stronie 12

Zabezpieczenia wyszukiwania korporacyjnego



Integracja niestandardowej analizy tekstu



Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu

Odsyłacze pokrewne

“Typy źródeł danych obsługiwane w wyszukiwaniu korporacyjnym”

Typy źródeł danych obsługiwane w wyszukiwaniu korporacyjnym

Wyszukiwanie korporacyjne zawiera predefiniowaną obsługę wyszukiwania różnych typów źródeł danych.

Po zainstalowaniu produktu IBM OmniFind Enterprise Edition można rozpocząć zbieranie danych z następujących typów źródeł danych:

- różne typy elementów programu IBM DB2 Content Manager (dokumenty, zasoby i elementy)
- bazy danych produktu IBM DB2
- bazy danych produktu IBM Domino Document Manager (poprzednio zwanego Domino.Doc)
- bazy danych produktu IBM Lotus Notes
- bazy danych produktu IBM Lotus QuickPlace
- biblioteki treści IBM Lotus Quickr
- repozytoria produktu IBM WebSphere Information Integrator Content Edition, w tym Documentum, FileNet Panagon Content Services, FileNet P8 Content Manager, Hummingbird Document Management (DM), Microsoft SharePoint, OpenText Livelink Enterprise Server i WebSphere Portal Document Manager (PDM)
- tabele pseudonimów produktu IBM WebSphere Information Integrator dla wielu typów systemów baz danych, w tym dla baz danych IBM DB2 for z/OS, IBM Informix, Microsoft SQL Server, Oracle i Sybase
- serwisy produktu IBM WebSphere Portal
- serwisy produktu IBM Workplace Web Content Management
- foldery publiczne produktu Microsoft Exchange Server

- bazy danych Microsoft SQL Server
- systemy plików Microsoft Windows
- grupy dyskusyjne protokołu NNTP (Network News Transfer Protocol)
- bazy danych Oracle
- systemy plików UNIX
- serwisy WWW w Internecie lub w intranecie

Możliwe jest również dodanie obsługi wyszukiwania w następujących typach źródeł zewnętrznych (bez dodawania dokumentów z tych źródeł do indeksu wyszukiwania korporacyjnego):

- Bazy danych obsługujące protokół JDBC (Java Database Connectivity) (tylko systemy baz danych DB2 i Oracle). Dla każdej tabeli udostępnionej do wyszukiwania tworzone jest osobne źródło zewnętrzne.
- Serwery LDAP (Lightweight Directory Access Protocol). Dla każdego serwera LDAP tworzone jest jedno źródło zewnętrzne.

Najnowsze informacje na temat obsługiwanych typów źródeł danych i obsługiwanych wersji produktów można znaleźć na stronie wymagań systemowych w serwisie WWW OmniFind Enterprise Edition Support.

Pojęcia pokrewne

Obsługa źródeł zewnętrznych

“Przeływ danych w wyszukiwaniu korporacyjnym” na stronie 12

Przegląd komponentów wyszukiwania korporacyjnego

Komponenty wyszukiwania korporacyjnego zbierają dane z całego przedsiębiorstwa, analizują i klasyfikują informacje, a także tworzą indeks, który może być przeszukiwany przez użytkowników.

Kolekcja wyszukiwania korporacyjnego reprezentuje zestaw źródeł, w których użytkownicy mogą wyszukiwać dane przy użyciu pojedynczych zapytań. Podczas tworzenia kolekcji użytkownik określa źródła, które mają zostać włączone do kolekcji, i konfiguruje opcje określające dostępne dla użytkowników metody wyszukiwania zindeksowanych danych.

Istnieje możliwość utworzenia wielu kolekcji i w każdej kolekcji mogą znajdować się dane pochodzące z różnych źródeł. Na przykład można utworzyć kolekcję, w której znajdują się dokumenty pochodzące z baz danych programów IBM DB2 Universal Database, IBM Lotus Notes i IBM DB2 Content Manager. Gdy użytkownicy przeszukują tę kolekcję, w wynikach wyszukiwania potencjalnie uwzględnione są dokumenty występujące w każdym z tych źródeł danych.

Obsługa stowarzyszonego wyszukiwania umożliwia użytkownikom wyszukiwanie w wielu kolekcjach za pomocą jednego zapytania. Wyniki wyszukiwania mogą zawierać dokumenty ze wszystkich kolekcji i źródeł zewnętrznych w systemie wyszukiwania korporacyjnego.

Tworzenie i administrowanie kolekcją obejmuje następujące działania:

Gromadzenie danych

Komponenty *przeszukiwacza* zbierają dokumenty pochodzące ze źródeł danych ciągle lub zgodnie z harmonogramem określonym przez użytkownika. Regularne przeszukiwanie danych zapewnia użytkownikom dostęp do najnowszych informacji.

Analizowanie danych

Komponenty *analizatora składni* wyodrębniają tekst z dokumentów i wykonują

analizę lingwistyczną, a także inne rodzaje analizy każdego przeszukiwanego przez przeszukiwacz dokumentu. Szczegółowa analiza treści zapewnia lepszą jakość wyników wyszukiwania.

Indeksowanie danych

Komponenty *indeksowania* są regularnie uruchamiane zgodnie z harmonogramem w celu dodania do indeksu informacji o nowych i zmodyfikowanych dokumentach. Komponenty indeksowania wykonują także analizę globalną dokumentów w kolekcji w celu poprawienia jakości wyników wyszukiwania.

Wyszukiwanie danych

Komponenty *wyszukiwania* przeszukują indeks i współpracują z używanymi aplikacjami wyszukiwującymi w celu przetworzenia żądań wyszukiwania i zwrócenia wyników wyszukiwania.

Inne komponenty produktu OmniFind Enterprise Edition umożliwiają określenie preferencji zabezpieczeń, monitorowanie działania systemu oraz rozwiązywanie występujących problemów. W produkcie tym dostępna jest także przykładowa działająca aplikacja wyszukiwująca, która może posłużyć jako szablon podczas tworzenia nowych aplikacji wyszukiwujących.

Pojęcia pokrewne

Administrowanie przeszukiwaczem

Administrowanie analizatorem składni

Administrowanie indeksem

Administrowanie serwerem wyszukiwania

“Przeływ danych w wyszukiwaniu korporacyjnym” na stronie 12

Przeszukiwacze wyszukiwania korporacyjnego

Przeszukiwacze wyszukiwania korporacyjnego gromadzą dokumenty pochodzące ze źródeł danych, aby umożliwić ich analizowanie, indeksowanie i wyszukiwanie.

W komponencie przeszukiwacza dostarczonym z programem OmniFind Enterprise Edition dostępne są następujące funkcje:

- Podczas konfigurowania przeszukiwacza procesy *wykrywania* znajdują informacje na temat źródeł dostępnych do przeszukiwania, takie jak nazwy wszystkich widoków i folderów w bazie danych Lotus Notes lub nazwy wszystkich systemów plików na serwerze UNIX.
- Po wybraniu źródeł, które mają zostać przeszukane, i uruchomieniu przeszukiwacza, komponenty przeszukiwacza gromadzą dane pochodzące ze źródeł, aby umożliwić ich przeanalizowanie i poindeksowanie.

Z pojedynczą kolekcją może być związanych wiele przeszukiwaczy, z których każdy jest zaprojektowany do zbierania danych z określonego typu źródła danych. Na przykład można utworzyć trzy przeszukiwacze, aby w pojedynczej kolekcji składać dane pochodzące z systemów plików, baz danych Notes i relacyjnych baz danych. Można także utworzyć kilka przeszukiwaczy tego samego typu i zdefiniować dla nich różne harmonogramy przeszukiwania odpowiednio do częstotliwości zmian przeszukiwanych danych.

Przeszukiwacz sieci WWW, gdy zostanie uruchomiony, działa w sposób ciągły. Po określeniu adresów URL, które mają być przeszukiwane, przeszukiwacz będzie okresowo do nich powracać, aby sprawdzić, czy pojawiły się nowe lub zmienione dane. Pozostałe typy przeszukiwaczy można uruchamiać lub zatrzymywać ręcznie albo można dla nich skonfigurować harmonogramy przeszukiwania. Przy definiowaniu harmonogramu

przeszukiwacza należy określić, kiedy ma zostać uruchomiony po raz pierwszy i jak często ma odwiedzać źródła danych w celu przeszukania nowych i zmienionych dokumentów.

Właściwości przeszukiwacza to zestaw reguł zarządzających działaniem określonego przeszukiwacza. Na przykład można zdefiniować reguły określające zasady wykorzystania zasobów systemu przez określony przeszukiwacz. Zestaw źródeł danych dla danego przeszukiwacza określa jego *przestrzeń przeszukiwania*. Po utworzeniu przeszukiwacza w dowolnym momencie można edytować jego właściwości w celu zmodyfikowania wykorzystywanych przez niego metod gromadzenia danych. Istnieje także możliwość zmodyfikowania ustawień przestrzeni przeszukiwania w celu określenia innego harmonogramu przeszukiwacza, dodania nowych źródeł lub usunięcia źródeł, które nie mają już być przeszukiwane.

Pojęcia pokrewne

Administrowanie przeszukiwaczem

Zadania pokrewne

Monitorowanie przeszukiwaczy

Analizatory składni wyszukiwania korporacyjnego

Analizator składni wyszukiwania korporacyjnego analizuje dokumenty zgromadzone przez przeszukiwacz i przygotowuje je do indeksowania.

Dostarczony z oprogramowaniem OmniFind Enterprise Edition komponent analizatora składni analizuje treść i metadane dokumentów. Komponent ten zapisuje wyniki analizy w składnicy danych, do której ma dostęp komponent indeksowania. Analizator składni wykonuje następujące zadania:

- Wyodrębnia tekst z dokumentu bez względu na jego format. Na przykład analizator składni wyodrębnia tekst ze znaczników w dokumentach XML i HTML. Jeśli jest stosowana technologia Oracle (uprzednio Stellent) Outside In Content Access for IBM OmniFind Enterprise Edition, analizator składni może także wyodrębniać tekst z formatów binarnych, takich jak dokumenty programu Microsoft Word dokumenty programu Adobe Acrobat (PDF).
- Wykrywa kodowanie zestawu znaków każdego dokumentu. Przed wykonaniem jakiegokolwiek analizy lingwistycznej analizator składni wykorzystuje te informacje podczas konwersji całego tekstu na kod Unicode.
- Wykrywa język źródłowy każdego dokumentu.
- Wyodrębnia tekst i dodaje leksemy w celu rozszerzenia możliwości pobierania danych. W tej fazie analizator składni wykonuje następujące zadania:
 - Normalizacja znaków (na przykład normalizacja wielkości liter i znaków diakrytycznych, takich jak niemiecki symbol umlaut).
 - Analiza struktury paragrafów, zdań, słów i znaków spacji, tabulacji lub nowego wiersza. Przy użyciu metod analizy lingwistycznej analizator składni rozkłada słowa złożone i przypisuje leksemy umożliwiające wyszukiwanie informacji w słownikach i wyszukiwanie synonimów.
- Stosuje określone dla kolekcji reguły analizowania składni. Konfigurując analizator składni, można skonfigurować następujące działania wykonywane w ramach analizy:

Reguły odwzorowywania pól dla dokumentów XML i HTML

Ta opcja umożliwia użytkownikom wyszukiwanie w dokumentach XML i HTML treści o określonej strukturze lub bez określonej struktury. Po odwzorowaniu elementów XML lub metadanych HTML na pola wyszukiwania w indeksie wyszukiwania korporacyjnego można podawać nazwy pól w zapytaniach i wyszukiwać dane w określonych częściach dokumentów XML i HTML. Za

pomocą zapytań, które wyszukują dane w konkretnych polach, można uzyskać dokładniejsze wyniki niż za pomocą zapytań pełnotekstowych, które wyszukują dane w całej treści dokumentów.

Kategorie

Ta opcja umożliwia użytkownikom wyszukiwanie w dokumentach należących do określonej kategorii. Użytkownicy mogą także wybrać kategorie w wynikach wyszukiwania i przeglądać tylko dokumenty należące do danej kategorii.

Tworząc kolekcję, można wybrać, czy ma być stosowane klasyfikowanie według kategorii. W przypadku zastosowania kategorii *opartych na regułach* dokumenty są związane z kategoriami zgodnie ze zdefiniowanymi regułami. Kategorie oparte na regułach mogą zostać skonfigurowane dla utworzonych przez użytkownika kolekcji wyszukiwania korporacyjnego i dla kolekcji zmigrowanych z portalu IBM WebSphere Portal.

Niestandardowa analiza tekstu

Twórcy aplikacji mogą tworzyć niestandardowe programy do złożonej analizy lingwistycznej danych do wyszukiwania. Programy takie można podłączyć do systemu wyszukiwania korporacyjnego i używać ich do adnotowania treści w kolekcjach. Poindeksowanie adnotacji umożliwia prowadzenie w kolekcjach wyszukiwania semantycznego.

Użytkownicy mogą na przykład szukać użytych w zapytaniu terminów, które występują blisko siebie lub które występują w tym samym zdaniu, mogą też szukać związków między terminami użytymi w zapytaniu. Na przykład użytkownikowi mogą być potrzebne dokumenty dotyczące osoby o nazwisku Kowalski będącej w firmie IBM sprzedawcą, a nie inżynierem.

Obsługa podziału na ciągi n-znakowe

Aby zwiększyć wydajność pobierania dokumentów w języku chińskim, japońskim i koreańskim, można włączyć podział na ciągi n-znakowe w analizie leksykalnej. W tej metodzie analizy spacje nie są separatorami wyrazów. Po utworzeniu kolekcji nie można już zmienić metody podziału.

Obsługa wyszukiwania w dokumentach XML za pomocą rodzimego wyszukiwania

XML Rodzime wyszukiwanie XML może zapewnić bardziej precyzyjne wyniki wyszukiwania dzięki wyszukiwaniu znaczników XML. Na przykład w zapytaniu można podać wyraz, który musi wystąpić w określonym elemencie XML.

Klasy kryteriów zmieniające względne wyniki ważności pól

Odwzorowując pola na klasy kryteriów można wpływać na sposób oceny dokumentów w wynikach wyszukiwania. Na przykład można uwydatnić wynik pól tytułowych, aby mieć pewność, że dokumenty, w tytule których występuje termin z zapytania, zostaną ocenione wyżej w wynikach wyszukiwania.


Pojęcia pokrewne

Praca z kategoriami

Pola wyszukiwania XML

Pola wyszukiwania HTML

 Integracja niestandardowej analizy tekstu

 Uwzględnianie analizy tekstu w wyszukiwaniu korporacyjnym

Zadania pokrewne

Monitorowanie analizatora składni

Indeksy wyszukiwania korporacyjnego

Komponenty indeksowania wyszukiwania korporacyjnego są regularnie uruchamiane zgodnie z harmonogramami w celu dodawania do indeksu informacji o nowych i zmodyfikowanych dokumentach.

Aby zapewnić użytkownikom dostęp do najnowszych informacji w przeszukiwanych źródłach danych, należy wykonać następujące dwa etapy procesu tworzenia indeksu:

Budowanie indeksu głównego

Podczas budowania indeksu głównego cały indeks jest przebudowywany, w wyniku czego jego struktura zostaje optymalnie zorganizowana. Procesy indeksowania odczytują wszystkie dane zebrane przez przeszukiwacze i przeanalizowane przez analizator składni.

Budowanie indeksów przyrostowych

Podczas budowania indeksu przyrostowego do indeksu zostają dodane informacje, które zostały przeszukane od czasu ostatniego budowania indeksu głównego.

Konfigurując opcje indeksu dla kolekcji, można określić harmonogramy budowania indeksu głównego i indeksów przyrostowych. Częstotliwość budowania indeksu zależy od zasobów systemowych oraz od tego, czy indeksowane źródła zawierają treść statyczną czy też dynamiczną.

Aby zapewnić dostępność nowych informacji, w harmonogramie należy skonfigurować częste budowanie indeksów przyrostowych. Okresowo należy zaplanować budowanie indeksu głównego, co pozwoli skonsolidować wszystkie nowe informacje, przeanalizować nową treść i zoptymalizować wydajność indeksu.

Procesy indeksowania mogą zostać uruchomione także bez użycia harmonogramów. Na przykład jeśli zmienione zostaną niektóre reguły analizowania, a zmiany te mają stać się dostępne dla aplikacji wyszukujących, to po ponownym przeszukaniu danych i ich przeanalizowaniu można samemu uruchomić budowanie indeksu głównego, zamiast czekać na jego planowe uruchomienie.

Kontrolowanie wykorzystania zasobów polega na kontrolowaniu liczby kolekcji, które mogą współużytkować procesy indeksowania i wprowadzać jednocześnie żądania budowania indeksu. Budowanie indeksów w sposób współbieżny sprawia, że budowanie dużego indeksu głównego nie musi blokować budowania indeksów przyrostowych dla innych kolekcji. Budowanie indeksu bywa procesem intensywnie korzystającym z zasobów, więc w przypadku dużych systemów należy monitorować ich obciążenie i dostosowywać do niego harmonogram budowania indeksów głównych i przyrostowych.

Podczas tworzenia indeksu w procesach indeksowania wykonywana jest globalna analiza dokumentów. W tej fazie używane są algorytmy identyfikowania duplikatów dokumentów do przeanalizowania struktury odsyłaczy dokumentów i wykonania specjalnego przetwarzania występującego w dokumentach WWW tekstu zakotwiczenia (tekstu opisującego docelową stronę w odsyłaczu hipertekstowym).

Można także określić opcje dla następujących operacji indeksowania:

- Aby umożliwić użytkownikom stosowanie znaków zastępczych, można wbudować w indeks obsługę rozwijania terminów użytych w zapytaniu lub określić, że terminy użyte w zapytaniu mają być rozwijane podczas przetwarzania zapytania. Taka decyzja wymaga kompromisu między wykorzystaniem zasobów a czasem odpowiedzi na zapytanie.
- Istnieje także możliwość skonfigurowania zasięgów. Zdefiniowanie *zasięgu* umożliwia ograniczenie dostępnych w określonej kolekcji danych, które mogą zostać wyświetlone przez użytkowników. Na przykład może zostać utworzony jeden zasięg obejmujący

identyfikatory URI dokumentów działu wsparcia technicznego w przedsiębiorstwie i drugi zasięg obejmujący identyfikatory URI dokumentów działu kadr. Jeśli używana aplikacja wyszukująca obsługuje zasięgi, użytkownicy mogą wyszukiwać i pobierać tylko dokumenty należące do określonych podzbiorów kolekcji.

- Istnieje możliwość określenia opcji zwijania dokumentów mających ten sam przedrostek URI w wynikach wyszukiwania. Ponadto można określić nazwę grupy, dzięki czemu dokumenty o różnych przedrostkach URI mogą być zwijane razem w wynikach wyszukiwania.
- Po utworzeniu indeksu można usunąć identyfikatory URI, aby zapobiec ich wyszukiwaniu.

Pojęcia pokrewne

Administrowanie indeksem

Znaki zastępcze w zapytaniach

Zasięgi

Zwinięte identyfikatory URI

Ocena dokumentów oparta na wzorcach identyfikatorów URI

Zadania pokrewne

Planowanie operacji budowania indeksu

Konfiguracja współbieżnych operacji budowania indeksu

Usuwanie z indeksu identyfikatorów URI

Monitorowanie aktywności indeksu dla kolekcji

Monitorowanie kolejki indeksu wyszukiwania korporacyjnego

Serwery wyszukiwania dla wyszukiwania korporacyjnego

Serwery wyszukiwania dla wyszukiwania korporacyjnego współpracują z aplikacjami wyszukującymi w celu przetwarzania zapytań, wyszukiwania danych w indeksie i zwracania wyników wyszukiwania.

Serwery wyszukiwania dla wyszukiwania korporacyjnego są instalowane podczas instalacji produktu OmniFind Enterprise Edition. Podczas konfigurowania serwerów wyszukiwania dla kolekcji można określić opcje dotyczące metod przeszukiwania kolekcji:

- Można skonfigurować pamięć podręczną wyszukiwania, tak aby były w niej przechowywane często wywoływane wyniki wyszukiwania. Użycie pamięci podręcznej może poprawić wydajność operacji wyszukiwania i pobierania.
- Istnieje możliwość wybrania domyślnego języka dla operacji wyszukiwania dokumentów w określonej kolekcji.
- Jeśli twórcy aplikacji utworzyli niestandardowe słowniki, można powiązać je z kolekcjami:
 - Podczas wprowadzania przez użytkowników zapytań dla kolekcji używającej *słownika synonimów*, w wynikach wyszukiwania uwzględniane są dokumenty zawierające synonimy terminów z zapytania.
 - Podczas wprowadzania przez użytkowników zapytań dla kolekcji używającej *słownika słów ignorowanych*, słowa ignorowane są usuwane z zapytania przed jego przetworzeniem.
 - Podczas wprowadzania przez użytkowników zapytań dla kolekcji używającej *słownika kryteriów*, ważność dokumentów zawierających słowa-kryteria jest zwiększana lub zmniejszana, w zależności do współczynnika kryterium powiązanego z danym słowem w słowniku.
- Jeśli zostanie ustawiona opcja określająca, że specyficzne dokumenty są istotne dla specyficznych zapytań, wówczas możliwe jest skonfigurowanie szybkich odsyłaczy. *Szybki odsyłacz* przypisuje konkretny identyfikator URI do określonych słów kluczowych i fraz.

Jeśli w zapytaniu występuje dowolne słowo kluczowe lub fraza skojarzona z definicją szybkiego odsyłacza, w wynikach wyszukiwania automatycznie zwracany jest określony w tej definicji identyfikator URI.

W konfiguracji wieloserwerowej dostępna jest funkcja zabezpieczająca przed awariami nie tylko na poziomie serwera, ale także na poziomie kolekcji. Jeśli z jakiegoś powodu kolekcja na jednym serwerze wyszukiwania jest niedostępna, zapytania dotyczące tej kolekcji są automatycznie kierowane do innego serwera wyszukiwania.

Pojęcia pokrewne

Aplikacje wyszukujące do wyszukiwania korporacyjnego

Pamięci podręczne wyszukiwania

Niestandardowe słowniki synonimów

Niestandardowe słowniki słów ignorowanych

Niestandardowe słowniki kryteriów

Szybkie odsyłacze

Zadania pokrewne

Monitorowanie serwerów wyszukiwania

Konsola administracyjna wyszukiwania korporacyjnego

Konsola administracyjna wyszukiwania korporacyjnego działa w przeglądarce. Oznacza to, że użytkownicy z uprawnieniami administracyjnymi mogą uzyskać do niej dostęp z dowolnego miejsca w dowolnym czasie. Mechanizmy bezpieczeństwa zapewniają, że dostęp do konsoli mają tylko użytkownicy z uprawnieniami dostępu do funkcji administracyjnych.

Konsola administracyjna wyszukiwania korporacyjnego jest instalowana na serwerach wyszukiwania podczas instalowania produktu OmniFind Enterprise Edition.

Konsola administracyjna udostępnia kreatorów ułatwiających wykonanie kilku podstawowych zadań administracyjnych. Na przykład Kreator kolekcji ułatwia tworzenie kolekcji i pozwala na zapisywanie wyników pracy w trybie roboczym. Kreatory przeszukiwaczy odpowiadają poszczególnym typom źródeł danych i pomagają w wyborze źródeł, które mają zostać udostępnione do wyszukiwania użytkownikom.

W przypadku pozostałych zadań administracyjnych użytkownik może wybrać poszczególne elementy, których ma dotyczyć czynność administracyjna. Na przykład podczas edytowania kolekcji można przejść do strony Indeks, aby zmienić harmonogram operacji na indeksie, albo do strony Analiza składni, aby zmodyfikować regułę analizowania składni dokumentów XML.

Pojęcia pokrewne

Administrowanie systemem wyszukiwania korporacyjnego

Role administracyjne

Zadania pokrewne

Logowanie się do konsoli administracyjnej

Monitorowanie systemu wyszukiwania korporacyjnego

Konsola administracyjna wyszukiwania korporacyjnego może być używana do monitorowania działań systemu i odpowiedniego dostosowania ustawień dotyczących operacji.

Po zainstalowaniu programu OmniFind Enterprise Edition i utworzeniu co najmniej jednej kolekcji użytkownik może wyświetlić szczegółowe informacje statystyczne dotyczące

najważniejszych działań (przeszukiwanie, analizowanie składni, indeksowanie i wyszukiwanie). Są to między innymi informacje o średnim czasie odpowiedzi i postępie, takie jak liczba dokumentów przeszukanych lub poindeksowanych podczas określonej sesji przeszukiwania lub budowania indeksu.

Istnieje możliwość zatrzymania i uruchomienia większości działań. Na przykład możliwe jest wstrzymanie operacji, zmodyfikowanie związanych z nią ustawień konfiguracyjnych lub rozwiązywanie zaistniałych problemów, a następnie zrestartowanie procesu przetwarzania w odpowiednim momencie, gdy możliwe będzie kontynuowanie operacji.

Możliwe jest również skonfigurowanie alertów, dzięki czemu będzie można otrzymywać wiadomości e-mail o niektórych podejmowanych działaniach, kiedy tylko wystąpi monitorowane zdarzenie. Na przykład alert może zostać wysłany w przypadku przekroczenia określonego czasu odpowiedzi wyszukiwania.

Jeśli dokument został usunięty z systemu wyszukiwania korporacyjnego, istnieje możliwość prześledzenia kiedy, gdzie i dlaczego został on usunięty. Na przykład mogło okazać się niemożliwe przeanalizowanie dokumentu przez analizator składni lub administrator mógł usunąć dokument z indeksu.

Pojęcia pokrewne

Monitorowanie działań związanych z wyszukiwaniem korporacyjnym

Uruchamianie i zatrzymywanie systemu wyszukiwania korporacyjnego

Pliki dzienników wyszukiwania korporacyjnego

Pliki dzienników tworzone są dla poszczególnych kolekcji i dla sesji na poziomie systemu.

Podczas konfigurowania opcji dotyczących rejestrowania dla kolekcji wyszukiwania korporacyjnego lub dla systemu należy określić typy komunikatów, które mają być rejestrowane, takie jak komunikaty o błędach lub ostrzeżenia. Należy także określić częstotliwość usuwania starszych plików dzienników w celu zwolnienia miejsca dla najnowszych komunikatów. Użytkownik może otrzymywać wiadomości e-mail informujące o wystąpieniu konkretnych komunikatów (w tym alertów) lub wszystkich komunikatów o błędach.

Podczas przeglądania plików dzienników należy wybrać plik dziennika, który ma zostać wyświetlony. W nazwie pliku dostępne są informacje o dacie i godzinie utworzenia określonego pliku oraz informacje o komponencie, który wygenerował określone komunikaty. Istnieje także możliwość określenia filtrów wyświetlania. Na przykład można wybrać opcję wyświetlenia tylko komunikatów o błędach lub tylko komunikatów pochodzących z określonej sesji wyszukiwania korporacyjnego.

Pojęcia pokrewne

Alerty i pliki dziennika

Alerty



Komunikaty dla wyszukiwania korporacyjnego

Zadania pokrewne

Konfiguracja plików dziennika

Konfigurowanie informacji o serwerze SMTP

Odbieranie wiadomości poczty elektronicznej dotyczących zarejestrowanych komunikatów

Wyświetlanie plików dziennika

Dostosowywanie wyszukiwania korporacyjnego

Aplikacyjne interfejsy programistyczne dla wyszukiwania korporacyjnego umożliwiają tworzenie niestandardowych aplikacji wyszukiwujących, niestandardowych aplikacji aktualizujących treść kolekcji, niestandardowych programów do analizy tekstu, a także niestandardowych słowników synonimów, słów ignorowanych i słów-kryteriów.

Po zainstalowaniu oprogramowania OmniFind Enterprise Edition dostępne są następujące typy interfejsów API rozszerzających kolekcje wyszukiwania korporacyjnego:

Interfejs API wyszukiwania i indeksowania (SI-API)

Ten interfejs API służy do budowania niestandardowych aplikacji wyszukiwujących oraz niestandardowego interfejsu administracyjnego.

Wtyczki przeszukiwacza

Interfejsy API w postaci wtyczek służą do dodawania metadanych do dokumentów podczas przeszukiwania tych dokumentów lub do tworzenia powiązań znaczników bezpieczeństwa, wymuszających przestrzeganie reguł biznesowych i reguł bezpieczeństwa danej organizacji.

Zintegrowanie niestandardowych programów do analizy lingwistycznej z kolekcjami wyszukiwania korporacyjnymi zapewnia lepszą wydajność pobierania informacji. Po dodaniu do systemu niestandardowych mechanizmów analizy tekstu można powiązać te mechanizmy z kolekcjami. Dzięki temu użytkownicy mogą korzystać podczas wyszukiwania z powiązań wyrazów, które zostały wprowadzone do indeksu przez programy niestandardowe. Na przykład użytkownicy mogą wyszukiwać pojęcia i związki między terminami, a nie tylko same terminy.

Wydajność pobierania informacji można także zwiększyć przez zintegrowanie niestandardowych słowników zawierających na przykład akronimy, skróty oraz terminologię specyficzną dla danej branży. Po dodaniu do słowników systemu można je powiązać z kolekcjami. Zapewnia to użytkownikom następujące korzyści podczas wprowadzania zapytań dla kolekcji:

- Jeśli zapytanie zawiera słowa zdefiniowane jako synonimy, dokumenty zawierające synonimy terminów użytych w zapytaniu zostaną uwzględnione w wynikach wyszukiwania.
- Jeśli zapytanie zawiera słowa ignorowane, zostaną one usunięte z zapytania, tak aby niewłaściwe dokumenty nie zostały zwrócone w wynikach wyszukiwania.
- Jeśli zapytanie zawiera słowa-kryteria, dokumenty zawierające te słowa zostaną ocenione wyżej lub niżej w wynikach wyszukiwania, w zależności od wartości kryterium powiązanej z danym słowem w słowniku.

Pojęcia pokrewne

Aplikacje wyszukiwujące do wyszukiwania korporacyjnego

Niestandardowe słowniki synonimów

Niestandardowe słowniki słów ignorowanych

Niestandardowe słowniki kryteriów



Przegląd interfejsu API wyszukiwania i indeksowania



Wtyczki przeszukiwacza

Przykładowa aplikacja wyszukiwująca dla wyszukiwania korporacyjnego

Przykładowej aplikacji wyszukiwującej dla wyszukiwania korporacyjnego można używać w roli szablonu do opracowywania niestandardowych aplikacji wyszukiwujących.

Przykładowa aplikacja wyszukiwująca jest instalowana podczas instalowania oprogramowania OmniFind Enterprise Edition. Przykładowa aplikacja wyszukiwująca zawiera większość funkcji wyszukiwania i pobierania dostępnych dla wyszukiwania korporacyjnego. Aplikacja ta jest również działającym przykładem umożliwiającym wyszukiwanie we wszystkich aktywnych kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Przy użyciu tej przykładowej aplikacji można testować nowe kolekcje i źródła zewnętrzne przed udostępnieniem ich użytkownikom.

Przykładowa aplikacja wyszukiwująca prezentuje obsługę wyszukiwania stowarzyszonego, umożliwiając jednocześnie wyszukiwanie w jednej lub w wielu kolekcjach i źródłach zewnętrznych.

W przypadku niektórych typów przeszukiwaczy za pomocą komponentu zarządzania tożsamościami dla wyszukiwania korporacyjnego można sprawdzać poprawność bieżących referencji, gdy użytkownicy uzyskują dostęp do aplikacji wyszukiwującej. Jeśli domena, w której ma być przeprowadzone wyszukiwanie, jest chroniona zabezpieczeniami opartymi na pojedynczym logowaniu (SSO), można użyć mechanizmów SSO do sprawdzania poprawności referencji użytkownika podczas całej sesji wyszukiwania. W przeciwnym razie komponent zarządzania tożsamościami może szyfrować i przechowywać referencje użytkownika w profilu i na podstawie tych referencji wykluczać z wyników wyszukiwania dokumenty zabronione.


Do dostosowania przykładowej aplikacji wyszukiwującej można użyć konfiguratora aplikacji wyszukiwującej, który jest graficznym interfejsem użytkownika umożliwiającym sprawdzanie efektów zmian w miarę ich wprowadzania. Aplikację wyszukiwującą można również dostosować, edytując jej plik konfiguracyjny.

Do utworzenia niestandardowej aplikacji wyszukiwującej należy użyć interfejsu API wyszukiwania i indeksowania przeznaczonego dla wyszukiwania korporacyjnego.

Pojęcia pokrewne

Aplikacje wyszukiwujące do wyszukiwania korporacyjnego

Funkcje przykładowej aplikacji wyszukiwującej

 Przegląd interfejsu API wyszukiwania i indeksowania

Zadania pokrewne

Uzyskiwanie dostępu do aplikacji wyszukiwujących

Edycja właściwości przykładowej aplikacji wyszukiwującej

Dostosowywanie aplikacji wyszukiwujących

Przepływ danych w wyszukiwaniu korporacyjnym

Komponenty wyszukiwania korporacyjnego zainstalowane razem z programem OmniFind Enterprise Edition ściśle współdziałają w celu zapewnienia właściwego przepływu danych w systemie.

Przeszukiwacze zbierają dokumenty z różnych źródeł danych w przedsiębiorstwie. Analizator składni wyodrębnia przydatne informacje z przeszukanych dokumentów i generuje leksemy, przy użyciu których możliwe jest na przykład przypisanie dokumentów do kategorii i łatwiejsze określenie ważności dokumentów względem terminów określonych w żądaniu wyszukiwania. Dane te są zapisywane w indeksie w celu poprawy wydajności pobierania.

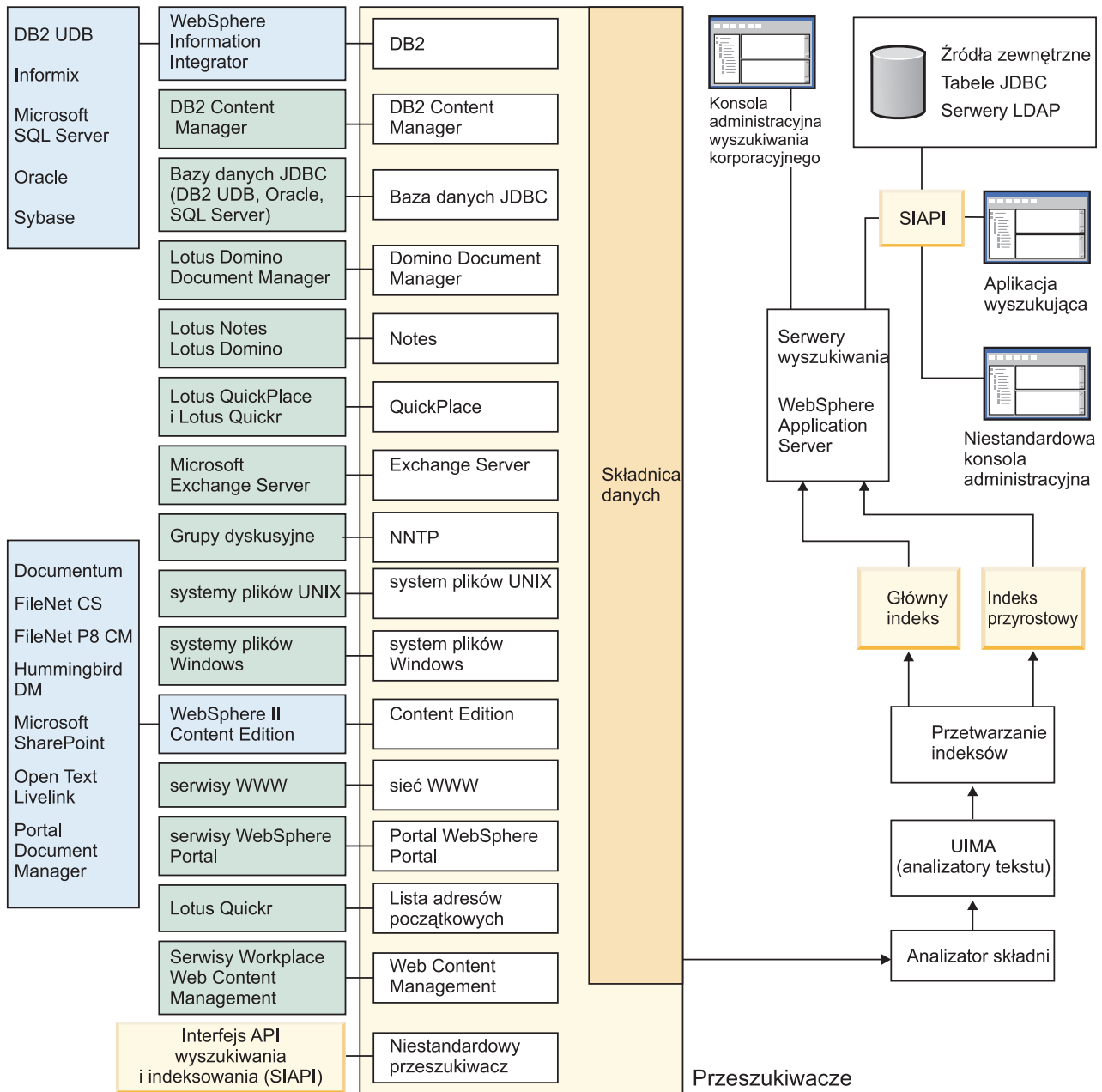
Korzystając z przeglądarki WWW i aplikacji wyszukiwującej, użytkownicy wyszukują dane w poindeksowanych kolekcjach i źródłach zewnętrznych. Aplikacja wyszukiwująca umożliwia wyświetlenie w przeglądarce listy wyników, na której użytkownicy mogą kliknąć

odpowiednią pozycję. Bardziej wyrafinowana aplikacja może zwrócić dynamicznie wygenerowaną treść utworzoną w oparciu o informacje dostępne w innych źródłach danych.

Na przykład w aplikacji wyszukującej dane w katalogu może istnieć możliwość dostosowania widoku produktów spełniających wymagania określone w żądaniu wyszukiwania. Pojedyncze zapytanie może służyć do wyszukiwania wśród dokumentów dostępnych w różnych typach źródeł danych, takich jak dokumenty w repozytoriach IBM DB2 Content Manager oraz Lotus Notes.

Typ zbieranych danych oraz metody przeszukiwania, analizowania, indeksowania i wyszukiwania tych danych określone są przez administratorów. Monitorując działania systemu, administratorzy dopasowują także odpowiednio ustawienia w celu zoptymalizowania przepustowości dla danych.

Na poniższym diagramie pokazano przepływ informacji w systemie wyszukiwania korporacyjnego.



Rysunek 1. Przepływ danych w systemie wyszukiwania korporacyjnego

Pojęcia pokrewne

“Co to jest wyszukiwanie korporacyjne?” na stronie 1

“Przegląd komponentów wyszukiwania korporacyjnego” na stronie 3

Odsyłacze pokrewne

“Typy źródeł danych obsługiwane w wyszukiwaniu korporacyjnym” na stronie 2

Administrowanie systemem wyszukiwania korporacyjnego

Konsola administracyjna wyszukiwania korporacyjnego służy do tworzenia kolekcji oraz źródeł zewnętrznych i administrowania nimi, uruchamiania i zatrzymywania komponentów, monitorowania działań w systemie, monitorowania plików dziennika, konfigurowania ustawień dotyczących użytkowników z uprawnieniami administracyjnymi, tworzenia powiązań aplikacji wyszukujących z kolekcjami i źródłami zewnętrznymi, a także do określania informacji umożliwiających wymuszanie stosowania zabezpieczeń.

Wskazówka: Kurs dostępny jest pod adresem <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html>. Kurs zawiera opis czynności wykonywanych podczas instalacji i konfiguracji, pokazuje w jaki sposób należy przeszukiwać różne typy źródeł danych, a także przedstawia sposoby rozszerzenia możliwości wyszukiwania korporacyjnego za pomocą aplikacyjnych interfejsów programistycznych produktu. Kurs odwołuje się do starszej wersji produktu OmniFind Enterprise Edition, ale większość pojęć i procedur jest nadal aktualna.


Szczegółowe przykłady konfigurowania przeszukiwaczy i włączania obsługi zabezpieczeń w małych, średnich i dużych organizacjach można znaleźć w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

Widok Kolekcje

Poniższe wytyczne mają za zadanie pomóc w utworzeniu pierwszej kolekcji i administrowaniu systemem.

Logowanie

Zaloguj się w konsoli administracyjnej wyszukiwania korporacyjnego. Widok Kolekcje to punkt wejścia dla tworzenia kolekcji i administrowania nimi.

Wskazówka: Aby uzyskać pomoc dotyczącą konsoli administracyjnej, należy w dowolnym momencie kliknąć przycisk **Pomoc** na pasku narzędzi lub przycisk **Pomoc do tej strony**. Jeśli dla komunikatów dostępne są szczegółowe wyjaśnienia i działania naprawcze, można kliknąć ikonę  **Więcej informacji** na końcu komunikatu, aby wyświetlić szczegóły.


Tworzenie kolekcji

W celu utworzenia nowej, pustej kolekcji należy wybrać jedną z następujących metod:

- Aby utworzyć kolekcję przy użyciu Kreatora kolekcji, kliknij opcję **Kreator kolekcji** i wykonaj czynności zgodnie z instrukcjami wyświetlonymi w kreatorze.
- Aby utworzyć kolekcję przy użyciu widoku Kolekcje, kliknij opcję **Utwórz kolekcję**, wypełnij odpowiednie pola na stronie Tworzenie kolekcji, a następnie kliknij przycisk **OK**.

Konfigurowanie kolekcji

Przygotuj także nową, pustą kolekcję, aby dodać do niej zawartość oraz określić opcje przeszukiwania danych i udostępniania danych do wyszukiwania.

Kliknij przycisk  **Edytuj** dla nowej kolekcji, a następnie wybierz stronę, na której zostaną podane opcje kolekcji.

Ważne: Aby przejść do wcześniej wyświetlonej strony lub odświeżyć informacje w konsoli administracyjnej, kliknij przycisk **Poprzednia** i **Odśwież** w konsoli administracyjnej wyszukiwania korporacyjnego. Jeśli zostanie kliknięty przycisk **Wstecz** lub **Odśwież** w przeglądarce WWW, mogą się pojawić niespójne wyniki i może dojść do utraty danych.

- Na stronie Ogólne można określić opcje mające zastosowanie do całej kolekcji:
 - Opcje ogólne można edytować i dzięki temu zmienić nazwę lub opis kolekcji bądź jej szacowaną wielkość.
 - Opcje, których nie można zmienić, można przeglądać. Jest to na przykład identyfikator kolekcji lub metoda oceny statycznej służąca do oceny dokumentów w wynikach wyszukiwania.
 - W przypadku kolekcji, dla których włączono zabezpieczenia w czasie tworzenia, można włączyć lub wyłączyć elementy sterujące zabezpieczeniami na poziomie dokumentu.
- Na stronie Przeszukiwanie należy skonfigurować przynajmniej jeden przeszukiwacz.

W pojedynczej kolekcji mogą znajdować się dane pochodzące z różnych źródeł danych. Dla każdego typu źródła danych, które ma zostać uwzględnione w kolekcji, należy skonfigurować co najmniej jeden przeszukiwacz. Po utworzeniu przeszukiwacza można go skonfigurować, korzystając z kreatora przeznaczonego dla określonego typu przeszukiwanych danych.
- Na stronie Analiza składni można skonfigurować opcje dotyczące sposobu analizowania przeszukiwanych danych, aby zapewnić ich wydajne wyszukiwanie:
 - Określenie, czy dokumenty XML mają być analizowane, aby umożliwić ich wyszukiwanie za pomocą rodzimego wyszukiwania XML.
 - Przypisanie dokumentów do kategorii, dzięki czemu użytkownicy mogą przeszukiwać podzbiór kolekcji lub przeglądać dokumenty będące wynikiem wyszukiwania według kategorii, do których należą.
 - Odwzorowanie elementów XML i elementów metadanych HTML na pola wyszukiwania w indeksie, co umożliwia użytkownikom określenie nazw pól w zapytaniach i wyszukiwanie konkretnych części dokumentów.
 - Jeśli do systemu wyszukiwania korporacyjnego dodano niestandardowy mechanizm analizy tekstu, można go wybrać, aby był używany dla kolekcji, a następnie podać opcje przetwarzania tekstu, aby zwiększyć możliwości pobierania informacji i włączyć obsługę wyszukiwania semantycznego.
 - Powiązanie pól z klasami kryteriów zapewniające wpływ na ocenę pól zgodnych z terminami użytymi w zapytaniu w wynikach wyszukiwania.
- Na stronie Indeks skonfiguruj harmonogramy budowania indeksu. Aby zapewnić użytkownikom dostęp do najnowszych informacji, należy zaplanować regularne wykonywanie operacji tworzenia indeksu. Opcjonalnie można także wykonać następujące czynności:
 - Umożliwienie użytkownikom stosowania znaków wieloznacznych w zapytaniach.
 - Skonfigurowanie zasięgów, aby użytkownicy przeszukiwali tylko ograniczoną część kolekcji zamiast wszystkich dokumentów w indeksie.
 - Zwinięcie wyników wyszukiwania, aby dokumenty z tego samego źródła były zwinięte.
 - Przypisanie współczynników kryteriów w celu oddziaływania na ocenę dokumentów zgodnych ze wzorcem identyfikatora URI w wynikach wyszukiwania.

- Usunięcie identyfikatorów URI z indeksu. Może to mieć na celu uniemożliwienie użytkownikom wyświetlenia określonych dokumentów po utworzeniu kolekcji.
- Na stronie Wyszukiwanie można określić opcje dotyczące wyszukiwania dokumentów w kolekcji:
 - Ustawienie wielkości pamięci podręcznej przeznaczonej dla wyników wyszukiwania i zmiana domyślnego języka kolekcji.
 - Jeśli do systemu wyszukiwania korporacyjnego zostały dodane niestandardowe słowniki synonimów, słów ignorowanych i słów-kryteriów, można wybrać te słowniki do używania podczas przeszukiwania kolekcji przez użytkowników.
 - Określenie długości podsumowania dokumentu wyświetlanego w wynikach wyszukiwania.
 - Skonfigurowanie szybkich odsyłaczy w celu automatycznego wyświetlania konkretnych identyfikatorów URI w wynikach wyszukiwania za każdym razem, gdy w zapytaniu wystąpią określone słowa kluczowe lub wyrażenia.
- Na stronie Rejestrowanie można wykonać następujące czynności:
 - Określenie opcji dotyczących typów komunikatów, które mają być umieszczane w dziennikach, oraz częstotliwości wtórnego przetwarzania plików dzienników.
 - Określenie opcji dotyczących otrzymywania alertów dotyczących działań związanych z kolekcją. Na przykład w alercie może zostać przesłana informacja o przekroczeniu średniego czasu odpowiedzi wyszukiwania.
 - Określenie opcji dotyczących otrzymywania wiadomości e-mail po zapisaniu w dzienniku określonych komunikatów lub komunikatów określonego typu.
 - Określenie opcji dotyczących rejestrowania informacji umożliwiających stwierdzenie kiedy, gdzie i z jakiego powodu dokument został usunięty z systemu wyszukiwania korporacyjnego.

Uruchamianie komponentów

Po określeniu źródeł danych, które mają zostać przeszukane, oraz opcji zbierania i wyszukiwania danych można uruchomić procesy tworzenia kolekcji. Bardzo ważna jest kolejność uruchamiania komponentów. Zanim będzie można analizować i indeksować dane, przeszukiwacze muszą do nich dotrzeć. Z kolei, zanim serwery wyszukiwania będą mogły przetwarzać żądania wyszukiwania, musi być zbudowany indeks główny.

Widok Źródła zewnętrzne

Jeśli wymagana jest możliwość wyszukiwania danych w źródłach bez ich przeszukiwania i indeksowania, należy kliknąć przycisk **Źródła zewnętrzne** na pasku narzędzi w celu określenia opcji, które pozwolą udostępnić wybrane źródło danych na potrzeby wyszukiwania. Konieczne jest określenie informacji umożliwiających dostęp do baz danych JDBC (Java Database Connectivity) oraz serwerów LDAP (Lightweight Directory Access Protocol) w procesie wyszukiwania korporacyjnego. Po powiązaniu źródeł zewnętrznych z aplikacjami wyszukującymi użytkownicy mogą wyszukiwać dane w tych źródłach jednocześnie z wyszukiwaniem prowadzonym w kolekcjach, których dane zostały wcześniej przeszukane, przeanalizowane i poindeksowane.

Widok System

Jeśli użytkownik ma przypisaną rolę administratora wyszukiwania korporacyjnego, może kliknąć ikonę **System** na pasku narzędzi, aby wykonać poniższe czynności. Administratorzy

kolekcji, operatorzy i osoby monitorujące mają dostęp do tego widoku tylko wtedy, gdy administrator wyszukiwania korporacyjnego nada im do tego uprawnienie.

- Dodanie do systemu niestandardowych mechanizmów analizy tekstu.
- Dodanie do systemu niestandardowych słowników synonimów, słów ignorowanych i słów-kryteriów.
- Określenie, ile kolekcji może budować indeksy równoległe oraz czy budowanie głównego indeksu dla jednej kolekcji może odbywać się współbieżnie z budowaniem indeksów różnicowych.
- Skonfigurowanie alertów dla zdarzeń systemowych.
- Określenie opcji rejestrowania komunikatów generowanych przez sesje systemowe.
- Określenie informacji o serwerze poczty w celu otrzymywania wiadomości e-mail dotyczących działań z zakresu wyszukiwania korporacyjnego.

Widok Zabezpieczenia


Jeśli użytkownik ma przypisaną rolę administratora wyszukiwania korporacyjnego, może kliknąć ikonę **Zabezpieczenia** na pasku narzędzi, aby określić opcje zabezpieczeń. Administratorzy kolekcji, operatorzy i monitory nie mogą uzyskać dostępu do tego widoku.

Jeśli na serwerze WebSphere Application Server włączona jest obsługa zabezpieczeń, to przy użyciu widoku Bezpieczeństwo można skonfigurować role administracyjne. Konfigurując role administracyjne, można zezwolić dodatkowym użytkownikom na administrowanie systemem i dla każdego z nich określić pewne ograniczenia dostępu do konkretnych funkcji i kolekcji.

Można także użyć widoku Bezpieczeństwo, aby skonfigurować opcje zarządzania tożsamością. Na przykład można określić opcje przechowywania referencji użytkownika w profilach, które mogą służyć do sprawdzania poprawności bieżących referencji użytkownika podczas przetwarzania zapytania. Jeśli źródło, w którym ma zostać przeprowadzone wyszukiwanie jest chronione zabezpieczeniami opartymi na pojedynczym logowaniu (SSO), można także określić opcje korzystania z metod uwierzytelniania SSO w celu sprawdzenia poprawności bieżących referencji użytkownika podczas przetwarzania zapytania.

Do czasu utworzenia własnych aplikacji wyszukujących można używać przykładowej aplikacji wyszukującej do wyszukiwania danych we wszystkich kolekcjach i źródłach zewnętrznych. Po utworzeniu niestandardowej aplikacji wyszukującej należy w widoku Zabezpieczenia powiązać ją z kolekcjami i źródłami zewnętrznymi, w których może ona prowadzić wyszukiwanie.

Widok Monitor

Można kliknąć ikonę  **Monitor**, aby w dowolnym momencie monitorować komponenty systemu lub kolekcji. Jeśli pozwalają na to ustawienia roli użytkownika, użytkownik może także uruchomić lub zatrzymać procesy komponentów podczas ich monitorowania.

Pojęcia pokrewne

“Administrowanie przeszukiwaczem” na stronie 37

“Monitorowanie działań związanych z wyszukiwaniem korporacyjnym” na stronie 301

Zadania pokrewne

“Uruchamianie systemu wyszukiwania korporacyjnego” na stronie 293

“Administrowanie serwerami wyszukiwania pracującymi w trybie autonomicznym” na stronie 298

“Zatrzymywanie systemu wyszukiwania korporacyjnego” na stronie 295

“Tworzenie kolekcji przy użyciu Kreatora kolekcji” na stronie 31

“Tworzenie kolekcji przy użyciu widoku kolekcji” na stronie 33

Logowanie się do konsoli administracyjnej

Aby administrować systemem wyszukiwania korporacyjnego, należy wprowadzić adres URL w przeglądarce WWW, po czym zalogować się w konsoli administracyjnej.

Zagadnienia wstępne

Zalogowanie się jest możliwe tylko przy użyciu identyfikatora, który ma prawa dostępu do konsoli administracyjnej wyszukiwania administracyjnego:

- Jeśli globalne zabezpieczenia w konfiguracji serwera WebSphere Application Server nie zostaną włączone, dostęp do konsoli administracyjnej będzie miał wyłącznie administrator wyszukiwania korporacyjnego wskazany podczas instalacji programu OmniFind Enterprise Edition.
- Jeśli włączono globalne zabezpieczenia serwera WebSphere Application Server, można użyć konsoli administracyjnej wyszukiwania korporacyjnego do konfigurowania ról administracyjnych. Konfigurowane identyfikatory użytkowników muszą figurować w rejestrze użytkowników serwera WebSphere Application Server. Skonfigurowanie ról administracyjnych umożliwia większej liczbie użytkowników logowanie się w konsoli administracyjnej. Można jednak przy tym zdefiniować funkcje i kolekcje dostępne dla poszczególnych administratorów.

Procedura

Aby zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego:

1. Wpisz adres URL konsoli administracyjnej w przeglądarce WWW. Na przykład:

`http://SerwerWyszukiwania.com/ESAdmin/`

SerwerWyszukiwania.com jest to nazwa hosta serwera wyszukiwania do wyszukiwania korporacyjnego.

W zależności od konfiguracji serwera WWW może być wymagane także podanie numeru portu. Na przykład:

`http://SerwerWyszukiwania.com:9080/ESAdmin/`

2. Na stronie powitania wpisz swój identyfikator użytkownika oraz hasło i kliknij przycisk **Zaloguj się**.

Wyświetlony zostanie widok Kolekcje będący punktem wejścia przy administrowaniu systemem i kolekcjami. W przypadku korzystania z ról administracyjnych zestaw widocznych kolekcji i możliwych do wykonania działań zależy od roli administracyjnej użytkownika.

Jeśli sesja przez pewien okres jest nieaktywna, system automatycznie przeprowadzi wylogowanie. Aby kontynuować administrowanie systemem, należy się ponownie zalogować.

Po skończeniu administrowania kolekcjami można kliknąć przycisk **Wyloguj się**, aby się wylogować z konsoli. Następnie można zalogować się przy użyciu innego identyfikatora użytkownika i hasła albo zamknąć okno przeglądarki WWW z konsolą administracyjną.

Pojęcia pokrewne

“Role administracyjne” na stronie 258

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji jednoserwerowej

Hasło administratora wyszukiwania korporacyjnego jest przechowywane w postaci zaszyfrowanej. Do zmiany hasła służy skrypt **eschangepw**.

Ograniczenia

W hasłach można używać następujących znaków specjalnych:

! @ # \$ % ^ & * () - _ = + , . / < > ?

Jeśli w systemach AIX, Linux i Solaris podane zostanie hasło zawierające znaki specjalne, to całe hasło trzeba umieścić w znakach pojedynczego cudzysłowu.

Na przykład: 'mypwd@\$%'

Jeśli w systemach Windows podane zostanie hasło zawierające znaki specjalne, to całe hasło trzeba umieścić w znakach podwójnego cudzysłowu.

Na przykład: "my?+!pwd"

Informacje o tym zadaniu

Hasło dla pierwszego identyfikatora administratora wyszukiwania korporacyjnego jest określone podczas instalacji produktu OmniFind Enterprise Edition.

Aby zmienić hasło, należy uruchomić skrypt **eschangepw**, który spowoduje wprowadzenie zmiany w całym systemie wyszukiwania korporacyjnego. Program instalacyjny tworzy dwie zmienne środowiskowe, które mogą zostać użyte w skrypcie **eschangepw**:

ES_INSTALL_ROOT

Katalog instalacyjny wyszukiwania korporacyjnego.

ES_NODE_ROOT

Katalog danych wyszukiwania korporacyjnego. Hasło administratora wyszukiwania korporacyjnego jest przechowywane w pliku **es.cfg** w tym katalogu.

Procedura

Aby zmienić hasło administratora wyszukiwania korporacyjnego w konfiguracji jednoserwerowej:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego:
esadmin system stopall

Ważne: W czasie, gdy system jest zatrzymany, użytkownicy nie mogą wprowadzać żądań wyszukiwania.

2. Za pomocą poleceń systemowych (w systemach AIX, Linux lub Solaris) lub za pomocą funkcji zmiany hasła (w systemach Windows) zmień hasło systemu dla identyfikatora użytkownika administratora wyszukiwania korporacyjnego.
3. Uruchom następujący skrypt, gdzie *nowaWartość* to hasło podane w kroku 2:

Systemy AIX, Linux lub Solaris
eschangepw.sh *nowaWartość*

Windows

eschangepw *nowaWartość*

4. Zrestartuj system wyszukiwania korporacyjnego:

esadmin system startall

Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 369

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji wieloserwerowej

Hasło administratora wyszukiwania korporacyjnego jest przechowywane w postaci zaszyfrowanej. Do zmiany hasła służy skrypt **eschangepw** pozwalający wprowadzić zmianę na wszystkich komputerach systemu wyszukiwania korporacyjnego.

Ograniczenia

W hasłach można używać następujących znaków specjalnych:

! @ # \$ % ^ & * () - _ = + , . / < > ?

Jeśli w systemach AIX, Linux i Solaris podane zostanie hasło zawierające znaki specjalne, to całe hasło trzeba umieścić w znakach pojedynczego cudzysłowu.

Na przykład: 'mypwd@\$%'

Jeśli w systemach Windows podane zostanie hasło zawierające znaki specjalne, to całe hasło trzeba umieścić w znakach podwójnego cudzysłowu.

Na przykład: "my?+!pwd"

Informacje o tym zadaniu

Hasło związane z identyfikatorem administratora wyszukiwania korporacyjnego, które podawane jest w czasie instalowania programu OmniFind Enterprise Edition, musi być takie samo na wszystkich serwerach wyszukiwania korporacyjnego.

Aby zmienić hasło i rozprzestrzenić tę zmianę we wszystkich serwerach wyszukiwania korporacyjnego, należy na każdym komputerze wykorzystywanym w systemie wyszukiwania korporacyjnego uruchomić skrypt **eschangepw**. Podana niżej procedura podaje zalecaną kolejność zmiany hasła na wszystkich serwerach. Przestrzeganie tej kolejności nie jest wymagane, ale konieczne jest wykonanie kroków opisanych dla każdego typu serwera.

Program instalacyjny tworzy dwie zmienne środowiskowe, które mogą zostać użyte w skrypcie **eschangepw**:

ES_INSTALL_ROOT

Katalog instalacyjny wyszukiwania korporacyjnego.

ES_NODE_ROOT

Katalog danych wyszukiwania korporacyjnego. Hasło administratora wyszukiwania korporacyjnego jest przechowywane w pliku **es.cfg** w tym katalogu.

Procedura

Aby zmienić hasło administratora wyszukiwania korporacyjnego w konfiguracji wieloserwerowej:

1. Na serwerze indeksowania wyszukiwania korporacyjnego zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego:
esadmin system stopall

Ważne: W czasie, gdy system jest zatrzymany, użytkownicy nie mogą wprowadzać żądań wyszukiwania.

- a. Za pomocą poleceń systemowych (w systemach AIX, Linux lub Solaris) lub za pomocą funkcji zmiany hasła (w systemach Windows) zmień hasło systemu dla identyfikatora użytkownika administratora wyszukiwania korporacyjnego.
- b. Uruchom następujący skrypt, gdzie *nowaWartość* to hasło podane w kroku 1a:

Systemy AIX, Linux lub Solaris

```
eschangepw.sh nowaWartość
```

Windows

```
eschangepw nowaWartość
```

2. Poniższe kroki wykonaj na pozostałych komputerach systemu wyszukiwania korporacyjnego:

- a. Zaloguj się jako administrator wyszukiwania korporacyjnego.
- b. Zatrzymaj warstwę CCL wyszukiwania korporacyjnego:

Systemy AIX, Linux lub Solaris

```
stopccl.sh
```

Wiersz komend systemu Windows

```
stopccl
```

Narzędzie administracyjne usług systemu Windows

- 1) Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
 - 2) Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Zatrzymaj**.
- c. Za pomocą poleceń systemowych (w systemach AIX, Linux lub Solaris) lub za pomocą funkcji zmiany hasła (w systemach Windows) zmień hasło systemu dla identyfikatora użytkownika administratora wyszukiwania korporacyjnego. Hasło musi być zgodne z hasłem podanym w kroku 1a.
 - d. Uruchom następujący skrypt, gdzie *nowaWartość* to hasło podane w kroku 1a:

Systemy AIX, Linux lub Solaris

```
eschangepw.sh nowaWartość
```

Windows

```
eschangepw nowaWartość
```

- e. Zrestartuj warstwę CCL:

Systemy AIX, Linux lub Solaris

```
startccl.sh -bg
```

Wiersz komend systemu Windows

```
startccl
```


Narzędzie administracyjne usług systemu Windows

Aby uruchomić warstwę CCL w tle:

- 1) Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.

- 2) Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Właściwości**.
 - 3) Kliknij kartę **Logowanie**.
 - 4) Zmień hasło, podając nową wartość, a następnie kliknij przycisk **OK**.
 - 5) Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
3. Na serwerze indeksowania wyszukiwania korporacyjnego zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego:
- ```
esadmin system startall
```

#### Odsyłacze pokrewne

-  Ustawianie jednakowego szyfrowanego hasła administratora na wszystkich serwerach “Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 369

## Numery portów TCP używane w wyszukiwaniu korporacyjnym

Przegląd domyślnych numerów portów stosowanych w systemie wyszukiwania korporacyjnego pozwalający uniknąć konfliktów portów w czasie konfigurowania zasobów lub przydzielania numerów portów innym aplikacjom.

Jeśli konfigurowany jest firewall, to trzeba w nim jawnie otworzyć określone numery portów. Oprócz tego trzeba się też upewnić, że serwery wyszukiwania korporacyjnego znajdują się po wewnętrznej stronie firewalla.

Tabela 1. Konfiguracja numerów portów wyszukiwania korporacyjnego

| Nazwa portu                                                                               | Numery portów                  | Miejsce konfigurowania                                                                                                        |
|-------------------------------------------------------------------------------------------|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Warstwa CCL                                                                               | 6002                           | Pliki ES_NODE_ROOT/nodeinfo/es.cfg i ES_NODE_ROOT/master_config/nodes.ini na wszystkich serwerach wyszukiwania korporacyjnego |
| Protokół HTTP na serwerach wyszukiwania                                                   | 80                             | Plik HTTP_SERVER_ROOT/conf/http.conf na serwerach wyszukiwania                                                                |
| Konsola administracyjna produktu WebSphere Application Server, wersja 6                   | 9060                           | Na serwerach wyszukiwania                                                                                                     |
| Konsola administracyjna produktu WebSphere Application Server, wersja 5.1                 | 9090                           | Na serwerach wyszukiwania                                                                                                     |
| Przeszukiwacz DB2                                                                         | 6000, 6001, 6002, 60003, 50000 | Na serwerze przeszukiwacza                                                                                                    |
| Centrum informacyjne                                                                      | 8888                           | Na serwerach wyszukiwania                                                                                                     |
| Anonimowe lub dynamiczne porty warstwy CCL, przesyłania plików (ESFTP) oraz kopii indeksu | od 49152 do 65535              | Na wszystkich serwerach wyszukiwania korporacyjnego                                                                           |
| Serwer sieciowy Apache Derby                                                              | 1527                           | Na serwerze przeszukiwacza                                                                                                    |
| Komunikacja niestandardowa                                                                | 8890                           | Na serwerze przeszukiwacza                                                                                                    |

Tabela 1. Konfiguracja numerów portów wyszukiwania korporacyjnego (kontynuacja)

| Nazwa portu                                                                                                                            | Numery portów             | Miejsce konfigurowania                                                              |
|----------------------------------------------------------------------------------------------------------------------------------------|---------------------------|-------------------------------------------------------------------------------------|
| Połączenia zdalnych klientów z serwerem DB2 (używane tylko w wersjach poprzedzających wersję 8.4 produktu OmniFind Enterprise Edition) | 50000                     | Na serwerze przeszukiwacza                                                          |
| Baza danych WebSphere Information Integrator Content Edition FastObjects                                                               | 6001 (alternatywnie 6002) | W konsoli administracyjnej systemu WebSphere Information Integrator Content Edition |
| Konektor proxy zdalnego wywołania metody (RMI) WebSphere Information Integrator Content Edition                                        | 1251 (port RMI)           | W konsoli administracyjnej systemu WebSphere Information Integrator Content Edition |

## Zmianianie numeru portu dla systemu wyszukiwania korporacyjnego

Jeśli numer portu używany przez system wyszukiwania korporacyjnego do komunikacji powoduje konflikt z numerem portu, który jest używany przez inny produkt, należy zmienić numer portu wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Numer portu wyszukiwania korporacyjnego podawany jest podczas instalowania systemu OmniFind Enterprise Edition. W konfiguracji z wieloma serwerami ten sam numer portu należy podawać na wszystkich serwerach.

Jeśli danego numeru portu nie da się użyć (na przykład może być przydzielony innej usłudze działającej na tym serwerze), to informacja o konflikcie pojawia się w pliku `CCLServer_data.log`, gdzie *data* oznacza dzień, w którym utworzony został plik dziennika.

```
FFQ00273W wystąpił błąd wewnętrzny - Komunikat wyjątku: {0}
at java.net.PlainSocketImpl.socketBind(Native Method)
at java.net.PlainSocketImpl.bind(PlainSocketImpl.java:357)
at java.net.ServerSocket.bind(ServerSocket.java:341)
at java.net.ServerSocket.<init>(ServerSocket.java:208)
at java.net.ServerSocket.<init>(ServerSocket.java:120)
```

### Procedura

Aby zmienić numer portu używany przez wyszukiwanie korporacyjne:

1. Przejdź do komputera, w którym konieczna jest zmiana portu, zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego:

```
esadmin system stopall
```

**Ważne:** W czasie, gdy system jest zatrzymany, użytkownicy nie mogą wprowadzać żądań wyszukiwania.

2. Dokonaj edycji pliku `ES_NODE_ROOT/nodeinfo/es.cfg`, odzyskaj w nim poniższą właściwość, przypisz jej nową wartość numeru portu, a następnie zapisz i zamknij plik:
 

```
CCLPort=nowy_numer_portu
```
3. Uruchom ponownie warstwę CCL (Common Communication Layer) wyszukiwania korporacyjnego:

#### AIX, Linux i Solaris

```
startccl.sh
```

#### Wiersz komend systemu Windows

```
startccl
```

#### Narzędzie administracyjne usług systemu Windows

Aby uruchomić warstwę CCL w tle:

- a. Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
  - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
4. Przejdź do serwera indeksowania, zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj warstwę CCL:

#### AIX, Linux i Solaris

```
stopccl.sh
```

#### Wiersz komend systemu Windows

```
stopccl
```

#### Narzędzie administracyjne usług systemu Windows

- a. Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
  - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Zatrzymaj**.
5. Dokonaj edycji pliku ES\_NODE\_ROOT/nodeinfo/es.cfg :
- a. Odszukaj następującą właściwość, gdzie *nazwa\_komputera* jest nazwą komputera, w którym w kroku 2 na stronie 24 zmodyfikowano numer portu. Wartość *N* we właściwości *nodeN* jest numerem określającym serwer wyszukiwania korporacyjnego.  

```
nodeN.destination=nazwa_komputera
```
  - b. Znajdź następującą właściwość podrzędną, podaj ten sam numer portu, który został określony dla serwera w kroku 2 na stronie 24, a następnie zapisz i zamknij plik:  

```
nodeN.port=nowy_numer_portu
```
6. Zrestartuj system wyszukiwania korporacyjnego:

```
esadmin system startall
```

#### Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 369

“Numery portów TCP używane w wyszukiwaniu korporacyjnym” na stronie 23

---

## Zmianie nazw hosta lub adresów IP serwera wyszukiwania korporacyjnego

Nazwy hosta i adresy IP, które według konfiguracji mają być używane przez serwery wyszukiwania korporacyjnego, można zmieniać.

Na przykład można zdecydować się na zmianę adresu IP, jeśli w każdym serwerze wyszukiwania korporacyjnego zainstalowano po kilka kart sieciowych, a okazało się, że serwer indeksowania jest skonfigurowany do używania powolnej sieci. Ustawienia w plikach konfiguracyjnych można zmienić tak, aby serwer indeksowania mógł używać szybszej sieci.

**Wskazówka:** Zamiast edytować pliki konfiguracyjne, można też ponownie zainstalować produkt OmniFind Enterprise Edition i w programie instalacyjnym określić nowe nazwy hosta lub adresy IP.

#### **Procedura:**

Aby zmienić nazwy hosta lub adresy IP:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W przypadku konfiguracji z wieloma serwerami można zalogować się na dowolnym serwerze wyszukiwania korporacyjnego.
2. Zatrzymaj sesję wyszukiwania korporacyjnego:  
`esadmin system stopall`
3. Zatrzymaj wspólną warstwę komunikacyjną (CCL). W przypadku konfiguracji z wieloma serwerami warstwę CCL należy zatrzymać na każdym serwerze wyszukiwania korporacyjnego jedną z następujących metod:

##### **AIX, Linux i Solaris**

```
stopccl.sh
```

##### **Wiersz komend systemu Windows**

```
stopccl
```

##### **Narzędzie administracyjne usług systemu Windows**

- a. Uruchom aplet Usługi systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
  - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Zatrzymaj**.
4. Wykonaj poniższą procedurę na każdym serwerze wyszukiwania korporacyjnego:
    - a. Poddaj edycji plik `ES_INSTALL_ROOT/configurations/ccl.properties` i określ nazwę hosta lub adres IP, które mają być używane dla danego serwera, w parametrze `es_server_hostName`.
    - b. Poddaj edycji plik `ES_NODE_ROOT/nodeinfo/es.cfg` i określ nazwę hosta lub adres IP, które mają być używane dla danego serwera, w parametrze `LocalHostName`.
    - c. Poddaj edycji plik `ES_NODE_ROOT/master_config/nodes.ini` i wszystkie wystąpienia parametru `nodeN.destination` zastąp nazwą hosta lub adresem IP, które mają być używane dla danego serwera.
    - d. Poddaj edycji plik `ES_NODE_ROOT/config/nodes.ini` i wszystkie wystąpienia parametru `nodeN.destination` zastąp nazwą hosta lub adresem IP, które mają być używane dla danego serwera.
  5. Zrestartuj warstwę CCL. W przypadku konfiguracji z wieloma serwerami warstwę CCL należy zrestartować na każdym serwerze wyszukiwania korporacyjnego jedną z następujących metod:

##### **Systemy AIX, Linux lub Solaris**

```
startccl.sh -bg
```

##### **Wiersz komend systemu Windows**

```
startccl
```

##### **Narzędzie administracyjne usług systemu Windows**

Aby uruchomić warstwę CCL w tle:

- a. Uruchom aplet Usługi systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.



- b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
6. Zrestartuj sesję wyszukiwania korporacyjnego:  
esadmin system startall

---

## Konfigurowanie obsługi podwójnych adresów IP

Jeśli serwery, na których jest instalowane wyszukiwanie korporacyjne, są skonfigurowane pod kątem obsługi podwójnych adresów IP, to należy ręcznie skonfigurować serwery wyszukiwania korporacyjnego, aby działały w tym środowisku.

Na przykład administrator komputera desktop może zainstalować program Microsoft Loopback Adapter, aby utworzyć sieć wirtualną realizującą wymagania sieciowe określonych produktów, takich jak Microsoft SQL Server.

### Procedura:

Aby skonfigurować system wyszukiwania korporacyjnego tak, aby działał na serwerach obsługujących podwójne adresy IP:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W przypadku konfiguracji z wieloma serwerami można zalogować się na dowolnym serwerze wyszukiwania korporacyjnego.
2. Zatrzymaj sesję wyszukiwania korporacyjnego:  
esadmin system stopall
3. Zatrzymaj wspólną warstwę komunikacyjną (CCL). W przypadku konfiguracji z wieloma serwerami warstwę CCL należy zatrzymać na każdym serwerze wyszukiwania korporacyjnego jedną z następujących metod:

#### AIX, Linux i Solaris

```
stopccl.sh
```

#### Wiersz komend systemu Windows

```
stopccl
```

#### Narzędzie administracyjne usług systemu Windows

- a. Uruchom aplet Usługi systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
  - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Zatrzymaj**.
4. Na każdym serwerze wyszukiwania korporacyjnego obsługującym podwójne adresy IP poddaj edycji plik ES\_NODE\_ROOT/nodeinfo/es.cfg i dodaj parametr o nazwie **LocalIPAddress**. Jako wartość należy przypisać adres IP, który może zostać przetłumaczony przez system DNS.
5. Zrestartuj warstwę CCL. W przypadku konfiguracji z wieloma serwerami warstwę CCL należy zrestartować na każdym serwerze wyszukiwania korporacyjnego jedną z następujących metod:

#### Systemy AIX, Linux lub Solaris

```
startccl.sh -bg
```

#### Wiersz komend systemu Windows

```
startccl
```

#### Narzędzie administracyjne usług systemu Windows

Aby uruchomić warstwę CCL w tle:

- a. Uruchom aplet Usługi systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
  - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
6. Zrestartuj sesję wyszukiwania korporacyjnego:  
esadmin system startall

---

## Włączanie obsługi protokołu IPv6

Aby włączyć obsługę adresów URL zgodnych z protokołem IP, wersja 6 (IPv6), należy zdefiniować zmienną środowiskową, która poinstruuje system wyszukiwania korporacyjnego, że ma używać tylko adresów gniazda IPv6.

### Zagadnienia wstępne

1. Upewnij się, że wartości właściwości **LocalHostName** oraz właściwości **LocalIPAddress** (jeśli są podane) w pliku %ES\_CFG% są nazwą hosta albo poprawnym adresem IPv6 komputera lokalnego. W razie konieczności wartości te można zmienić ręcznie, a następnie zapisać plik %ES\_CFG%.
2. Upewnij się, że wartość właściwości **DerbyServerHostName** w pliku %ES\_CFG% jest nazwą hosta. Wartość ta nie może być adresem IPv4 ani adresem IPv6. Jeśli właściwość ta zostanie zmieniona, plik %ES\_CFG% należy zapisać.
3. Upewnij się, że wartości właściwości **destination** i **serverhost** w pliku ES\_NODE\_ROOT/master\_config/nodes.ini na serwerze indeksowania są poprawnym adresem IPv6 albo poprawną nazwą hosta. Wartości te nie mogą być adresem IPv4. Jeśli do pliku zostaną wprowadzone zmiany, to należy je zapisać.

### Ograniczenia

Obsługa protokołu IPv6 jest dostępna tylko w systemach wyszukiwania korporacyjnego zainstalowanych na serwerach Windows 2003 Server. Po włączeniu obsługi adresów IPv6 wyszukiwanie korporacyjne nie będzie już korzystać z adresów IPv4 do jakiegokolwiek komunikacji gniazdowej.

### Obsługiwane przeglądarki

Przeglądarki Internet Explorer i Mozilla Firefox różnie obsługują adresy IPv6.

#### Mozilla Firefox

Aby uruchomić konsolę administracyjną wyszukiwania korporacyjnego lub aplikację wyszukiwującą, w adresie URL można określić adres IPv6 lub nazwę hosta. Na przykład:

```
http://[2001::db8]/ESAdmin
http://SerwerWyszukiwania.com/ESSearchApplication/
```

#### Internet Explorer

Aby uruchomić konsolę administracyjną wyszukiwania korporacyjnego lub aplikację wyszukiwującą, w adresie URL nie można określić adresu IPv6. Należy użyć następującego formatu i zapewnić, że nazwa hosta jest odwzorowywana na adres IPv6 w systemie DNS lub w pliku c:\windows\system32\etc\hosts:

```
http://SerwerWyszukiwania.com/ESAdmin/
http://SerwerWyszukiwania.com/ESSearchApplication/
```

Jeśli w adresie URL zostanie określony adres IPv6, wyświetlony zostanie komunikat Błąd składni. Dodatkowe informacje na temat tego ograniczenia można znaleźć w sekcji <http://support.microsoft.com/kb/325414>.

## Obsługa źródeł danych IPv4

Źródła danych, które działają na serwerze IPv4, są obsługiwane przez przeszukiwacze wyszukiwania korporacyjnego. W przypadku konfigurowania przeszukiwacza sieci WWW i określania początkowych adresów URL w formacie adresów IPv6 należy pamiętać o ujęciu adresów URL w nawiasy kwadratowe. Na przykład:

```
http://[2001:db8:0:1:0:0:0:1]
```

```
http://[2001:db8:0:1::1]
```

### Procedura

Aby włączyć obsługę protokołu IPv6:

1. Na serwerze indeksowania wyszukiwania korporacyjnego zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego:  
esadmin system stopall

**Ważne:** W czasie, gdy system jest zatrzymany, użytkownicy nie mogą wprowadzać żądań wyszukiwania.

2. Zatrzymaj serwer warstwy CCL na wszystkich serwerach wyszukiwania korporacyjnego:
  - a. Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
  - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Zatrzymaj**.
3. Zdefiniuj zmienną **ES\_IPV6=1** w zmiennych środowiskowych systemu. W instalacjach wieloserwerowych czynność tę należy wykonać na serwerach wyszukiwania.
4. Zaktualizuj niestandardowe właściwości wirtualnej maszyny języka Java (JVM) na serwerze WebSphere Application Server. W instalacjach wieloserwerowych czynność tę należy wykonać na serwerach wyszukiwania.
  - a. Zatrzymaj serwer aplikacji ESSearchServer (jeśli działa) na serwerze WebSphere Application Server.
  - b. Otwórz konsolę administracyjną serwera WebSphere Application Server i przejdź do panelu niestandardowych właściwości wirtualnej maszyna języka Java. Wybierz kolejno opcje **Serwery** → **Serwery aplikacji** → **ESSearchServer** → **Język Java i zarządzanie procesami** → **Definicja procesu** → **Wirtualna maszyna języka Java** → **Właściwości niestandardowe**.
  - c. Skonfiguruj następujące właściwości:

```
java.net.preferIPv4Stack=false
java.net.preferIPv6Addresses=true
```

Jeśli dana właściwość niestandardowa nie jest jeszcze wymieniona, należy utworzyć nową właściwość, wprowadzając nazwę właściwości w polu **Nazwa** i poprawną wartość w polu **Wartość**.
  - d. Jeśli uruchomiony jest program WebSphere Application Server, wersja 6.1, należy w pliku **ES\_NODE\_ROOT/nodeinfo/es.cfg** dodać właściwość **LocalIPAddress** określającą adres IPv6 serwera wyszukiwania.
  - e. Zrestartuj aplikację ESSearchServer.
5. Poniższe czynności należy wykonać, jeśli używany jest program WebSphere Application Server, wersja 6.1. W instalacjach wieloserwerowych czynności te należy wykonać na serwerach wyszukiwania.
  - a. Dokonaj edycji pliku **httpd.conf** dla serwera IBM HTTP Server.

- b. Usuń znak # z początku następującego wiersza, aby instrukcja nie była traktowana jak komentarz, co umożliwi serwerowi IBM HTTP Server nasłuchiwanie połączeń IPv6 na porcie 80:  
# Listen [::]:80
- c. Zrestartuj serwer IBM HTTP Server.
6. Uruchom ponownie serwer CCL na wszystkich serwerach wyszukiwania korporacyjnego:
  - a. Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
  - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
7. Zrestartuj system wyszukiwania korporacyjnego:  
esadmin system startall
8. Sprawdź w pliku `ES_NODE_ROOT/node/logs/ccl0.log`, czy serwer CCL został uruchomiony bez błędów. Jeśli protokół IPv6 został poprawnie włączony, zarejestrowane zostaną komunikaty podobne do następujących:  
INFO: serwer CCL gotowy do działania. Oczekiwanie na żądania zewnętrzne.  
Nazwa hosta CCL: fe80::250:56ff:feb4:27d1  
Adres dot.ip hosta CCL: fe80:0:0:0:250:56ff:feb4:27d1  
Numer portu serwera CCL: 6002  
Łączna ilość pamięci JVM Runtime: 33022 kB  
Bieżąca ilość wolnej pamięci: 21346 kB  
java.net.preferIPv6Addresses=true  
java.net.preferIPv4Stack=false  
INFO: Sesji "TraceDaemonSession" przypisano PID=4992
9. Wykonaj następującą procedurę, aby upewnić się, że aplikacje wyszukujące i administracyjne działają poprawnie:
  - a. Uruchom serwery wyszukiwania dla dowolnej kolekcji, a następnie otwórz przykładową aplikację wyszukującą i wprowadź zapytanie. Ten krok służy upewnieniu się, że przetwarzanie wprowadzania zapytań działa, mimo że w indeksie nie ma dokumentów do wyszukiwania.
  - b. Otwórz konsolę administracyjną i upewnij się, że można się zalogować bez problemów.

### Wyłączenie obsługi protokołu IPv6

1. Zapoznaj się z sekcją **Zanim rozpocziesz** i dla każdego wystąpienia adresu IPv6 wymień istniejącą wartość na nazwę hosta lub adres IPv4.
2. Wykonaj procedurę włączania obsługi protokołu IPv6 z następującymi wyjątkami:
  - W kroku 3 na stronie 29 usuń zmienną środowiskową **ES\_IPV6**.
  - W kroku 4 na stronie 29 usuń niestandardowe właściwości maszyny JVM ustawione dla aplikacji ESSearchServer na serwerze WebSphere Application Server.
3. Usuń właściwość **LocalIPAddress** z pliku `ES_NODE_ROOT/nodeinfo/es.cfg`, jeśli zawiera ona adres IPv6.
4. Wykonaj procedurę z kroku 9, aby sprawdzić aplikacje wyszukujące i administracyjne.

---

## Kolekcje wyszukiwania korporacyjnego

W kolekcji wyszukiwania korporacyjnego dostępny jest cały zestaw źródeł, w których użytkownicy mogą przeprowadzać wyszukiwanie za pomocą pojedynczego zapytania. Dzięki stowarzyszeniu użytkownicy mogą wyszukiwać w wielu kolekcjach za pomocą jednego zapytania.

Podczas tworzenia kolekcji należy określić opcje dotyczące całej kolekcji. Kolekcja pozostaje pusta do czasu dodania do niej treści.

Istnieją dwa sposoby dodawania kolekcji do wyszukiwania korporacyjnego:

- Jeśli użytkownik nie używa zbyt często konsoli administracyjnej wyszukiwania korporacyjnego lub jeśli jego wiedza na temat współdziałania komponentów kolekcji nie jest zbyt szeroka, wówczas do utworzenia kolekcji zalecane jest użycie kreatora kolekcji. Kreator kolekcji pomaga wykonać kolejne zadania i podczas tworzenia kolekcji umożliwia zapisywanie wyników pracy jako kolekcji roboczej.
- Użytkownik zaznajomiony z konsolą administracyjną może utworzyć kolekcję, wybierając odpowiednie opcje na stronach dostępnych w widoku Kolekcje.

Po utworzeniu kolekcji można za pomocą pól sterujących dostępnych w widoku Kolekcje edytować i monitorować kolekcje, system wyszukiwania korporacyjnego oraz opcje zabezpieczeń.

### Stowarzyszenie kolekcji

Jeśli aplikacja wyszukująca zapewnia obsługę stowarzyszeń, użytkownicy mogą jednocześnie wyszukiwać w wielu kolekcjach. Stowarzyszenie umożliwia również obejście ograniczenia wielkości kolekcji do 20 000 000 dokumentów. Na przykład użytkownicy mogą wykonywać wyszukiwanie w dwóch kolekcjach, z których każda zawiera 20 000 000 dokumentów.

Jakość wyszukiwania zależy od generowanych przez poszczególne kolekcje wyników, które następnie są scalane w celu otrzymania ostatecznej tabeli wynikowej. Otrzymane wyniki są takie same jak po wprowadzeniu dwóch oddzielnych zapytań oraz scaleniu i sklasyfikowaniu uzyskanych wyników.

#### Zadania pokrewne

“Monitorowanie kolekcji” na stronie 302

---

## Tworzenie kolekcji przy użyciu Kreatora kolekcji

Użytkownicy rozpoczynający dopiero pracę z systemem wyszukiwania korporacyjnego mogą przy tworzeniu kolekcji skorzystać z pomocy kreatora. Kreator dostarcza szczegółowe informacje na temat poszczególnych etapów procesu i umożliwia zapisanie wprowadzanych ustawień.

### Zagadnienia wstępne

Aby utworzyć kolekcję, wymagana jest rola administratora wyszukiwania korporacyjnego.

Aby dodać zawartość do kolekcji lub podać opcje określające sposób analizowania, indeksowania lub wyszukiwania w kolekcji, należy być administratorem systemu wyszukiwania korporacyjnego lub administratorem tej kolekcji.

## Ograniczenia

W celu utworzenia następujących typów przeszukiwaczy można skorzystać z kreatora kolekcji:


- Content Edition
- DB2
- DB2 Content Manager
- Exchange Server
- Notes
- System plików UNIX
- Sieć WWW
- System plików Windows

## Informacje o tym zadaniu

Podczas tworzenia kolekcji możliwe jest zapisanie jej w postaci roboczej. Kolekcja w postaci roboczej może być modyfikowana przez każdego administratora dysponującego uprawnieniami do administrowania kolekcjami. Na przykład administrator kolekcji mający doświadczenie w konfigurowaniu źródeł Lotus Notes może skonfigurować przeszukiwacz Notes. Następnie administrator kolekcji mający doświadczenie z systemami UNIX może zmienić wersję roboczą kolekcji i skonfigurować przeszukiwacz systemu plików UNIX.

## Procedura

Aby utworzyć kolekcję przy użyciu Kreatora kolekcji, należy wykonać poniższe czynności:

1. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.
2. Kliknij opcję **Kreator kolekcji**.
3. Postępuj zgodnie z instrukcjami wyświetlanymi na ekranach kreatora w celu utworzenia pustej kolekcji i dodania do niej treści.  
Należy skonfigurować ogólne informacje o kolekcji i utworzyć przynajmniej jeden przeszukiwacz. Dla pozostałych opcji konfiguracyjnych można zaakceptować ustawienia domyślne lub określić indywidualne wartości dla nowej kolekcji.
4. Aby zapisać kolekcję przed jej ostatecznym utworzeniem, kliknij opcję **Zapisz jako robocze**.  
Utworzona kolekcja będzie wymieniona wraz z innymi kolekcjami roboczymi w widoku Kolekcje. Jeśli zabezpieczenie kolekcji jest włączone, obok nazwy kolekcji jest wyświetlana ikona  **Zabezpieczenie kolekcji jest włączone**.
5. Aby powrócić do kolekcji będącej w trakcie tworzenia, kliknij opcję **Powrót do kreatora** w widoku Kolekcje.
6. Kliknij przycisk **Zakończ**, aby utworzyć kolekcję.  
Nowo utworzona kolekcja będzie wymieniona wraz z innymi kolekcjami w widoku Kolekcje.

Po utworzeniu kolekcji należy uruchomić procesy odpowiedzialne za przeszukiwanie, analizę składni, indeksowanie i wyszukiwanie danych w kolekcji. Dopóki kolekcja nie będzie gotowa do powiązania z aplikacjami wyszukującymi, które mogą w niej wyszukiwać, do wyszukiwania w nowej kolekcji można wykorzystać przykładową aplikację wyszukującą o nazwie Default.

---

## Tworzenie kolekcji przy użyciu widoku kolekcji

Widok Kolekcje służy do tworzenia pustej kolekcji. Następnie można uruchomić edycję kolekcji w celu dodania do niej danych i umożliwienia wyszukiwania w niej.

### Zagadnienia wstępne

Aby utworzyć kolekcję, wymagana jest rola administratora wyszukiwania korporacyjnego.

Aby dodać zawartość do kolekcji lub podać opcje określające sposób analizowania, indeksowania lub wyszukiwania w kolekcji, należy być administratorem systemu wyszukiwania korporacyjnego lub administratorem tej kolekcji.

### Informacje o tym zadaniu

Informacje na temat wartości dozwolonych do użycia w odniesieniu do nowej kolekcji można uzyskać, klikając opcję **Pomoc** podczas tworzenia kolekcji.

### Procedura

Aby utworzyć kolekcję przy użyciu widoku Kolekcje, należy wykonać poniższe czynności:


1. W widoku Kolekcje kliknij opcję **Utwórz kolekcję**.
2. Na stronie Utwórz kolekcję wprowadź informacje lub dokonaj wyboru opcji w następujących polach:
  - **Nazwa kolekcji**. Służy do określania opisowej nazwy zawartości lub przeznaczenia kolekcji.
  - **Zabezpieczenie kolekcji**. Określa, czy dla kolekcji mają być włączone zabezpieczenia. Po utworzeniu kolekcji tego ustawienia już nie można zmienić. Jeśli zabezpieczenie jest włączone, można później określać opcje wymuszania kontroli dostępu na poziomie dokumentów.
  - **Ważność dokumentu (statyczny model oceniania)**. Służy do określania strategii przypisywania statycznego współczynnika oceniania, który będzie wykorzystywany do oceny dokumentów w wynikach wyszukiwania. Wartości tej nie można zmieniać po utworzeniu kolekcji.
  - **Typ klasyfikacji**. Służy do określenia, czy będzie możliwe wyszukiwanie dokumentów według kategorii, do których należą.
  - **Używany język**. Służy do wybrania języka domyślnego dla operacji wyszukiwania dokumentów w kolekcji.
3. W poniższych polach należy zaakceptować wartości domyślne lub wprowadzić wartości stosowne dla tworzonej kolekcji:
  - **Opis**. Domyślnie opis nie jest tworzony.
  - **Szacowana liczba dokumentów**. Domyślna szacowana wielkość kolekcji to 1 000 000 dokumentów. Wartość ta jest wykorzystywana przez system do oszacowania zasobów pamięciowych i dyskowych dla kolekcji, a nie do ograniczania wielkości kolekcji.
  - **Położenie kolekcji danych**. Domyślną lokalizacją plików powiązanych z kolekcją jest serwer indeksowania. Wartości tej nie można zmieniać po utworzeniu kolekcji.
  - **Identyfikator kolekcji**. Domyślny identyfikator kolekcji tworzony jest na podstawie nazwy kolekcji. Wartości tej nie można zmieniać po utworzeniu kolekcji. Jeśli podany zostanie niestandardowy identyfikator kolekcji, aplikacje wyszukujące będą odwoływały się do kolekcji za pomocą tego identyfikatora zamiast tworzonych przez system identyfikatora, który może być niezrozumiały.
  - **Podział na ciągi n-znakowe**. Domyślna metoda podziału oparta jest o zestaw znaków Unicode z podziałem wg znaków spacji. Aby korzystać z podziału na ciągi n-znakowe,



opcję tę należy wybrać tylko wtedy, gdy kolekcja zawiera dokumenty w języku chińskim, japońskim lub koreańskim i analizator składni ma używać podziału na ciągi n-znakowe do oddzielania wyrazów. Wartości tej nie można zmieniać po utworzeniu kolekcji.

Więcej informacji na temat konfigurowania obsługi pełnej analizy n-znakowej i podziału na leksemy w kolekcjach wyszukiwania korporacyjnego można znaleźć w dokumencie <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088>.

#### 4. Kliknij przycisk **OK**.

W widoku Kolekcje nowa kolekcja zostanie wyświetlona na liście wraz z innymi kolekcjami należącymi do systemu wyszukiwania korporacyjnego. Jeśli zabezpieczenie kolekcji jest włączone, obok nazwy kolekcji jest wyświetlana ikona  **Zabezpieczenie kolekcji jest włączone**.

Kolekcja pozostaje pusta do czasu dodania do niej treści. W celu dodania treści do nowej kolekcji należy wybrać kolekcję w widoku Kolekcje, poddać ją edycji, utworzyć przynajmniej jeden przeszukiwacz i określić opcje dotyczące sposobu analizowania składni, indeksowania i wyszukiwania danych.

Następnie należy uruchomić procesy przeszukiwania, analizy składni, indeksowania i wyszukiwania w kolekcji. Do momentu przygotowania własnych aplikacji wyszukiwania w nowej kolekcji można wyszukiwać za pomocą przykładowej aplikacji wyszukiwania.

---

## Edytowanie kolekcji

Edycja kolekcji ma na celu określenie informacji o dokumentach, które mają zostać uwzględnione w kolekcji.

### Zagadnienia wstępne


Aby edytować kolekcję, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

### Informacje o tym zadaniu

Podczas edytowania kolekcji należy określić opcje przeszukiwania źródeł danych, przetwarzania dokumentów, budowania indeksu, wyszukiwania w zawartości indeksu oraz rejestrowania komunikatów o błędach. Podczas tworzenia kolekcji należy ją poddać edycji w celu dodania do niej treści. Później kolekcję można edytować w celu zaktualizowania treści bądź zmiany sposobu przeszukiwania, analizowania, indeksowania, wyszukiwania i rejestrowania informacji.

### Procedura

Aby zmodyfikować kolekcję, należy wykonać poniższe czynności:

1. Kliknij opcję **Kolekcje** na pasku narzędzi, aby otworzyć widok Kolekcje.
2. Na liście kolekcji znajdź kolekcję, która ma zostać zmodyfikowana i kliknij ikonę  **Edytuj**.
3. Wprowadź zmiany na dowolnych z poniższych stron:

#### Przeszukiwanie

Umożliwia określenie źródeł danych, które mają zostać przeszukane, oraz opcje dotyczących sposobu przeszukiwania treści. Każda kolekcja musi zawierać co najmniej jeden przeszukiwacz i może zawierać dane z wielu typów źródeł danych. Wymagane jest skonfigurowanie przynajmniej jednego przeszukiwacza dla każdego typu źródeł danych, który ma być uwzględniony w kolekcji.

### **Analiza składni**

Umożliwia określenie sposobu analizowania składni dokumentów poddanych przeszukaniu. Można tu skonfigurować kategorie umożliwiające użytkownikom wyszukiwanie w podzbiorach kolekcji oraz reguły pozwalające na wyszukiwanie w określonych częściach dokumentów XML i HTML. Jeśli do systemu wyszukiwania korporacyjnego dodano niestandardowe mechanizmy analizy tekstu, można wybrać jeden z nich do analizowania i dodawania adnotacji do treści w danej kolekcji. Można też powiązać pola z klasami kryteriów i w ten sposób wpłynąć na pozycje dokumentów w wynikach wyszukiwania.

**Indeks** Umożliwia zdefiniowanie harmonogramów budowania całego indeksu oraz aktualizowania go informacjami o nowej i zmienionej treści. Można tu także skonfigurować opcje związane z: wykorzystywaniem znaków wieloznacznych w zapytaniach, ograniczaniem widoków kolekcji do zakresu identyfikatorów URI, zwiżaniem wyników wyszukiwania z tej samej witryny sieci WWW i usuwaniem identyfikatorów URI z indeksu.

### **Wyszukiwanie**

Umożliwia określanie opcji procesów wyszukiwania, takich jak konfigurowanie pamięci podręcznej wyszukiwania i wybór języka wyszukiwania. Można ponadto konfigurować szybkie odsyłacze, co gwarantuje zwracanie wcześniej określonych identyfikatorów URI, gdy tylko użytkownik użyje w zapytaniu pewnych słów lub fraz. Jeśli do systemu wyszukiwania korporacyjnego zostały dodane słowniki niestandardowe, można wybrać słowniki, które mają być używane do wyszukiwania w danej kolekcji.

### **Dziennik**

Umożliwia określanie typów komunikatów objętych rejestrowaniem i opcji dotyczących tworzenia i rotacji plików dzienników. Można skonfigurować alerty umożliwiające powiadamianie użytkownika o określonych zdarzeniach oraz podać opcje odbierania wiadomości e-mail w przypadku zarejestrowania określonych komunikatów lub typów komunikatów. Można również określić opcje rejestrowania informacji o dokumentach usuniętych z systemu wyszukiwania korporacyjnego.

### **Ogólne**

Służy do podania ogólnych informacji o kolekcji i wyświetlenia ustawień, których nie można zmieniać. Jeśli podczas tworzenia kolekcji zostało włączone zabezpieczenie, można konfigurować opcje zabezpieczeń na poziomie dokumentu.

---

## **Usuwanie kolekcji**

Całkowite usunięcie kolekcji powoduje usunięcie wszystkich informacji o tej kolekcji z systemu wyszukiwania korporacyjnego.

### **Zagadnienia wstępne**

Aby usunąć kolekcję, wymagana jest rola administratora wyszukiwania korporacyjnego.

Przed usunięciem kolekcji należy zatrzymać wszystkie powiązane z nią procesy.


### **Informacje o tym zadaniu**

Proces usuwania kolekcji może być czasochłonny. Po otrzymaniu potwierdzenia, że kolekcja ma zostać usunięta, system usuwa wszystkie dane związane z tą kolekcją.

**Wskazówka:** Może zostać wyświetlony komunikat o przekroczeniu limitu czasu przez żadaną operację, mimo że proces nadal działa w tle. Aby dowiedzieć się, czy zadanie zostało ukończone, kliknij przycisk **Odśwież** w konsoli administracyjnej (nie klikaj przycisku **Odśwież** w przeglądarce WWW). Proces usuwania można uznać za zakończony, gdy nazwa kolekcji zniknie z listy kolekcji.

### Procedura

Aby usunąć kolekcję, należy wykonać poniższe czynności:

1. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.
2. Na liście kolekcji odszukaj przeznaczoną do usunięcia kolekcję i kliknij przycisk  **Usuń**.

---

## Określanie identyfikatora kolekcji

W wielu zadaniach administracyjnych potrzebna jest znajomość identyfikatora kolekcji.

### Zagadnienia wstępne

Aby zobaczyć identyfikator kolekcji, należy mieć rolę administratora wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Gdy tworzona jest kolekcja, można samemu określić wartość identyfikatora albo umożliwić przypisanie identyfikatora przez system. W celu określenia identyfikatora kolekcji, gdy kolekcja już została utworzona, można się posłużyć konsolą administracyjną lub przejrzeć plik konfiguracyjny.

### Procedura

1. Aby określić identyfikator kolekcji przy użyciu konsoli administracyjnej:
  - a. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.
  - b. Na stronie Ogólne kliknij opcję **Wyświetl ustawienia kolekcji**.  
Strona Ustawienia kolekcji przedstawia identyfikator kolekcji, pełną ścieżkę do miejsca, w którym przechowywane są dane kolekcji, oraz model rankingu statycznego używanego z dokumentami w danej kolekcji.
2. Aby określić identyfikator kolekcji, przeglądając plik konfiguracyjny:
  - a. Otwórz plik `ES_NODE_ROOT/master_config/collections.ini`. Aby ułatwić sobie przeglądanie, posortuj ten plik. W następujących przykładowych danych wyjściowych identyfikatorem kolekcji jest `coll1`:

```
% sort $ES_NODE_ROOT/master_config/collections.ini | more
collection1.configfile=coll_config.ini
collection1.datadir=/home/esearch/node/data/coll
collection1.description=
collection1.displayname=Collection1
collection1.flags=0
collection1.id=coll
collection1.sectiontype=collection
collection1.type=1
...
```

---

## Administrowanie przeszukiwaczem

Konfiguracja przeszukiwaczy dla różnych typów danych, które mają zostać włączone do kolekcji. Pojedyncza kolekcja może zawierać dowolną liczbę przeszukiwaczy.

**Wskazówka:** Kurs dostępny jest pod adresem <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html>. Kurs zawiera opis czynności wykonywanych podczas instalacji i konfiguracji, pokazuje w jaki sposób należy przeszukiwać różne typy źródeł danych, a także przedstawia sposoby rozszerzenia możliwości wyszukiwania korporacyjnego za pomocą aplikacyjnych interfejsów programistycznych produktu. Kurs odwołuje się do starszej wersji produktu OmniFind Enterprise Edition, ale większość pojęć i procedur jest nadal aktualna.

Szczegółowe przykłady konfigurowania przeszukiwaczy i włączania obsługi zabezpieczeń w małych, średnich i dużych organizacjach można znaleźć w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

### Konfiguracja przeszukiwaczy

Do tworzenia, edytowania i usuwania przeszukiwaczy służy konsola administracyjna wyszukiwania korporacyjnego. Zazwyczaj przeszukiwacz konfiguruje specjalista w zakresie obsługi typu danych, który będzie przeszukiwany. Na przykład przeszukiwacz do obsługi źródeł danych Lotus Notes powinien zostać skonfigurowany przez administratora kolekcji, który jest administratorem Notes lub bliskim współpracownikiem osoby kompetentnej w dziedzinie przeszukiwanych baz danych.

Podczas tworzenia przeszukiwacza kreator związany z typem przeszukiwanych danych ułatwia użytkownikowi określenie właściwości sterujących wykorzystaniem zasobów systemu przez ten przeszukiwacz. Kreator ten ułatwia także wybranie źródeł, które mają być przeszukiwane.

Istnieje także możliwość zmodyfikowania istniejących przeszukiwaczy. W razie potrzeby można dokonać edycji właściwości przeszukiwacza lub części przestrzeni przeszukiwania. Kreatory przeszukiwaczy ułatwiają wprowadzanie również takich zmian.

### Konfigurowanie nowego przeszukiwacza przy użyciu wartości podstawowych

Przeszukiwacz można utworzyć przy użyciu domyślnych wartości systemowych lub kopiując wartości określone dla innego przeszukiwacza tego samego typu. Użycie istniejącego przeszukiwacza jako podstawy dla nowego przeszukiwacza pozwala szybko utworzyć wiele przeszukiwaczy o podobnych właściwościach, które można następnie skonfigurować na przykład do przeszukiwania różnych źródeł lub działania na podstawie różnych harmonogramów przeszukiwania.

Kopiując przeszukiwacz, użytkownik może rozdzielić obciążenie przeszukiwaniem między wiele przeszukiwaczy wykorzystujących te same reguły przeszukiwania. Można na przykład skopiować przeszukiwacz Notes w celu użycia tych samych właściwości i reguł przeszukiwania pól, ale z innym serwerem Lotus Notes. Różnice między poszczególnymi przeszukiwaczami mogą dotyczyć jedynie przeszukiwanych baz danych i ustawień zabezpieczeń na poziomie dokumentu.

## Łączenie typów przeszukiwaczy w kolekcji

Przeszukiwacze wyszukiwania korporacyjnego są zaprojektowane do zbierania informacji z określonych typów źródeł danych. Podczas konfigurowania przeszukiwaczy dla kolekcji użytkownik musi zdecydować, w jaki sposób połączyć różne typy źródeł danych, aby użytkownicy mogli łatwo przeszukiwać dane w przedsiębiorstwie. Na przykład, aby zapewnić użytkownikom możliwość przeszukiwania systemów plików Microsoft Windows i folderów publicznych Microsoft Exchange Server za pomocą pojedynczego zapytania, należy utworzyć kolekcję zawierającą przeszukiwacze systemu plików Windows i przeszukiwacze Exchange Server.

Podczas łączenia wielu typów przeszukiwaczy w pojedynczej kolekcji należy upewnić się, że wszystkie przeszukiwacze mogą używać tej samej statycznej metody oceniania. (Statyczna metoda oceniania jest określana podczas tworzenia kolekcji). Na przykład, jeśli jednocześnie używane są źródła danych WWW (w których jako czynnik oceny stopnia ważności wykorzystywane są występujące w dokumentach odsyłacze) i źródła danych NNTP (w których zazwyczaj jako czynnik oceny stopnia ważności wykorzystywana jest data dokumentu), wówczas jakość wyników wyszukiwania może być niska.

## Konfigurowanie zabezpieczeń na poziomie dokumentu

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentu. Dla każdego przeszukiwacza istnieje możliwość przypisania tokenów zabezpieczeń do przeszukiwanych dokumentów. Jeśli podczas konfigurowania przeszukiwacza użytkownik określi, że chce korzystać z zabezpieczeń na poziomie dokumentu, przeszukiwacz przypisze określone tokeny zabezpieczeń do każdego dokumentu i doda je indeksu razem z dokumentami.

Jeśli w używanych niestandardowych aplikacjach wyszukujących zostanie włączona funkcja zabezpieczeń, w aplikacjach tych możliwe będzie użycie tokenów zabezpieczeń przypisanych przez przeszukiwacz do dokumentów w celu uwierzytelniania użytkowników. Dzięki temu można ograniczyć dostęp do niektórych dokumentów w kolekcji i udostępnić inne dokumenty do przeszukiwania. Na przykład w jednej kolekcji wszystkie dokumenty w folderach publicznych serwera Microsoft Exchange Server mogą zostać udostępnione wszystkim użytkownikom, a dokumenty w bazach danych Lotus Notes mogą zostać udostępnione tylko użytkownikom o określonych identyfikatorach.

Do określania wartości tokenów zabezpieczeń można stosować niestandardowe reguły biznesowe zakodowane w klasie Języka Java. Podczas konfigurowania właściwości przeszukiwacza należy określić nazwę wtyczki, której przeszukiwacz ma używać podczas przeszukiwania dokumentów. Tokeny zabezpieczeń dodawane przez wtyczkę są przechowywane w indeksie i mogą być używane do kontrolowania dostępu do dokumentów.

Podczas konfigurowania określonych typów przeszukiwaczy można określić dodatkowe elementy sterujące zabezpieczeniami. Na przykład można określić, że podczas przetwarzania zapytania sprawdzane mają być referencje użytkownika. Po włączeniu tej opcji referencje użytkownika są porównywane z bieżącymi listami kontroli dostępu przechowywanymi w źródłach danych, w których wykonywane jest wyszukiwanie. Takie sprawdzanie poprawności bieżących referencji może być wykonywane obok lub zamiast sprawdzania poprawności opartego na tokenach zabezpieczeń w indeksie wyszukiwania korporacyjnego.

### Pojęcia pokrewne

“Zabezpieczenia na poziomie dokumentu” na stronie 261

### Zadania pokrewne

“Monitorowanie przeszukiwaczy” na stronie 304

### Odsyłacze pokrewne

## Tworzenie przeszukiwacza

Podczas tworzenia przeszukiwacza określany jest typ przeszukiwacza, który ma być utworzony. Kreator zapewnia pomoc w określaniu informacji o danych, które mają zostać włączone do kolekcji.

### Zagadnienia wstępne

Aby utworzyć przeszukiwacz, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należy przeszukiwacz.

### Informacje o tym zadaniu

Dla każdej kolekcji należy utworzyć przynajmniej jeden przeszukiwacz. Typ tworzonego przeszukiwacza zależy od typu danych, które mają zostać uwzględnione w kolekcji. Kreator właściwy dla typu tworzonego przeszukiwacza pomaga w definiowaniu opcji przeszukiwacza. Na przykład kreator zapewnia pomoc w określaniu opcji decydujących o sposobie wykorzystania przez przeszukiwacz zasobów systemowych. Kreator pomaga także w wybraniu źródeł danych, które mają zostać włączone do kolekcji.

### Procedura

Aby utworzyć przeszukiwacz, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Przeszukiwanie i kliknij opcję **Utwórz przeszukiwacz**.
2. Wybierz typ przeszukiwacza i odpowiednie wartości podstawowe:
  - a. Wybierz typ przeszukiwacza obsługujący typ danych, które mają zostać przeszukane, takich jak serwisy WWW, bazy danych Lotus Notes lub systemy plików UNIX.  
Po wybraniu typu przeszukiwacza wyświetlane są opcje dotyczące sposobu tworzenia tego przeszukiwacza.
  - b. Wybierz wartości podstawowe dla przeszukiwacza:

#### **Użyj dla nowego przeszukiwacza systemowych wartości domyślnych**

Wprowadzane są początkowe ustawienia przeszukiwacza zgodnie z domyślnymi wartościami instalacyjnymi.

Po wybraniu tej opcji kliknij przycisk **Dalej**, aby przystąpić do konfigurowania nowego przeszukiwacza.

#### **Klonuj wartości ustawień istniejącego przeszukiwacza dla nowego przeszukiwacza**

Wprowadzane są początkowe ustawienia przeszukiwacza zgodnie z wartościami skonfigurowanymi dla innego przeszukiwacza tego typu.

Po wybraniu tej opcji zostaje wyświetlona lista istniejących przeszukiwaczy tego samego typu. Wybierz przeszukiwacz, który ma być użyty do utworzenia nowego przeszukiwacza i kliknij przycisk **Dalej**, aby przystąpić do konfigurowania nowego przeszukiwacza.

Zostanie otwarty kreator odpowiedni dla typu tworzonego przeszukiwacza. Postępuj zgodnie z instrukcjami dotyczącymi tworzenia przeszukiwacza wyświetlanymi na ekranie kreatora. Aby uzyskać więcej informacji na temat opcji dotyczących danego typu przeszukiwacza, kliknij przycisk **Pomoc** na dowolnej stronie kreatora.



Nowo utworzony przeszukiwacz będzie wymieniony na stronie Przeszukiwanie wraz z innymi przeszukiwaczami należącymi do kolekcji. W dowolnie wybranym momencie można użyć opcji pozwalających zmodyfikować przestrzeń przeszukiwania i inne właściwości przeszukiwacza.

---

## Edycja właściwości przeszukiwacza

Istnieje możliwość zmiany informacji o przeszukiwaczu oraz sposobie przeszukiwania przez niego danych. Na przykład można zmienić sposób, w jaki przeszukiwacz korzysta z zasobów systemowych.

### Zagadnienia wstępne


Aby edytować właściwości przeszukiwacza, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należy przeszukiwacz.

### Informacje o tym zadaniu

Podczas edycji właściwości przeszukiwacza można kliknąć przycisk **Pomoc**, aby uzyskać więcej informacji na temat typu zmian, jakie można wprowadzić. Właściwości dostępne do edycji zależą od typu przeszukiwacza.

### Procedura

Aby zmodyfikować właściwości przeszukiwacza, należy wykonać poniższe czynności:

1. Otwórz kolekcję do edycji, wybierz stronę Przeszukiwanie, odszukaj przeszukiwacz, który ma być modyfikowany, i kliknij ikonę  **Właściwości przeszukiwacza**.
2. Zmodyfikuj właściwości przeszukiwacza i kliknij przycisk **OK**.
3. Aby zmiany zostały uwzględnione, zatrzymaj i ponownie uruchom przeszukiwacz. Jeśli zmiana dotyczyła wyłącznie opisu przeszukiwacza, nie trzeba go restartować.

---

## Edytowanie przestrzeni przeszukiwania

Istnieje możliwość zmiany informacji dotyczących źródeł danych przeszukiwanych przez przeszukiwacz. Na przykład można dodawać i usuwać źródła danych, zmieniać harmonogram przeszukiwania oraz reguły przeszukiwania dokumentów z określonego źródła danych.

### Zagadnienia wstępne


Aby edytować przestrzeń przeszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należy przeszukiwacz.

### Informacje o tym zadaniu

Aby w trakcie edycji przestrzeni przeszukiwania uzyskać więcej informacji na temat zmian, jakie można wprowadzić dla administrowanego typu przeszukiwacza, kliknij przycisk **Pomoc**.

### Procedura

Aby zmodyfikować przestrzeń przeszukiwania, należy wykonać poniższe czynności:

1. Otwórz kolekcję do edycji, wybierz stronę Przeszukiwanie, odszukaj przeszukiwacz, który ma być modyfikowany, i kliknij ikonę  **Przestrzeń przeszukiwania**.
2. Zmodyfikuj przestrzeń przeszukiwania, wybierając opcje, które mają zostać zmienione.



Dostępne do wybrania opcje zależą od typu przeszukiwacza. W przypadku niektórych operacji, takich jak dodawanie źródeł danych do kolekcji, otwierany jest kreator właściwy dla typu przeszukiwacza umożliwiający zmianę przestrzeni przeszukiwania.

3. Aby zmiany zostały uwzględnione, zatrzymaj i ponownie uruchom przeszukiwacz.

---

## Usuwanie przeszukiwacza

Usunięcie przeszukiwacza powoduje usunięcie wszystkich informacji o przeszukiwaczu z całego systemu wyszukiwania korporacyjnego. Informacje, które były uprzednio przeszukiwane przez przeszukiwacz, pozostają w indeksie do czasu następnej operacji budowania indeksu głównego.

### Zagadnienia wstępne

Aby usunąć przeszukiwacz, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należy przeszukiwacz.


### Informacje o tym zadaniu

Usuwanie przeszukiwacza może być czasochłonnym procesem. Po potwierdzeniu chęci usunięcia przeszukiwacza z systemu usuwane są wszystkie dane powiązane z przeszukiwaczem.

**Wskazówka:** Ponieważ wykonanie tego zadania wymaga czasu, może pojawić się komunikat o przekroczeniu limitu czasu przez żadaną operację - nawet wtedy, gdy proces nadal działa w tle. Aby sprawdzić, czy zadanie zostało zakończone, należy co pewien czas klikać opcję **Odśwież** na konsoli administracyjnej (nie należy klikać opcji **Odśwież** w przeglądarce WWW). Proces usuwania należy uznać za zakończony, gdy nazwa przeszukiwacza zniknie z listy przeszukiwaczy.

### Procedura

Aby usunąć przeszukiwacz, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji i wybierz stronę Przeszukiwanie.
2. Znajdź przeszukiwacz, który ma zostać usunięty, i kliknij ikonę  **Usuń**.

---

## Harmonogramy przeszukiwacza

Przeszukiwacze utworzone dla źródeł WWW działają w sposób ciągły. Po uruchomieniu przeszukiwacza WWW zazwyczaj nie ma potrzeby zatrzymywania go, o ile nie zostanie zmieniona jego konfiguracja. W przypadku wszystkich pozostałych typów przeszukiwaczy harmonogram przeszukiwania określa się podczas konfigurowania przeszukiwacza.

W przypadku niektórych typów źródeł danych do sterowania wizytami przeszukiwacza we wszystkich źródłach danych w przestrzeni przeszukiwania używany jest jeden harmonogram. Dla pozostałych typów źródeł danych można określić różne harmonogramy dla poszczególnych źródeł. Na przykład można określić różne harmonogramy przeszukiwania dla poszczególnych baz danych Lotus Notes przeszukiwanych przez ten przeszukiwacz.

Konfigurując harmonogram określa się typ przeszukiwania, które ma zostać wykonane. W harmonogramie można uwzględnić pełne przeszukiwanie wszystkich dokumentów w przestrzeni przeszukiwania, przeszukiwanie obejmujące wszystkie aktualizacje w przestrzeni przeszukiwania (nowe, zmodyfikowane i usunięte dokumenty) lub przeszukiwanie obejmujące wyłącznie nowe i zmodyfikowane dokumenty. Pełne przeszukiwanie zajmuje

najwięcej czasu. Przeszukiwanie usuwające dokumenty, które zostały usunięte, trwa dłużej niż przeszukiwanie ignorujące usunięte dokumenty.

Podczas edycji przestrzeni przeszukiwania danego przeszukiwacza istnieje możliwość określenia drugiego harmonogramu przeszukiwania. Na przykład można skonfigurować jeden harmonogram do przeszukiwania wszystkich dokumentów w przestrzeni przeszukiwania w każdą sobotę wieczorem oraz drugi harmonogram o większej częstotliwości do przeszukiwania nowych i zmodyfikowanych dokumentów.

Tworząc wiele harmonogramów przeszukiwania, użytkownik może lepiej sterować wizytami konkretnych przeszukiwaczy w docelowych źródłach danych. Na przykład podczas przeszukiwania baz danych w różnych strefach czasowych, można zaplanować wykonywanie przeszukiwania w godzinach wieczornych, po godzinach pracy w przedsiębiorstwie.

---

## Przeszukiwacze Content Edition

Aby do kolekcji wyszukiwania korporacyjnego włączyć repozytoria produktu IBM WebSphere Information Integrator Content Edition, należy skonfigurować przeszukiwacz Content Edition.

Przeszukiwacz Content Edition może służyć do przeszukiwania następujących typów repozytoriów:

- Documentum
- FileNet P8 Content Manager
- FileNet Panagon Content Services
- Hummingbird Document Management (DM)
- Microsoft SharePoint
- OpenText Livelink
- Portal Document Manager (PDM)

Podczas konfigurowania przeszukiwacza należy określić opcje przeszukiwania wszystkich repozytoriów w określonej przestrzeni przeszukiwania. Należy również wybrać klasy elementów, które mają być przeszukiwane w każdym repozytorium.

### Wskazówka:

Szczegółowe przykłady konfigurowania konektorów i bezpiecznego przeszukiwacza Content Edition można znaleźć w scenariuszu dla średniej organizacji w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

## Konfiguracja serwera przeszukiwacza

Sposób przygotowania do przeszukiwania repozytoriów zależy od tego, czy do połączenia z danymi do przeszukania ma być używany tryb bezpośredni, czy tryb serwera. W przypadku trybu bezpośredniego należy skonfigurować konektor w programie WebSphere Information Integrator Content Edition. W trybie serwera konieczne jest uruchomienie odpowiedniego skryptu na serwerze przeszukiwacza. Skrypt ten jest dostarczany razem z programem OmniFind Enterprise Edition i umożliwia przeszukiwaczowi Content Edition komunikowanie się z serwerami WebSphere Information Integrator Content Edition.

W przypadku trybu serwera przed utworzeniem przeszukiwacza Content Edition wykonaj zadanie, które jest odpowiednie dla danego środowiska:

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu WebSphere II Content Edition” na stronie 45.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows dla programu WebSphere II Content Edition” na stronie 46.

Szczegółowe instrukcje dotyczące konfigurowania systemu wyszukiwania korporacyjnego do wyszukiwania w repozytoriach WebSphere Information Integrator Content Edition zawiera artykuł na stronie IBM developerWorks *Search WebSphere Portal Document Manager using WebSphere Information Integrator OmniFind Edition* pod adresem <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0606lee/>.

## Przegląd konfiguracji

Aby utworzyć lub zmienić przeszukiwacz Content Edition, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich repozytoriów w przestrzeni przeszukiwania.
- Określenie, czy w celu uzyskania dostępu do repozytoriów przeszukiwacz używa trybu bezpośredniego czy też trybu serwera. W przypadku trybu serwera należy również określić informacje umożliwiające przeszukiwaczowi dostęp do serwera aplikacji WWW.
- Wybór repozytoriów do przeszukiwania.
- Określenie identyfikatorów użytkowników i haseł umożliwiających przeszukiwaczowi uzyskanie dostępu do treści wybranych repozytoriów.
- Skonfigurowanie harmonogramu przeszukiwania repozytoriów.
- Wybór klas elementów, które mają zostać przeszukane w każdym z repozytoriów.
- Określenie opcji udostępnienia właściwości klas elementów dla operacji przeszukiwania. Na przykład można wykluczyć określone typy dokumentów z przestrzeni przeszukiwania albo wybrać konkretną wersję repozytorium do przeszukania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie list kontroli dostępu lub tokenów zabezpieczeń. Dla klas elementów Documentum, FileNet Panagon Content Services, Hummingbird DM, Portal Document Manager i SharePoint można także wybrać opcję sprawdzania referencji użytkownika w momencie, gdy wprowadza on zapytanie. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych. Ten rodzaj sprawdzania bieżących referencji nie jest dostępny dla innych typów repozytoriów.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

## Dostęp do repozytoriów Content Edition w trybie bezpośrednim

Przeszukiwacz Content Edition można skonfigurować pod kątem dostępu do repozytoriów produktu WebSphere Information Integrator Content Edition w trybie bezpośrednim.

## Informacje o tym zadaniu

W trybie bezpośrednim przeszukiwacz używa konektora WebSphere Information Integrator Content Edition zainstalowanego na serwerze przeszukiwacza podczas instalacji produktu OmniFind Enterprise Edition. Przeszukiwacz korzysta z funkcji API integracji treści w celu uzyskania bezpośredniego połączenia z repozytoriami do przeszukania. Nie wszystkie funkcje serwera integracji treści są dostępne, gdy działa on w trybie bezpośrednim. Dokumentacja produktu WebSphere Information Integrator Content Edition zawiera informacje dotyczące uruchamiania serwera integracji treści w trybie bezpośrednim oraz różnicy funkcjonalności w porównaniu z serwerem integracji treści działającym w trybie serwera.

Procedura ta podsumowuje czynności wymagane do skonfigurowania dostępu za pomocą trybu bezpośredniego. Szczegółowe instrukcje zawiera artykuł na stronie IBM developerWorks *Search WebSphere Portal Document Manager using WebSphere Information Integrator OmniFind Edition* pod adresem <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0606lee/>.

## Procedura

Aby tak skonfigurować system, żeby przeszukiwacz miał dostęp do repozytoriów w trybie bezpośrednim:

1. Upewnij się, że zmienne środowiskowe VBR\_HOME i JAVA\_HOME w pliku *główny\_katalog\_instalacji\_iice/bin/config.sh* (w systemie UNIX) lub w pliku *główny\_katalog\_instalacji\_iice/bin/config.bat* (w systemie Microsoft Windows) określają poprawny katalog.
2. Aby skonfigurować konsolę administracyjną WebSphere Information Integrator Content Edition do działania w trybie bezpośrednim, dodaj właściwość systemową Java **-Dvbr.as.operationMode=direct** do pliku *iice\_install\_root/bin/Admin.sh* (w systemach UNIX) lub pliku *iice\_install\_root/bin/Admin.bat* (w systemach Windows). Na przykład:

### Plik Admin.sh

```
java -classpath \
"$VBR_CLASSPATH" \
-Dvbr.home="$VBR_HOME" \
-Dvbr.as.operationMode=direct \
-Dlog4j.category.com.venetica.vbr.tools.admin=WARN \
com.venetica.vbr.tools.admin.AdminFrame $1 $2 $3 $4
```

### Plik Admin.bat

```
java -classpath "%VBR_CLASSPATH%" ^
-Dvbr.home="%VBR_HOME%" ^
-Dvbr.as.operationMode=direct ^
-Dlog4j.category.com.venetica.vbr.tools.admin=WARN ^
com.venetica.vbr.tools.admin.AdminFrame %*
```

3. Uruchom konsolę administracyjną WebSphere Information Integrator Content Edition w trybie bezpośrednim i skonfiguruj konektor dla serwera przeszukiwacza OmniFind Enterprise Edition
4. Do skonfigurowania przeszukiwacza Content Edition podczas korzystania z konsoli administrowania wyszukiwaniem korporacyjnym wybierz opcję trybu bezpośredniego.

## Dostęp do repozytoriów WebSphere II Content Edition w trybie serwera

Przeszukiwacz Content Edition można skonfigurować pod kątem dostępu do repozytoriów w trybie serwera.

W trybie serwera konektor WebSphere Information Integrator Content Edition używany przez przeszukiwacz w celu uzyskiwania dostępu do danych jest zainstalowany jako aplikacja

korporacyjna na serwerze WebSphere Application Server, a przeszukiwacz ma dostęp do repozytoriów za pośrednictwem tego serwera. Takie podejście umożliwia korzystanie z zalet środowisk serwerów aplikacji J2EE.

Przed skonfigurowaniem przeszukiwacza pod kątem dostępu do repozytoriów WebSphere Information Integrator Content Edition w trybie serwera konieczne jest uruchomienie odpowiedniego skryptu na serwerze przeszukiwacza. Skrypt ten, dostarczany razem z produktem OmniFind Enterprise Edition, umożliwia przeszukiwaczowi Content Edition dostęp do repozytoriów na serwerze.

Wykonaj odpowiednie zadanie dla używanego środowiska:

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu WebSphere II Content Edition”.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows dla programu WebSphere II Content Edition” na stronie 46.

## **Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu WebSphere II Content Edition**

Jeśli program OmniFind Enterprise Edition zostanie zainstalowany na komputerze z systemem IBM AIX, Linux lub ze Środowiskiem Operacyjnym Solaris, a przeszukiwacz Content Edition zostanie skonfigurowany do używania trybu serwera podczas uzyskiwania dostępu do repozytoriów, należy uruchomić skrypt konfiguracyjny serwera przeszukiwacza. Ten skrypt umożliwia przeszukiwaczowi Content Edition uzyskanie dostępu do repozytoriów programu WebSphere Information Integrator Content Edition.

### **Informacje o tym zadaniu**

Przeszukiwacz Content Edition korzysta z bibliotek języka Java programu WebSphere Information Integrator Content Edition jako klient Java. W trybie serwera te biblioteki Java wymagają bibliotek Java związanych z komponentami EJB serwera WebSphere Application Server. Aby upewnić się, że przeszukiwacz Content Edition może współpracować z bibliotekami Java, należy uruchomić skrypt konfiguracyjny, który program OmniFind Enterprise Edition udostępnia na serwerze przeszukiwacza po zainstalowaniu serwera WebSphere Application Server.

Program WebSphere Information Integrator Content Edition jest instalowany na serwerze przeszukiwacza podczas instalacji produktu OmniFind Enterprise Edition. Aby móc korzystać z przeszukiwacza Content Edition w trybie serwera, należy skopiować plik `vbr_access_services.jar` z serwera WebSphere Information Integrator Content Edition na serwer przeszukiwacza.

### **Procedura**

Aby skonfigurować serwer przeszukiwacza na potrzeby przeszukiwania repozytoriów programu WebSphere Information Integrator Content Edition:

1. Jeśli program OmniFind Enterprise Edition został zainstalowany na wielu serwerach, zainstaluj i powiąż biblioteki Java serwera WebSphere Application Server.
2. Na serwerze przeszukiwacza uruchom skrypt konfiguracyjny przeszukiwacza Content Edition:
  - a. Zaloguj się jako administrator wyszukiwania korporacyjnego.
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu `$ES_INSTALL_ROOT/bin`, i odpowiedz na pytania:  
`escrvbr.sh`

3. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
esadmin system startall
```

4. Skopiuj plik `vbr_access_services.jar` z serwera WebSphere Information Integrator Content Edition na serwer przeszukiwacza.

**Kopiuj z:**

Plik `vbr_access_services.jar` znajduje się w następującym domyślnym położeniu:

```
katalog_instalacyjny_was/installedApps/nazwa_serwera/nazwa_aplikacji
```

*katalog\_instalacyjny\_was* jest katalogiem instalacyjnym serwera WebSphere Application Server, *nazwa\_serwera* jest nazwą serwera określoną przez użytkownika, a *nazwa\_aplikacji* jest określoną przez użytkownika nazwą aplikacji WebSphere Information Integrator Content Edition na serwerze WebSphere Application Server.

**Kopiuj do:**

Katalog docelowy na serwerze przeszukiwacza to *katalog\_instalacyjny\_iice/lib*, gdzie *katalog\_instalacyjny\_iice* jest katalogiem instalacyjnym programu WebSphere Information Integrator Content Edition na serwerze przeszukiwacza.

## Konfigurowanie serwera przeszukiwacza w systemie Windows dla programu WebSphere II Content Edition

W przypadku zainstalowania programu OmniFind Enterprise Edition na komputerze z systemem Microsoft Windows i skonfigurowania przeszukiwacza Content Edition do korzystania z trybu serwera podczas uzyskiwania dostępu do repozytoriów, należy uruchomić skrypt konfiguracyjny serwer przeszukiwacza. Ten skrypt umożliwi przeszukiwaczowi Content Edition uzyskanie dostępu do repozytoriów programu WebSphere Information Integrator Content Edition.

### Informacje o tym zadaniu

Przeszukiwacz Content Edition korzysta z bibliotek języka Java programu WebSphere Information Integrator Content Edition jako klient Java. W trybie serwera te biblioteki Java wymagają bibliotek Java związanych z komponentami EJB serwera WebSphere Application Server. Aby upewnić się, że przeszukiwacz Content Edition może współpracować z bibliotekami Java, należy uruchomić skrypt konfiguracyjny, który program OmniFind Enterprise Edition udostępnia na serwerze przeszukiwacza po zainstalowaniu serwera WebSphere Application Server.

Program WebSphere Information Integrator Content Edition jest instalowany na serwerze przeszukiwacza podczas instalacji produktu OmniFind Enterprise Edition. Aby móc korzystać z przeszukiwacza Content Edition w trybie serwera, należy skopiować plik `vbr_access_services.jar` z serwera WebSphere Information Integrator Content Edition na serwer przeszukiwacza.

### Procedura

Aby skonfigurować serwer przeszukiwacza na potrzeby przeszukiwania repozytoriów programu WebSphere Information Integrator Content Edition:

1. Jeśli program OmniFind Enterprise Edition został zainstalowany na wielu serwerach, zainstaluj i powiąz biblioteki Java serwera WebSphere Application Server.



2. Na serwerze przeszukiwacza uruchom skrypt konfiguracyjny przeszukiwacza Content Edition:
  - a. Zaloguj się przy użyciu identyfikatora administratora wyszukiwania korporacyjnego (ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu %ES\_INSTALL\_ROOT%\bin, i odpowiedz na pytania:  
escrivr.vbs
3. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
esadmin system startall
```

4. Skopiuj plik vbr\_access\_services.jar z serwera WebSphere Information Integrator Content Edition na serwer przeszukiwacza.

**Kopiuj z:**

Plik vbr\_access\_services.jar znajduje się w następującym domyślnym położeniu:

*katalog\_instalacyjny\_was*/installedApps/*nazwa\_serwera*/*nazwa\_aplikacji*

*katalog\_instalacyjny\_was* jest katalogiem instalacyjnym serwera WebSphere Application Server, *nazwa\_serwera* jest nazwą serwera określoną przez użytkownika, a *nazwa\_aplikacji* jest określoną przez użytkownika nazwą aplikacji WebSphere Information Integrator Content Edition na serwerze WebSphere Application Server.

**Kopiuj do:**

Katalog docelowy na serwerze przeszukiwacza to *katalog\_instalacyjny\_iice*\lib, gdzie *katalog\_instalacyjny\_iice* jest katalogiem instalacyjnym programu WebSphere Information Integrator Content Edition na serwerze przeszukiwacza.

---

## Przeszukiwacze DB2

Przeszukiwacz DB2 służy do włączenia baz danych IBM DB2 do kolekcji.

Jeśli używany jest program IBM WebSphere Information Integrator w celu stowarzyszenia i utworzenia tabel pseudonimów dla poniższych rodzajów baz danych, można użyć przeszukiwacza DB2 do przeszukania tabel za pomocą pseudonimów:

- CA-Datacom
- IBM DB2 for z/OS
- DB2 for iSeries
- IBM Informix
- IMS
- Oracle
- Microsoft SQL Server
- Software AG Adabas
- Sybase
- VSAM

Należy skonfigurować oddzielne przeszukiwacze dla każdego serwera bazy danych, który ma zostać przeszukany. Podczas konfigurowania przeszukiwacza należy określić opcje dotyczące wykorzystywanej w przeszukiwaczu metody przeszukiwania wszystkich baz danych na określonym serwerze. Należy również określić tabele, które mają zostać przeszukane w każdej z baz danych.



Tabele wybrane do przeszukiwania muszą być tabelami baz danych, tabelami pseudonimów lub widokami. Przeszukiwacz DB2 nie obsługuje połączonych tabel.

#### **Wskazówka:**

Szczegółowe przykłady konfigurowania bezpiecznego przeszukiwacza DB2 można znaleźć w scenariuszu dla dużej organizacji w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

## **Konfiguracja serwera przeszukiwacza**

Aby można było przeszukiwać tabele baz danych, należy na serwerze przeszukiwacza zainstalować klienta DB2 Administration Client. Następnie należy uruchomić skrypt na serwerze przeszukiwacza. Skrypt ten jest dostarczany razem z programem OmniFind Enterprise Edition i umożliwia przeszukiwaczowi DB2 komunikowanie się z serwerami baz danych.

Przed skonfigurowaniem przeszukiwacza DB2 za pomocą konsoli administracyjnej wyszukiwania korporacyjnego należy wykonać zadanie odpowiednie dla używanego środowiska:

- “Konfigurowanie serwera przeszukiwania w systemie UNIX do przeszukiwania źródeł DB2” na stronie 49.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł DB2” na stronie 50.

## **Publikowanie zdarzeń**

Jeśli używany jest produkt WebSphere Information Integrator Event Publisher Edition, a bazy danych do przeszukiwania powiązane z mapami kolejek publikowania, przeszukiwacz DB2 może użyć tych map do przeszukiwania aktualizacji tabel baz danych.

Mapa kolejek publikowania identyfikuje kolejkę WebSphere MQ, która otrzymuje komunikaty XML, gdy publikowane są aktualizacje tabeli bazy danych. Przeszukiwacz nasłuchuje w kolejce informacji na temat opublikowanych zdarzeń i aktualizuje przestrzeń przeszukiwania w przypadku aktualizacji tabel (przy pierwszym przeszukiwaniu tabeli przez przeszukiwacz przeszukiwane są wszystkie dokumenty).

Publikowanie zdarzeń umożliwia szybsze udostępnianie do wyszukiwania nowych i zmienionych dokumentów niż w przypadku dokumentów przeszukiwanych przez przeszukiwacz zgodnie z harmonogramem.

Jeśli niektóre lub wszystkie tabele są skonfigurowane do używania funkcji publikowania zdarzeń, podczas konfigurowania przeszukiwacza można określić informacje, które umożliwią mu dostęp do produktu WebSphere MQ oraz map kolejek publikowania.

Należy również upewnić się, że na serwerze, który będzie przeszukiwany, skonfigurowane są produkty WebSphere MQ i WebSphere Information Integrator Event Publisher Edition oraz że na serwerze przeszukiwacza skonfigurowany jest moduł klienta WebSphere MQ. Aby używać funkcji publikowania zdarzeń z przeszukiwaczem DB2, należy wykonać następujące zadania:

- “Konfigurowanie programu WebSphere MQ dla przeszukiwacza DB2” na stronie 54.
- “Konfigurowanie programu WebSphere Information Integrator Event Publisher Edition dla przeszukiwacza DB2” na stronie 51.

## Przegląd konfiguracji

Aby utworzyć lub zmienić przeszukiwacz DB2, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania przez niego wszystkich baz danych na określonym serwerze bazy danych.

- Określenie informacji o typach baz danych, które mają zostać przeszukane.

Jeśli mają być przeszukiwane zdalne bazy danych, które nie są wpisane do katalogu na lokalnym serwerze baz danych, należy na serwerze zdalnym uruchomić serwer DB2 Administration Server, zanim będzie można używać przeszukiwacza DB2 w celu ich przeszukiwania. Podczas konfigurowania przeszukiwacza należy podać nazwę hosta i numer portu na zdalnym serwerze bazy danych.

- Określ bazy danych do przeszukania.
- Określenie identyfikatorów użytkowników i haseł pozwalających przeszukiwaczowi uzyskać dostęp do baz danych, w których używane są funkcje kontroli dostępu.
- Skonfigurowanie harmonogramu przeszukiwania baz danych.
- Wybór tabel, które mają zostać przeszukane w każdej bazie danych.

**Ważne:** Aby zoptymalizować wydajność procesów wykrywania (i zapobiec przekroczeniu limitu czasu przez proces konfigurowania przeszukiwacza), opcję przeszukiwania wszystkich tabel należy wybierać tylko wtedy, gdy baza danych nie zawiera zbyt wielu tabel lub gdy tabele nie zawierają zbyt wielu kolumn. Tabele do przeszukania można wybrać od razu, natomiast edycja przestrzeni przeszukiwania i dodanie dodatkowych tabel do kolekcji będzie możliwe również w późniejszym czasie.

- Wybór tabel, które mają być przeszukiwane w przypadku opublikowania dla nich aktualizacji w kolejce publikowania zdarzeń, a także określenie informacji umożliwiających przeszukiwaczowi dostęp do kolejki publikowania zdarzeń.
- Określenie opcji udostępnienia kolumn w określonych tabelach dostępnych dla operacji przeszukiwania. Na przykład można uwzględnić pewne kolumny w zapytaniach parametrycznych lub można określić, które kolumny mogą być zwracane w wynikach wyszukiwania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

## Konfigurowanie serwera przeszukiwania w systemie UNIX do przeszukiwania źródeł DB2

W przypadku zainstalowania programu OmniFind Enterprise Edition na komputerze z systemem IBM AIX, Linux lub w Środowisku Operacyjnym Solaris należy uruchomić odpowiedni skrypt w celu skonfigurowania serwera przeszukiwacza. Skrypt umożliwia przeszukiwaczowi DB2 komunikowanie się z serwerami bazy danych. W przypadku korzystania z funkcji publikowania zdarzeń skrypt umożliwia też przeszukiwaczowi dostęp do menedżerów kolejek WebSphere MQ oraz do samych kolejek.

### Informacje o tym zadaniu

Aby mieć pewność, że przeszukiwacz DB2 może przeszukiwać tabele bazy danych, należy uruchomić skrypt konfiguracyjny `escrdb2.sh` instalowany razem z programem OmniFind Enterprise Edition na serwerze przeszukiwania.

Przed uruchomieniem skryptu należy się upewnić, że na serwerze przeszukiwacza jest zainstalowany klient DB2 Administration Client.

W przypadku korzystania z funkcji publikowania zdarzeń należy zainstalować moduły programu WebSphere MQ 5.3 dla środowiska Java Messaging na serwerze przeszukiwacza, dzięki czemu przeszukiwacz DB2 uzyska dostęp do menedżerów kolejek i samych kolejek programu WebSphere MQ. Po instalacji modułów programu WebSphere MQ należy uruchomić skrypt `escrdb2.sh`.

### Procedura

Aby skonfigurować serwer przeszukiwacza i umożliwić działanie przeszukiwaczom DB2:

1. Opcjonalnie: jeśli zamierzasz korzystać z funkcji publikowania zdarzeń, na serwerze przeszukiwacza zainstaluj moduły programu WebSphere MQ 5.3 komponentu Java Messaging:
  - a. Zaloguj się jako administrator i wprowadź następującą komendę:

```
export LD_ASSUME_KERNEL=2.4.19
```
  - b. Włóż dysk CD z programem WebSphere MQ.
  - c. Przejdź do katalogu, w którym znajdują się moduły MQ dla usługi Java Messaging.
  - d. Wprowadź następującą komendę w celu zainstalowania modułów:

```
rpm -i MQSeriesJava-5.3.0-1.i386.rpm
```
2. Na serwerze przeszukiwacza uruchom skrypt konfiguracyjny przeszukiwacza DB2:
  - a. Zaloguj się jako administrator wyszukiwania korporacyjnego (z identyfikatorem użytkownika określonym podczas instalowania programu OmniFind Enterprise Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu `$ES_INSTALL_ROOT/bin`, i odpowiedz na zachęty:

```
escrdb2.sh
```
3. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
esadmin system startall
```

## Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł DB2

W przypadku instalowania programu OmniFind Enterprise Edition na komputerze z systemem Microsoft Windows należy uruchomić skrypt w celu skonfigurowania serwera przeszukiwacza. Skrypt umożliwia przeszukiwaczowi DB2 komunikowanie się z serwerami bazy danych. W przypadku korzystania z funkcji publikowania zdarzeń skrypt umożliwia też przeszukiwaczowi dostęp do menedżerów kolejek WebSphere MQ oraz do samych kolejek.

### Informacje o tym zadaniu

Aby upewnić się, że przeszukiwacz DB2 może przeszukiwać tabele bazy danych, należy uruchomić skrypt konfiguracyjny `escrdb2.vbs` instalowany razem z programem OmniFind Enterprise Edition na serwerze przeszukiwania.

Przed uruchomieniem skryptu należy się upewnić, że na serwerze przeszukiwacza jest zainstalowany klient DB2 Administration Client.

W przypadku korzystania z funkcji publikowania zdarzeń należy zainstalować moduły programu WebSphere MQ 5.3 dla środowiska Java Messaging na serwerze przeszukiwacza, dzięki czemu przeszukiwacz DB2 uzyska dostęp do menedżerów kolejek i samych kolejek programu WebSphere MQ. Po instalacji modułów programu WebSphere MQ należy uruchomić skrypt `escrdb2.vbs`.

### Procedura

Aby skonfigurować serwer przeszukiwacza i umożliwić działanie przeszukiwaczom DB2:

1. Opcjonalnie: jeśli zamierzasz korzystać z funkcji publikowania zdarzeń, na serwerze przeszukiwacza zainstaluj moduły programu WebSphere MQ 5.3 komponentu Java Messaging:
  - a. Włóż dysk CD z programem WebSphere MQ.
  - b. Uruchom instalator programu WebSphere MQ.
  - c. W oknie wyboru opcji produktu wybierz opcję instalacyjną **Java Messaging**.
2. Na serwerze przeszukiwacza uruchom skrypt konfiguracyjny przeszukiwacza DB2:
  - a. Zaloguj się przy użyciu identyfikatora administratora wyszukiwania korporacyjnego (ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu `%ES_INSTALL_ROOT%\bin`, i odpowiedz na pytania:  
`escrdb2.vbs`
3. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
esadmin system startall
```

## Konfigurowanie programu WebSphere Information Integrator Event Publisher Edition dla przeszukiwaczy DB2

Przed skonfigurowaniem przeszukiwacza DB2 do użycia funkcji publikowania zdarzeń należy się upewnić, że na serwerze do przeszukania został skonfigurowany program IBM WebSphere Information Integrator Event Publisher Edition.

### Informacje o tym zadaniu

Podczas konfigurowania programu WebSphere Information Integrator Event Publisher Edition do użycia z przeszukiwaczem DB2 należy zastosować się do następujących wytycznych:

- Do opublikowania muszą zostać wybrane zarówno zmienione, jak i niezmienione kolumny w tabelach źródłowych.
- Do opublikowania muszą zostać wybrane usunięte wiersze w tabelach źródłowych.
- Kolejka publikacji zdarzeń nie może być współużytkowana przez wiele baz danych.
- Pojedyncza baza danych może mieć wiele kolejek i map kolejek.

- Tabela powinna mieć jedną publikację XML powiązaną z jedną mapą kolejek publikowania. (Tabela nie powinna mieć więcej niż jednej publikacji XML powiązanej z pojedynczą mapą kolejek publikowania. Może mieć więcej niż jedną publikację XML, jeśli każda z nich jest powiązana z inną mapą kolejek publikowania).

## Procedura

Aby skonfigurować serwer bazy danych w celu umożliwienia przeszukiwaczowi DB2 dostępu do aktualizacji tabel publikowanych w kolejce publikowania zdarzeń, należy wykonać następujące czynności. (W dokumentacji programu WebSphere Information Integrator Publisher Edition został podany szczegółowy opis tych czynności).

1. Zainstaluj program WebSphere Information Integrator Event Publisher Edition na serwerze bazy danych, który ma być przeszukiwany.
2. Uruchom Starter Centrum replikacji:

**AIX, Linux i Solaris**  
db2rc

### Wiersz komend systemu Windows

Kliknij polecenia **Start** → **Centrum replikacji IBM DB2**.

3. Utwórz tabele sterujące kolejkowego programu przechwytyjącego zmiany:
  - a. Wybierz opcję **Publikowanie zdarzeń** jako widok startera, wybierz opcję **Utwórz tabele sterujące kolejkowym programem przechwytyjącego zmiany**, a następnie kliknij przycisk **Dalej**.
  - b. W polu **Kolejkowy serwer przechwytywania** wybierz z listy dostępnych serwerów baz danych serwer, który ma być używany jako kolejkowy serwer przechwytywania, a następnie kliknij przycisk **OK**.
  - c. Określ identyfikator i hasło użytkownika, który może uzyskiwać dostęp do wybranego kolejkowego serwera przechwytywania. Zmień schemat kolejkowego programu przechwytyjącego zmiany lub zaakceptuj domyślną nazwę schematu i kliknij przycisk **Dalej**.
  - d. Określ nazwę menedżera kolejek, kolejki administracyjnej i kolejki restartu podane podczas konfiguracji programu WebSphere MQ na tym serwerze bazy danych, a następnie kliknij przycisk **Dalej**.
  - e. Kliknij przycisk **Zakończ**. Po wyświetleniu strony z komunikatami i skryptami SQL kliknij przycisk **Zamknij**.
  - f. Jako opcję przetwarzania wybierz pozycję **Uruchom teraz** i kliknij przycisk **OK**. Po wyświetleniu komunikatu informującego o zakończeniu wykonywania skryptów SQL kliknij przycisk **Zamknij**.
4. Utwórz publikację XML:
  - a. W Starterze Centrum replikacji wybierz opcję **Publikowanie zdarzeń** jako widok startera, wybierz opcję **Utwórz publikację XML**, a następnie kliknij przycisk **Dalej**.
  - b. Na stronie Uruchom kliknij przycisk **Dalej**.
  - c. Na stronie Mapa serwerów i kolejek potwierdź poprawność kolejkowego serwera przechwytywania i schematu kolejkowego programu przechwytyjącego zmiany, kliknij opcję obok pola **Mapa kolejek publikowania** i kliknij przycisk **Nowa**, aby utworzyć mapę kolejek publikowania.
  - d. Na stronie Ogólne wpisz nazwę mapy kolejek.
  - e. Na stronie Właściwości określ nazwę kolejki nadawczej (taką jak nazwa kolejki danych określona podczas konfigurowania programu WebSphere MQ na tym serwerze), wybierz opcję **Operacja na wierszach** lub **Transakcja** jako typ treści komunikatu, usuń zaznaczenie pól wyboru wysyłania komunikatów pulsu i dodawania nagłówków komunikatów JMS, a następnie kliknij przycisk **OK**.

- f. Po wyświetleniu strony z komunikatami i skryptami SQL kliknij przycisk **Zamknij**.
  - g. Jako opcję przetwarzania wybierz pozycję **Uruchom teraz** i kliknij przycisk **OK**. Po wyświetleniu komunikatu informującego o zakończeniu wykonywania skryptów SQL kliknij przycisk **Zamknij**.
  - h. Na stronie Wybierz mapę kolejek publikowania wybierz utworzoną mapę kolejek i kliknij przycisk **OK**.
  - i. Na stronie Serwer i mapa kolejek potwierdź poprawność nazwy mapy kolejek i kliknij przycisk **Dalej**.
  - j. Na stronie Tabela źródłowa kliknij przycisk **Dodaj**, kliknij przycisk **Pobierz wszystkie**, wybierz tabelę, dla której ma zostać włączone publikowanie zdarzeń, kliknij przycisk **OK**, a następnie przycisk **Dalej**.
  - k. Na stronie Kolumny i wiersze wybierz kolumny, które ma przeszukiwać przeszukiwacz DB2 (lub wszystkie kolumny) i wybierz kolumny kluczy. Na stronie, na której zostały wybrane wiersze do przeszukania (lub wszystkie wiersze), wybierz opcję publikowania operacji usuwania tabeli źródłowej. Po zakończeniu konfiguracji tych opcji kliknij przycisk **Dalej**.
  - l. Na stronie Treść komunikatu wybierz opcję włączania do danych kolumn zarówno kolumn zmienionych, jak i niezmienionych, i wybierz opcję dotyczącą tylko nowych wartości danych. Upewnij się, że pole wyboru dla automatycznego uruchamiania publikacji XML jest zaznaczone, i kliknij przycisk **Dalej**.
  - m. Na stronie Przejrzyj i zakończ publikację XML kliknij przycisk **Dalej**.
  - n. Na stronie Podsumowanie kliknij przycisk **Zakończ**. Po wyświetleniu strony z komunikatami i skryptami SQL kliknij przycisk **Zamknij**.
  - o. Jako opcję przetwarzania wybierz pozycję **Uruchom teraz** i kliknij przycisk **OK**. Po wyświetleniu komunikatu informującego o zakończeniu wykonywania skryptów SQL kliknij przycisk **Zamknij**.
5. Uruchom kolejkowy serwer przechwytywania:
- a. Zamknij Starter Centrum replikacji i uruchom Centrum replikacji.
  - b. W drzewie obiektów kliknij kolejno pozycje: **Replikacja kolejkowa** → **Definicje** → **Kolejkowe serwery przechwytywania**.
  - c. Kliknij prawym przyciskiem myszy ikonę skonfigurowanego kolejkowego serwera przechwytywania i wybierz opcję **Włącz replikację kolejkową bazy danych**.
  - d. Po wyświetleniu komunikatu ostrzegawczego kliknij przycisk **OK**.
  - e. Po wyświetleniu strony z komunikatami DB2 kliknij przycisk **Zamknij**.
  - f. W drzewie obiektów kliknij prawym przyciskiem myszy ikonę kolejkowego serwera przechwytywania i wybierz opcję **Uruchom kolejkowy program przechwytyjący zmiany**.
  - g. Jako opcję przetwarzania wybierz pozycję **Uruchom teraz**, określ nazwę systemu, identyfikator i hasło użytkownika bazy danych DB2, ścieżkę do katalogu, w którym są przechowywane dzienniki, oraz nazwę instancji bazy danych DB2, a następnie kliknij przycisk **OK**.
  - h. Po wyświetleniu komunikatu informującego o wprowadzeniu żądania kliknij przycisk **Zamknij**.
  - i. W drzewie obiektów kliknij prawym przyciskiem myszy ikonę kolejkowego serwera przechwytywania i wybierz opcję **Sprawdź status**.  
Zostanie wyświetlony status kolejkowego serwera przechwytywania. Jeśli wystąpią błędy, komunikat o statusie będzie informował, że serwer jest prawdopodobnie wyłączony. Aby sprawdzić dzienniki i określić przyczynę błędów, wprowadź następującą komendę w wierszu komend:  
`asnqcap Capture_Server=nazwa serwera przechwytywania LOGSTDOUT=y`



## Konfigurowanie programu WebSphere MQ dla przeszukiwaczy DB2

Przed skonfigurowaniem przeszukiwacza DB2 do korzystania z funkcji publikowania zdarzeń upewnij się, że na serwerze, którego przeszukiwacz będzie nasłuchiwać, skonfigurowany jest program IBM WebSphere MQ.

### Zagadnienia wstępne

Należy upewnić się, że na docelowym serwerze bazy danych zostały zainstalowane programy DB2, WebSphere Information Integrator Event Publisher Edition i WebSphere MQ.

### Ograniczenia

Jeśli docelowy serwer bazy danych jest zainstalowany na komputerze z systemem Linux, wszyscy użytkownicy programów DB2, WebSphere MQ oraz OmniFind Enterprise Edition muszą zdefiniować ustawienie następującej zmiennej środowiskowej:

```
export LD_ASSUME_KERNEL=2.4.19
```

Ta zmienna środowiskowa umożliwi wyeksportowanie implementacji wątków LinuxThread z dowolnej powłoki, w której została przeprowadzona instalacja, zostały wprowadzone komendy sterujące programu WebSphere MQ lub zostały uruchomione aplikacje programu WebSphere MQ. Program WebSphere MQ wymaga wyeksportowania tej zmiennej środowiskowej.

Przeszukiwacze DB2, które korzystają z publikowania zdarzeń, nawiązują połączenie klienckie z kolejkami programu WebSphere MQ. Aby umożliwić nawiązywanie połączeń klienckich, zaloguj się jako administrator programu WebSphere MQ i uruchom następującą komendę w celu ustawienia identyfikatora CCSID menedżera kolejki na wartość 819:

```
runmqsc nazwa_menedzera_kolejki
ALTER QMGR CCSID(819)
END
```

### Informacje o tym zadaniu

Przeszukiwacz DB2 obsługuje tryb połączeń klientów z serwerem programu WebSphere MQ. Przeszukiwacz nasłuchuje komunikatów XML, które są publikowane w kolejce publikowania zdarzeń. Przeszukiwacz nie może nasłuchiwać komunikatów XML przesyłanych przez więcej niż jedną kolejkę.

Po skonfigurowaniu programu WebSphere MQ przeszukiwacz DB2 korzysta z nazwy menedżera kolejek, nazwy kolejki, nazwy hosta serwera, numeru portu serwera oraz nazwy kanału serwera w celu uzyskania komunikatów XML z kolejki publikowania. Przeszukiwacz analizuje komunikat i aktualizuje przestrzeń przeszukiwania informacjami o zaktualizowanych tabelach.

### Procedura

W celu skonfigurowania serwera bazy danych w taki sposób, aby przeszukiwacz DB2 mógł nasłuchiwać kolejki publikowania zdarzeń, należy wykonać następujące czynności. (W dokumentacji programu WebSphere MQ został podany szczegółowy opis tych czynności).

1. Zaloguj się jako administrator programu WebSphere MQ i wprowadź następujące komendy, aby utworzyć menedżera kolejek i kolejki.
  - a. W wierszu komend wprowadź następującą komendę:

```
crtmqm QM1
```



- b. Po wyświetleniu komunikatu wprowadź następującą komendę:  
strmqm QM1
  - c. Po wyświetleniu komunikatu Uruchomiono menedżera kolejek 'QM1' wprowadź następującą komendę:  
runmqsc QM1
  - d. Po wyświetleniu komunikatu Uruchamianie poleceń MQSC dla menedżera kolejek QM1 wprowadź następującą komendę, aby utworzyć kolejkę administracyjną:  
DEFINE QLOCAL('ASN.QM1.ADMINQ')
  - e. Po wyświetleniu komunikatu Utworzono kolejkę komunikatów programu WebSphere wprowadź następującą komendę, aby utworzyć kolejkę restartowania:  
DEFINE QLOCAL(' ASN.QM1.RESTARTQ')
  - f. Po ponownym wyświetleniu komunikatu Utworzono kolejkę komunikatów programu WebSphere wprowadź następującą komendę, aby utworzyć kolejkę danych:  
DEFINE QLOCAL(' ASN.QM1.DATAQ')
  - g. Po ponownym wyświetleniu komunikatu Utworzono kolejkę komunikatów programu WebSphere wprowadź następującą komendę, aby wyjść:  
end
2. Wprowadź następującą komendę, aby uruchomić program MQ Listener na serwerze bazy danych (program MQ Listener musi być uruchomiony podczas tworzenia przeszukiwacza DB2, który korzysta z funkcji publikowania zdarzeń). W tym przykładzie 1414 to numer portu serwera. Używany jest domyślny kanał SYSTEM.DEF.SVRCONN:  
runmqtsr -m QM1 -t TCP -p 1414 &
  3. Wprowadź następującą komendę, aby autoryzować dostęp użytkownika DB2 do kolejek i menedżera kolejek za pomocą interfejsu MQI (Message Queuing Interface) dla publikowania zdarzeń (w tym przykładzie identyfikator użytkownika to db2inst1):  
  
setmqaut -m QM1 -t qmgr -p db2inst1 +allmqi  
setmqaut -m QM1 -t queue -n ASN.QM1.DATAQ -p db2inst1 +allmqi  
setmqaut -m QM1 -t queue -n ASN.QM1.ADMINQ -p db2inst1 +allmqi  
setmqaut -m QM1 -t queue -n ASN.QM1.RESTARTQ -p db2inst1 +allmqi
  4. Wprowadź następujące komendy dla identyfikatora użytkownika, który został użyty do utworzenia i uruchomienia przeszukiwacza DB2 z publikowaniem zdarzeń. Te komendy autoryzują dostęp ID użytkownika do menedżera kolejek i kolejek za pomocą interfejsu MQI w celu publikowania zdarzeń. W tym przypadku identyfikator użytkownika to esuser:  
  
setmqaut -m ASN.QM1.QM2 -t qmgr -p esuser +allmqi  
setmqaut -m ASN.QM1.QM2 -t queue -n ASN.QM1.DATAQ  
-p esuser +allmqi

## Przeszukiwanie baz danych DB2 na klasycznym serwerze źródła danych

Przeszukiwacz DB2 może przeszukiwać bazę danych DB2 na klasycznym serwerze źródła danych za pośrednictwem produktu WebSphere Information Integrator Classic Federation.

### Informacje o tym zadaniu

Aby możliwe było przeszukiwanie bazy danych DB2 na klasycznym serwerze źródła danych, baza danych musi być stowarzyszona z bazą danych DB2 na nieklasycznym serwerze źródła danych przy użyciu opakowania ODBC produktu WebSphere Information Integrator Classic Federation.

### **Procedura**

Aby stowarzyszyć bazę danych:

1. Zainstaluj produkt WebSphere Information Integrator Classic Federation na klasycznym serwerze źródła danych.
2. Zainstaluj moduł klienta produktu WebSphere Information Integrator Classic Federation na nieklasycznym serwerze źródła danych, zawierającym bazę danych stowarzyszoną z bazą danych na klasycznym serwerze źródła danych.
3. Skonfiguruj sterownik ODBC produktu WebSphere Information Integrator Classic Federation, aby umożliwić nawiązywanie połączeń z klasycznym serwerem źródła danych.
4. Nawiąż połączenie ze stowarzyszoną bazą danych i utwórz opakowanie ODBC, aby zdefiniować stowarzyszenie z bazą danych na klasycznym serwerze źródła danych.
5. Skonfiguruj przeszukiwacz DB2, aby przeszukiwać stowarzyszoną bazę danych na nieklasycznym serwerze źródła danych. Pozwoli to na przeszukiwanie bazy danych na klasycznym serwerze źródła danych za pośrednictwem produktu WebSphere Information Integrator Classic Federation.

---

## **Przeszukiwacze DB2 Content Manager**

Aby w kolekcji wyszukiwania korporacyjnego zostały uwzględnione typy elementów produktu IBM DB2 Content Manager, należy skonfigurować przeszukiwacz DB2 Content Manager.

### **Konfiguracja serwera przeszukiwacza**

Aby umożliwić przeszukiwanie serwera DB2 Content Manager, konieczne jest uruchomienie odpowiedniego skryptu na serwerze przeszukiwacza. Skrypt ten jest dostarczany razem z programem OmniFind Enterprise Edition i umożliwia przeszukiwaczowi DB2 Content Manager komunikowanie się z serwerami DB2 Content Manager.

Przed skonfigurowaniem przeszukiwacza DB2 Content Manager za pomocą konsoli administracyjnej wyszukiwania korporacyjnego należy wykonać zadanie odpowiednie dla używanego środowiska:

- “Konfigurowanie serwera przeszukiwaczy w systemie UNIX do współpracy z przeszukiwaczami DB2 Content Manager” na stronie 57.
- “Konfigurowanie serwera przeszukiwaczy w systemie Windows do współpracy z przeszukiwaczami DB2 Content Manager” na stronie 59.

### **Wskazówka:**

Szczegółowe przykłady konfigurowania bezpiecznego przeszukiwacza DB2 Content Manager można znaleźć w scenariuszu dla dużej organizacji w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

## Przegląd konfiguracji

Przeszukiwacz DB2 Content Manager może być używany do przeszukania dowolnej liczby serwerów DB2 Content Manager. Podczas konfigurowania przeszukiwacza należy określić metody, które mają być wykorzystywane w odniesieniu do wszystkich serwerów DB2 Content Manager w określonej przestrzeni przeszukiwania. Należy także wybrać konkretne typy elementów, które mają zostać przeszukane na każdym serwerze.

Aby utworzyć lub zmienić przeszukiwacz DB2 Content Manager, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich typów elementów na wszystkich serwerach DB2 Content Manager w przestrzeni przeszukiwania.
- Wybór serwerów DB2 Content Manager, które mają zostać przeszukane.
- Określenie identyfikatorów użytkowników i haseł pozwalających przeszukiwaczowi na uzyskanie dostępu do treści na serwerach DB2 Content Manager.
- Skonfigurowanie harmonogramu przeszukiwania serwerów.
- Wybór typów elementów, które mają zostać przeszukane na każdym z serwerów DB2 Content Manager.
- Określenie opcji udostępniania atrybutów niektórych typów elementów dla operacji wyszukiwania. Na przykład istnieje możliwość wykluczenia określonych typów dokumentów z przestrzeni przeszukiwania i określenia atrybutów, które mogą być zwracane w wynikach wyszukiwania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

## Konfigurowanie serwera przeszukiwacza w systemie UNIX do współpracy z przeszukiwaczami DB2 Content Manager

W przypadku zainstalowania programu OmniFind Enterprise Edition na komputerze z systemem IBM AIX, Linux lub w Środowisku Operacyjnym Solaris należy uruchomić odpowiedni skrypt w celu skonfigurowania serwera przeszukiwacza. Skrypt ten umożliwia komunikację przeszukiwacza DB2 Content Manager z serwerami IBM DB2 Content Manager.

### Informacje o tym zadaniu

Przeszukiwacz DB2 Content Manager uzyskuje dostęp do serwerów DB2 Content Manager przy użyciu konektora Java dla programu DB2 Content Manager, wersja 8. Konektor ten jest instalowany podczas instalowania na serwerze przeszukiwacza następujących produktów:

- IBM DB2 Information Integrator for Content, wersja 8.3, dla systemów AIX, Solaris lub Linux
- IBM DB2 Information Integrator for Content, wersja 8.2, dla systemów AIX lub Solaris
- IBM DB2 Content Manager Toolkit, wersja 8.2, dla systemu Linux

Aby zyskać pewność, że przeszukiwacz DB2 Content Manager może współpracować z produktem DB2 Content Manager, należy uruchomić skrypt konfiguracyjny dostępny w produkcie OmniFind Enterprise Edition. Skrypt należy uruchomić na serwerze przeszukiwacza po zainstalowaniu konektora.

### Procedura

Aby skonfigurować serwer przeszukiwacza tak, aby umożliwić przeszukiwanie serwerów DB2 Content Manager, należy wykonać poniższe czynności:

1. Zainstaluj konektor Java dla programu DB2 Content Manager, wersja 8, na serwerze przeszukiwacza:
  - a. Zaloguj się na serwerze przeszukiwacza jako użytkownik root:  
su - root
  - b. Uruchom plik db2profile. Na przykład:  
. /home/instancja\_db2/sqllib/db2profile
  - c. Wyeksportuj zmienną środowiskową JAVAHOME. Na przykład:  
export JAVAHOME=/usr/IBMJava2-141
  - d. Dodaj katalog Java do zmiennej środowiskowej PATH:  
export PATH=\$PATH:\$JAVAHOME/bin
  - e. Włóż do napędu instalacyjny dysk CD programu DB2 Information Integrator for Content i uruchom kreatora instalacji.
  - f. W oknie Wybór komponentów wykonaj następujące działania. (Jeśli jest używany program DB2 Information Integrator for Content, wersja 8.3, w oknie Wybór komponentów widoczna jest opcja instalacji niestandardowej).
    - 1) Z listy **Składniki** wybierz pozycję **Konektory lokalne**, a następnie pozycję **Konektor Content Manager, wersja 8** z listy **Składniki podrzędne**.
    - 2) Z listy **Składniki** wybierz pozycję **Pakiety narzędzi i przykładów do programowania konektorów**, a następnie pozycję **Konektor Content Manager, wersja 8** z listy **Składniki podrzędne**.
  - g. Określ nazwę bazy danych, nazwę użytkownika i hasło dla biblioteki DB2 Content Manager i zaakceptuj ustawienia domyślne w pozostałych oknach.
2. Zaloguj się na serwerze przeszukiwacza przy użyciu identyfikatora użytkownika należącego do grupy administracyjnej DB2.
3. Wpisz do katalogu zdalną bazę danych serwera bibliotecznego DB2 Content Manager i sprawdź, czy serwer przeszukiwacza może nawiązać połączenie z serwerem DB2 Content Manager:

```
db2 catalog tcpip node nazwa_węzła remote nazwa_hosta server port
db2 catalog database nazwa_bazy_danych as alias at node
nazwa_węzła
```

Gdzie:

*nazwa\_węzła*

Jest to krótka nazwa hosta serwera DB2 Content Manager (na przykład ibmes).

*nazwa\_hosta*

Jest to pełna nazwa hosta serwera DB2 Content Manager (na przykład ibmes.ibm.com).

*port*

Jest to numer portu serwera DB2 Content Manager.

*nazwa\_bazy\_danych*

Jest to nazwa bazy danych DB2 Content Manager (na przykład ICMNLSDB).

*alias*

Jest to alias bazy danych DB2 Content Manager (na przykład CMSVR).

4. Opcjonalnie: Zaloguj się jako użytkownik root i przetestuj połączenie z bazą danych:

```
. kat_inst_Information_Integrator_for_Content/bin/cmbenv81.sh
cd kat_inst_Information_Integrator_for_Content/samples/java/icm
javac *.java
java SConnectDisconnectICM nazwa_bazy_danych_ICM
 identyfikator_administradora_CM
 hasło_administradora_CM
```

5. Na serwerze przeszukiwacza uruchom skrypt konfiguracyjny przeszukiwacza DB2 Content Manager:

- a. Przejdź do katalogu ES\_INSTALL\_ROOT/bin:

```
cd $ES_INSTALL_ROOT/bin
```

- b. Uruchom poniższy skrypt i odpowiadaj na komunikaty:

```
escrm.sh
```

6. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
```

```
esadmin system startall
```

## Konfigurowanie serwera przeszukiwaczy w systemie Windows do współpracy z przeszukiwaczami DB2 Content Manager

W przypadku instalowania programu OmniFind Enterprise Edition na komputerze z systemem Microsoft Windows należy uruchomić skrypt w celu skonfigurowania serwera przeszukiwacza. Skrypt ten umożliwia komunikację przeszukiwacza DB2 Content Manager z serwerami IBM DB2 Content Manager.

### Informacje o tym zadaniu

Przeszukiwacz DB2 Content Manager uzyskuje dostęp do serwerów DB2 Content Manager przy użyciu konektora Java dla programu DB2 Content Manager, wersja 8. Konektor ten jest instalowany podczas instalowania na serwerze przeszukiwacza programu IBM DB2 Information Integrator for Content w wersji 8.2 lub wersji 8.3 dla systemu Windows. Aby zyskać pewność, że przeszukiwacz DB2 Content Manager może współpracować z produktem DB2 Content Manager, należy uruchomić skrypt konfiguracyjny dostępny w produkcie OmniFind Enterprise Edition. Skrypt należy uruchomić na serwerze przeszukiwacza po zainstalowaniu konektora.

### Procedura

Aby skonfigurować serwer przeszukiwacza tak, aby umożliwić przeszukiwanie serwerów DB2 Content Manager, należy wykonać poniższe czynności:

1. Zainstaluj konektor Java dla programu DB2 Content Manager, wersja 8, na serwerze przeszukiwacza:
  - a. Włóż do napędu instalacyjny dysk CD programu DB2 Information Integrator for Content. Program instalacyjny zostanie uruchomiony automatycznie.  
Otwarty zostanie kreator instalacji programu DB2 Content Manager Enterprise Information Portal.
  - b. W oknie Wybór komponentów wykonaj następujące działania. (Jeśli jest używany program DB2 Information Integrator for Content, wersja 8.3, w oknie Wybór komponentów widoczna jest opcja instalacji niestandardowej).
    - 1) Z listy **Składniki** wybierz pozycję **Konektory lokalne**, a następnie pozycję **Konektor Content Manager, wersja 8** z listy **Składniki podrzędne**.
    - 2) Z listy **Składniki** wybierz pozycję **Pakiety narzędzi i przykładów do programowania konektorów**, a następnie pozycję **Konektor Content Manager, wersja 8** z listy **Składniki podrzędne**.
  - c. Określ nazwę bazy danych, nazwę użytkownika i hasło dla biblioteki DB2 Content Manager i zaakceptuj ustawienia domyślne w pozostałych oknach.
2. Wpisz do katalogu zdalną bazę danych serwera bibliotecznego DB2 Content Manager i sprawdź, czy serwer przeszukiwacza może nawiązać połączenie z serwerem DB2 Content Manager. Wprowadź następujące komendy w wierszu komend na serwerze przeszukiwacza:

```
db2 catalog tcpip node nazwa_węzła remote nazwa_hosta server port
db2 catalog database nazwa_bazy_danych as alias at node
nazwa_węzła
```

Gdzie:

*nazwa\_węzła*

Jest to krótka nazwa hosta serwera DB2 Content Manager (na przykład *ibmes*).

*nazwa\_hosta*

Jest to pełna nazwa hosta serwera DB2 Content Manager (na przykład *ibmes.ibm.com*).

*port*

Jest to numer portu serwera DB2 Content Manager.

*nazwa\_bazy\_danych*

Jest to nazwa bazy danych DB2 Content Manager (na przykład *ICMNLSDB*).

*alias*

Jest to alias bazy danych DB2 Content Manager (na przykład *CMSVR*).

3. Opcjonalne: Przetestuj połączenie z bazą danych, otwierając wiersz komend i wprowadzając następujące komendy:

```
cmbenv81.bat
cd kat_inst_Information_Integrator_for_Content\samples\java\icm
javac *.java
java SConnectDisconnectICM nazwa_bazy_danych_ICM
identyfikator_administradora_CM
haslo_administradora_CM
```

4. Na serwerze przeszukiwacza uruchom skrypt konfiguracyjny przeszukiwacza DB2 Content Manager:
  - a. Przejdź do katalogu *ES\_INSTALL\_ROOT\bin*:



```
cd %ES_INSTALL_ROOT%\bin
```

b. Uruchom poniższy skrypt i odpowiadaj na komunikaty:

```
escrm.vbs
```

5. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
```

```
esadmin system startall
```

---

## Przeszukiwacze Domino Document Manager

Aby do kolekcji wyszukiwania korporacyjnego włączyć biblioteki i szafki Domino Document Manager, należy skonfigurować przeszukiwacz Domino Document Manager.

### Konfiguracja serwera przeszukiwacza

Jeśli na serwerze Domino Document Manager, który ma być przeszukiwany, używany jest protokół zdalnego wywołania procedury Notes (NRPC), na serwerze przeszukiwacza trzeba uruchomić skrypt konfiguracyjny. Skrypt ten jest dostarczany razem z produktem OmniFind Enterprise Edition i umożliwia przeszukiwaczowi Domino Document Manager komunikowanie się z serwerami używającymi protokołu NRPC.

Jeśli na serwerze Domino Document Manager, który ma być przeszukiwany, używany jest protokół DIIOP (Domino Internet Inter-ORB Protocol), na serwerze przeszukiwacza nie trzeba uruchamiać skryptu konfiguracyjnego. Konieczne jest jednak skonfigurowanie serwera Domino Document Manager, tak aby przeszukiwacz Domino Document Manager mógł uzyskać do niego dostęp.

**Ważne:** Jeśli serwer Domino Document Manager używa DIIOP i przeszukiwacz został skonfigurowany do używania protokołów HTTPS lub DIIOP przez SSL (dzięki czemu transmisja między przeszukiwaczem i serwerem jest szyfrowana), należy skopiować plik `TrustedCerts.class` (na przykład `c:\certs` lub `/data/certs`) z serwera Domino Document Manager do serwera przeszukiwacza. W konfiguracji dwuserwerowej i czteroserwerowej, należy również skopiować plik `TrustedCerts.class` na serwery, na których zainstalowany jest komponent wyszukiwania. Należy upewnić się, że plik ten znajduje się w tym samym katalogu na serwerze przeszukiwacza i na serwerze wyszukiwania. Ścieżka dla pliku `TrustedCerts.class` określana jest podczas konfigurowania przeszukiwacza.

Jeśli produkt OmniFind Enterprise Edition został zainstalowany w systemie IBM AIX, należy upewnić się, że na serwerze przeszukiwacza jest zainstalowany i dostępny moduł portu zakończenia `we/wy`.

Przed użyciem konsoli administracyjnej wyszukiwania korporacyjnego do skonfigurowania przeszukiwacza Domino Document Manager należy wykonać zadania odpowiednie dla używanego środowiska:

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 76.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 78.
- “Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 80.
- “Konfigurowanie portu zakończenia `we/wy` w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 81.



## Zabezpieczenia na poziomie dokumentu

Jeśli włączone są zabezpieczenia kolekcji, a serwer, który ma być przeszukiwany używa protokołu NRPC, należy na serwerze przeszukiwacza skonfigurować zaufany serwer Lotus Domino. Zaufany serwer służy do wymuszania kontroli dostępu na poziomie dokumentu. Przed udostępnieniem kolekcji użytkownikom do wyszukiwania należy wykonać następujące zadania:

- Skonfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika.
- Włączenie globalnych zabezpieczeń na serwerze WebSphere Application Server i skonfigurowanie aplikacji wyszukiwującej, aby z nich korzystała. Czynność ta gwarantuje, że referencje logowania są sprawdzane w momencie, gdy użytkownicy próbują korzystać z aplikacji wyszukiwującej. Serwery wyszukiwania korzystają z referencji do weryfikacji uprawnień użytkownika do uzyskiwania dostępu do dokumentów Lotus Domino.

## Przegląd konfiguracji

Przeszukiwacz Domino Document Manager może być używany do przeszukiwania dowolnej liczby bibliotek Domino Document Manager. Podczas tworzenia przeszukiwacza wybiera się biblioteki do przeszukiwania na pojedynczym serwerze Domino Document Manager. Później, w trakcie edycji przestrzeni przeszukiwania, można dodać do niej dokumenty z innych serwerów Domino Document Manager, które mają zostać włączone do tej samej przestrzeni przeszukiwania. Podczas tworzenia lub edycji przeszukiwacza można określić, czy przeszukiwane mają być wszystkie szafki w bibliotekach wybranych do przeszukiwania czy tylko konkretne szafki.

Aby utworzyć lub zmienić przeszukiwacz Domino Document Manager, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich dokumentów w przestrzeni przeszukiwania.
- Określenie serwera Domino Document Manager oraz protokołu komunikacyjnego.
- Jeśli serwer jest skonfigurowany do korzystania z protokołu DIOP, można określić, jak przeszukiwacz ma łączyć się z obiektami Domino. Na przykład można określić opcje korzystania z protokołu HTTPS lub Secure Sockets Layer (SSL) do szyfrowania komunikacji.
- Wybór bibliotek do przeszukiwania.
- Skonfigurowanie harmonogramu przeszukiwania bibliotek.
- Wybranie dokumentów, które mają zostać przeszukane. Przeszukiwacz może przeszukiwać wszystkie szafki w bibliotece lub też tylko dokumenty znajdujące się w wybranych szafkach.
- Określenie opcji udostępniających do wyszukiwania pola w różnych bibliotekach i szafkach. Na przykład można wykluczyć niektóre pola z przestrzeni przeszukiwania i określić opcje wyszukiwania załączników.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukiwujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Pojęcia pokrewne**

“Wskazówki dotyczące przeszukiwania baz danych Lotus Domino” na stronie 75

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino” na stronie 282

#### **Zadania pokrewne**

“Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 76

“Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 78

“Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 80

“Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 81

“Konfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika” na stronie 282

---

## **Przeszukiwacze Exchange Server**

Aby umożliwić włączenie publicznych folderów serwera Microsoft Exchange Server do kolekcji wyszukiwania korporacyjnego, należy skonfigurować przeszukiwacz serwera Exchange Server.

Przeszukiwacz serwera Exchange Server może służyć do przeszukiwania dowolnej liczby folderów i podfolderów na serwerach folderów publicznych Exchange Server. Podczas tworzenia przeszukiwacza należy określić treść, która ma być przeszukiwana na serwerze folderów publicznych. Edycja przestrzeni przeszukiwania w celu dodania treści z innego serwera folderów publicznych będzie możliwa również w późniejszym czasie.

Aby utworzyć lub zmienić przeszukiwacz Exchange Server, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich podfolderów na wszystkich serwerach w przestrzeni przeszukiwania.
- Określenie informacji o serwerze folderów publicznych Exchange Server, który ma zostać przeszukany.  
Należy podać identyfikator użytkownika i hasło, aby przeszukiwacz mógł uzyskać dostęp do treści na określonym serwerze. Jeśli na serwerze używany jest protokół SSL (Secure Sockets Layer), można określić opcje pozwalające przeszukiwaczowi na uzyskanie dostępu do pliku kluczy na serwerze przeszukiwacza.
- Skonfigurowanie harmonogramu przeszukiwania serwera folderów publicznych.
- Wybranie folderów i podfolderów, które mają zostać przeszukane.

- Określenie opcji udostępniania dokumentów w podfolderach dla operacji przeszukiwania. Na przykład istnieje możliwość wykluczenia określonych typów dokumentów z przestrzeni przeszukiwania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Zadania pokrewne**

“Sprawdzanie dostępu do zabezpieczonych dokumentów serwera Exchange” na stronie 281

---

## **Przeszukiwacze baz danych JDBC**

Przeszukiwacz JDBC umożliwia włączanie baz danych, do których można uzyskać dostęp za pomocą protokołu Java Database Connectivity (JDBC) w kolekcji wyszukiwania korporacyjnego.

Należy skonfigurować oddzielny przeszukiwacz dla każdego rodzaju systemu bazy danych, który ma zostać przeszukany. Po utworzeniu przeszukiwacza należy określić opcje przeszukiwania jednej bazy danych. Następnie do przestrzeni przeszukiwania można dodać bazy danych tego samego typu.

Każdy wiersz w tabeli bazy danych jest traktowany jak dokument, a wartości kolumn bazy danych są analizowane oraz indeksowane jako pola, w których można prowadzić wyszukiwanie. Przeszukiwacz można skonfigurować tak, aby przeszukiwał wiele tabel o określonej strukturze, przez powiązanie z przeszukiwaczem wtyczki podczas konfigurowania właściwości przeszukiwacza. Dzięki tej wtyczce wiersze, które pochodzą z wielu tabel w relacyjnej bazie danych i mają takie same pola kluczowe, można łączyć i traktować jak pojedynczy dokument. Gdy użytkownik przeprowadzi wyszukiwanie w bazie danych, dane ze złączonych tabel będą widoczne jako dodatkowe pola przy wyświetlaniu dokumentu w wynikach wyszukiwania.

### **Obsługiwane systemy i sterowniki baz danych**

Aby użyć protokołu JDBC do przeszukiwania tabel w bazie danych, na serwerze przeszukiwacza musi istnieć odpowiedni sterownik JDBC. Przeszukiwacz JDBC obsługuje następujące systemy baz danych i sterowniki JDBC typu 4:

| <b>System bazy danych</b>                                                                                                      | <b>Nazwa sterownika JDBC typu 4</b> | <b>Standardowe ścieżki klas sterownika JDBC</b>                                                                                        |
|--------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| IBM DB2 Universal Database, wersja 8.2, oraz IBM DB2 Enterprise Server Edition, wersja 9.1, dla systemów Linux, UNIX i Windows | com.ibm.db2.jcc.DB2Driver           | <i>główny_katalog_instalacyjny_db2/java/db2jcc.jar</i><br><i>główny_katalog_instalacyjny_db2</i><br><i>/java/db2jcc_license_cu.jar</i> |
| Oracle 9i i 10g                                                                                                                | Oracle.jdbc.driver.OracleDriver     | <i>oracle_home/jdbc/lib/ojdbc14.jar</i>                                                                                                |

| System bazy danych        | Nazwa sterownika JDBC typu 4                                                                                                       | Standardowe ścieżki klas sterownika JDBC                                                                                                                                         |
|---------------------------|------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Microsoft SQL Server 2000 | com.microsoft.jdbc.sqlserver.SQLServerDriver                                                                                       | <i>katalog_główny_mssql_jdbc/lib/mssqlserver.jar</i><br><i>katalog_główny_mssql_jdbc/lib/msbase.jar</i><br><i>katalog_główny_mssql_jdbc/lib/msutil.jar</i>                       |
| Microsoft SQL Server 2005 | com.microsoft.sqlserver.jdbc.SQLServerDriver<br>(Sterownik JDBC dla serwera SQL Server 2005 nie jest obsługiwany w systemach AIX). | <i>katalog_instalacyjny/sqljdbc_1.0/język/sqljdbc.jar</i> gdzie <i>język</i> oznacza ustawienia narodowe, na przykład<br><i>katalog_instalacyjny/sqljdbc_1.0/enu/sqljdbc.jar</i> |

## Przeszukiwacze JDBC a przeszukiwacze DB2

Jeśli obecnie używasz przeszukiwacza DB2, możesz nadal go używać. Nie można migrować danych zapisanych dla przeszukiwacza DB2 do przeszukiwacza JDBC.

Zamiast przeszukiwacza JDBC należy używać przeszukiwacza DB2 w następujących sytuacjach:

- Mają być przeszukiwane bazy danych DB2 z zastosowaniem sterownika interfejsu JDBC typu 2.
- Mają być przeszukiwane bazy danych Oracle i SQL Server stowarzyszone z bazą danych DB2. Za pomocą przeszukiwacza DB2 można uzyskać dostęp do wszystkich tych typów baz danych poprzez pseudonim.
- Mają być przeszukiwane bazy danych DB2 for z/OS, DB2 for iSeries, Informix, Sybase, VSAM, IMS, CA-Datcom lub Software AG Adabas. Bazy danych tych typów należy stowarzyszyć z bazą danych DB2, a dostęp do nich realizować za pomocą przeszukiwacza DB2 przy użyciu pseudonimu.
- Korzystanie z publikowania zdarzeń w celu aktualizacji indeksu wyszukiwania korporacyjnego podczas publikowania aktualizacji bazy danych.

## Przegląd konfiguracji

Aby utworzyć lub zmienić przeszukiwacz JDBC, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich baz danych w przestrzeni przeszukiwania.
- Określenie typu bazy danych, która ma zostać przeszukana.
- Wybranie bazy danych do przeszukania i, w razie potrzeby, podanie identyfikatora użytkownika i hasła umożliwiającego przeszukiwaczowi uzyskanie dostępu do bazy danych.
- Skonfigurowanie harmonogramu przeszukiwania bazy danych.
- Wybranie tabel, które mają zostać przeszukane.

**Ważne:** Aby zoptymalizować wydajność procesów wykrywania i zapobiec przekroczeniu limitu czasu przez proces konfigurowania przeszukiwacza, opcję przeszukiwania wszystkich tabel należy wybierać tylko wtedy, gdy baza danych nie zawiera zbyt wielu tabel lub gdy tabele nie zawierają zbyt wielu kolumn. Tabele do przeszukania można wybrać od razu, natomiast edycja przestrzeni przeszukiwania i dodanie dodatkowych tabel do kolekcji będzie możliwe również w późniejszym czasie.

- Określenie opcji udostępnienia kolumn w określonych tabelach dostępnych dla operacji przeszukiwania. Na przykład można uwzględnić pewne kolumny w zapytaniach parametrycznych lub można określić, które kolumny mogą być zwracane w wynikach wyszukiwania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

## Odwzorowania związków dla baz danych JDBC

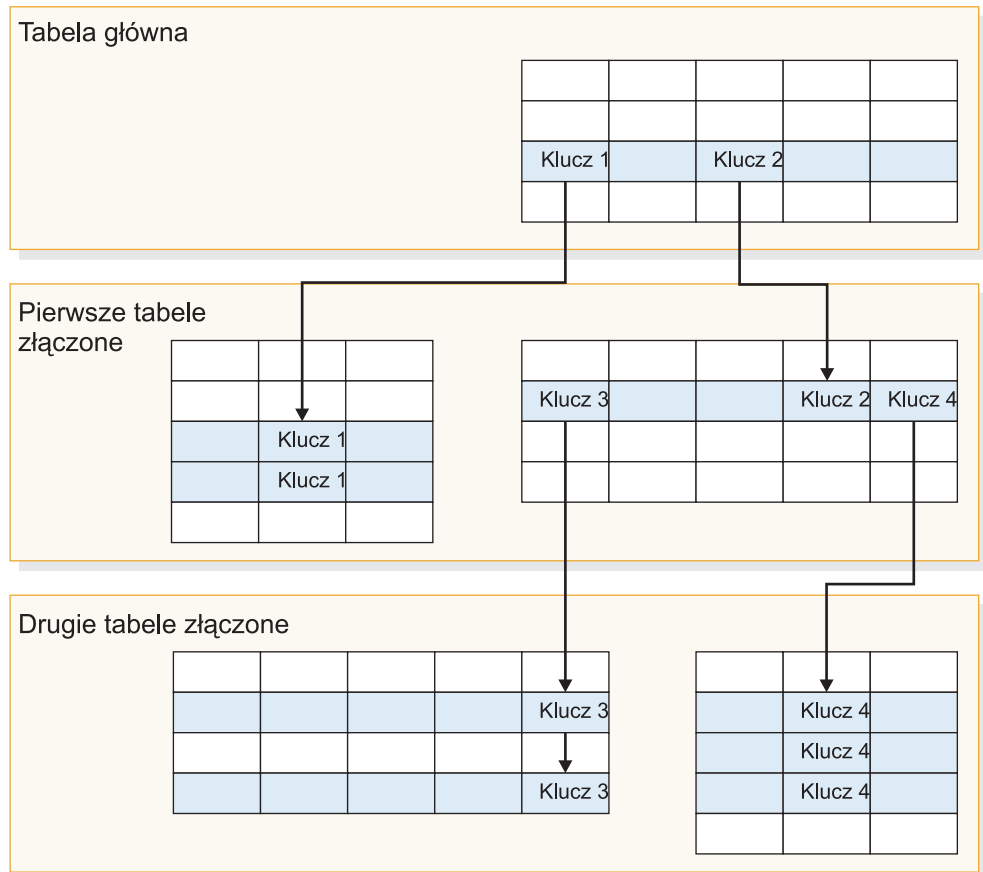
W przypadku tworzenia reguły dla wtyczki przeszukującej jednocześnie wiele tabel bazy danych JDBC o określonej strukturze określa się informacje o tabeli głównej oraz sposób łączenia tabel nadrzędnej i podrzędnej.

Wtyczka udostępniana na potrzeby wyszukiwania korporacyjnego umożliwia przeszukiwaczowi JDBC łączenie wielu tabel o określonej strukturze. Wtyczkę tworzy się przez określenie reguł w pliku `ES_INSTALL_ROOT/default_config/crawler_rdb_plugin.xml`. Gdy przeszukiwacz zostanie już skonfigurowany pod kątem używania tej wtyczki, wiersze z tabel o takich samych polach kluczowych będą łączone i traktowane jak pojedynczy dokument. Gdy użytkownik przeprowadzi wyszukiwanie w bazie danych, dane ze złączonych tabel będą widoczne jako dodatkowe pola przy wyświetlaniu dokumentu w wynikach wyszukiwania.

### Łączenie tabel przy użyciu kolumn kluczy

Poniższy rysunek pokazuje sposób budowania odwzorowania związków dla wielu tabel. Przeszukiwacz JDBC skanuje tabelę główną w bazie danych. Niektóre kolumny w tabeli służą jako pola kluczowe, których można użyć do łączenia tej tabeli z innymi tabelami. Kolumn w tak złączonych tabelach można potem użyć jako kluczy do łączenia dodatkowych tabel. Wiersze w wielu tabelach są traktowane jako pojedynczy dokument w przestrzeni przeszukiwania. Wartości kolumn są traktowane jako metadane dokumentu. Tabela główna jest *nadrzędna* w związku, a tabela złączona na pierwszym poziomie jest *podrzędna*. Tabele podrzędne na pierwszym poziomie również mogą być nadrzędne: względem tabel złączonych na drugim poziomie.

W tym przykładzie kolumny Klucz 1 i Klucz 2 w tabeli głównej (nadrzędnej) są polami kluczowymi umożliwiającymi łączenie danej tabeli z tabelami podporządkowanymi, które również mają kolumny Klucz 1 i Klucz 2. Jedna tabela złączona na tym pierwszym poziomie ma pola kluczowe, Klucz 3 i Klucz 4, które umożliwiają łączenie tej tabeli z dodatkowymi tabelami.

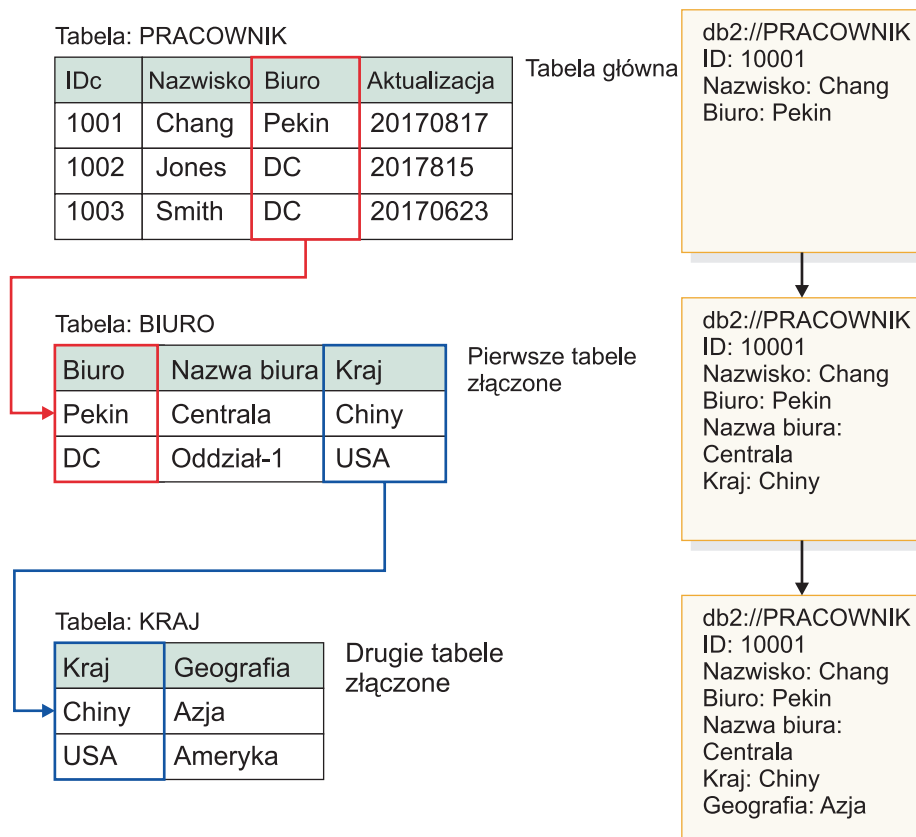


Rysunek 2. Tabele bazy danych JDBC złączone przy użyciu pól kluczowych

## Wyświetlanie wyników wyszukiwania

Następujący rysunek pokazuje sposób, w jaki dane z wielu tabel o określonej strukturze są w wynikach wyszukiwania przedstawiane jako pojedynczy dokument. Bez wtyczki użytkownikowi, który prowadził wyszukiwanie w tabeli PRACOWNIK, w wynikach wyszukiwania wyświetlany jest wiersz z tabeli głównej i tylko wartości z kolumn tabeli PRACOWNIK (ID, Nazwisko, Biuro).

Z wtyczką przeszukiwacz może złączyć tabelę PRACOWNIK z tabelą BIURO, używając kolumny Biuro jako klucza. Kolumna Kraj w tabeli BIURO służy jako klucz do łączenia tej tabeli z tabelą KRAJ. Po złączeniu tabel użytkownikom, którzy prowadzą wyszukiwanie w tabeli PRACOWNIK, w wynikach wyszukiwania jako dodatkowe pola wyświetlane są wartości z kolumn w tabelach BIURO i KRAJ.



Rysunek 3. Wartości ze złączonych tabel JDBC są wyświetlane w wynikach wyszukiwania

## Przeszukiwanie wielu ustrukturyzowanych tabel bazy danych JDBC

Przeszukiwacz JDBC można skonfigurować tak, aby łączył wiele tabel o określonej strukturze, które mają takie same pola kluczowe.

### Zagadnienia wstępne

Aby wykonać to zadanie, trzeba mieć uprawnienie do logowania się jako administrator wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

W przypadku konfigurowania właściwości przeszukiwacza JDBC można określić wtyczkę do przeszukiwania wielu tabel o określonej strukturze, które są wzajemnie powiązane poprzez pola kluczowe. Bez tej wtyczki wiersze w tabeli bazy danych są traktowane jak osobne dokumenty, a wartości kolumn bazy danych podlegają wyszukiwaniu jako osobne pola. Dzięki tej wtyczce wiersze, które pochodzą z wielu tabel w relacyjnej bazie danych o takich samych polach kluczowych, są łączone i traktowane jak pojedynczy dokument. Przeszukiwacz dodaje pobierane ze złączonych tabel dane do metadanych oryginalnego wiersza tabeli bazy danych. Gdy użytkownik prowadzi wyszukiwanie w bazie danych, te dodatkowe dane są widoczne jako dodatkowe pola przy wyświetlaniu dokumentu w wynikach wyszukiwania.

### Ograniczenia



### **Typy danych, których nie można przeszukiwać**

Przeszukiwacz nie może przeszukiwać pól w złączonych tabelach, które zawierają następujące typy danych binarnych:

BLOB  
CHARACTER FOR BIT DATA  
VARCHAR FOR BIT DATA  
LONG VARCHAR FOR BIT DATA

### **Ograniczenia zasięgu przestrzeni przeszukiwania**

Tabele do złączenia muszą się znajdować w tej samej relacyjnej bazie danych. Nie można łączyć tabel między różnymi bazami danych.

Jeśli tabela w bazie danych jest skonfigurowana do łączenia jej z innymi tabelami, ustawienie to jest uniwersalne dla wszystkich przeszukiwaczy w kolekcji mogących używać wtyczki. Niemniej jednak można tworzyć wiele kolekcji i konfigurować osobne przeszukiwacze do przeszukiwania różnych tabel głównych i łączenia różnych tabel.

### **Ograniczenia w używaniu innych wtyczek**

Jeśli przeszukiwacz jest skonfigurowany tak, aby używał wtyczki do przeszukiwania wielu tabel o określonej strukturze, to z tym przeszukiwaczem nie można już powiązać innej wtyczki. Na przykład nie można określić wtyczki niestandardowej do stosowania reguł biznesowych i reguł bezpieczeństwa. Z jednym przeszukiwaczem nie można powiązać więcej niż jednej wtyczki.

### **Ograniczenia w liczbie tabel, wierszy, pól i kluczy**

Na jedną bazę danych może przypadać maksymalnie pięć złączonych tabel, a suma liczby wierszy w tych tabelach musi być mniejsza niż jeden milion. Z jednej tabeli można odczytać maksymalnie 10 pól. Do łączenia tabel służy para kluczy. Oznacza to, że nie jest możliwe łączenie tabel przy użyciu wielu kluczy.

### **Zapewnianie przeszukiwania zmian w złączonych tabelach**

Jeśli wiersze w tabeli głównej nie zmieniają się między przeszukaniem, a przeszukiwacz nie jest skonfigurowany do wykonywania pełnego przeszukiwania, to niezmienione wiersze są ignorowane przez przeszukiwacz. Jeśli wiersze w tabeli, która jest złączona z tabelą główną, zmieniają się, mimo że nie zachodzą zmiany w tabeli głównej, to dla zapewnienia, że zmiany będą wykrywane i przeszukiwane, należy wykonać jedną z następujących czynności:

- Tabela główna w docelowej bazie danych powinna mieć pole datownika. Docelową bazę danych należy skonfigurować tak, aby miała pole datownika aktualizowane w chwili, gdy zmienia się wiersz w tabeli głównej lub gdy zmieniają się wiersze w dowolnej ze złączonych tabel podrzędnych. Podczas konfigurowania przeszukiwacza JDBC należy pamiętać o określeniu tego pola datownika jako pola, na podstawie którego przeszukiwacz ma ustalać, czy w tabelach wystąpiły zmiany.
- Podczas konfigurowania harmonogramu przeszukiwacza należy określić, że przeszukiwacz ma wykonywać pełne przeszukiwanie. Ta opcja zapewni, że za każdym razem będą przeszukiwane wszystkie tabele bez względu na to, czy wystąpiły zmiany.

### **Procedura**

Aby skonfigurować przeszukiwacz JDBC tak, aby przeszukiwał wiele tabel o określonej strukturze:

1. Zaloguj się na serwerze przeszukiwacza jako administrator wyszukiwania korporacyjnego i przez skopiowanie pliku `ES_INSTALL_ROOT/default_config/crawler_rdb_plugin.xml` utwórz plik `ES_NODE_ROOT/master_config/crawler_rdb_plugin.xml`.


2. Dokonaj edycji pliku ES\_NODE\_ROOT/master\_config/crawler\_rdb\_plugin.xml w edytorze tekstu obsługującym kodowanie UTF-8.
  - a. Poddaj edycji element `<Server DBURL="jdbc:db2://db_server_url:50000/SAMPLE">` i zastąp wartość `jdbc:db2://db_server_url:50000/SAMPLE` adresem URL bazy danych JDBC do przeszukiwania. Podczas konfigurowania przeszukiwacza należy pamiętać o określeniu tego samego adresu URL dla bazy danych do przeszukiwania.
  - b. Jeśli baza danych do przeszukiwania nie jest bazą danych DB2, poddaj edycji element `<JDBCdriver>com.ibm.db2.jcc.DB2Driver</JDBCdriver>` i zastąp wartość `com.ibm.db2.jcc.DB2Driver` odpowiednim sterownikiem JDBC. Podczas konfigurowania przeszukiwacza należy pamiętać o określeniu tego samego sterownika dla bazy danych do przeszukiwania.
  - c. Poddaj edycji element `<User>username</User>` i zastąp wartość `username` identyfikatorem użytkownika z uprawnieniem zapewniającym dostęp do bazy danych do przeszukiwania.
  - d. Poddaj edycji element `<Password Encryption="True">encrypted_password</Password>` i zastąp wartość `encrypted_password` zaszyfrowanym hasłem dla określonego identyfikatora użytkownika. Zaszyfrowane hasło można skopiować z pliku ES\_NODE\_ROOT/master\_config/col\_nazwa\_kolekcji.JDBC\_nazwa\_przeszukiwacza/jdbccrawler.xml i wkleić je w tym miejscu. Jeśli hasło nie musi być zaszyfrowane, element `Encryption="True"` należy zastąpić przez `Encryption="False"`, a wartość `encrypted_password` zastąpić hasłem w postaci jawnej.
  - e. Jeśli element `<Delimiters Use="True">` zostanie pozostawiony bez zmian, wiele terminów w jednej kolumnie jest oddzielanych za pomocą znaków (,) zdefiniowanych w elemencie `<Delimiter>`. Zestawy terminów przypadających na jedną tabelę są oddzielane znakami (;) zdefiniowanymi w elemencie `<SecondDelimiter>`. Jeśli element `<Delimiters Use="True">` zostanie ustawiony na `<Delimiters Use="False">`, to znaki ograniczników nie będą używane, a wielokrotne pola metadanych o tej samej nazwie pola będą dodawane jako metadane dokumentu.
  - f. W przypadku użycia elementu `<Delimiters Use="True">` poddaj edycji elementy `<Delimiter>`, `</Delimiter>` i `<SecondDelimiter>`; `</SecondDelimiter>`, aby określić znaki, które mają służyć jako separatory wartości.
  - g. Poddaj edycji element `<RelationMap Root="DB2INST1.TABLE_0">` i zastąp wartość `DB2INST1.TABLE_0` nazwą tabeli głównej, która ma być przeszukiwana.
  - h. Dokonaj edycji elementu `<Relation Parent="DB2INST1.TABLE_0" ParentAlias="T0" ParentKey="ID" Child="DB2INST1.TABLE_1" ChildAlias="T1" ChildKey="ID"/>`.
    - Zastąp wartość `Parent="DB2INST1.TABLE_0"` nazwą tabeli, która jest nadrzędna w danej relacji.
    - Zastąp wartość `ParentKey="T0"` aliasem tabeli nadrzędnej. Alias powinien być unikalny i nie może być zduplikowany w pliku `crawler_rdb_plugin.xml`.
    - Zastąp wartość `ParentKey="ID"` nazwą kolumny użytej jako pole kluczowe w relacji.
    - Zastąp wartości `Child="DB2INST1.TABLE_1"` `ChildAlias="T1"` `ChildKey="ID"` informacjami o tabeli podrzędnej do przeszukiwania.

Ta struktura definiuje sposób, w jaki tabele mają być łączone. Na przykład następujące odwzorowanie związków określa, że ma być przeszukiwana tabela główna o nazwie `DB2INST1.TABLE_A`. Tabele `DB2INST1.TABLE_B` i `DB2INST1.TABLE_C` są złączone na podstawie warunku `DB2INST1.TABLE_A.ID=DB2INST1.TABLE_B.ID AND DB2INST1.TABLE_B.ID=DB2INST1.TABLE_C.ID`.

```

<RelationMap Root="DB2INST1.TABLE_A">
<Relation Parent="DB2INST1.TABLE_A" ParentAlias="TA" ParentKey="ID"
 Child="DB2INST1.TABLE_B" ChildAlias="TB" ChildKey="ID"/>
<Relation Parent="DB2INST1.TABLE_B" ParentAlias="TB" ParentKey="ID"
 Child="DB2INST1.TABLE_C" ChildAlias="TC" ChildKey="ID"/>

```

- i. Powtórz krok 2h na stronie 70, aby utworzyć elementy <Relation> dla wszystkich relacji łączących tabele z tabelą główną.
  - j. Poddaj edycji element <Target TableAlias="T1"> i zastąp wartość TableAlias="T1" wartością ChildAlias zdefiniowaną w kroku 2h na stronie 70.
  - k. Dokonaj edycji elementu <Field Name="ID" FieldName="ID\_1" Enabling="True" Searchable="True" FieldSearchable="True" IsContent="True"/>.
    - Zastąp wartość Name="ID" nazwą kolumny w dokumentach do przeszukiwania.
    - Zastąp wartość FieldName="ID\_1" nazwą pola metadanych w dokumentach do przeszukiwania. Ta wartość posłuży jako nazwa wyświetlana kolumny w konsoli administracyjnej wyszukiwania korporacyjnego i w wynikach wyszukiwania.
    - Zastąp wartość Enabling="True" wartością "False", jeśli dana kolumna ma nie być uwzględniana w metadanych dokumentu.
    - Zastąp wartość Searchable="True" wartością "False", aby uniemożliwić użytkownikom wyszukiwanie w danej kolumnie z użyciem zapytania pełnotekstowego.
    - Zastąp wartość FieldSearchable="True" wartością "False", aby uniemożliwić użytkownikom wyszukiwanie w danej kolumnie na podstawie nazwy kolumny.
    - Zastąp wartość IsContent="True" wartością "False", aby wskazać, że kolumna nie zawiera treści podlegającej wyszukiwaniu. Jeśli zostaną określone wartości Searchable="True" oraz IsContent="True", to wartość danej kolumny posłuży do wykrywania duplikatów dokumentów i wejdzie w skład dynamicznego podsumowania dokumentu w wynikach wyszukiwania.
  - l. Powtórz krok 2k, aby utworzyć elementy <Field> dla wszystkich kolumn, które mają być przeszukiwane.
  - m. Powtórz kroki 2j i 2k, aby utworzyć elementy <Target> i <Field> dla wszystkich tabel podrzędnych, do których istnieje odwołanie w odwzorowaniu związków (<RelationMap>).
  - n. Powtórz kroki od 2g na stronie 70 do 2m, aby utworzyć wiele odwzorowań związków dla wielu tabel głównych.
  - o. Powtórz kroki od 2a na stronie 70 do 2n, aby skonfigurować odwzorowanie związków dla kolejnej bazy danych.
3. Skonfiguruj przeszukiwacz tak, aby korzystał z wtyczki:
- a. Otwórz konsolę administracyjną wyszukiwania korporacyjnego, dokonaj edycji kolekcji i wybierz stronę Przeszukiwanie.
  - b. Utwórz przeszukiwacz JDBC lub znajdź przeszukiwacz, który chcesz zmienić, i kliknij opcję  **Właściwości przeszukiwacza**.
  - c. W polu **Nazwa klasy wtyczki** wpisz nazwę wtyczki do przeszukiwania wielu tabel o określonej strukturze:  
com.ibm.es.plugin.rdb.RDBPlugin
  - d. W polu **Ścieżka klasy wtyczki** wpisz pełną ścieżkę do wtyczki i sterowników JDBC używanych przez wtyczkę. Na przykład ścieżką dla sterownika JDBC dla bazy danych DB2 w systemie Windows może być:  
C:\Program Files\IBM\es\lib\plugin\_rdb.jar;C:\Program Files\IBM\SQLLIB\java\db2jcc.jar;C:\Program Files\IBM\SQLLIB\java\db2jcc\_license\_cu.jar;
  - e. Kliknij przycisk **Dalej**, aby kontynuować tworzenie przeszukiwacza, albo kliknij przycisk **OK**, aby zapisać zmiany.

4. Aby wdrożyć plik `crawler_rdb_plugin.xml` do konfiguracji systemu, zrestartuj system wyszukiwania korporacyjnego:

```
esadmin system stop
esadmin system start
```

---

## Przeszukiwacze NNTP

Aby włączyć artykuły z grup dyskusyjnych NNTP do kolekcji wyszukiwania korporacyjnego, należy skonfigurować przeszukiwacz NNTP.

Przeszukiwacz NNTP może być używany do przeszukiwania dowolnej liczby serwerów NNTP. Podczas konfigurowania przeszukiwacza wybiera się grupy dyskusyjne do przeszukania na jednym serwerze NNTP. Później, podczas edycji przestrzeni przeszukiwania, można dodać inne serwery NNTP, które także mają być przeszukiwane przez ten przeszukiwacz.

Podczas identyfikowania grup dyskusyjnych do przeszukania można wybrać te, które zostaną włączone do przestrzeni przeszukiwania oraz te, które zostaną z niej wykluczone. Dzięki temu można łatwo umożliwić przeszukiwaczowi przeszukiwanie większości grup dyskusyjnych na serwerze i zabronić mu przeszukania tych kilku grup, których użytkownicy mają nie przeszukiwać.

Na przykład można podać reguły powodujące włączenie wszystkich grup dyskusyjnych dostępnych na określonym serwerze NNTP, a następnie określić, że mają zostać wykluczone dostępne na tym serwerze grupy dyskusyjne, których nazwy zawierają łańcuch **prywatne**.

Aby utworzyć lub zmienić przeszukiwacz NNTP, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają metodę przeszukiwania wszystkich grup dyskusyjnych w przestrzeni przeszukiwania.
- Zidentyfikowanie serwera NNTP do przeszukania. Jeśli dany serwer jest chroniony hasłem, należy określić także identyfikator użytkownika i hasło, których przeszukiwacz ma używać w celu uzyskania dostępu do grup dyskusyjnych na serwerze.
- Skonfigurowanie harmonogramu przeszukiwania serwera.
- Określenie wzorców dla grup dyskusyjnych, które mają zostać włączone, oraz wzorców dla grup dyskusyjnych, które mają zostać wykluczone z przestrzeni przeszukiwania.
- Określenie, czy przeszukiwacz powinien automatycznie wykrywać język i stronę kodową przeszukiwanych artykułów, czy też ma używać konkretnego języka i strony kodowej.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

---

## Przeszukiwacze Notes

Aby włączyć do kolekcji wyszukiwania korporacyjnego bazy danych programu IBM Lotus Notes, należy skonfigurować przeszukiwacz Notes.

### Wskazówka:

Szczegółowe przykłady konfigurowania bezpiecznego przeszukiwacza Notes można znaleźć w scenariuszu dla dużej organizacji w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

## Konfiguracja serwera przeszukiwacza

Jeśli na serwerze Lotus Notes, który ma być przeszukiwany, używany jest protokół zdalnego wywołania procedury Notes (NRPC), na serwerze przeszukiwacza trzeba uruchomić skrypt konfiguracyjny. Skrypt ten jest dostarczany razem z produktem OmniFind Enterprise Edition i umożliwia przeszukiwaczowi Notes komunikowanie się z serwerami używającymi protokołu NRPC.

Jeśli na serwerze Lotus Notes, który ma być przeszukiwany, używany jest protokół DIIOP (Domino Internet Inter-ORB Protocol), na serwerze przeszukiwacza nie trzeba uruchamiać skryptu konfiguracyjnego. Konieczne jest jednak skonfigurowanie serwera Lotus Notes w celu zezwolenia przeszukiwaczowi Notes na uzyskanie dostępu do tego serwera.

**Ważne:** Jeśli serwer Lotus Notes używa DIIOP i przeszukiwacz został skonfigurowany do używania protokołów HTTPS lub DIIOP przez SSL (dzięki czemu transmisja między przeszukiwaczem i serwerem jest szyfrowana), należy skopiować plik `TrustedCerts.class` (na przykład `c:\certs` lub `/data/certs`) z serwera Lotus Notes do serwera przeszukiwacza. W konfiguracji dwuserwerowej i czteroserwerowej, należy również skopiować plik `TrustedCerts.class` na serwery, na których zainstalowany jest komponent wyszukiwania. Należy upewnić się, że plik ten znajduje się w tym samym katalogu na serwerze przeszukiwacza i na serwerze wyszukiwania. Ścieżka dla pliku `TrustedCerts.class` określana jest podczas konfigurowania przeszukiwacza.

Jeśli produkt OmniFind Enterprise Edition został zainstalowany w systemie IBM AIX, należy upewnić się, że na serwerze przeszukiwacza jest zainstalowany i dostępny moduł portu zakończenia `we/wy`.

Przed użyciem konsoli administracyjnej wyszukiwania korporacyjnego do skonfigurowania przeszukiwacza Notes należy wykonać zadania odpowiednie dla używanego środowiska:

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 76.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 78.
- “Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 80.
- “Konfigurowanie portu zakończenia `we/wy` w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 81.

## Zabezpieczenia na poziomie dokumentu

Jeśli włączone są zabezpieczenia kolekcji, a serwer, który ma być przeszukiwany używa protokołu NRPC, należy na serwerze przeszukiwacza skonfigurować zaufany serwer Lotus



Domino. Zaufany serwer służy do wymuszania kontroli dostępu na poziomie dokumentu. Przed udostępnieniem kolekcji użytkownikom do wyszukiwania należy wykonać następujące zadania:

- Skonfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika.
- Włączenie globalnych zabezpieczeń na serwerze WebSphere Application Server i skonfigurowanie aplikacji wyszukiwującej, aby z nich korzystała. Czynność ta gwarantuje, że referencje logowania są sprawdzane w momencie, gdy użytkownicy próbują korzystać z aplikacji wyszukiwującej. Serwery wyszukiwania korzystają z referencji do weryfikacji uprawnień użytkownika do uzyskiwania dostępu do dokumentów Lotus Domino.

## Przegląd konfiguracji

Przeszukiwacz Notes może być używany do przeszukiwania dowolnej liczby standardowych baz danych Lotus Notes (plików .nsf). Podczas tworzenia przeszukiwacza należy wybrać bazy danych lub katalogi do przeszukania na pojedynczym serwerze Lotus Notes. Później, w trakcie edycji przestrzeni przeszukiwania, można dodać dokumenty z innych serwerów Lotus Notes, które mają zostać włączone do tej samej przestrzeni wyszukiwania. Podczas tworzenia lub edycji przeszukiwacza można określić, czy przeszukiwane mają być wszystkie bazy danych lub katalogi na określonym serwerze, czy tylko konkretne bazy danych, widoki i foldery.

Aby utworzyć lub zmienić przeszukiwacz Notes, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich dokumentów w przestrzeni przeszukiwania.
- Określenie nazwy hosta, numeru portu i protokołu komunikacyjnego serwera Lotus Notes.
- Jeśli serwer jest skonfigurowany do korzystania z protokołu DIOP, można określić, jak przeszukiwacz ma łączyć się z obiektami Domino. Na przykład można określić opcje korzystania z protokołu HTTPS lub Secure Sockets Layer (SSL) do szyfrowania komunikacji.
- Wybranie baz danych lub katalogów, które mają zostać przeszukane. W przypadku przeszukiwania katalogów można określić wzorce uwzględniania lub wykluczania baz danych, co może pomóc w podziale zadania przeszukiwania dużych katalogów na wiele przeszukiwaczy.
- Skonfigurowanie harmonogramu przeszukiwania baz danych lub katalogów.
- Wybranie dokumentów, które mają zostać przeszukane. Przeszukiwane mogą być wszystkie dokumenty w katalogu, wszystkie dokumenty w bazie danych lub tylko dokumenty w wybranych widokach i folderach bazy danych.
- Określenie opcji udostępniania pól w różnych bazach danych, widokach i folderach dla operacji przeszukiwania. Na przykład można wykluczyć niektóre pola z przestrzeni przeszukiwania i określić opcje wyszukiwania załączników.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukiwujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Pojęcia pokrewne**

“Wskazówki dotyczące przeszukiwania baz danych Lotus Domino”

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino” na stronie 282

#### **Zadania pokrewne**

“Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 76

“Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 78

“Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 80

“Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 81

“Konfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika” na stronie 282

## **Wskazówki dotyczące przeszukiwania baz danych Lotus Domino**

Przed skonfigurowaniem przeszukiwacza Notes należy przejrzeć wytyczne dotyczące przeszukiwania baz danych Lotus Domino.

- Bazy danych Notes oparte na standardowych szablonach (takie jak baza danych dyskusji) są najlepszym typem baz danych do przeszukiwania.
- Przeszukiwacz Notes stosuje następujące reguły odwzorowywania pól:
  - Podstawowe nazwy pól ze standardowych szablonów Domino są początkowo rejestrowane.
  - Wartości z pól Notes określonych w tabeli reguł odwzorowywania są używane jako podsumowania dokumentów w wynikach wyszukiwania.
  - Wartości z pól Notes nieokreślonych w tabeli reguł odwzorowywania nie są używane w podsumowaniach dokumentów.
  - Wartości z pól Notes odwzorowanych na pole Title są używane jako tytuł dokumentu w wynikach wyszukiwania.
  - Pola w następującej tabeli są domyślnie odwzorowywane na nazwy pól wyszukiwania:

*Tabela 2. Domyślne reguły odwzorowywania pól*

| <b>Nazwa pola bazy danych Notes</b> | <b>Nazwa pola wyszukiwania</b> |
|-------------------------------------|--------------------------------|
| Title                               | Title                          |
| EventTitle                          | Title                          |
| Subject                             | Title                          |
| Body                                | Body                           |
| Mission                             | Body                           |
| From                                | Creator                        |
| Author                              | Creator                        |



Tabela 2. Domyślne reguły odwzorowywania pól (kontynuacja)

| Nazwa pola bazy danych Notes | Nazwa pola wyszukiwania |
|------------------------------|-------------------------|
| Keywords                     | Categories              |
| Categories                   | Categories              |
| TeamRoomName                 | Organization            |
| TeamName                     | Organization            |
| Department                   | Organization            |

- Przeszukiwacz Notes może przeszukiwać wszystkie typy pól za wyjątkiem pól obliczanych przed wyświetleniem.
- Statyczny tekst i obrazy umieszczone w formularzach Notes nie są przeszukiwane.
- Podczas konfigurowania przeszukiwacza należy zaznaczyć pole wyboru **Wszystkie** w opcji **Przeszukiwanie**, aby przeszukiwać wszystkie pola i zmaksymalizować dane pola do przeszukania (można użyć pola **Przeszukaj wszystkie pola z wyjątkiem**, aby ograniczyć liczbę pól do przeszukania).

Aby zminimalizować przeszukiwanie zbędnych pól, należy anulować zaznaczenie pola wyboru **Przeszukiwanie** dla wszystkich pól za wyjątkiem tych odwzorowanych na pola wyszukiwania.

#### Pojęcia pokrewne

“Przeszukiwacze Domino Document Manager” na stronie 61

“Przeszukiwacze Notes” na stronie 73

“Przeszukiwacze QuickPlace” na stronie 82

## Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino

W przypadku zainstalowania programu OmniFind Enterprise Edition na komputerze z systemem IBM AIX, Linux lub ze Środowiskiem Operacyjnym Solaris, gdy jest planowane przeszukiwanie serwerów używających protokołu NRPC (Notes remote procedure call), należy uruchomić skrypt konfiguracyjny serwera przeszukiwacza. Ten skrypt umożliwi przeszukiwaczom Notes, QuickPlace i Domino Document Manager komunikację z serwerami baz danych.

### Ograniczenia

Serwer Domino nie może być uruchomiony w tym samym czasie na tym samym komputerze z przeszukiwaczem Notes, QuickPlace lub Domino Document Manager, który skonfigurowano do korzystania z protokołu NRPC. Uruchomienie jednego z tych przeszukiwaczy w trakcie pracy serwera Domino powoduje wystąpienie błędu i zatrzymanie przeszukiwacza.

### Informacje o tym zadaniu

Przeszukiwacz używający protokołu NRPC korzysta z bibliotek Domino jako klient. Biblioteki te są instalowane podczas instalacji serwera Lotus Domino na serwerze przeszukiwania. Aby mieć pewność, że przeszukiwacze mogą współpracować z bibliotekami Domino, należy uruchomić skrypt konfiguracyjny, który program OmniFind Enterprise Edition udostępnia na serwerze przeszukiwacza po zainstalowaniu bibliotek Domino.

### Procedura

Aby skonfigurować serwer przeszukiwacza tak, aby umożliwić przeszukiwanie serwerów Lotus Notes, Lotus QuickPlace i Domino Document Manager, należy wykonać poniższe czynności:

1. Utwórz użytkownika **server** i grupę **notes** na serwerze przeszukiwacza:
  - a. Zaloguj się jako administrator:  
`su - root`
  - b. Dodaj użytkownika:  
`useradd server`
  - c. Dodaj hasło tego użytkownika:  
`passwd server`  
Zostanie wyświetlone żądanie zmiany tego hasła.
2. Zainstaluj serwer Lotus Domino na serwerze przeszukiwacza:
  - a. Włóż do napędu dysk instalacyjny serwera Domino Server i podłącz go. (W razie braku dysku CD można pobrać jego obraz).
  - b. Przejdź do katalogu właściwego dla systemu operacyjnego.  
  
System AIX: `cd /mnt/cdrom/aix`  
System Linux: `cd /mnt/cdrom/linux`  
Środowisko Solaris: `cd /mnt/cdrom/solaris`
  - c. Uruchom program instalacyjny:  
`./install`
  - d. Odpowiedz na komunikaty i zaakceptuj wartości domyślne lub podaj preferowane ustawienia instalacyjne (takie jak ścieżki do katalogu instalacyjnego i katalogu danych).  
  
Więcej informacji na temat instalacji serwera Domino można znaleźć w dokumentacji tego produktu.
  - e. Upewnij się, że identyfikator administratora wyszukiwania korporacyjnego ma uprawnienie dostępu do katalogu `home/server`. Ten identyfikator administratora został określony podczas instalowania produktu OmniFind Enterprise Edition.
3. Na serwerze przeszukiwania wykonaj skrypt konfiguracyjny OmniFind Enterprise Edition:
  - a. Zaloguj się jako administrator wyszukiwania korporacyjnego (z identyfikatorem użytkownika określonym podczas instalowania programu OmniFind Enterprise Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu `$ES_INSTALL_ROOT/bin`:  
`escrnote.sh`
  - c. Odpowiedz na komunikaty:
    - Na przedstawiony poniżej komunikat udziel odpowiedzi Y, jeśli serwer Domino jest instalowany w katalogu domyślnym, lub N, jeśli nie jest:  
`Odnaleziono ścieżkę katalogu programu Lotus Notes  
/opt/lotus/notes/latest/linux.  
Czy to jest poprawna ścieżka katalogu programu Lotus Notes?`  
  
Domyślna ścieżka dla systemu AIX to `/opt/lotus/notes/latest/ibmpow`.  
Domyślna ścieżka dla systemu Linux to `/opt/lotus/notes/latest/linux`.  
Domyślna ścieżka dla środowiska Solaris to `/opt/lotus/notes/latest/sunspa`.
    - Jeśli serwer Domino nie jest instalowany w domyślnym katalogu na serwerze przeszukiwacza, określ właściwą ścieżkę do katalogu Domino w odpowiedzi na poniższy komunikat:

Wprowadź ścieżkę katalogu programu Lotus Notes.

Na przykład na komputerze z systemem Linux można wprowadzić ścieżkę /opt/lotus/notes/latest/linux.

- Na przedstawiony poniżej komunikat udziel odpowiedzi Y, jeśli katalog danych serwera Domino został umieszczony w domyślnym katalogu, lub N w przeciwnym razie:

Odnaleziono ścieżkę katalogu danych programu Lotus Notes /local/notesdata. Czy jest to poprawna ścieżka katalogu danych programu Lotus Notes?

Domyślna ścieżka to /local/notesdata.

- Jeśli katalog danych serwera Domino nie znajduje się w domyślnej lokalizacji na serwerze przeszukiwacza, określ ścieżkę do katalogu danych w odpowiedzi na poniższy komunikat:

Wprowadź ścieżkę katalogu danych programu Lotus Notes.

4. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
esadmin system startall
```

#### Pojęcia pokrewne

“Przeszukiwacze Domino Document Manager” na stronie 61

“Przeszukiwacze Notes” na stronie 73

“Przeszukiwacze QuickPlace” na stronie 82

## Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino

W przypadku zainstalowania programu OmniFind Enterprise Edition na komputerze z systemem Microsoft Windows, gdy jest planowane przeszukiwanie serwerów używających protokołu NRPC (Notes remote procedure call), należy uruchomić skrypt konfiguracyjny serwera przeszukiwacza. Ten skrypt umożliwi przeszukiwaczom Notes, QuickPlace i Domino Document Manager komunikację z serwerami baz danych.

### Ograniczenia

Serwer Lotus Domino Server oraz klient Lotus Notes nie mogą być uruchomione w tym samym czasie i na tym samym komputerze z przeszukiwaczem Notes, QuickPlace lub Domino Document Manager, który skonfigurowano do korzystania z protokołu NRPC. Uruchomienie jednego z tych przeszukiwaczy w trakcie pracy serwera Domino powoduje wystąpienie błędu i zatrzymanie przeszukiwacza.

### Informacje o tym zadaniu

Przeszukiwacze korzystające z protokołu NRPC wymagają dostępu do bibliotek klienta Lotus Domino. Biblioteki te są instalowane podczas instalacji serwera Lotus Domino na serwerze przeszukiwania. Aby mieć pewność, że przeszukiwacze mogą współpracować z bibliotekami Domino, należy uruchomić skrypt konfiguracyjny, który program OmniFind Enterprise Edition udostępni na serwerze przeszukiwacza po zainstalowaniu bibliotek Domino.

### Procedura

Aby skonfigurować serwer przeszukiwacza tak, aby umożliwić przeszukiwanie serwerów Lotus Notes, Lotus QuickPlace i Domino Document Manager, należy wykonać poniższe czynności:

1. Zaloguj się na serwerze przeszukiwacza przy użyciu identyfikatora użytkownika należącego do grupy Administratorzy. Upewnij się, że użytkownik o podanym identyfikatorze ma uprawnienia pozwalające na instalację programu Lotus Notes.
2. Zainstaluj program Lotus Notes:
  - a. Włóż do napędu dysk instalacyjny serwera Domino Server. (W razie braku dysku CD można pobrać jego obraz).
  - b. Uruchom program instalacyjny: `setup.exe`.
  - c. Odpowiedz na komunikaty i zaakceptuj wartości domyślne lub podaj preferowane ustawienia instalacyjne (takie jak ścieżki do katalogu instalacyjnego i katalogu danych).

Więcej informacji można znaleźć w dokumentacji produktu Lotus Domino.
  - d. Upewnij się, że identyfikator administratora wyszukiwania korporacyjnego ma uprawnienie dostępu do katalogu danych produktu Domino. Ten identyfikator administratora został określony podczas instalowania produktu OmniFind Enterprise Edition.
3. Na serwerze przeszukiwacza uruchom skrypt konfiguracyjny dostarczony z programem OmniFind Enterprise Edition:
  - a. Zaloguj się przy użyciu identyfikatora administratora wyszukiwania korporacyjnego (ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu `%ES_INSTALL_ROOT%\bin`:  
`escrnote.vbs`
  - c. Odpowiedz na komunikaty:
    - Na przedstawiony poniżej komunikat udziel odpowiedzi Y, jeśli program Lotus Notes jest instalowany w katalogu domyślnym, lub N w przeciwnym razie:  
`Odnaleziono ścieżkę katalogu programu Lotus Notes c:\lotus\notes.`  
Czy to jest poprawna ścieżka katalogu programu Lotus Notes?  
  
Domyślna ścieżka instalacji na komputerze z systemem Windows to `c:\lotus\notes` lub `c:\lotus\domino`.
    - Jeśli program Lotus Notes nie jest instalowany w domyślnym katalogu na serwerze przeszukiwacza, określ właściwą ścieżkę do katalogu Lotus Notes w odpowiedzi na poniższy komunikat:  
Wprowadź ścieżkę katalogu programu Lotus Notes.
    - Na przedstawiony poniżej komunikat udziel odpowiedzi Y, jeśli katalog danych programu Lotus Notes został umieszczony w domyślnej lokalizacji, lub N w przeciwnym razie:  
`Odnaleziono ścieżkę katalogu danych programu Lotus Notes`  
`c:\lotus\notes\data.`  
Czy jest to poprawna ścieżka katalogu danych programu Lotus Notes?  
  
Typowa ścieżka na komputerze z systemem Windows to `c:\lotus\notes\data` lub `c:\lotus\domino\data`.
    - Jeśli katalog danych programu Lotus Notes nie znajduje się w domyślnej lokalizacji na serwerze przeszukiwacza, określ ścieżkę do katalogu danych w odpowiedzi na poniższy komunikat:  
Wprowadź ścieżkę katalogu danych programu Lotus Notes.

4. Na serwerze przeszukiwacza zatrzymaj, a następnie uruchom ponownie system wyszukiwania korporacyjnego:

```
esadmin system stopall
esadmin system startall
```

#### Pojęcia pokrewne

“Przeszukiwacze Domino Document Manager” na stronie 61

“Przeszukiwacze Notes” na stronie 73

“Przeszukiwacze QuickPlace” na stronie 82

## Konfigurowanie serwerów korzystających z protokołu DIIOP

Aby przeszukiwać serwery korzystające z protokołu Domino Internet Inter-ORB Protocol (DIIOP), należy skonfigurować serwer tak, aby przeszukiwacze Notes, QuickPlace i Domino Document Manager mogły używać tego protokołu.

### Zagadnienia wstępne

Serwer, który ma zostać przeszukany, musi obsługiwać zadania protokołów DIIOP i HTTP.

### Procedura

Aby skonfigurować serwery korzystające z protokołu DIIOP, należy wykonać poniższe czynności:

1. Skonfiguruj dokument serwera:
  - a. Otwórz dokument **server** na serwerze Lotus Notes, Lotus QuickPlace, lub Domino Document Manager, który ma zostać przeszukany. Dokument ten jest przechowywany w katalogu Domino.
  - b. Na stronie Konfiguracja rozwiń sekcję **serwer**.
  - c. Na stronie Zabezpieczenia w obszarze **Ograniczenia dotyczące programowania** w następujących polach określ odpowiednie ograniczenia dotyczące zabezpieczeń dla danego środowiska:
    - **Uruchamianie ograniczonych agentów Lotus Script/Java**
    - **Uruchamianie ograniczonego kodu Java/Javascript/COM**
    - **Uruchamianie nieograniczonego kodu Java/Javascript/COM**Na przykład można wprowadzić gwiazdkę (\*), aby zezwolić na nieograniczony dostęp agentom Lotus Script/Java, lub wpisać nazwy użytkowników zarejestrowanych w katalogu Domino, aby objąć ich ograniczeniami dotyczącymi kodu Java/Javascript/COM.  
**Ważne:** Przeszukiwacz skonfigurowany do przeszukiwania danego serwera z protokołem DIIOP musi mieć możliwość korzystania z nazw użytkowników określonych w tych polach.
  - d. Otwórz stronę Protokół internetowy, a następnie stronę HTTP i ustaw dla opcji **Zezwalaj klientom HTTP na przeglądanie bazy danych** wartość **Tak**.
2. Skonfiguruj dokument użytkownika:
  - a. Otwórz dokument **user** na serwerze Lotus Notes, Lotus QuickPlace lub Domino Document Manager, który ma zostać przeszukany. Dokument ten jest przechowywany w katalogu Domino.
  - b. Na stronie Dane podstawowe w polu **Hasło internetowe** określ hasło.  
W przypadku korzystania z konsoli administracyjnej wyszukiwania korporacyjnego podczas konfigurowania opcji przeszukiwania tego serwera należy podać wskazany

tutaj identyfikator i hasło użytkownika na stronie, na której określany jest serwer do przeszukania. Przeszukiwacz używa tych referencji do uzyskania dostępu do serwera.

3. Zrestartuj zadanie DIIOP na serwerze.

#### Pojęcia pokrewne

“Przeszukiwacze Domino Document Manager” na stronie 61

“Przeszukiwacze Notes” na stronie 73

“Przeszukiwacze QuickPlace” na stronie 82

## Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino

Zanim można będzie użyć przeszukiwaczy Notes, QuickPlace lub Domino Document Manager w systemie IBM AIX, należy zainstalować moduł portu zakończenia we/wy (IOCP - I/O completion port) i skonfigurować go do użycia przez przeszukiwacz.

### Informacje o tym zadaniu

Bez modułu IOCP procesy wykrywania będą się kończyć niepowodzeniem podczas próby utworzenia przeszukiwacza. Zostanie wyświetlony następujący komunikat o błędzie:

FFQM0105E Otrzymano informację o błędzie z serwera -  
Komunikat: FFQG0024E Przechwycono nieoczekiwany wyjątek: discover

Poniższy komunikat, który zawiera błąd ENOEXEC, zostanie zapisany do pliku \$ES\_NODE\_ROOT/logs/system\_rrrmdd.log. (Część tekstu komunikatu jest dzielona na kilka wierszy w celu poprawienia czytelności).

```
5/20/05 18:08:52.423 JST [Error] [ES_ERR_EXCEPTION_DEFAULT_MESSAGE] [] [discovery]
iies10.yamato.ibm.com:0:2108088751:control:ComponentDiscoveryW.java:
com.ibm.es.control.discovery.server.ComponentDiscoveryW.discover:86
FFQ00277E An exception was caught with the detail 'java.lang.UnsatisfiedLinkError:
/opt/lotus/notes/65010/ibmpow/liblsxbe_r.a:
load ENOEXEC on shared library(s) /opt/lotus/notes/latest/ibmpow/libnotes_r.a'
and a stack trace of 'java.lang.UnsatisfiedLinkError:
/opt/lotus/notes/65010/ibmpow/liblsxbe_r.a:
load ENOEXEC on shared library(s) /opt/lotus/notes/latest/ibmpow/libnotes_r.a
at java.lang.ClassLoader$NativeLibrary.load(Native Method)
at java.lang.ClassLoader.loadLibrary0(ClassLoader.java:2120)
at java.lang.ClassLoader.loadLibrary(ClassLoader.java:1998)
at java.lang.Runtime.loadLibrary0(Runtime.java:824)
at java.lang.System.loadLibrary(System.java:908)
at lotus.domino.NotesThread.load(NotesThread.java:306)
at lotus.domino.NotesThread.checkLoaded(NotesThread.java:327)
at lotus.domino.NotesThread.sinitThread(NotesThread.java:181)
at com.ibm.es.crawler.discovery.notes.NotesLibrary$NotesOperation.discover
(Unknown Source)
at com.ibm.es.crawler.discovery.api.DiscoveryAPI.discover(Unknown Source)
at com.ibm.es.control.discovery.server.ComponentDiscoveryW.discover
(ComponentDiscoveryW.java:72)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:85)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:58)
at sun.reflect.DelegatingMethodAccessorImpl.invoke
(DelegatingMethodAccessorImpl.java:60)
at java.lang.reflect.Method.invoke(Method.java:391)
at com.ibm.es.ccl.sessionwrapper.CallThread.run(CallThread.java:77)
```

### Procedura

Aby zainstalować moduł IOCP i upewnić się, że został on poprawnie zainstalowany na serwerze przeszukiwacza, należy wykonać poniższe czynności:

1. Zainstaluj moduł IOCP (`bos.iocp.rte`) z dysku CD produktu AIX na serwerze przeszukiwacza.  
Po zainstalowaniu modułu IOCP i przed utworzeniem przeszukiwacza Notes, QuickPlace lub Domino Document Manager należy zastosować do modułu poprawkę oprogramowania. Instrukcje są dostępne po skorzystaniu z następującego odsyłacza:  
<http://www.ibm.com/support/docview.wss?uid=swg21086556>
2. Wprowadź następującą komendę, aby upewnić się, że moduł IOCP został zainstalowany na serwerze przeszukiwacza:  
`$ lspp -l bos.iocp.rte`  
Dane wyjściowe komendy `lspp` powinny być analogiczne do poniższego przykładu:

| Fileset                                 | Level    | State     | Description              |
|-----------------------------------------|----------|-----------|--------------------------|
| -----                                   |          |           |                          |
| Path: /usr/lib/objrepos<br>bos.iocp.rte | 5.2.0.10 | COMMITTED | I/O Completion Ports API |
| Path: /etc/objrepos<br>bos.iocp.rte     | 5.2.0.10 | COMMITTED | I/O Completion Ports API |
3. Wprowadź następującą komendę, aby upewnić się, że port IOCP ma status **Available** (dostępny):  
`$ lsdev -Cc iocp`  
Dane wyjściowe komendy `lsdev` powinny być zgodne z poniższym przykładem:  
`iocp0 Available I/O Completion Ports`
4. Jeśli port IOCP ma status **Defined**, należy go zmienić na **Available**:
  - a. Zaloguj się na serwerze przeszukiwacza jako administrator i wprowadź następującą komendę:  
`# smit iocp`
  - b. Wybierz opcję **Zmień/pokaż charakterystykę portów I/O Completion Ports** i zmień wartość opcji **STAN konfigurowany w momencie ładowania systemu** z **Zdefiniowane** na **Dostępne**.
  - c. Zrestartuj serwer przeszukiwacza.
  - d. Ponownie wprowadź komendę `lsdev` i potwierdź, że status portu IOCP zmienił się na **Available**.

#### Pojęcia pokrewne

“Przeszukiwacze Domino Document Manager” na stronie 61

“Przeszukiwacze Notes” na stronie 73

“Przeszukiwacze QuickPlace”

---

## Przeszukiwacze QuickPlace

Aby do kolekcji wyszukiwania korporacyjnego włączyć lokalizacje i pomieszczenia Lotus QuickPlace, należy skonfigurować przeszukiwacz QuickPlace.

Przeszukiwacza QuickPlace można również użyć do przeszukiwania lokalizacji zarządzanych z użyciem usług Lotus Quickr dla serwera Lotus Domino. Jeśli używane są usługi Lotus Quickr dla serwera WebSphere Portal, to dokumenty przechowywane w bibliotekach Lotus Quickr należy dodawać do kolekcji wyszukiwania korporacyjnego za pomocą przeszukiwacza listy adresów początkowych.

#### Wskazówka:

Szczegółowe przykłady konfigurowania bezpiecznego przeszukiwacza QuickPlace można znaleźć w scenariuszu dla małej organizacji w dokumentacji technicznej (redbook) firmy IBM



zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

## Konfiguracja serwera przeszukiwacza

Jeśli na serwerze, który ma być przeszukiwany, używany jest protokół zdalnego wywołania procedury Notes (NRPC), na serwerze przeszukiwacza trzeba uruchomić skrypt konfiguracyjny. Skrypt ten jest dostarczany razem z produktem OmniFind Enterprise Edition i umożliwia przeszukiwaczowi QuickPlace komunikowanie się z serwerami używającymi protokołu NRPC.

Jeśli na serwerze, który ma być przeszukiwany, używany jest protokół DIIOP (Domino Internet Inter-ORB Protocol), na serwerze przeszukiwacza nie trzeba uruchamiać skryptu konfiguracyjnego. Konieczne jest jednak skonfigurowanie serwera docelowego tak, aby przeszukiwacz QuickPlace miał do tego serwera dostęp.

Jeśli na serwerze, który ma być przeszukiwany, używany jest serwer LDAP (Lightweight Directory Access Protocol), to serwer docelowy musi zostać skonfigurowany do używania protokołu DIIOP (przeszukiwacz QuickPlace nie może używać protokołu NRPC do przeszukiwania danych LDAP). Należy również skonfigurować bazę danych Directory Assistance oraz skonfigurować serwer docelowy tak, aby używał serwera LDAP jako dodatkowego serwera Domino.

**Ważne:** Jeśli serwer docelowy używa protokołu DIIOP, a przeszukiwacz został skonfigurowany do używania protokołu HTTPS lub DIIOP przez SSL (dzięki czemu transmisja między przeszukiwaczem a serwerem jest szyfrowana), należy skopiować plik `TrustedCerts.class` (na przykład `c:\certs` lub `/data/certs`) z serwera docelowego na serwer przeszukiwacza. W konfiguracji dwuserwerowej i czteroserwerowej, należy również skopiować plik `TrustedCerts.class` na serwery, na których zainstalowany jest komponent wyszukiwania. Należy upewnić się, że plik ten znajduje się w tym samym katalogu na serwerze przeszukiwacza i na serwerze wyszukiwania. Ścieżka dla pliku `TrustedCerts.class` określana jest podczas konfigurowania przeszukiwacza.

Gdy konfigurowany jest przeszukiwacz i określany jest identyfikator użytkownika, którego przeszukiwacz ma używać, należy wybrać identyfikator, do którego przypisano dostateczne uprawnienia dostępu do wszystkich lokalizacji QuickPlace w domenie Domino. Aby to zapewnić, należy przypisać ten identyfikator użytkownika do zastrzeżonej grupy o nazwie `QuickPlaceAdministratorsSUGroup`.

Jeśli produkt OmniFind Enterprise Edition został zainstalowany w systemie IBM AIX, należy upewnić się, że na serwerze przeszukiwacza jest zainstalowany i dostępny moduł portu zakończenia `we/wy`.

Przed użyciem konsoli administracyjnej wyszukiwania korporacyjnego do skonfigurowania przeszukiwacza QuickPlace należy wykonać zadania odpowiednie dla używanego środowiska:

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 76.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 78.
- “Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 80.
- “Konfigurowanie serwera QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego” na stronie 284.
- “Konfigurowanie opcji Directory Assistance na serwerze QuickPlace” na stronie 284.

- “Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 81.

## Zabezpieczenia na poziomie dokumentu

Jeśli włączone są zabezpieczenia kolekcji, a serwer, który ma być przeszukiwany używa protokołu NRPC, należy na serwerze przeszukiwacza skonfigurować zaufany serwer Lotus Domino. Zaufany serwer służy do wymuszania kontroli dostępu na poziomie dokumentu. Przed udostępnieniem kolekcji użytkownikom do wyszukiwania należy wykonać następujące zadania:

- Skonfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika.
- Włączenie globalnych zabezpieczeń na serwerze WebSphere Application Server i skonfigurowanie aplikacji wyszukującej, aby korzystała z zabezpieczeń. Czynność ta gwarantuje, że referencje logowania są sprawdzane w momencie, gdy użytkownicy próbują korzystać z aplikacji wyszukującej. Serwery wyszukiwania korzystają z referencji do weryfikacji uprawnień użytkownika do uzyskiwania dostępu do dokumentów Lotus Domino.

## Przeszukiwanie załączników

W programie Lotus QuickPlace można importować i publikować dokumenty Microsoft Office (wśród opcji dostępnych podczas importowania są między innymi Imported Page, Microsoft Word Page, Microsoft Excel Page, Microsoft PowerPoint Page i Multiple Imported Pages). Przeszukiwacz QuickPlace może przeszukiwać te typy importowanych dokumentów jako załączniki, tylko jeśli spełnione są poniższe warunki:

- Serwer używa protokołu DIIOP.
- Przeszukiwanie załączników zostało włączone podczas konfigurowania opcji przeszukiwacza.
- Skonfigurowano przeszukiwacz do przeszukiwania pola "\$FILE" lub wszystkich pól.

## Przegląd konfiguracji

Przeszukiwacz QuickPlace może być używany do przeszukiwania dowolnej liczby lokalizacji QuickPlace. Podczas tworzenia przeszukiwacza wybiera się lokalizacje do przeszukiwania na pojedynczym serwerze QuickPlace. Później, w trakcie edycji przestrzeni przeszukiwania, można dodać do niej dokumenty z innych serwerów QuickPlace, które mają zostać włączone do tej samej przestrzeni przeszukiwania. Podczas tworzenia lub edycji przeszukiwacza można określić, czy przeszukiwane mają być wszystkie lokalizacje na serwerze czy tylko te, które zostały określone, oraz czy mają być przeszukiwane wszystkie pomieszczenia w lokalizacjach przeznaczonych do przeszukiwania czy tylko te, które zostały określone.

**Ograniczenie:** Określając identyfikator użytkownika, którego ma używać przeszukiwacz, należy wybrać użytkownika mającego uprawnienia dostępu do wszystkich lokalizacji QuickPlace w domenie Domino. Można to zrobić, konfigurując serwer QuickPlace i przypisując identyfikator użytkownika do zastrzeżonej grupy o nazwie QuickPlaceAdministratorsSUGroup. Należy mieć na uwadze, że nazwa grupy nie zawiera wbudowanych spacji i rozróżniana jest w niej wielkość liter.

Aby utworzyć lub zmienić przeszukiwacz QuickPlace, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich dokumentów w przestrzeni przeszukiwania.
- Określenie serwera QuickPlace oraz protokołu komunikacyjnego.
- Jeśli serwer jest skonfigurowany do korzystania z protokołu DIIOP, można określić, jak przeszukiwacz ma łączyć się z obiektami Domino. Na przykład można określić opcje korzystania z protokołu HTTPS lub Secure Sockets Layer (SSL) do szyfrowania komunikacji.
- Określenie informacji na temat katalogu użytkownika powiązanego z serwerem (przeszukiwacz potrzebuje tych informacji, aby możliwe było wymuszanie kontroli dostępu podczas wyszukiwania danych w kolekcji przez użytkowników).
- Wybór lokalizacji do przeszukiwania.
- Skonfigurowanie harmonogramu przeszukiwania lokalizacji.
- Wybranie dokumentów, które mają zostać przeszukane. Przeszukiwacz może przeszukiwać wszystkie pomieszczenia w lokalizacji lub też tylko dokumenty znajdujące się w wybranych pomieszczeniach.
- Określenie opcji udostępniających do wyszukiwania pola w różnych lokalizacjach i pomieszczeniach. Na przykład można wykluczyć niektóre pola z przestrzeni przeszukiwania i określić opcje wyszukiwania załączników.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Pojęcia pokrewne**

“Wskazówki dotyczące przeszukiwania baz danych Lotus Domino” na stronie 75

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino” na stronie 282

#### **Zadania pokrewne**

“Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 76

“Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 78

“Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 80

“Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 81

“Konfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika” na stronie 282

“Konfigurowanie serwera QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego” na stronie 284

“Konfigurowanie opcji Directory Assistance na serwerze QuickPlace” na stronie 284

---

## Przeszukiwacze listy adresów początkowych

Jeśli używane są usługi IBM Lotus Quickr dla produktu WebSphere Portal, przeszukiwacza listy adresów początkowych można użyć do dodawania dokumentów z bibliotek Lotus Quickr do kolekcji wyszukiwania korporacyjnego.

Biblioteka Lotus Quickr jest kontenerem na pliki dokumentów. Przeszukiwacz listy adresów początkowych nie obsługuje przeszukiwania treści z interfejsem WWW, takich jak serwisy wiki i blogi.

Jeśli używane są usługi Lotus Quickr dla produktu Lotus Domino, do dodawania dokumentów Lotus Quickr do kolekcji należy użyć przeszukiwacza QuickPlace.

### Konfiguracja serwera WebSphere Portal

Jeśli program Lotus Quickr jest zainstalowany na serwerze WebSphere Portal, wersja 6, to za pomocą przeszukiwacza listy adresów początkowych można przeszukiwać dokumenty biblioteki Lotus Quickr. Opcje przeszukiwania tych dokumentów można skonfigurować niezależnie od opcji określonych dla serwisów portali przeszukiwanych przez przeszukiwacz WebSphere Portal.

Przed utworzeniem przeszukiwacza listy adresów początkowych należy wykonać procedury konfigurowania wyszukiwania korporacyjnego w portalu WebSphere Portal. Aby skonfigurować środowisko wyszukiwania korporacyjnego, należy uruchomić skrypt (`wp6_install.sh` w systemie AIX, Linux lub Solaris albo `wp6_install.bat` w systemie Windows) udostępniany wraz z produktem OmniFind Enterprise Edition na serwerach wyszukiwania.

Łańcuch agenta użytkownika identyfikuje przeglądarkę lub robota uzyskującego dostęp do serwera. Przeszukując serwer Lotus Quickr, przeszukiwacz listy adresów początkowych używa łańcucha agenta użytkownika OmniFind SeedlistCrawler/1.0.

### Przegląd konfiguracji

Przeszukiwacz listy adresów początkowych może służyć do przeszukiwania dowolnej liczby dokumentów Lotus Quickr. Konfigurując ten przeszukiwacz, należy określić serwer do przeszukiwania. Przeszukiwacz będzie wtedy przeszukiwać wszystkie dokumenty zawarte w bibliotekach Lotus Quickr na tym serwerze.

Dokumenty do przeszukiwania muszą być dostępne z użyciem tego samego identyfikatora i hasła administratora produktu Lotus Quickr. Aby przeszukiwać serwisy, które korzystają z innych referencji, należy skonfigurować oddzielny przeszukiwacz listy adresów początkowych.

Aby utworzyć lub zmienić przeszukiwacz listy adresów początkowych, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza wpływają na sposób, w jaki dany przeszukiwacz przeszukuje dokumenty Lotus Quickr.
- Określenie serwera do przeszukiwania oraz informacji umożliwiających przeszukiwaczowi nawiązywanie połączenia z serwerem.

Podczas tworzenia lub edycji przeszukiwacza można przetestować jego zdolność do łączenia się z dokumentami, które mają być przeszukiwane. Przed uruchomieniem przeszukiwacza komunikaty informują użytkownika, czy przeszukiwacz ma dostęp do zadanych dokumentów.

- Określenie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Umożliwiają one kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

**Ważne:** Aby wyszukiwać w zabezpieczonych dokumentach Lotus Quickr, należy wprowadzać wyszukiwanie za pomocą portletu wyszukiwania korporacyjnego z portalu WebSphere Portal. Wyszukiwania wprowadzone z przykładowej aplikacji wyszukującej ESSearchApplication nie mają właściwych referencji i nie umożliwiają sprawdzenia uprawnień użytkownika do dostępu do dokumentów.

- Określenie informacji umożliwiających przeszukiwaczowi komunikowanie się z serwerem proxy, jeśli strony są dostarczane przy użyciu serwera proxy.
- Określenie referencji do pojedynczego logowania umożliwiających przeszukiwaczowi dostęp do dokumentów na serwerze, jeśli do ochrony serwera WebSphere Portal i dokumentów Lotus Quickr używany jest inny produkt (na przykład IBMTivoli Access Manager WebSEAL lub CA SiteMinder SSO Agent for PeopleSoft).
- Określenie informacji o pliku magazynu kluczy, aby przeszukiwacz mógł korzystać z protokołu Secure Sockets Layer (SSL) do łączenia się z serwerem.
- Określenie języka i strony kodowej dokumentów, które mają być przeszukiwane.
- Określenie opcji przeszukiwania i wyszukiwania metadanych w dokumentach Lotus Quickr.
- Określenie harmonogramów przeszukiwania dokumentów Lotus Quickr.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Pojęcia pokrewne**

“Integracja z produktem WebSphere Portal” na stronie 343

#### **Zadania pokrewne**

“Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 6” na stronie 350

“Konfigurowanie portletu wyszukiwania korporacyjnego dla produktu Lotus Quickr” na stronie 356

---

## **Przeszukiwacze systemu plików UNIX**

Aby do kolekcji wyszukiwania korporacyjnego włączyć dokumenty zapisane w systemie plików AIX, Linux lub Solaris, należy skonfigurować przeszukiwacz systemu plików UNIX.

Przeszukiwacz systemu plików UNIX może być używany do przeszukiwania dowolnej liczby systemów plików. Podczas konfigurowania przeszukiwacza należy wybrać lokalne i zdalne katalogi oraz podkatalogi, które mają zostać przeszukane.

Jeśli serwer przeszukiwacza jest zainstalowany na komputerze z systemem Windows, nie może być on używany do przeszukiwania źródeł w systemie plików AIX, Linux lub Solaris (przeszukiwacz systemu plików UNIX nie jest dostępny na liście dostępnych typów przeszukiwaczy).

Przeszukiwacz systemu plików UNIX przeszukuje dokumenty zgodnie z uprawnieniami odczytu określonymi dla administratora wyszukiwania korporacyjnego.

Aby utworzyć lub zmienić przeszukiwacz systemu plików UNIX, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają metodę przeszukiwania wszystkich podkatalogów w przestrzeni przeszukiwania.
- Skonfigurowanie harmonogramu przeszukiwania systemów plików.
- Wybór podkatalogów i poziomów podkatalogów, które mają zostać przeszukane przez określony przeszukiwacz.
- Określenie opcji udostępniania dokumentów w podkatalogach dla operacji przeszukiwania. Na przykład istnieje możliwość wykluczenia określonych typów dokumentów z przestrzeni przeszukiwania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

---

## Przeszukiwacze sieci WWW

Aby do kolekcji wyszukiwania korporacyjnego włączyć strony z serwisów WWW, należy skonfigurować przeszukiwacz sieci WWW.

Przeszukiwacz sieci WWW służy do przeszukiwania dowolnej liczby serwerów HTTP (Hypertext Transfer Protocol) i serwerów HTTPS (chroniony HTTP). Przeszukiwacz odwiedza serwis WWW i odczytuje dostępne w tym serwisie dane. Następnie podaje za odsyłaczami występującymi w dokumentach, aby przeszukać dodatkowe dokumenty. Przeszukiwacz sieci WWW może przeszukiwać i wyodrębniać odsyłacze z poszczególnych stron lub *zestawów ramek* (stron HTML zawierających ramki).

Przeszukane dane mogą być zapisane w jednym z wielu powszechnie używanych formatów i mogą pochodzić z różnych źródeł w sieci intranet lub Internet. Powszechnie używanymi formatami są: HTML, PDF, Microsoft Word, Lotus WordPro, Extensible Markup Language (XML) itd.

### Wskazówka:

Szczegółowe przykłady konfigurowania przeszukiwacza sieci WWW można znaleźć w scenariuszu dla średniej organizacji w dokumentacji technicznej (redbook) firmy IBM



zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

Aby utworzyć lub zmienić przeszukiwacz sieci WWW, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich serwisów WWW w przestrzeni przeszukiwania.
- Określenie reguł umożliwiających lub uniemożliwiających odwiedzanie określonych serwisów WWW. Po określeniu reguł przeszukiwania można je przetestować i zweryfikować, czy przeszukiwacz może uzyskać dostęp do serwisów, które mają zostać włączone do przestrzeni przeszukiwania.
- Określenie opcji dotyczących uwzględniania pewnych typów plików i wykluczania plików z wybranymi rozszerzeniami.
- Określenie reguł obsługi stron błędów programowych przez przeszukiwacz sieci WWW.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.
- Określenie opcji przeszukiwania serwisów WWW chronionych hasłem (serwery WWW, które mają zostać przeszukane, muszą używać podstawowego uwierzytelniania HTTP lub formularzy HTML monitorujących o wprowadzenie hasła).
- Określenie opcji przeszukiwania serwisów WWW obsługiwanych przez serwer proxy.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

## Konfiguracja agenta użytkownika

Aby przeszukać serwis WWW używający protokołu Robots Exclusion, należy sprawdzić, czy plik `robots.txt` umieszczony w serwisie WWW zezwala na dostęp do serwisu danemu agentowi użytkownika skonfigurowanemu dla przeszukiwacza sieci WWW.

Po uruchomieniu systemu wyszukiwania korporacyjnego przeszukiwacz sieci WWW ładuje skonfigurowaną dla niego nazwę agenta użytkownika. Zanim przeszukiwacz pobierze stronę z serwisu WWW, którego jeszcze nie odwiedzał (lub nie odwiedzał od pewnego czasu), będzie najpierw próbował pobrać plik o nazwie `robots.txt`. Plik ten znajduje się w katalogu głównym serwisu WWW.

Jeśli plik `robots.txt` nie istnieje, serwis WWW może być przeszukiwany bez żadnych ograniczeń. Jeśli plik istnieje, określa, jakie obszary serwisu (katalogi) są niedostępne dla przeszukiwaczy. Plik `robots.txt` określa uprawnienia dla przeszukiwaczy w zależności od określonej przez użytkownika nazwy agenta.

Protokół Robots Exclusion jest nieobowiązkowy, ale korporacyjny przeszukiwacz sieci WWW próbuje postępować zgodnie z nim:



- Jeśli w pliku `robots.txt` znajduje się wpis dla nazwy agenta użytkownika, która została skonfigurowana dla przeszukiwacza sieci WWW, wtedy przeszukiwacz podlega ograniczeniom określonym dla tego agenta.
- Jeśli nazwa agenta użytkownika nie występuje w pliku `robots.txt`, ale ostatni wiersz to wpis `User-agent: *` (który odnosi się do dowolnego agenta użytkownika) i występuje ograniczenie `Disallow: /` (które oznacza, że zabronione jest przeszukiwanie, począwszy od katalogu głównego serwisu WWW), wtedy przeszukiwacz sieci WWW nie może przeszukiwać tego serwisu.
- Jeśli nazwa agenta użytkownika nie występuje w pliku `robots.txt`, ale w ostatni wpis to `User-agent: *` i występuje ograniczenie `Allow: /`, wtedy przeszukiwacz sieci WWW może przeszukiwać dany serwis.

Administratorzy serwisów WWW często specjalnie dodają ostatni wpis uniemożliwiający dostęp wszystkim przeszukiwaczom, którym nie przydzielono jawnie praw dostępu. Jeśli w trakcie konfigurowania nowego przeszukiwacza sieci WWW wiadomo, że pewne serwisy WWW, które mają zostać przeszukane, używają protokołu Robots Exclusion, należy poprosić administratorów takich serwisów o dodanie wpisu w plikach `robots.txt` dla skonfigurowanego przeszukiwacza.

Należy pamiętać, żeby określić taką samą nazwę agenta użytkownika we właściwościach przeszukiwacza sieci WWW oraz we wszystkich plikach `robots.txt` w serwisach, których to dotyczy.

Jeśli żaden serwis WWW, który ma zostać przeszukany, nie używa protokołu Robots Exclusion, to wartość określona dla właściwości agent użytkownika z reguły nie ma znaczenia. Jednak pewne serwery aplikacji, strony JSP i serwlety dostosowują swoje odpowiedzi do nazwy agenta użytkownika. Możliwość wysyłania różnych odpowiedzi jest na przykład wykorzystywana do obsługi niezgodności między przeglądarkami internetowymi. W takich przypadkach określona nazwa agenta użytkownika może mieć znaczenie, nawet gdy protokół Robots Exclusion nie jest używany. W celu przeszukania serwisów tego typu należy skonsultować się z administratorem serwisu WWW, aby zapewnić odpowiedni dostęp dla przeszukiwacza sieci WWW.

## Jak przeszukiwacz sieci WWW korzysta z protokołu Robots Exclusion

Przeszukiwacz sieci WWW próbuje postępować zgodnie z protokołem Robots Exclusion i nie przeszukiwać serwisów, jeśli reguły w pliku `robots.txt` serwera nie zezwalają na przeszukiwanie.

Pomyślne pobranie ma miejsce, gdy przeszukiwacz może pobrać plik `robots.txt` z serwera WWW lub potwierdzić, że plik `robots.txt` nie istnieje. Kiedy przeszukiwacz nie może uzyskać reguł ani potwierdzić, że plik `robots.txt` istnieje, pobieranie uznaje się za zakończone niepowodzeniem.

Pobieranie zakończone powodzeniem nie oznacza, że przeszukiwacz ma uprawnienie do przeszukiwania, ponieważ reguły w pliku `robots.txt` mogą nie zezwalać na przeszukiwanie. Niepowodzenie pobierania tymczasowo uniemożliwia przeszukiwanie, ponieważ przeszukiwacz nie może określić reguł.

Poniżej przedstawiono czynności, jakie przeszukiwacz wykonuje, próbując pobrać plik `robots.txt`:

1. Kiedy przeszukiwacz wykryje nowy serwis, próbuje uzyskać adres IP serwera. Jeśli ta próba się nie powiedzie, przeszukiwanie nie będzie możliwe.

2. Kiedy dostępny jest przynajmniej jeden adres IP, przeszukiwacz próbuje pobrać plik robots.txt za pomocą komendy **HTTP (lub HTTPS) GET**.
3. Jeśli nastąpi przekroczenie limitu czasu połączenia na gnieździe, połączenie zostanie zerwane lub wystąpi inny błąd niskiego poziomu (np. problem z certyfikatem SSL), przeszukiwacz zarejestruje problem i powtórzy próbę dla każdego znanego adresu IP serwera docelowego.
4. Jeśli po wypróbowaniu wszystkich adresów nie zostanie nawiązane połączenie, przeszukiwacz odczeka dwie sekundy, a następnie wypróbuje każdy z adresów jeszcze raz.
5. W przypadku nawiązania połączenia i wymiany nagłówków HTTP zostanie przeanalizowany status powrotu. Jeśli kod statusu to 500 lub więcej, przeszukiwacz zinterpretuje to jako niską jakość połączenia i wypróbuje pozostałe adresy IP. W przypadku każdego innego statusu przeszukiwacz zaprzestanie próbowania alternatywnych adresów IP i postąpi zgodnie z kodem statusu.

Po odebraniu przez przeszukiwacz kodu statusu niższego niż 500 lub po dwukrotnym wypróbowaniu wszystkich adresów IP przeszukiwacz postępuje w następujący sposób:

1. Jeśli nie odebrano statusu HTTP poniżej 500, serwis jest tymczasowo dyskwalifikowany.
2. Jeśli odebrano status HTTP 400, 404 lub 410, serwis kwalifikuje się do przeszukiwania bez użycia reguł.
3. Jeśli odebrano status HTTP od 200 do 299, kolejne działanie zależy od następujących warunków:
  - Jeśli treść została obciążona, serwis jest tymczasowo dyskwalifikowany.
  - Jeśli podczas analizy treści nie znaleziono błędów, serwis kwalifikuje się do przeszukiwania zgodnie ze znalezionymi regułami.
  - Jeśli podczas analizy treści znaleziono błędy, serwis kwalifikuje się do przeszukiwania bez użycia reguł.
4. Jeśli został zwrócony inny status HTTP, serwis jest tymczasowo dyskwalifikowany.

Kiedy przeszukiwacz próbuje pobrać plik robots.txt dla serwisu, aktualizuje on dla niego trwały znacznik czasu zwany datą pliku robots. Jeśli serwis jest zdyskwalifikowany, ponieważ informacje z pliku robots.txt są niedostępne, liczba trwałych niepowodzeń pliku robots jest zwiększana o jeden.

Po upłygnięciu określonego czasu przeszukiwacz ponownie próbuje odtworzyć informacje w pliku robots.txt dla danego serwisu. Jeśli liczba kolejnych niepowodzeń osiągnie maksymalną dozwoloną liczbę niepowodzeń, przeszukiwacz przestaje próbować odtworzyć plik robots.txt dla serwisu i dyskwalifikuje serwis do przeszukiwania.

Po zakwalifikowaniu serwisu do przeszukiwania (sprawdzenie reguł w pliku robots.txt powiedzie się) licznik niepowodzeń jest zerowany. Przeszukiwacz korzysta z wyników pobrania do momentu, gdy upłynie odstęp czasu sprawdzania reguł. Po tym czasie należy ponownie zakwalifikować serwis.

#### **Wskazówka:**

- Jeśli serwer zwróci treść zawierającą błędy składni lub używa protokołu robots innego niż w wersji 1994 albo jeśli treść zawiera coś innego niż reguły protokołu robots (na przykład stronę błędu programowego), przeszukiwacz będzie działał tak, jakby serwer nie miał odpowiedniego pliku reguł i przeszuka serwis. Działanie takie jest zazwyczaj poprawne, ponieważ administratorzy kolekcji nie mają kontroli nad treścią serwisu ani domyślnym działaniem serwera. Jeśli administrator serwera WWW nie chce, aby serwis był przeszukiwany i nie chce zainstalować zgodnego pliku reguł, administrator kolekcji może

zablokować serwis WWW przed przeszukiwaczem, określając domenę, adres IP lub przedrostek HTTP w regułach przeszukiwacza.

- Jeśli serwer zwróci kod statusu 302 lub inny kod przekierowania, przeszukiwacz zinterpretuje go w ten sposób, że serwis ma plik `robots.txt`, którego należy użyć, ale nie znajduje się on w standardowym miejscu (katalog główny serwisu). Administrator serwera WWW musi przenieść plik w prawidłowe miejsce, aby przeszukiwacz mógł postąpić zgodnie z określonymi w nim regułami.
- Jeśli występują problemy z certyfikatem (na przykład jest on nieaktualny, ośrodek certyfikacji nie jest zaufany albo certyfikat jest samopodpisany, a przeszukiwacz nie jest skonfigurowany do akceptowania takich certyfikatów), przeszukiwacz interpretuje problem jako błąd połączenia z serwisem i dyskwalifikuje go. Te problemy prawdopodobnie i tak uniemożliwiłyby przeszukiwanie innych stron w tym serwisie. Aby umożliwić przeszukiwanie serwisu, administrator kolekcji musi włączyć samopodpisane certyfikaty, dodać uprawnienie serwisu do pliku magazynu kluczy zaufanych lub poprosić administratora serwera WWW o uzyskanie aktualnego certyfikatu.
- Przeszukiwacz sieci WWW może być skonfigurowany do korzystania z podstawowego uwierzytelniania HTTP (w tym podstawowego uwierzytelniania HTTP za pomocą serwera proxy). Jeśli uwierzytelnianie jest poprawnie skonfigurowane, będzie ono wymagane również do pobrania plików `robots.txt`. Kod statusu 403, 407 lub inne odpowiedzi związane z uwierzytelnianiem oznaczają problemy z autoryzacją i przeszukiwacz dyskwalifikuje taki serwis. (Obsługiwane jest wtedy tylko podstawowe uwierzytelnianie HTTP).
- Jeśli plik `robots.txt` dla serwisu jest większy niż maksymalna dopuszczalna długość strony `robots`, administrator kolekcji może zwiększyć skonfigurowaną wartość domyślną (równą jednemu milionowi bajtów), która jednak powinna być wystarczająca.

Aby rozwiązać problemy, można zażądać raportu dla serwisu podczas monitorowania przeszukiwacza. Należy w tym celu wybrać opcje wyświetlenia treści pliku `robots.txt` (aby zobaczyć, czy reguły uniemożliwiają przeszukiwaczowi dostęp do serwisu), daty i godziny ostatniej próby pobrania pliku `robots.txt` (przeszukiwacz nie będzie ponawiał próby, dopóki nie upłynie odstęp czasu między kolejnymi próbami) oraz liczby kolejnych prób, jakie przeszukiwacz podjął w celu pobrania pliku `robots.txt`, które jednak się nie powiodły. Monitorując przeszukiwacz, kliknij przycisk **Pomoc**, aby dowiedzieć się więcej o opcjach raportowania serwisów i sposobach interpretacji wyników.

Więcej informacji na temat protokołu Robots Exclusion można znaleźć pod adresem <http://www.robotstxt.org/wc/exclusion.html>.

## Obsługa języka JavaScript

Przeszukiwacz sieci WWW systemu wyszukiwania korporacyjnego może napotkać pewne odsyłacze (adresy URL), które są zawarte w skryptach JavaScript dokumentów WWW.

Przeszukiwacz sieci WWW wykrywa odsyłacze zarówno względne, jak i bezwzględne. Jeśli dokument HTML zawiera element `BASE`, przeszukiwacz używa tego elementu do wykrywania odsyłaczy względnych. W pozostałych przypadkach przeszukiwacz używa adresu URL samego dokumentu.

Obsługa języka JavaScript jest ograniczona do ekstrakcji odsyłaczy. Przeszukiwacz nie analizuje składni skryptów JavaScript, nie tworzy modelu DOM (Document Object Model) i nie interpretuje ani nie wykonuje instrukcji JavaScript. Przeszukiwacz wyszukuje łańcuchy w treści dokumentu (także w częściach zawierających skrypty JavaScript), które prawdopodobnie są adresami URL zawartymi w instrukcjach języka JavaScript. Oznacza to, że:

- Zostaną znalezione pewne adresy URL, które są pomijane przez bardziej rygorystyczne analizatory składni HTML. Przeszukiwacz odrzuci całą treść, która nie stanowi poprawnego składniowo adresu URL, jednak niektóre poprawne adresy URL zwrócone w trakcie skanowania mogą nie być szczególnie przydatne do wyszukiwania.
- Treść dokumentu wygenerowana przez kod JavaScript, na przykład wtedy, gdy przeglądarka uruchomi skrypt w języku JavaScript w momencie, gdy użytkownik wyświetli w niej stronę, nie może zostać wykryta przez przeszukiwacz, a zatem nie zostanie poindeksowana.

Ponieważ przeszukiwacz sieci WWW nie analizuje kodu JavaScript w plikach HTML, adresy URL w kodzie JavaScript nie są przeszukiwane. Aby przeszukiwaczowi sieci WWW umożliwić przeszukiwanie adresów URL w kodzie JavaScript, można wykonać jedną z następujących czynności:

- W konsoli administracyjnej wyszukiwania korporacyjnego poddaj edycji przeszukiwacz sieci WWW i na stronie Przestrzeń przeszukiwania w sieci WWW dodaj adresy URL do listy adresów URL, od których przeszukiwacz ma rozpocząć dodawanie adresów URL do kolekcji (**Początkowe adresy URL**). Aby zmiany zostały uwzględnione, zrestartuj przeszukiwacz sieci WWW (nie trzeba uruchamiać pełnego przeszukiwania).
- Używając znacznika zakotwiczenia (<a href="..">), określ adresy URL jako odsyłacze hipertekstowe w pliku HTML.

## Reguły ograniczania przestrzeni przeszukiwania w sieci WWW

Aby zapewnić dostęp użytkownikom tylko do serwisów WWW, w których mogą oni wykonywać przeszukiwanie, należy zdefiniować reguły określające elementy, które mogą być przeszukiwane przy użyciu przeszukiwacza sieci WWW.

Gdy przeszukiwacz sieci WWW przeszukuje serwis WWW, wykrywa odsyłacze do innych serwisów i umieszcza je w kolejce serwisów, które zostaną przeszukane w dalszej kolejności. Przeszukiwanie i wykrywanie odsyłaczy może być powtarzane, dopóki pozwala na to czas i dostępne zasoby pamięci. Podczas konfigurowania przeszukiwacza sieci WWW należy określić miejsce, w którym przeszukiwacz ma rozpocząć przeszukiwanie. Wychodząc od tych wstępnych adresów URL (określanych również jako *początkowe adresy URL*) przeszukiwacz sieci WWW może uzyskać dostęp do każdego dokumentu w sieci WWW, który jest powiązany przy użyciu bezpośredniego lub pośredniego odsyłacza.

Aby ograniczyć przestrzeń przeszukiwania, należy skonfigurować przeszukiwacz sieci WWW tak, aby określone adresy URL były dokładnie przeszukiwane, a odsyłacze wskazujące na dokumenty spoza obszaru zainteresowania - ignorowane. Ponieważ przeszukiwacz sieci WWW domyślnie akceptuje każdy wykryty adres URL, należy określić reguły identyfikujące adresy URL, które mają być uwzględniane w kolekcji i reguły eliminujące pozostałe serwisy WWW.

Jest kilka sposobów określenia elementów, które mają zostać przeszukane przy użyciu przeszukiwacza sieci WWW i elementów, które mają być pomijane. Są to następujące sposoby:

- Użycie listy początkowych adresów URL, od których przeszukiwacz ma rozpoczynać przeszukiwanie.
- Użycie jednego z trzech typów reguł przeszukiwania: reguły dotyczące domeny, adresu IP (Internet Protocol) i przedrostka URL.
- Użycie listy typów MIME dokumentów, które mają zostać włączone do przestrzeni wyszukiwania.
- Użycie listy rozszerzeń plików dokumentów, które mają zostać wykluczone.
- Określenie maksymalnej liczby katalogów w ścieżce URL.

Format reguł przeszukiwania jest następujący:  
działanie typ element\_docelowy

działanie określa jedno z dwóch działań: zabroń lub zezwól; typ to jeden z trzech typów: domena, adres IP lub przedrostek URL (HTTP albo HTTPS); element\_docelowy jest zależny od wartości typ. W ograniczonym zakresie można używać znaku wieloznacznego w formie gwiazdki (\*) służącego do określenia elementów docelowych zgodnych z podanym wzorcem.

## Reguły dotyczące domen

Element docelowy dla reguły dotyczącej domeny to nazwa domeny DNS. Na przykład można określić opcję przeszukiwania całej domeny `www.ibm.com`:

```
allow domain www.ibm.com
```

Można określić gwiazdkę jako znak wieloznaczny, co powoduje zastosowanie reguły do dowolnej nazwy hosta, która będzie zgodna z resztą wzorca. Na przykład można określić, że nie należy przeszukiwać domen zaczynających się od słowa `server` i kończących frazą `ibm.com`:

```
forbid domain server*.ibm.com
```

Podczas dopasowywania nazw hostów uwzględniane są wielkości liter - bez względu na to, czy zostanie określona jawna nazwa domeny, czy wzorec nazwy domeny. Na przykład z wzorcem `*.user.ibm.com` jest zgodna nazwa `joe.user.ibm.com` i nazwa `mary.smith.user.ibm.com`, lecz nie jest zgodna nazwa `joe.user.IBM.com`.

Reguła domeny, w której nie określono portu, dotyczy wszystkich portów w danej domenie. W poniższym przykładzie dozwolone są wszystkie porty w domenie `sales`:

```
allow domain sales.ibm.com
```

Jeśli w regule domeny został określony numer portu, wtedy reguła ma zastosowanie tylko do tego portu. W poniższym przykładzie dozwolony jest tylko port 443 w domenie `sales`:

```
allow domain sales.ibm.com:443
```

## Reguły dotyczące przedrostka

Reguła dotycząca przedrostka służy do sterowania przeszukiwaniem adresów URL, w których na początku nazwy występuje określony łańcuch. Elementem docelowym jest pojedynczy adres URL, w którym zazwyczaj występuje jedna lub więcej gwiazdek określających wzorec. Na przykład gwiazdka występuje często na końcu łańcucha przedrostka.

Reguła dotycząca przedrostka umożliwia przeszukanie całego serwisu WWW lub jego części. Użytkownik może określić nazwę ścieżki do katalogu lub wzorec nazwy, a następnie zezwolić lub zabronić dostępu do wszystkich elementów występujących dalej w drzewie katalogów. Na przykład następujące reguły są używane łącznie w celu zezwolenia przeszukiwaczowi na przeszukiwanie wszystkich elementów w katalogu `public` na serwerze `sales.ibm.com`, a zarazem w celu uniemożliwienia przeszukiwaczowi dostępu do jakichkolwiek innych stron w tym serwisie:

```
allow prefix http://sales.ibm.com/public/*
forbid prefix http://sales.ibm.com/*
```

Określając regułę dotyczącą przedrostka, można użyć wielu gwiazdek w dowolnym miejscu łańcucha przedrostka, a nie tylko na ostatniej pozycji tego łańcucha. Na przykład poniższa reguła zabrania przeszukiwaczowi przeszukiwania jakichkolwiek dokumentów w katalogu najwyższego poziomu serwisu `sales.ibm.com`, jeśli nazwa katalogu kończy się na `fs`. (Na

przykład można używać pozycji dołączania systemu plików, w których nie występują żadne informacje, które mogłyby być przydatne w indeksie wyszukiwania).

```
forbid http://sales.ibm.com/*fs/*
```

## Reguły dotyczące adresów

Reguła dotycząca adresu umożliwia sterowanie przeszukiwaniem całych hostów lub sieci przez określenie adresu IP i maski sieci jako elementu docelowego. Na przykład:

```
IPv4 allow address 9.0.0.0 255.0.0.0
```

**IPv6** Jeśli wyszukiwanie korporacyjne działa na serwerze Windows 2003, a dla systemu wyszukiwania korporacyjnego włączono protokół IP w wersji 6 (IPv6), adres należy ująć w nawiasy kwadratowe.

```
allow address [2001:db8:0:1:0:0:0:1]
```

Maska sieci umożliwia użycie dopasowywania wzorca. Aby reguła dotycząca adresu została zastosowana względem napotkanego adresu IP, adres IP określony w regule i napotkany adres IP muszą być identyczne, za wyjątkiem składowych, które w regule są w masce sieci zastąpione zerami. W regule dotyczącej adresu zdefiniowany jest wzorzec, a w masce sieci zdefiniowane są istotne bity we wzorcu adresu. Zero w masce sieci odpowiada znakowi wieloznacznemu i określa, że napotkany adres jest zgodny bez względu na wartość występującą na tej samej pozycji bitu w adresie.

W powyższym przykładzie reguła allow ma zastosowanie dla każdego adresu IP, w którym występuje wartość 9 w pierwszym oktecie i jakakolwiek wartość w pozostałych trzech oktetach.

Poniższą regułę warto uwzględnić jako końcowy adres na liście reguł adresowych. Według tej reguły zgodne są wszystkie adresy IP, ponieważ w masce sieci wszystkie bity są określone jako nieistotne (użycie tej reguły powoduje zabronienie dostępu do wszystkich adresów, do których dostęp nie został zezwolony przy użyciu reguł występujących na poprzednich pozycjach na liście reguł).

```
IPv4 forbid address 0.0.0.0 0.0.0.0
```

**IPv6**

```
forbid address :: ::
```

**Ograniczenia dotyczące serwerów proxy:** Jeśli planowane jest przeszukiwanie serwisów WWW obsługiwanych przez serwer proxy, nie należy określać reguł adresów IP. Serwer proxy jest zwykle używany, gdy agent użytkownika (przeglądarka lub przeszukiwacz) nie ma bezpośredniego dostępu do sieci, w których znajdują się serwery WWW. Na przykład serwer proxy HTTP może przekazywać żądania HTTP od przeszukiwacza do serwera WWW i dostarczać odpowiedzi z powrotem do przeszukiwacza.

Jeśli przeszukiwacz sieci WWW używa serwera proxy, adres IP tego serwera proxy jest jedynym adresem IP innego hosta dostępnym dla przeszukiwacza. Jeśli reguły związane z adresami IP są stosowane w celu ograniczenia działania przeszukiwacza do podsieci adresów IP, ograniczenie to powoduje, że prawie wszystkie adresy URL klasyfikowane są z kodem statusu 760 (co wskazuje, że są one zabronione przez przestrzeń WWW).

## Porządek reguł przeszukiwania

Przeszukiwacz stosuje reguły przeszukiwania na różnych etapach procesu wykrywania i przeszukiwania adresów URL. Porządek reguł ma istotne znaczenie, lecz dotyczy to wyłącznie reguł tego samego typu. Porządek, w jakim występują dwie reguły dotyczące



adresu, ma decydujące znaczenie, lecz nie ma znaczenia, czy reguła dotycząca adresu występuje przed, czy po regule dotyczącej przedrostka, ponieważ przeszukiwacz nie stosuje tych reguł w tym samym momencie.

Przeszukiwacz porównuje napotkaną domenę, adres lub adres URL z odpowiednimi elementami określonymi w każdej z reguł w ramach zestawu reguł tego samego typu - począwszy od pierwszej określonej reguły do ostatniej - zgodnie z ich porządkiem na liście reguł, dopóki nie zostanie znaleziona reguła, która może zostać zastosowana. Wykonywane jest działanie określone w pierwszej napotkanej regule, która może zostać użyta.

Z tej zależności od porządku reguł wynika typowa struktura większości reguł przeszukiwania:

- Zestaw reguł domenowych zazwyczaj rozpoczyna się od reguł zabraniających, które służą do wyeliminowania pojedynczych domen z przestrzeni przeszukiwania. Na przykład administrator kolekcji może określić, że w pewnych domenach nie ma przydatnych informacji.
- Zazwyczaj po liście reguł zabraniających występuje szereg reguł zezwalających (ze znakami wieloznacznymi) umożliwiających przeszukiwanie wszystkich domen, których nazwy zawierają na końcu nazwy domen wysokiego poziomu wchodzących w skład sieci intranet w przedsiębiorstwie (takie jak \*.ibm.com i \*.lotus.com).

Na końcu zestawu reguł dotyczących domen powinna występować następująca domyślna reguła służąca do wyeliminowania domen, które nie zostały dopuszczone przez wcześniejszą regułę:

```
forbid domain *
```

Ostatnia reguła jest regułą o znaczeniu krytycznym, ponieważ uniemożliwia włączenie całej sieci Internet do przestrzeni przeszukiwania.

- Zestaw reguł dotyczących adresów zaczyna się zazwyczaj od niewielkiej liczby reguł zezwalających przeszukiwaczowi na przeszukanie sieci wysokiego poziomu (klasy A, klasy B lub klasy C), które występują w sieci intranet w przedsiębiorstwie.

We wcześniejszym omówieniu reguł adresowych zapoznaj się z przykładami pokazującymi, jak określić końcową regułę na liście reguł adresowych, aby przeszukiwacz nie przeszukiwał serwisów WWW spoza sieci korporacyjnej.

- Zestaw reguł dotyczących przedrostków jest zazwyczaj największy, ponieważ znajdują się w nim dowolnie zdefiniowane specyfikacje regionów dostępnych i niedostępnych dla operacji przeszukiwania, które określone są jako drzewa i poddrzewa. Poprawna strategia polega na zezwoleniu lub zabronieniu wykonywania przeszukiwania najpierw w bardziej zawężonych lokalizacjach, a następnie określeniu reguł o przeciwnym działaniu, w których występują bardziej ogólne wzorce, służących do uniknięcia lub wymuszenia przeszukiwania całego pozostałego obszaru sieci.

Nie istnieje żadna typowa reguła występująca na końcu sekcji reguł dotyczących przedrostków. Sugerowane reguły końcowe dotyczące domen i adresów pozwalają skuteczniej zapobiec przeszukiwaniu obszaru poza siecią przedsiębiorstwa niż metoda testowania przedrostków adresów URL.

Przeszukiwacz może zastosować reguły dotyczące przedrostków bardziej wydajnie, jeśli reguły te zostaną pogrupowane według działań (grupy reguł zabraniających i zezwalających). Na przykład zamiast określania krótkich sekwencji różnych reguł zezwalających i zabraniających, lepiej określić długą sekwencję reguł związanych z określonym działaniem, a następnie długą sekwencję reguł związanych z innym działaniem. Reguły zezwalające i zabraniające można przeplatać, aby osiągnąć pewne cele w określonej przestrzeni przeszukiwania. Jednak zgrupowanie reguł zezwalających, a następnie zabraniających może poprawić wydajność przeszukiwacza.



## Rozszerzenia plików, typy MIME i maksymalna głębokość przeszukiwania

Opcje te służą do określenia dodatkowych ustawień związanych z treścią w określonej przestrzeni przeszukiwania. Można wykluczyć niektóre typy dokumentów w oparciu o rozszerzenie pliku dokumentu lub włączyć określone typy dokumentów na podstawie typów MIME dokumentów. Jeśli zostaną określone typy MIME, które mają zostać przeszukane przez określony przeszukiwacz, należy uwzględnić fakt, że typ MIME dokumentów w sieci WWW jest często niepoprawny.

Maksymalna głębokość przeszukiwania to liczba ukośników w adresie URL liczona od katalogu głównego serwisu. Opcja ta pozwala na uniknięcie przeszukiwania występujących w systemie plików rekurencyjnych struktur o nieskończonej głębokości. Głębokość przeszukiwania nie odpowiada liczbie poziomów odwiedzanych przez przeszukiwacz, gdy przechodzi on między poszczególnymi dokumentami przy użyciu odsyłaczy.

### Początkowe adresy URL

Początkowe adresy URL to adresy URL, od których przeszukiwacz zaczyna przeszukiwanie; są one stosowane w operacjach przeszukiwania przy każdym uruchomieniu przeszukiwacza. Jeśli początkowe adresy URL zostały już wykryte, nie zostaną one przeszukane ani przeszukane ponownie wcześniej niż serwisy WWW, których przeszukiwanie zostało dozwolone w regułach przeszukiwania.

Początkowy adres URL jest istotny przy pierwszym uruchomieniu przeszukiwacza sieci WWW, gdy przestrzeń przeszukiwania jest pusta. Początkowy adres URL jest istotny również wówczas, gdy użytkownik doda do listy początkowych adresów URL w istniejącej przestrzeni przeszukiwania adres URL, który nie został uprzednio wykryty.

Początkowe adresy URL muszą być pełnymi adresami URL, a nie tylko nazwami domen. Należy określić protokół oraz numer portu, jeśli nie jest używany port numer 80.

Następujące adresy URL są poprawnymi początkowymi adresami URL:

```
http://w3.ibm.com/
http://sales.ibm.com:9080/
```

Następujący adres URL nie jest poprawnym początkowym adresem URL:

```
www.ibm.com
```

Początkowe adresy URL należy włączyć do używanych reguł przeszukiwania. Przeszukiwacz nie może na przykład rozpocząć przeszukiwania od określonego początkowego adresu URL, jeśli reguły przeszukiwania nie zezwalają na jego przeszukiwanie.

**Obsługa adresów IPv6:** Jeśli wyszukiwanie korporacyjne działa na serwerze Windows 2003, a dla systemu wyszukiwania korporacyjnego włączono protokół IP w wersji 6 (IPv6), początkowe adresy URL należy ująć w nawiasy kwadratowe. Na przykład:

```
http://[2001:db8:0:1:0:0:0:1]
http://[2001:db8:0:1::1]
```

#### Zadania pokrewne

“Włączanie obsługi protokołu IPv6” na stronie 28

## Testowanie połączeń URL za pomocą przeszukiwacza sieci WWW

Po określeniu dla przeszukiwacza sieci WWW adresów URL do przeszukiwania można przetestować konfigurację reguł przeszukiwania.

Określając domeny, przedrostki HTTP lub adresy IP do przeszukiwania, można kliknąć opcję **Test** lub wybrać stronę Testuj adresy URL, aby przetestować możliwości przeszukiwacza w zakresie łączenia się ze startowymi adresami URL oprócz podanych.

Wyniki testu pokazują, czy przeszukiwacz może uzyskać dostęp do adresów URL przy użyciu nazwy agenta użytkownika określonej we właściwościach przeszukiwacza. Wyniki testu pokazują ponadto, czy adres URL nie może być przeszukiwany z powodu reguł wykluczania (na przykład dokument może nie być przeszukiwany, ponieważ ma rozszerzenie pliku zgodne z rozszerzeniem wykluczonym z przestrzeni przeszukiwania).

Gdy serwis zostanie przeszukany co najmniej raz, można przetestować adresy URL w celu uzyskania dodatkowych informacji. Na przykład w raporcie testu może być dostępny najbardziej aktualny kod statusu HTTP (który wskazuje, czy powiodło się przeszukanie danego adresu URL), informacje o dacie i godzinie ostatniego przeszukiwania adresu URL oraz dacie i godzinie zaplanowanego ponownego przeszukiwania pod tym adresem, a także informacje o tym, czy agent użytkownika korzysta z bieżącego pliku robots.txt na serwerze WWW.

## Ustawienia przeszukiwacza sieci WWW określające odstęp czasu przed ponownym przeszukiwaniem

Aby określić częstotliwość wykonywania ponownych odwiedzin adresów URL przez przeszukiwacz sieci WWW, należy określić odpowiednie opcje właściwości przeszukiwacza sieci WWW.

Większość pozostałych typów przeszukiwaczy używanych w wyszukiwaniu korporacyjnym jest uruchamiana zgodnie z harmonogramami zdefiniowanymi przez administratora. Natomiast przeszukiwacz uruchomiony ręcznie działa zazwyczaj ciągle. Aby sterować częstotliwością ponownego przeszukiwania adresów URL, które zostały już przeszukane, należy określić minimalny i maksymalny odstęp czasu między kolejnymi przeszukiwaniami.

Gdy do utworzenia przeszukiwacza sieci WWW lub edytowania jego właściwości używana jest konsola administracyjna wyszukiwania korporacyjnego, użytkownik może wybrać opcję konfigurowania właściwości zaawansowanych. Na stronie Zaawansowane właściwości przeszukiwacza sieci WWW użytkownik może określić opcje określające minimalny i maksymalny odstęp czasu między kolejnymi przeszukiwaniami. Określone wartości są używane przez przeszukiwacz sieci WWW do obliczenia odstępu czasu między ponownymi operacjami przeszukiwania danych.

Podczas pierwszego przeszukania określonej strony przeszukiwacz wykorzystuje datę i godzinę przeszukania tej strony oraz wartość średnią obliczoną przy użyciu minimalnego i maksymalnego odstępu czasu przed ponownym przeszukaniem do obliczenia daty ponownego przeszukiwania tej strony. Strona ta nie zostanie przeszukana ponownie przed tym terminem. Czas przeszukania tej strony po tym terminie zależy od obciążenia przeszukiwacza i bilansu nowych i starych adresów URL w przestrzeni przeszukiwania.

Podczas ponownego przeszukiwania strony przeszukiwacz sprawdza zmiany treści na tej stronie. Jeśli treść uległa zmianie, następny odstęp czasu przed ponownym przeszukiwaniem będzie krótszy od poprzedniego, lecz nigdy nie będzie krótszy niż określony minimalny odstęp czasu przed ponownym przeszukiwaniem. Jeśli treść nie uległa zmianie, wówczas

następny odstęp czasu przed ponownym przeszukiwaniem będzie dłuższy od poprzedniego, lecz nigdy nie będzie dłuższy niż określony maksymalny odstęp czasu przed ponownym przeszukiwaniem.

## Opcje odwiedzania adresów URL przy użyciu przeszukiwacza sieci WWW

Istnieje możliwość wymuszenia przeszukania konkretnych adresów URL przez przeszukiwacz sieci WWW w najbliższym możliwym terminie.

Gdy zaistnieje potrzeba odświeżenia przestrzeni przeszukiwania przy użyciu informacji dostępnych w określonych serwisach WWW, użytkownik może użyć funkcji monitorowania przeszukiwacza, wybrać opcję **Adresy URL do natychmiastowego odwiedzenia lub ponownego odwiedzenia**, a następnie określić adresy URL lub wzorce adresów URL stron, które mają zostać przeszukane lub przeszukane ponownie.

Na przykład, jeśli w Dziale Komunikacyjnym w sieci intranet strona WWW została dodana lub zaktualizowana przy użyciu nowych informacji w wyniku istotnych zmian strategii, można określić adres URL tej nowej lub zmodyfikowanej strony. Jeśli przeszukiwacz jest uruchomiony, określony adres URL zostanie umieszczony w kolejce stron do przeszukania i pozostanie w tej kolejce do następnego sprawdzenia listy stron oczekujących na przeszukanie (zazwyczaj kolejka ta przeglądana jest co 10 minut). Jeśli przeszukiwacz nie jest uruchomiony, określony adres URL zostanie umieszczony w kolejce i zostanie przeszukany po następnym uruchomieniu przeszukiwacza.

Należy upewnić się, że wśród reguł przeszukiwania znajduje się reguła zezwalająca przeszukiwaczowi na odwiedzenie określonych adresów URL. Przeszukiwacz będzie mógł przeszukać określone adresy URL wcześniej niż zwykle. Aby jednak adres URL mógł zostać w ogóle przeszukany, musi istnieć odpowiednia reguła przeszukiwania zezwalająca na przeszukanie określonego adresu URL.

Nowo przeszukane dane staną się dostępne do wyszukiwania po kolejnej operacji budowania indeksu głównego.

## Jak przeszukiwacz sieci WWW obsługuje strony błędów programowych

Przeszukiwacz sieci WWW można skonfigurować w taki sposób, aby obsługiwał niestandardowe strony tworzone przez administratorów serwisów WWW, które są zwracane zamiast standardowych kodów błędów w odpowiedzi na żądania dostępu do pewnych stron.

Jeśli serwer HTTP nie może zwrócić strony, której żąda klient, zazwyczaj zwraca odpowiedź zawierającą nagłówek z kodem statusu. Kod statusu określa, na czym polega problem (na przykład błąd 404, który wskazuje, że nie można było znaleźć pliku). Niektórzy administratorzy serwisów WWW tworzą specjalne strony, na których problemy są wyjaśnione w bardziej szczegółowy sposób, i konfiguruje serwer HTTP tak, aby zwracał te strony zamiast kodów błędów. Takie niestandardowe strony są nazywane *stronami błędów programowych*.

Strony błędów programowych mogą powodować przekłamanie wyników działania przeszukiwacza sieci WWW. Zamiast na przykład odebrać nagłówek wskazujący problem, przeszukiwacz odbiera stronę błędu programowego i kod statusu 200, który informuje o pomyślnym pobraniu poprawnej strony HTML. Jednak pobrana strona błędu programowego nie jest związana z żądanym adresem URL, jej treść jest niemal identyczna za każdym razem, gdy jest zwracana zamiast żądanej strony. Te nieistotne i prawie identyczne strony zniekształcają indeks i wyniki wyszukiwania.

W celu obsługi tego typu sytuacji można w trakcie konfigurowania przeszukiwacza sieci WWW określić opcje obsługi stron błędów programowych. Przeszukiwacz wymaga podania następujących informacji o każdym serwisie WWW, który zwraca strony błędów programowych:

- Wzorzec adresu URL serwisu, który używa stron błędów programowych. Wzorzec URL składa się z protokołu (HTTP lub HTTPS), nazwy hosta, numeru portu (jeśli jest inny niż standardowy) oraz nazwy ścieżki. Można używać gwiazdki (\*) jako znaku wieloznacznego dla jednego lub większej liczby znaków, aż do wystąpienia pierwszego znaku, który nie jest znakiem wieloznacznym. We wzorcach rozróżniana jest wielkość liter.
- Wzorzec tytułu tekstu, który odpowiada znacznikowi <TITLE> dokumentu HTML. Aby określić ten wzorzec, można użyć gwiazdki (\*) jako znaku wieloznacznego. W tym wzorcu rozróżniana jest wielkość liter.
- Wzorzec treści tekstu, który odpowiada treści dokumentu HTML. Jeśli znacznik <BODY> jest obecny, treść nie jest po prostu jego zawartością. Treść to wszystko, co jest umieszczone za nagłówkiem HTTP w pliku. Aby określić ten wzorzec, można użyć gwiazdki (\*) jako znaku wieloznacznego. W tym wzorcu rozróżniana jest wielkość liter.
- Liczba całkowita, która reprezentuje kod statusu, jaki ma być używany względem dokumentów zgodnych z podanym wzorcem adresu URL, tytułu i zawartości.

## Przykład

Następująca konfiguracja powoduje, że przeszukiwacz porównuje wszystkie poprawne strony HTML (kod statusu 200), które zwrócono z serwisu WWW `http://www.mojserwis.pl/ka/*` z podanymi wzorcami tytułu i zawartości. Jeśli znacznik <TITLE> strony zaczyna się od słów "Sorry, the page", a dokument zawiera dowolną treść (\*), przeszukiwacz traktuje stronę w taki sam sposób, jak traktowałby kod 404 (nie znaleziono strony).

Tabela 3. Przykład strony błędu programowego.

| Wzorzec adresu URL                        | Wzorzec tytułu                | Wzorzec treści | Kod statusu HTTP |
|-------------------------------------------|-------------------------------|----------------|------------------|
| <code>http://www.mojserwis.pl/ka/*</code> | <code>Sorry, the page*</code> | <code>*</code> | 404              |

Można utworzyć wiele pozycji dla tego samego serwisu WWW w celu obsługi różnych kodów statusu. Każdy kod statusu z tego samego serwisu WWW wymaga odrębnego wpisu w konfiguracji przeszukiwacza.

## Stosowanie znaków wieloznacznych

Wzorce adresu URL, tytułu oraz treści nie są wyrażeniami regularnymi. Znak gwiazdki zastępuje dowolne znaki, aż do wystąpienia pierwszego znaku niebędącego znakiem wieloznacznym. Na przykład:

`*404` zastępuje *dowolne znaki*  
`404: *` zastępuje *404: dowolne znaki*  
`http://*.mojserwis.pl/*` zastępuje *http://dowolny host.mojserwis.pl/dowolny plik*  
`*` zastępuje *dowolne znaki*

## Wpływ na wydajność

Skonfigurowanie opcji obsługi stron błędów programowych wydłuża czas przeszukiwania, gdyż wszystkie pomyślnie przeszukane strony muszą zostać sprawdzone. W celu dopasowania wzorca i określenia, czy należy zwrócić stronę, czy kod statusu zamiennika wymagany jest dłuższy czas przetwarzania.

## Przeszukiwanie zabezpieczonych serwisów WWW

Określenie referencji w konsoli administracyjnej wyszukiwania korporacyjnego umożliwia wyszukiwaczowi sieci WWW dostęp do zabezpieczonych treści (np. dokumentów wymagających podania hasła dostępu).

Jeśli serwer WWW używa podstawowego uwierzytelniania HTTP lub uwierzytelniania opartego na formularzu HTML w celu ograniczenia dostępu do serwisów WWW, można określić referencje w konfiguracji przeszukiwacza sieci WWW, które umożliwią przeszukanie serwisów WWW zabezpieczonych hasłem. Istnieje także możliwość ręcznego skonfigurowania plików z informacjami cookie.

### Serwisy WWW zabezpieczone przez podstawowe uwierzytelnianie HTTP

Jeśli serwer WWW używa podstawowego uwierzytelniania HTTP w celu ograniczenia dostępu do serwisów WWW, można określić referencje uwierzytelniania, które umożliwią wyszukiwaczowi sieci WWW dostęp do stron zabezpieczonych hasłem.

Wiele serwerów WWW używa schematu uwierzytelniania klienta nazywanego podstawowym uwierzytelnianiem HTTP w celu określenia tożsamości użytkownika i ustalenia, czy użytkownik (lub aplikacja klienta) ma wystarczające uprawnienia dostępu do stron w serwisie WWW. Zazwyczaj takie uwierzytelnianie ma charakter interaktywny:

- Kiedy agent HTTP użytkownika (np. przeglądarka WWW) żąda dostępu do strony zabezpieczonej za pomocą podstawowego uwierzytelniania HTTP, serwer WWW odpowiada kodem statusu 401, który oznacza, że agent nie ma dostępu do żądanej strony.
- Serwer WWW wzywa także agenta do przedstawienia referencji, które mogą służyć do zweryfikowania, czy użytkownik ma dostęp do zastrzeżonej treści.
- W przeglądarce WWW wyświetlane jest okno dialogowe, w którym należy podać nazwę użytkownika, hasło oraz inne wymagane informacje umożliwiające określenie referencji użytkownika.
- Przeglądarka WWW koduje referencje i wykorzystuje je ponownie przy kolejnym żądaniu dostępu do zabezpieczonej strony.
- Jeśli referencje są poprawne, serwer WWW odpowiada kodem powrotu 200 i udostępnia treść żądanej strony.
- Kolejne żądania stron z tego samego serwera WWW zazwyczaj zawierają te same referencje, co umożliwia uprawnionemu użytkownikowi uzyskanie dostępu do dodatkowej zastrzeżonej treści bez podawania dodatkowych referencji.

Po ustaleniu tożsamości użytkownika serwer WWW i agent użytkownika HTTP zazwyczaj wymieniają tokeny zwane *informacjami cookie*, dzięki którym znany jest status zaogłowania użytkownika między żądaniami HTTP.

Ponieważ przeszukiwacz sieci WWW nie działa w sposób interaktywny, referencje umożliwiające przeszukiwanie stron zabezpieczonych hasłem należy określić przed rozpoczęciem przeszukiwania. Podczas tworzenia przeszukiwacza sieci WWW lub edycji przestrzeni przeszukiwania należy określić informacje o każdym zabezpieczonym serwisie WWW, który ma zostać przeszukany.

Informacje te należy określić wspólnie z administratorami serwisów WWW lub serwerów WWW używających uwierzytelniania podstawowego HTTP. Administratorzy powinni udostępnić wymagania bezpieczeństwa dla serwisów WWW, które mają zostać przeszukane, w tym wszystkie informacje niezbędne do uwierzytelnienia tożsamości przeszukiwacza sieci WWW i określenia jego uprawnień do przeszukiwania zabezpieczonych stron.

Jeśli w trakcie tworzenia kolekcji zostały włączone zabezpieczenia, to w ramach konfigurowania przeszukiwacza można określić tokeny zabezpieczeń (np. identyfikatory

użytkowników, identyfikatory grup lub role użytkowników) w celu kontroli dostępu do dokumentów. Przeszukiwacz sieci WWW powiąże te tokeny zabezpieczeń z każdym przeszukiwanym dokumentem w drzewie systemu plików dla określonego głównego adresu URL. Te tokeny używane są oprócz tokenów zabezpieczeń na poziomie dokumentu, które zostały skonfigurowane dla całej przestrzeni przeszukiwania sieci WWW.

Istotne znaczenia ma kolejność adresów URL. Po dodaniu informacji o serwisie zabezpieczonym hasłem należy określić kolejność, w jakiej przeszukiwacz będzie przetwarzał adresy URL. Szczegółowo określone adresy URL powinny znaleźć się na początku listy, a bardziej ogólne na końcu. Przeszukiwacz sieci WWW podczas oceny danego adresu URL użyje danych uwierzytelnienia określonych dla pierwszego adresu URL, który odpowiada temu adresowi URL.

## **Serwisy WWW zabezpieczone przez uwierzytelnianie oparte na formularzu**

Jeśli serwer WWW używa formularzy HTML w celu ograniczenia dostępu do serwisów WWW, można określić referencje uwierzytelniania, które umożliwią przeszukiwaczowi sieci WWW dostęp do stron zabezpieczonych hasłem.

Wiele serwerów WWW używa formularzy HTML w celu określenia tożsamości użytkownika i ustalenia, czy użytkownik (lub aplikacja klienta) ma wystarczające uprawnienia dostępu do stron w serwisie WWW. Zazwyczaj takie uwierzytelnianie ma charakter interaktywny:

- Kiedy agent HTTP użytkownika (np. przeglądarka WWW) żąda dostępu do strony zabezpieczonej za pomocą uwierzytelniania opartego na formularzu, serwer WWW sprawdza, czy żądanie zawiera informację cookie potwierdzającą tożsamość użytkownika.
- Jeśli nie ma takiej informacji, serwer WWW wyświetla zachętę do wprowadzenia w formularzu danych bezpieczeństwa. Po wysłaniu formularza przez użytkownika serwer WWW zwraca wymagane informacje cookie i zezwala na żądanie dostępu do zabezpieczonej hasłem strony.
- Kolejne żądania, które zawierają wymagane informacje cookie, również mają dostęp do zabezpieczonych stron. Autoryzowany użytkownik uzyskuje dostęp do dodatkowych zabezpieczonych treści bez konieczności wypełniania formularza i podawania referencji przy każdym żądaniu.

Ponieważ przeszukiwacz sieci WWW nie działa w sposób interaktywny, referencje umożliwiające przeszukiwanie stron zabezpieczonych hasłem należy określić przed rozpoczęciem przeszukiwania. Podczas tworzenia przeszukiwacza sieci WWW lub edycji przestrzeni przeszukiwania należy określić informacje o każdym zabezpieczonym serwisie WWW, który ma zostać przeszukany.

Określone pola muszą odpowiadać polom, które użytkownik wypełnia w trybie interaktywnym w przeglądarce WWW. W celu pomyślnego logowania należy określić również wszystkie pola ukryte oraz statyczne.

Informacje te należy określić wspólnie z administratorami serwisów WWW lub serwerów WWW używających uwierzytelniania opartego na formularzu. Administratorzy powinni udostępnić wymagania bezpieczeństwa dla serwisów WWW, które mają zostać przeszukane, w tym wszystkie informacje niezbędne do uwierzytelnienia tożsamości przeszukiwacza sieci WWW i określenia jego uprawnień do przeszukiwania zabezpieczonych stron.

Istotne znaczenia ma kolejność wzorców adresów URL. Po dodaniu informacji o serwisie zabezpieczonym hasłem należy określić kolejność, w jakiej przeszukiwacz będzie przetwarzał adresy URL. Szczegółowo określone wzorce adresów URL powinny znaleźć się na początku



listy, a bardziej ogólne na końcu. Przeszukiwacz sieci WWW podczas oceny danego adresu URL użyje danych formularza określonych dla pierwszego wzorca adresu URL, który odpowiada temu adresowi URL.

### **Korzystanie z wtyczki do przeszukiwania bezpiecznych serwisów WebSphere Portal**

Jeśli na serwerze WebSphere Application Server włączone są globalne zabezpieczenia i wymagane jest przeszukiwanie bezpiecznych serwisów produktu WebSphere Portal przy użyciu przeszukiwacza WWW, konieczne jest utworzenie wtyczki przeszukiwacza do obsługi żądań uwierzytelniania opartego na formularzu. Omówienie uwierzytelniania opartego na formularzu oraz przykładowy program, który można dostosować na potrzeby swojej niestandardowej wtyczki przeszukiwacza WWW, można znaleźć pod adresem <http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani>.

Wtyczka ta jest wymagana, gdy przeszukiwacz WWW jest używany do przeszukiwania dowolnych serwisów za pośrednictwem produktu WebSphere Portal, w tym serwisów Workplace Web Content Management oraz serwisów Lotus Quickr.

## **Serwisy WWW obsługiwane przez serwery proxy**

Jeśli przeszukiwacz sieci WWW nie ma zezwolenia na bezpośredni dostęp do sieci, można skonfigurować go, aby korzystał z serwera proxy HTTP w celu uzyskania dostępu do treści, która ma zostać przeszukana.

Jeśli na komputerze, na którym ma zostać uruchomiony przeszukiwacz sieci WWW, nie ma dostępu do sieci TCP/IP lub dostęp jest ograniczony do procesów uprzywilejowanych, można skonfigurować przeszukiwacz do korzystania z serwera HTTP proxy. HTTP proxy to proces, który nasłuchuje żądań HTTP na określonym porcie podanego hosta. Serwer proxy przekazuje żądania do serwera WWW i odpowiedzi z tego serwera do klienta wysyłającego żądania (przeszukiwacz sieci WWW). Serwer proxy może działać na tym samym komputerze, na którym działa przeszukiwacz sieci WWW, lub na innym.

W trakcie przeszukiwania bez pośrednictwa serwera proxy żądania adresu URL są wysyłane bezpośrednio do hosta. W trakcie przeszukiwania za pośrednictwem serwera proxy żądania są wysyłane do serwera proxy.

Podczas tworzenia przeszukiwacza sieci WWW lub edycji przestrzeni przeszukiwania należy określić informacje o serwerach proxy, których przeszukiwacz używa podczas przeszukiwania stron w domenie serwera proxy. Przed dodaniem serwera proxy do przestrzeni przeszukiwania należy uzyskać nazwy domen obsługiwanych przez ten serwer proxy, jego nazwę hosta lub adres IP oraz używany przez niego numer portu.

Jeśli serwer proxy wymaga uwierzytelnienia, uzyskaj nazwę użytkownika i hasło, dzięki któremu przeszukiwacz sieci WWW będzie mógł uzyskać dostęp do stron udostępnianych przez serwer proxy. Przeszukiwacz sieci WWW obsługuje tylko podstawowe uwierzytelnianie proxy HTTP opisane w dokumencie RFC2616 (<http://rfc.net/rfc2616.html>). Inne rodzaje autoryzacji, takie jak Windows NT LAN Manager (NTLM), nie są obsługiwane.

Po dodaniu serwera proxy należy go wybrać i określić kolejność, w jakiej przeszukiwacz będzie z niego korzystał. Szczegółowo określone domeny powinny znaleźć się na początku listy, a bardziej ogólne na końcu. Gdy przeszukiwacz sieci WWW ocenia dany adres URL, używa danych serwera proxy, które są określone dla pierwszej domeny z listy, która odpowiada temu adresowi URL. Zakłada się, że adresy URL, które nie pasują do żadnej reguły proxy, są bezpośrednio dostępne dla przeszukiwacza.



## Administrowanie informacjami cookie

Zazwyczaj administrowanie informacjami cookie odbywa się w sposób automatyczny bez konieczności podejmowania żadnych działań ze strony administratora wyszukiwania korporacyjnego. Jednak w razie potrzeby można ręcznie określić informacje cookie dla sesji przeszukiwania sieci WWW.

Informacje cookie to osłonięte tokeny, które serwer WWW zwraca do agenta użytkownika jako część nagłówka odpowiedzi HTTP. Informacje cookie są ważne jedynie dla tego serwera WWW, które je wystawił, i są używane do utrzymywania stanu między kolejnymi żądaniami HTTP. Na przykład podczas uwierzytelniania klienta serwer WWW może zwracać informacje cookie, które pozwalają serwerowi określić, czy uwierzytelniony użytkownik jest już zalogowany. Obecność informacji cookie umożliwia użytkownikowi wysyłanie dodatkowych żądań dotyczących stron na tym serwerze WWW bez konieczności ponownego logowania.

Przeszukiwacz sieci WWW zachowuje informacje cookie otrzymane z serwerów WWW i używa ich w czasie działania instancji przeszukiwacza. Informacje cookie są przechowywane w pliku `cookies.ini`, który jest ponownie zapisywany na koniec każdej sesji przeszukiwacza. Po zatrzymaniu działania przeszukiwacza sieci WWW wszystkie informacje cookie, które nie wygasły, są zapisywane i przeladowywane na początku następnej sesji.

Ręcznie określone informacje cookie należy zapisać w oddzielnym pliku, a następnie w razie potrzeby scalić je z informacjami cookie z pliku `cookies.ini`. Przeszukiwacz nie usuwa informacji cookie, które nie wygasły, ale w razie problemu uniemożliwiającego zapisanie całej kolekcji informacji cookie informacje wprowadzone ręcznie nie zostaną utracone. Przed rozpoczęciem sesji przeszukiwania należy ręcznie scalić własne informacje cookie z informacjami, które są obsługiwane automatycznie przez przeszukiwacz.

### Format pliku cookie

Informacje cookie, które mają być scalone z plikiem `cookies.ini` systemu wyszukiwania korporacyjnego, muszą mieć ściśle określony format.

- Każda informacja cookie musi się znajdować w osobnym wierszu. Dopuszczalne jest użycie pustych wierszy i komentarzy, które jednak nie zostaną zachowane w pliku `cookies.ini`.
- Każda informacja cookie musi mieć następujący format:

```
CookieN(długość_cookie,długość_URL)tekst_cookie,adres_URL_sprawdzania_poprawności
```

#### Cookie

Wymagany parametr definiujący początek wpisu informacji cookie.

Parametr Cookie nie może zawierać spacji i musi mieć dołączoną pojedynczą cyfrę: 0, 1 lub 2. Cyfra określa typ informacji cookie: wersja 0 (Netscape), wersja 1 (RFC2109) lub wersja 2 (RFC2965). Informacje cookie typu RFC2965 nie obsługują list portów.

#### *długość\_cookie*

Długość tekstu cookie, wyrażona liczbą znaków.

#### *długość\_URL*

Długość adresu URL sprawdzania poprawności, wyrażona liczbą znaków.

#### *tekst\_cookie*

Zawartość informacji cookie, która ma zostać wysłana do źródłowego serwera WWW. Łańcuch ten (odpowiadający prawej stronie dyrektywy Set-Cookie w nagłówku odpowiedzi HTTP) określa parę nazwa-wartość oraz wszelkie inne dane (na przykład ścieżkę, ustawienia zabezpieczeń itp.), które będą wysyłane razem z informacją cookie. Po tym łańcuchu następuje separator w postaci przecinka (,).

#### *adres\_URL\_sprawdzania\_poprawności*

Adres URL, pod którym dana informacja cookie została odkryta. Na podstawie tego

adresu ustala się, gdzie należy wysłać informację cookie (na przykład przez podanie nazwy domeny i ścieżki). Adres URL sprawdzania poprawności musi spełniać wymagania dotyczące zabezpieczeń i poufności dla informacji cookie obowiązujące na źródłowym serwerze WWW.

W poniższym przykładzie wprowadzono podział na dwa wiersze dla zwiększenia czytelności. Informacje cookie określane przez użytkowników muszą w rzeczywistości mieścić się w jednym wierszu:

```
Cookie0(53,40)ASPSESSIONIDQSQTACSD=SLNSIDFNLSIDNFLSIDNFLSNL;path=/,
https://www.ibm.com:443/help/solutions/
```

## Konfigurowanie informacji cookie dla przeszukiwacza sieci WWW

Informacje cookie można skonfigurować ręcznie na potrzeby sesji przeszukiwania WWW oraz scalić je z informacjami cookie przechowywanymi przez przeszukiwacz WWW.

### Zagadnienia wstępne

Aby ręcznie skonfigurować informacje cookie dla przeszukiwacza sieci WWW, trzeba być administratorem wyszukiwania korporacyjnego.

### Procedura

Aby ręcznie skonfigurować informacje cookie dla przeszukiwacza sieci WWW:

1. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom monitorowanie kolekcji, dla której chcesz skonfigurować informacje cookie, i zatrzymaj przeszukiwacz WWW.
2. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerze przeszukiwania (ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition).
3. Przejdź do katalogu danych przeszukiwacza, który ma zostać skonfigurowany, gdzie *identyfikator\_sesji\_przeszukiwacza* to identyfikator przypisany sesji przeszukiwacza przez system wyszukiwania korporacyjnego. Na przykład:  
ES\_NODE\_ROOT/data/col\_56092.WEB\_88534
4. Zmodyfikuj plik `cookies.ini`, dołącz zdefiniowane ręcznie pozycje cookie do pozycji obecnych już w pliku, a następnie zapisz i zamknij plik. Upewnij się, że nowe informacje cookie nie przesłoniły tych, które już były w pliku obecne.
5. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom ponownie zatrzymany wcześniej przeszukiwacz sieci WWW.

## Konfiguracja globalnej przestrzeni przeszukiwania sieci WWW

Możliwość konfigurowania globalnej przestrzeni przeszukiwania dla przeszukiwaczy sieci WWW zapewnia lepszą kontrolę usuwania adresów URL z indeksu.

Każdy przeszukiwacz sieci WWW jest skonfigurowany za pomocą przestrzeni przeszukiwania, która określa adresy URL, które mają lub nie mają być przeszukiwane. Adresy URL znajdujące się w przestrzeni przeszukiwania są zachowywane (w bazie danych) w celu późniejszego przeszukania; adresy URL, które nie znajdują się w przestrzeni przeszukiwania, są usuwane. Jeśli przeszukiwacz zostanie uruchomiony z pustą bazą danych, definicja przestrzeni przeszukiwania oraz baza danych pozostają spójne w trakcie działania przeszukiwacza.

Czasami przeszukiwacz zostaje zatrzymany, a jego przestrzeń przeszukiwania zmniejszona (na przykład przez nowe reguły zabraniające przeszukiwania pewnych stron). W takiej sytuacji po zrestartowaniu przeszukiwacza definicja jego przestrzeni przeszukiwania i baza

danych stają się niespójne. Baza danych zawiera wówczas adresy URL (zarówno te przeszukane, jak i nieprzeszukane), których nie ma już w nowej, mniejszej przestrzeni przeszukiwania.

Jeśli kolekcja zawiera tylko jeden przeszukiwacz sieci WWW, wtedy przeszukiwacz może przywrócić spójność, zmieniając kody statusu HTTP dla tych adresów URL na 760 (co określa, że adresy mają zostać wykluczone) i żądając usunięcia nowo wykluczonych stron z indeksu.

Jeśli przestrzeń przeszukiwania jest podzielona między dwa lub więcej przeszukiwaczy sieci WWW (na przykład w celu zapewnienia, że pewne strony będą przeszukiwane częściej niż pozostałe), to każdy przeszukiwacz sieci WWW obsługuje niezależne tabele w bazie danych (początkowo puste) i każdy przeszukuje inną część przestrzeni przeszukiwania sieci WWW. Pierwotna przestrzeń przeszukiwania przeszukiwacza zostaje pomniejszona o wszystkie części, które mają być przeszukiwane przez pozostałe przeszukiwacze. Problemy pojawiają się, kiedy pierwotny przeszukiwacz próbuje przywrócić spójność, usuwając z indeksu przeniesione strony. Ponieważ przeniesione strony są obecnie przeszukiwane przez inne przeszukiwacze, powinny pozostać w indeksie.

Konfiguracja przestrzeni globalnej, która jest przestrzenią przeszukiwania wyższego poziomu, umożliwia określenie adresów URL, które mają nie być przeszukiwane przez pierwotny przeszukiwacz, ale nie powinny być także usunięte z indeksu. Adresy URL, których nie ma w żadnej przestrzeni przeszukiwania, dalej są oznaczone jako wykluczone z procesu wykrywania i zostaną usunięte z indeksu po ponownym przeszukaniu.

Globalna przestrzeń przeszukiwania jest określona za pomocą pliku konfiguracyjnego o nazwie `global.rules`. Plik ten musi istnieć w katalogu konfiguracyjnym przeszukiwacza (istnienie pliku `global.rules` uaktywnia funkcję obsługi globalnej przestrzeni przeszukiwania). Jeśli plik istnieje, jest odczytywany podczas inicjowania przeszukiwacza. Jeśli plik nie istnieje, przeszukiwacz działa w jednopoziomowej przestrzeni przeszukiwania, a dokumenty są usuwane z indeksu w celu zapewnienia spójności między definicją przestrzeni przeszukiwania i bazą danych.

Jeśli istnieje globalna przestrzeń przeszukiwania, obowiązują dalej te same reguły włączania i wyłączania adresów URL, ale żądania usunięcia adresu URL z indeksu będą realizowane tylko, gdy adres URL nie istnieje w żadnej przestrzeni przeszukiwania sieci WWW.

Składnia pliku `global.rules` jest taka sama, jak w przypadku pliku reguł lokalnych `crawl.rules`, jednak plik globalny może zawierać tylko reguły nazw domen. To ograniczenie umożliwia partycjonowanie przestrzeni przeszukiwania między przeszukiwacze wyłącznie na podstawie nazw hostów DNS, a nie na podstawie adresów IP lub wzorców przedrostków HTTP. Globalna przestrzeń przeszukiwania nie ma wpływu na adresy URL, które zostały wykluczone za pomocą reguł dotyczących adresów IP lub przedrostków adresów URL w lokalnej przestrzeni przeszukiwania (określonych w pliku `crawl.rules`); takie adresy URL dalej pozostaną wykluczone.

Globalna przestrzeń przeszukiwania jest używana tylko do zapobiegania usuwaniu z indeksu adresów URL, które zostały wykluczone z przestrzeni przeszukiwania jednego z przeszukiwaczy za pomocą lokalnej reguły domeny. Reguły są stosowane w następującym porządku:

1. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły przedrostków lub reguły adresów, adresowi URL zostaje przypisany kod 760 i adres jest usuwany z indeksu. Adres URL nie będzie ponownie przeszukiwany.

2. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły domeny, a nie istnieje globalna przestrzeń przeszukiwania, wtedy adresowi URL zostaje przypisany kod 760 i adres jest usuwany z indeksu. Adres URL nie będzie ponownie przeszukiwany.
3. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły domeny, ale jest jawnie dozwolony przez regułę w globalnej przestrzeni przeszukiwania, wtedy adresowi URL zostaje przypisany kod 761. Adres URL nie będzie ponownie przeszukiwany, ale nie zostanie usunięty z indeksu (zakłada się, że adres URL znajduje się w lokalnej przestrzeni przeszukiwania innego przeszukiwacza).
4. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły domeny i nie jest jawnie dozwolony przez regułę w globalnej przestrzeni przeszukiwania, wtedy adresowi URL zostaje przypisany kod 760 i adres jest usuwany z indeksu.

Ponieważ globalna przestrzeń przeszukiwania jest sprawdzana tylko w celu zapobiegania usuwaniu adresów URL wykluczonych już z lokalnej przestrzeni przeszukiwania, to domyślne zabrania ona przeszukiwania adresów URL, o ile dla tych adresów nie zostały zastosowane inne reguły.

Plik `global.rules` musi istnieć w katalogu `master_config` każdego przeszukiwacza, który współużytkuje globalną przestrzeń przeszukiwania. Należy starannie przeprowadzić edycję wszystkich kopii pliku `global.rules` oraz poszczególnych plików lokalnych `crawl.rules`, aby zapewnić ich wzajemną spójność.

## Dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie

Jakość wyszukiwania można poprawić, określając dyrektywy dla przeszukiwacza sieci WWW, które decydują o tym, czy należy podążać za odsyłaczami znajdującymi się na stronie i czy strona ma być indeksowana.

Pewne strony WWW zawierają dyrektywy blokujące wykorzystywanie odsyłaczy lub indeksowanie, które informują roboty (takie jak przeszukiwacz sieci WWW) o tym, aby nie podążać za odsyłaczami znalezionymi na stronie, nie uwzględniać zawartości strony w indeksie albo nie wykonywać żadnego z tych działań.

Sterowanie tymi ustawieniami może poprawić jakość przeszukiwania. Na przykład pewne strony katalogów mogą zawierać tysiące odsyłaczy, ale żadnej użytecznej treści. Odsyłacze z takich stron powinny zostać przeszukane, ale poindeksowanie samych stron nie przyniesie żadnych korzyści.

Czasami użytkownik chce, aby przeszukiwacz nie przeszukiwał dokumentów poniżej pewnego poziomu w hierarchii, jednak żądane strony-liście zawierają odsyłacze, ale nie zawierają dyrektyw blokujących ich wykorzystywanie. Ponieważ wiele takich stron jest generowanych automatycznie, nie mają one właścicieli, którzy mogliby wstawić wymagane dyrektywy.

W celu określenia reguł przeszukiwania takich stron można utworzyć bądź wyedytować plik konfiguracyjny o nazwie `followindex.rules`. Poniżej przedstawiono wskazówki pomocne przy określaniu reguł w tym pliku:

- Konfigurowane reguły muszą zawierać przedrostki URL (nie można określać serwisów WWW za pomocą adresu IP lub nazwy hosta DNS).
- Przedrostki URL mogą zawierać gwiazdki (\*) jako znaki zastępcze, co umożliwia sformułowanie zezwolenia lub zakazu dla wielu serwisów o podobnych adresach URL.

- Kolejność reguł ma znaczenie (przeszukiwacz stosuje pierwszą regułę, która pasuje do sprawdzanego adresu URL).
- Reguły, które jawnie zezwalają lub zabraniają korzystania z odsyłaczy lub indeksowania, zastępują inne ustawienia (także ustawienia znajdujące się w dokumencie docelowym).

## Ignorowanie dyrektyw blokujących wykorzystywanie odsyłaczy i indeksowanie na stronach WWW

W pliku konfiguracyjnym można określić reguły decydujące o tym, czy przeszukiwacz WWW ma podążać za odsyłaczami do stron lub indeksować strony, które zawierają dyrektywy blokujące takie działania.

### Zagadnienia wstępne

Aby skonfigurować dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie dla przeszukiwacza sieci WWW, trzeba być administratorem wyszukiwania korporacyjnego. Dyrektywy te przesłaniają dyrektywy zawarte w kodzie przeszukiwanych stron.

### Procedura

Aby przesłonić dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie:

1. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom monitorowanie kolekcji, dla której chcesz skonfigurować reguły, i zatrzymaj przeszukiwacz sieci WWW.
2. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerze przeszukiwania. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
3. Przejdź do katalogu konfiguracyjnego przeszukiwacza, który ma zostać skonfigurowany, gdzie *identyfikator\_sesji\_przeszukiwacza* to identyfikator przypisany sesji przeszukiwacza przez system wyszukiwania korporacyjnego. Na przykład:  
ES\_NODE\_ROOT/master\_config/col\_56092.WEB\_88534
4. Utwórz lub zmodyfikuj plik o nazwie `followindex.rules`.
5. Wprowadź reguły dla przeszukiwacza w poniższym formacie, gdzie *przedrostek\_URL* to początkowe znaki w adresach serwisów WWW, których mają dotyczyć reguły zezwalające lub zakazujące wykorzystywania odsyłaczy lub indeksowania:  

```
forbid follow przedrostek_URL
allow follow przedrostek_URL
forbid index przedrostek_URL
allow index przedrostek_URL
```
6. Zapisz i zamknij plik.
7. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom ponownie zatrzymany wcześniej przeszukiwacz sieci WWW.

## Konfigurowanie daty używanej przez przeszukiwacze sieci WWW dla przeszukanych dokumentów

Istnieje możliwość określenia opcji w pliku konfiguracyjnym określającej datę używaną przez przeszukiwacze sieci WWW jako daty dla przeszukanych dokumentów.

### Zagadnienia wstępne

Domyślnie przeszukiwacze sieci WWW ustawiają wartość w polu `Data` dla przeszukanych dokumentów na datę przeszukania. Można jednak określić ustawianie tej daty na datę ostatniej modyfikacji zwracanej przez serwer WWW.

Data ostatniej modyfikacji może nie być dostępna dla wszystkich dokumentów. Jeśli przeszukiwacz zostanie skonfigurowany do używania tej wartości i serwer WWW nie zwróci daty ostatniej modyfikacji dla dokumentu, wtedy przeszukiwacz ustawi wartość pola Data dla przeszukanego dokumentu na datę i godzinę przeszukania.

Tabela 4. W jaki sposób przeszukiwacze sieci WWW konfiguruje pole daty dla przeszukanych dokumentów

| Konfiguracja przeszukiwacza                         | Data zwracana przez serwer WWW                                                   | Pole metadanych Data dla przeszukanych dokumentów       |
|-----------------------------------------------------|----------------------------------------------------------------------------------|---------------------------------------------------------|
| Konfiguracja domyślna                               | Serwer WWW zwraca wartość ostatniej modyfikacji dla przeszukanego dokumentu      | Data i godzina przeszukania                             |
|                                                     | Serwer WWW nie zwraca wartości ostatniej modyfikacji dla przeszukanego dokumentu | Data i godzina przeszukania                             |
| Skonfigurowano do użycia daty ostatniej modyfikacji | Serwer WWW zwraca wartość ostatniej modyfikacji dla przeszukanego dokumentu      | Wartość ostatniej modyfikacji zwracana przez serwer WWW |
|                                                     | Serwer WWW nie zwraca wartości ostatniej modyfikacji dla przeszukanego dokumentu | Data i godzina przeszukania                             |

Aby określić datę używaną przez przeszukiwacze sieci WWW do ustawiania metadanych dla przeszukanych dokumentów, trzeba być administratorem wyszukiwania korporacyjnego.

### Procedura

Aby skonfigurować przeszukiwacz sieci WWW, aby data przeszukanych dokumentów była datą i godziną ostatniej modyfikacji zwracanej przez serwery WWW:

1. Zaloguj się na serwerze przeszukiwacza jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Otwórz w edytorze plik `crawl.properties` dla przeszukiwacza sieci WWW, który ma być skonfigurowany. Parametr `id_sesji_przeszukiwacza` jest identyfikatorem przypisanym do sesji przeszukiwacza przez serwer wyszukiwania korporacyjnego.  
ES\_NODE\_ROOT/master\_config/id\_sesji\_przeszukiwacza/crawl.properties

Na przykład:

```
/home/esadmin/master_config/col_00112.WEB_23344/crawl.properties
C:\Program Files\IBM\esadmin\master_config\col_55667.WEB_78899\crawl.properties
```

3. Dodaj następujący wiersz i zapisz plik:  
`which_date=LastModified`
4. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom ponownie przeszukiwacza sieci WWW.

Jeśli serwer WWW zwraca datę ostatniej modyfikacji dla dokumentów odczytanych przez przeszukiwacza, będzie ona używana do wypełniania pola Data dla przeszukanych dokumentów.



---

## Przeszukiwacze Web Content Management

Aby w kolekcji wyszukiwania korporacyjnego zostały uwzględnione dokumenty produktu IBM Workplace Web Content Management, należy skonfigurować przeszukiwacz Web Content Management.

### Konfiguracja serwera WebSphere Portal

Po zainstalowaniu produktu Workplace Web Content Management na serwerze WebSphere Portal, wersja 6 można używać przeszukiwacza Web Content Management do przeszukiwania serwisów Web Content Management. Opcje przeszukiwania tych serwisów można skonfigurować niezależnie od opcji określonych dla serwisów portali przeszukiwanych przez przeszukiwacz WebSphere Portal.

Przed utworzeniem przeszukiwacza Web Content Management należy wykonać procedury konfigurowania wyszukiwania korporacyjnego w portalu WebSphere Portal. Aby skonfigurować środowisko wyszukiwania korporacyjnego, należy uruchomić skrypt (`wp6_install.sh` w systemie AIX, Linux lub Solaris albo `wp6_install.bat` w systemie Windows) udostępniany wraz z produktem OmniFind Enterprise Edition na serwerach wyszukiwania.

### Format adresu URL

Określając adresy URL do przeszukiwania, należy używać następującego formatu:

```
protokół_http://nazwa_hosta_portalu:port/przedrostek_portalu
/ścieżka_do_serwletu_zarodka_wyszukiwania_WCM
/searchseed?site=nazwa_serwisu_WCM&lib=nazwa_biblioteki_WCM
```

W poniższym przykładzie pokazano adres URL serwisu znajdującego się w domyślnej ścieżce instalacyjnej produktu Workplace Web Content Management na portalu WebSphere Portal:

```
http://portal.server.ibm.com:80/wps/wcmsearchseed/searchseed?site=SiteTest01
&lib=Web+Content
```

Jeśli nazwa serwisu lub biblioteki zawiera spację, należy je zastąpić znakiem plus (+). Na przykład wyrażenie **Web Content** należy zastąpić wyrażeniem **Web+Content**.

### Przegląd konfiguracji

Przeszukiwacz Web Content Management może służyć do przeszukiwania dowolnej liczby serwisów Web Content Management. Podczas konfigurowania przeszukiwacza użytkownik określa adresy URL serwisów do przeszukiwania. Następnie przeszukiwacz pobiera strony, do których znajdują się odsyłacze w podanych serwisach.

Serwisy do przeszukiwania muszą być dostępne za pomocą tego samego identyfikatora administratora i hasła do portalu WebSphere Portal. Aby przeszukiwać serwisy, które korzystają z innych referencji, należy skonfigurować oddzielny przeszukiwacz Web Content Management.

Aby utworzyć lub zmienić przeszukiwacz Web Content Management, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają metodę przeszukiwania wszystkich dokumentów w przestrzeni przeszukiwania.

- Określenie adresów URL serwisów do przeszukiwania oraz informacji umożliwiających przeszukiwaczowi nawiązanie połączenia z danymi serwisami.

Podczas tworzenia lub edycji przeszukiwacza można przetestować jego zdolność do łączenia się z podanymi adresami URL. Przed uruchomieniem przeszukiwacza komunikaty informują użytkownika, czy przeszukiwacz ma dostęp do zadanych dokumentów.

- Określenie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Umożliwiają one kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

**Ważne:** Aby wyszukiwać w zabezpieczonych stronach Web Content Management, należy wprowadzać wyszukiwanie za pomocą portletu wyszukiwania korporacyjnego z portalu WebSphere Portal. Wyszukiwania wprowadzone z przykładowej aplikacji wyszukującej ESSearchApplication nie będą miały właściwych referencji i nie umożliwią sprawdzenia uprawnień użytkownika do dostępu do dokumentów.

- Określenie informacji umożliwiających przeszukiwaczowi komunikowanie się z serwerem proxy, jeśli serwisy Web Content Management korzystają z serwera proxy do udostępniania dokumentów.
- Określenie referencji do pojedynczego logowania umożliwiających przeszukiwaczowi dostęp do dokumentów na serwerze, jeśli do ochrony serwera WebSphere Portal i serwisów WWW używany jest inny produkt (na przykład IBMTivoli Access Manager WebSEAL lub CA SiteMinder SSO Agent for PeopleSoft).
- Określenie informacji o pliku magazynu kluczy, aby przeszukiwacz mógł korzystać z protokołu Secure Sockets Layer (SSL) do łączenia się z serwisami Web Content Management.
- Określenie języka i strony kodowej dokumentów, które mają być przeszukiwane.
- Określenie opcji przeszukiwania i wyszukiwania metadanych w dokumentach Web Content Management.
- Określenie harmonogramów przeszukiwania serwisów Web Content Management.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Pojęcia pokrewne**

“Integracja z produktem WebSphere Portal” na stronie 343

#### **Zadania pokrewne**

“Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 6” na stronie 350

---

## **Przeszukiwacze WebSphere Portal**

Aby do kolekcji wyszukiwania korporacyjnego włączyć strony serwisu IBM WebSphere Portal, należy skonfigurować przeszukiwacz WebSphere Portal.

## Konfiguracja serwera WebSphere Portal

Przed utworzeniem przeszukiwacza WebSphere Portal należy uruchomić skrypt w celu skonfigurowania środowiska wyszukiwania korporacyjnego w portalu WebSphere Portal. Dla różnych wersji portalu WebSphere Portal wymagane są różne skrypty. Są one instalowane na serwerach wyszukiwania podczas instalowania produktu OmniFind Enterprise Edition.

- W przypadku portalu WebSphere Portal w wersji 5.1.0 lub nowszej należy uruchomić skrypt `wp5_install.sh` w systemie AIX, Linux lub Solaris albo skrypt `wp5_install.bat` w systemie Windows.
- W przypadku portalu WebSphere Portal, wersja 6 należy uruchomić skrypt `wp6_install.sh` w systemie AIX, Linux lub Solaris albo skrypt `wp6_install.bat` w systemie Windows.

### Wskazówka:

Szczegółowe przykłady konfigurowania bezpiecznego przeszukiwacza WebSphere Portal można znaleźć w scenariuszu dla średniej organizacji w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

## Przegląd konfiguracji

Przeszukiwacz WebSphere Portal może być używany do przeszukiwania pojedynczego serwisu WebSphere Portal. Podczas konfigurowania przeszukiwacza użytkownik określa adres URL serwisu portalu do przeszukiwania. Następnie przeszukiwacz pobiera portlety i strony, do których znajdują się odsyłacze z podanych adresów URL portali. Aby przeszukać inny serwis portalu, należy utworzyć nowy przeszukiwacz.

Aby utworzyć lub zmienić przeszukiwacz WebSphere Portal, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania przez niego wszystkich stron serwisu.
- Określenie adresu URL serwisu portalu do przeszukiwania oraz informacji umożliwiających przeszukiwaczowi nawiązanie połączenia z danym serwisem. Ponieważ adresy URL tego typu mogą być długie i zawierać zakodowane znaki spoza zestawu znaków ASCII, taki adres URL wygodnie jest skopiować z serwera WebSphere Portal i wkleić w konsoli administracyjnej wyszukiwania korporacyjnego.

Podczas tworzenia lub edycji przeszukiwacza można przetestować jego zdolność do łączenia się z podanym adresem URL. Przed uruchomieniem przeszukiwacza komunikaty informują użytkownika, czy przeszukiwacz ma dostęp do zadanych dokumentów.

- Określenie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Umożliwiają one kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

**Ważne:** Aby wyszukiwać w zabezpieczonych stronach portalu WebSphere Portal, należy wprowadzać wyszukiwanie za pomocą portletu wyszukiwania korporacyjnego z portalu WebSphere Portal. Wyszukiwania wprowadzone z przykładowej aplikacji wyszukującej ESSearchApplication nie mają właściwych referencji i nie umożliwiają sprawdzenia uprawnień użytkownika do dostępu do dokumentów.

- Określenie informacji umożliwiających przeszukiwaczowi komunikowanie się z serwerem proxy, jeśli serwis WebSphere Portal korzysta z serwera proxy do udostępniania stron.
- Określenie referencji do pojedynczego logowania umożliwiających przeszukiwaczowi dostęp do dokumentów na serwerze, jeśli do ochrony serwera WebSphere Portal i serwisów WWW używany jest inny produkt (na przykład IBM Tivoli Access Manager WebSEAL lub CA SiteMinder SSO Agent for PeopleSoft).
- Określenie informacji o pliku magazynu kluczy, aby przeszukiwacz mógł korzystać z protokołu Secure Sockets Layer (SSL) do łączenia się z serwisem WebSphere Portal.
- Określenie języka i strony kodowej dokumentów, które mają być przeszukiwane.
- Określenie opcji przeszukiwania i wyszukiwania metadanych w dokumentach portalu WebSphere Portal.
- Określenie harmonogramów przeszukiwania serwisów portalu WebSphere Portal.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Pojęcia pokrewne**

“Integracja z produktem WebSphere Portal” na stronie 343

#### **Zadania pokrewne**

“Kopiowanie adresu URL do przeszukania z portalu WebSphere Portal”

“Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 6” na stronie 350

“Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 5.1” na stronie 345

## **Kopiowanie adresu URL do przeszukania z portalu WebSphere Portal**

Aby zmniejszyć możliwość wpisania niepoprawnego adresu URL, podczas konfigurowania przeszukiwacza WebSphere Portal lub Web Content Management można do odpowiedniego pola skopiować adres URL serwisu, który ma zostać przeszukany.

### **Informacje o tym zadaniu**

Podczas tworzenia przeszukiwacza WebSphere Portal lub Web Content Management należy podać adres URL serwisu na serwerze WebSphere Portal, który ma zostać przeszukany. Ponieważ takie adresy URL są długie i zawierają zwykle zakodowane znaki spoza zestawu znaków ASCII, wygodne może okazać się użycie tej procedury w celu skopiowania adresu URL z serwera WebSphere Portal i wklejenia go w konsoli administracyjnej wyszukiwania korporacyjnego.

### **Procedura**

Aby podać adres URL, który przeszukiwacz ma przeszukać:

1. Przed określeniem adresu lub adresów URL w konsoli administracyjnej wyszukiwania korporacyjnego należy upewnić się, że serwer WebSphere Portal jest uruchomiony, a następnie zalogować się do niego jako administrator.

2. W przypadku portalu WebSphere Portal, wersja 5.1 wykonaj następujące czynności na serwerze WebSphere Portal:
  - a. Kliknij opcję **Administrowanie** w prawym górnym rogu.
  - b. Kliknij opcję **Ustawienia portalu** w obszarze nawigacyjnym z lewej strony, a następnie kliknij opcję **Administrowanie wyszukiwaniem**.
  - c. Na stronie Zarządzaj kolekcjami wyszukiwania kliknij opcję **Kolekcja portalu** w obszarze Kolekcje wyszukiwania. Jeśli są dostępne inne kolekcje, można wybrać inną.
  - d. W obszarze Źródła treści w kolekcji kliknij opcję **Dodaj źródło treści**.
  - e. Dla opcji **Typ przeszukiwanego źródła** wybierz ustawienie **Serwis portalu**. Adres URL serwisu zostanie wyświetlony w polu **Pobieraj powiązane dokumenty, zaczynając od tego adresu URL**.
  - f. Skopiuj ten adres URL do schowka Na przykład podświetl adres URL, a następnie, trzymając klawisz Ctrl, naciśnij klawisz Insert.
3. W przypadku portalu WebSphere Portal, wersja 6 wykonaj następujące czynności na serwerze WebSphere Portal:
  - a. W lewym dolnym rogu kliknij opcję **Administracja**.
  - b. Kliknij opcję **Zarządzaj wyszukiwaniem** w obszarze nawigacyjnym z lewej strony.
  - c. Na stronie Zarządzaj wyszukiwaniem kliknij odsyłacz **Kolekcje wyszukiwania**.
  - d. Na stronie Zarządzaj wyszukiwaniem kliknij opcję **Domyślna usługa wyszukiwania w portalu** z opcji usługi Wyszukiwanie.
  - e. W tabeli Kolekcje wyszukiwania kliknij kolekcję **Treść portalu**.
  - f. W tabeli Źródła treści kliknij pierwszą z prawej ikonę Edycja (z ołówkiem) znajdującą się obok opcji **Źródło treści portalu**.
  - g. Dla opcji **Typ źródłowy treści** wybierz ustawienie **Serwis portalu**. Adres URL serwisu zostanie wyświetlony w polu **Pobieraj powiązane dokumenty, zaczynając od tego adresu URL**.
  - h. Skopiuj ten adres URL do schowka Na przykład podświetl adres URL, a następnie, trzymając klawisz Ctrl, naciśnij klawisz Insert.
4. Wróć do konsoli administracyjnej wyszukiwania korporacyjnego i wklej adres URL skopiowany z pola adresu URL serwisu.

#### Pojęcia pokrewne

“Przeszukiwacze WebSphere Portal” na stronie 111

---

## Przeszukiwacze systemu plików Windows

Aby do kolekcji wyszukiwania korporacyjnego włączyć dokumenty zapisane w systemie plików Microsoft Windows, należy skonfigurować przeszukiwacz systemu plików Windows.

Przeszukiwacz systemu plików Windows może być używany do przeszukiwania dowolnej liczby systemów plików Windows. Podczas konfigurowania przeszukiwacza należy wybrać lokalne i zdalne katalogi oraz podkatalogi, które mają zostać przeszukane.

W przypadku zainstalowania serwera przeszukiwacza na komputerze z systemem AIX, Linux lub Solaris nie można go używać do przeszukiwania źródeł w systemach plików Windows (przeszukiwacz systemu plików Windows nie jest wyświetlany na liście dostępnych typów przeszukiwaczy).

#### Wskazówka:

Szczegółowe przykłady konfigurowania bezpiecznego przeszukiwacza systemu plików Windows można znaleźć w scenariuszu dla małej organizacji w dokumentacji technicznej

(redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

## Przeszukiwanie współużytkowanych katalogów sieciowych

Przeszukiwacz systemu plików Windows przeszukuje dokumenty zgodnie z uprawnieniami odczytu określonymi dla administratora wyszukiwania korporacyjnego. Administratorowi odpowiada konto usług OmniFind Enterprise Edition.

Można określić identyfikator użytkownika i hasło dla katalogów do przeszukania. Jednak będą one używane tylko do łączenia się ze współużytkowanymi katalogami sieciowymi. Przeszukiwacz przeszukuje pliki zgodnie z uprawnieniami do odczytu ustawionymi dla tego użytkownika we współużytkowanych katalogach sieciowych, a nie napędach lokalnych.

Połączenia z katalogami sieciowymi nie zostają rozłączone, dopóki nie zostanie zrestartowana usługa OmniFind Enterprise Edition. Po nawiązaniu połączenia możliwy jest dostęp do katalogu z niepoprawnym identyfikatorem użytkownika i hasłem. Jednak to połączenie jest dozwolone tylko dla tych sesji wykrywania i przeszukiwacza systemu plików Windows, które są sterowane przez system wyszukiwania korporacyjnego. Aby zapobiec ewentualnemu naruszeniu bezpieczeństwa, należy upewnić się, że autoryzacja dla konta administratora wyszukiwania korporacyjnego (z którym działa usługa OmniFind Enterprise Edition) jest ustawiona prawidłowo.

Aby w przyszłości uniknąć problemów z łączeniem się z katalogiem sieciowym, należy podać ten sam identyfikator użytkownika i hasło dla danego katalogu sieciowego. W przypadku podania błędnego identyfikatora użytkownika i hasła, a następnie zrestartowania usługi OmniFind Enterprise Edition, przeszukiwanie za pomocą przeszukiwacza systemu plików Windows może się nie powieść, ponieważ będzie on próbował połączyć się z katalogiem, używając niepoprawnych referencji. Przeszukiwanie może powieść się, jeśli połączenie sieciowe zostanie nawiązane przez inny przeszukiwacz systemu plików Windows, używający poprawnego identyfikatora użytkownika i hasła.

## Przegląd konfiguracji

Aby utworzyć lub zmienić przeszukiwacz systemu plików Windows, należy zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego. Użytkownik musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub musi on być administratorem kolekcji, do której należy określony przeszukiwacz.

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają metodę przeszukiwania wszystkich podkatalogów w przestrzeni przeszukiwania.
- Skonfigurowanie harmonogramu przeszukiwania systemów plików.
- Wybór podkatalogów do przeszukania.

Można zdefiniować liczbę poziomów katalogów przeszukiwanych przez określony przeszukiwacz. Aby umożliwić przeszukiwanie zdalnych systemów plików, należy podać także identyfikator użytkownika i hasło umożliwiające przeszukiwaczowi uzyskanie dostępu do danych.

- Określenie opcji udostępniania dokumentów w podkatalogach dla operacji przeszukiwania. Na przykład można wykluczyć z przestrzeni przeszukiwania określone typy dokumentów albo określić identyfikator użytkownika i hasło, które umożliwi przeszukiwaczowi dostęp do plików w konkretnym podkatalogu.



- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

Do wymuszenia zabezpieczeń na poziomie dokumentu konieczne jest poprawne skonfigurowanie informacji na temat konta użytkownika i domeny na serwerze przeszukiwacza.

Aby uzyskać więcej informacji na temat pól w kreatorze i sposobu określania informacji potrzebnych do przeszukiwania danych, podczas tworzenia przeszukiwacza należy kliknąć opcję **Pomoc**.

#### **Pojęcia pokrewne**

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows” na stronie 285

“Bezpieczne wyszukiwanie w zaufanych domenach systemu Windows” na stronie 287

---

## **Konfiguracja obsługi aplikacji nasłuchiwania danych**

Istnieje możliwość rozszerzenia wyszukiwania korporacyjnego przez wykorzystanie funkcji API nasłuchiwania danych do utworzenia przeszukiwacza zewnętrznego. Niestandardowe aplikacje nasłuchiwania danych mogą dodawać dane do kolekcji lub wskazywać sposób, w jaki przeszukiwacz sieci WWW ma odwiedzać adresy URL.

### **Zagadnienia wstępne**

Aby konfigurować aplikacje nasłuchiwania danych, wymagana jest rola administratora wyszukiwania korporacyjnego.

**Ważne:** Nasłuchiwanie danych nie będzie obsługiwane w przyszłych wersjach. Na potrzeby tworzenia aplikacji klienckich dla wyszukiwania korporacyjnego należy używać interfejsów API wyszukiwania i indeksowania (SI-API) zamiast interfejsów API nasłuchiwania danych. Poniższe informacje są przeznaczone dla użytkowników, którzy utworzyli już wcześniej aplikacje nasłuchiwania danych.

### **Informacje o tym zadaniu**




Aplikacja kliencka nasłuchiwania danych umożliwi przeszukiwanie typów danych, które nie mogą być przeszukiwane przez domyślne przeszukiwacze w wyszukiwaniu korporacyjnym. Przed rozpoczęciem korzystania z aplikacji nasłuchiwania danych należy skonfigurować referencje, aby aplikacja mogła uzyskać dostęp do kolekcji i mogła ją aktualizować.

Gdy aplikacja kliencka nasłuchiwania danych łączy się z aplikacją nasłuchiwania danych, musi przekazać swój identyfikator i hasło oraz identyfikator aktualizowanej kolekcji. Informacje te muszą być zgodne z informacjami skonfigurowanymi dla aplikacji w konsoli administracyjnej.

Proces nasłuchiwania danych jest uruchamiany automatycznie przy uruchamianiu systemu wyszukiwania korporacyjnego. Jeżeli po skonfigurowaniu aplikacji za pomocą konsoli administracyjnej zostanie zmieniony numer portu, należy zatrzymać i ponownie uruchomić aplikację nasłuchiwania danych.

## Procedura

Aby skonfigurować aplikację nasłuchiwania danych, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Nasłuchiwanie danych kliknij opcję **Konfiguruj aplikacje nasłuchiwania danych**.
4. Na stronie Aplikacje nasłuchiwania danych określ liczbę wątków, które aplikacja nasłuchiwania danych może utworzyć podczas przetwarzania żądań z aplikacji klienckich, i numer portu, na którym aplikacja nasłuchiwania danych będzie nasłuchiwać żądań. Określ także maksymalną liczbę dokumentów przypadających na kolekcję, które mogą być przechowywane w pamięci tymczasowej, zanim analizator składni rozpocznie ich analizowanie.
5. Kliknij opcję **Dodaj aplikację nasłuchiwania danych**, aby dodać informacje o aplikacji klienckiej.
6. Na stronie Dodaj aplikację nasłuchiwania danych określ informacje uwierzytelniające, które pozwolą aplikacjom klienckim nasłuchiwania danych na dostęp do kolekcji wyszukiwania korporacyjnego. Identyfikatory klientów nasłuchiwania danych muszą być unikalne w obrębie systemu wyszukiwania korporacyjnego.
7. Wybierz kolekcje, które aplikacja nasłuchiwania danych będzie mogła aktualizować:
  - Kliknij opcję **Wszystkie kolekcje**, jeśli aplikacja ma aktualizować wszystkie kolekcje.
  - Kliknij opcję **Konkretne kolekcje**, jeśli aplikacja ma aktualizować tylko wskazane kolekcje.    Po wybraniu tej opcji wyświetlona zostaje lista nazw kolekcji. Zaznacz pole wyboru **Wybierz** dla każdej kolekcji, którą aplikacja może aktualizować.
8. Kliknij przycisk **OK**.
9. Jeśli zmieniony został port nasłuchiwania danych lub liczba dokumentów, jakie mogą być przechowywane w pamięci tymczasowej, uruchom ponownie aplikację nasłuchiwania danych:
  - a. Kliknij ikonę  **Monitoruj**, aby przejść do widoku monitorowania systemu.
  - b. Na stronie Nasłuchiwanie danych kliknij ikonę  **Restartuj**.

### Zadania pokrewne

“Monitorowanie nasłuchiwania danych” na stronie 318

---

## Niestandardowe wtyczki przeszukiwacza

Podczas konfigurowania właściwości przeszukiwaczy można określić klasę języka Java, która ma być stosowana do wymuszania kontroli dostępu na poziomie dokumentów. Klasy języka Java można też użyć do aktualizowania indeksu przez dodawanie, modyfikowanie lub usuwanie metadanych i treści dokumentów. Poprzez utworzenie wtyczki można również rozszerzyć możliwości przeszukiwacza o przeszukiwanie plików archiwum.

Wtyczka zawiera klasę Java wywoływaną dla każdego dokumentu przeszukiwanego przez przeszukiwacz. Do klasy języka Java przekazywany jest identyfikator dokumentu pobrany z indeksu wyszukiwania korporacyjnego, znaczniki bezpieczeństwa, metadane oraz treść dokumentu. Klasa może zwrócić nowe lub zmodyfikowane znaczniki zabezpieczeń, metadane i treść, ale może też je usunąć.

Po jednokrotnym przeszukaniu wszystkich dokumentów w przestrzeni przeszukiwania wtyczka jest wywoływana tylko dla nowych lub zmodyfikowanych dokumentów. Aby zmienić znaczniki bezpieczeństwa, metadane lub treść dokumentów znajdujących się w indeksie wyszukiwania korporacyjnego, które nie zostały zaktualizowane w oryginalnym

źródle danych, należy uruchomić przeszukiwanie wszystkich dokumentów znajdujących się w przestrzeni przeszukiwania, a następnie ponownie zbudować indeks główny.

## **Wtyczki wymuszające stosowanie zabezpieczeń**

Zabezpieczenia na poziomie dokumentu są wymuszane poprzez powiązanie jednego lub więcej tokenów zabezpieczeń (łańcuchów rozdzielanych przecinkami) z każdym dokumentem przeszukiwanym przez przeszukiwacz. Jako tokeny zabezpieczeń są często używane identyfikatory grupy.

Domyślnie każdy dokument ma przypisany publiczny znacznik, który powoduje udostępnienie dokumentu wszystkim użytkownikom. Publiczny token można zastąpić wartością podaną przez administratora lub wartością wyodrębnioną z pola przeszukiwanego dokumentu.

Wtyczka umożliwi zastosowanie własnych reguł biznesowych w celu określenia wartości tokenów zabezpieczeń dla przeszukiwanych dokumentów. Tokeny zabezpieczeń powiązane z każdym dokumentem są zapisywane w indeksie. Są one używane do filtrowania dokumentów zgodnych z tokenami zabezpieczeń i zapewniają, że w wynikach wyszukiwania są zwracane tylko dokumenty, do których przeglądania użytkownik ma uprawnienia.

## **Wtyczki przeznaczone do dodawania, modyfikowania lub usuwania metadanych**

Metadane dokumentu, takie jak data ostatniej modyfikacji dokumentu, są tworzone dla wszystkich przeszukiwanych dokumentów. Wtyczka przeszukiwacza umożliwi zastosowanie własnych reguł biznesowych w celu określenia wartości metadanych, które mają zostać poindeksowane dla każdego dokumentu.

Metadane są tworzone jako para nazwa-wartość. Użytkownicy mogą przeszukiwać metadane za pomocą zapytań pełnotekstowych lub zapytań określających nazwę pola metadanych.

## **Wtyczki przeznaczone do dodawania, modyfikowania lub usuwania treści dokumentów**

Na treść dokumentów składają się te części, które zawierają treść poddawaną wyszukiwaniu oraz treść, która w wynikach wyszukiwania może stać się częścią dynamicznego podsumowania dokumentu. Wtyczka przeszukiwacza pozwala na zastosowanie własnych reguł biznesowych określających treść, która ma być poddana indeksowaniu w poszczególnych dokumentach.

## **Wtyczki przeszukiwacza WWW**

Za pomocą aplikacyjnych interfejsów programistycznych przeszukiwacza sieci WWW można sterować sposobem przeszukiwania dokumentów i sposobem ich przygotowywania do analizy. Na przykład można dodać pola do nagłówka żądania HTTP, który zostanie użyty, gdy przeszukiwacz zażąda dokumentu. Po przeszukaniu dokumentu, a przed jego przeanalizowaniem i podzieleniu na leksemy, można zmienić jego treść, tokeny zabezpieczeń i metadane. Można również wyłączyć wysyłanie dokumentu do analizatora składni.

Omówienie uwierzytelniania opartego na formularzu oraz przykładowy program, który można dostosować na potrzeby swojej niestandardowej wtyczki przeszukiwacza WWW, można znaleźć pod adresem <http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani>.

## Wtyczki dla plików archiwum


Poprzez utworzenie wtyczki można rozszerzyć możliwości przeszukiwacza i włączyć obsługę przeszukiwania plików archiwum w formatach innych niż ZIP i TAR. Można na przykład utworzyć wtyczkę do obsługi przeszukiwania dokumentów w formacie LZH.

### Tryb bez ochrony

W przypadku konfigurowania przeszukiwacza innego niż przeszukiwacz sieci WWW można wybrać opcję uruchamiania wtyczki w trybie bez ochrony. W tym trybie proces wtyczki działa wewnątrz procesu przeszukiwacza, dzięki czemu wtyczka może mieć lepszą wydajność.

**Ważne:** Po napotkaniu przez wtyczkę problemu, po którym nie można wykonać odtwarzania, gdy działa ona w tym trybie, proces przeszukiwacza może zostać zakończony.

#### Pojęcia pokrewne

 Wtyczki przeszukiwaczy

---

## Obsługa przeszukiwania plików archiwum

Przeszukiwacze wyszukiwania korporacyjnego mogą wyodrębniać pliki z pliku archiwum (na przykład ZIP lub TAR) w celu indeksowania i wyszukiwania pojedynczych plików w archiwum.

### Obsługiwane formaty plików archiwów

Obsługiwane są następujące formaty plików archiwów:

*Tabela 5. Formaty plików archiwów obsługiwane przez przeszukiwacze wyszukiwania korporacyjnego*

| Rozszerzenie pliku | Typ MIME        | Typ danych | Uwagi                                                                                                                                                                                                                                                                                              |
|--------------------|-----------------|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| .zip, .ZIP         | application/zip | zip        | <ul style="list-style-type: none"><li>• W zależności od możliwości pakietu <code>java.util.zip</code></li><li>• Obsługuje kompresję metodą deflacji (metoda 8):<ul style="list-style-type: none"><li>– Brak obsługi plików zaszyfrowanych</li><li>– Brak obsługi formatu zip64</li></ul></li></ul> |
| .tar               | application/tar | tar        | Obsługiwane formaty archiwów tar: <ul style="list-style-type: none"><li>• GNU tar 1.13</li><li>• POSIX 1003.1-1998 (ustar)</li><li>• POSIX 1003.1-2001 (pax)</li></ul>                                                                                                                             |
| .tar, .gz, .tgz    | nie dotyczy     | tgz        | W zależności od możliwości pakietu <code>java.util.zip</code>                                                                                                                                                                                                                                      |

### Ograniczenia i wytyczne

Automatyczne wykrywanie strony kodowej nie jest dostępne dla plików wyodrębnianych z pliku archiwum. Podczas wyodrębniania plików przeszukiwacz używa ustawień strony kodowej skonfigurowanych do użycia z plikami tekstowymi i nieznanymi typami dokumentów. Używając konsoli administracyjnej wyszukiwania korporacyjnego do skonfigurowania ustawień języka i strony kodowej dla przeszukiwacza, można określić stronę kodową, której przeszukiwacz powinien używać w odniesieniu do dokumentów tekstowych i tych, których strony kodowej nie można wykryć automatycznie.

Aby określić, kiedy pliki w pliku archiwum należy ponownie przeszukać, przeszukiwacz odczytuje datę modyfikacji w danych nagłówka archiwum dla każdego pliku. Podczas monitorowania przeszukiwacza statystyki dotyczące przeszukanych dokumentów, włączając w to statystyki dotyczące wstawionych, zaktualizowanych i usuniętych dokumentów, zawierają informacje o plikach wyodrębnionych z plików archiwum.

Aby umożliwić przeszukiwaczom przeszukiwanie plików archiwum w innych formatach archiwum, na przykład LZH, należy utworzyć wtyczkę przeszukiwacza i skonfigurować przeszukiwacz do jej używania.

## Migracja

Aby przeszukiwacze istniejące przed zainstalowaniem wersji 8.4 programu OmniFind Enterprise Edition mogły przeszukiwać pliki archiwów, należy dokonać edycji przestrzeni przeszukiwania, upewniając się, że ustawienia typów MIME dotyczące wykluczonych rozszerzeń plików nie zawierają odwołań do plików .zip, .tar, .tgz ani .gz.

**Ważne:** W przypadku zmiany tych ustawień dla przeszukiwacza Content Edition lub DB2 Content Manager należy ponownie przeszukać dokumenty, aby zmiany mogły zostać zastosowane.

---

## Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego

Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

Identyfikatory URI lub wzorce URI można określić podczas konfigurowaniu kategorii, zasięgów oraz szybkich odsyłaczy do kolekcji. Można również określić identyfikator URI w razie konieczności usunięcia dokumentów z indeksu lub wyświetlenia szczegółowych informacji o statusie konkretnego identyfikatora URI.

Przeszukaj kolekcję w celu określenia identyfikatorów URI lub ich wzorców dla dokumentu. Kliknięcie identyfikatorów URI w wynikach wyszukiwania powoduje pobranie interesujących użytkownika dokumentów. Identyfikator URI można skopiować z wyników wyszukiwania i korzystać z niego na konsoli administracyjnej wyszukiwania korporacyjnego. Na przykład można określić wzorzec identyfikatora URI do automatycznego powiązania dokumentów, które odpowiadają temu wzorcowi identyfikatora URI, z kategorią wyszukiwania korporacyjnego.

### Pliki archiwów

Format identyfikatora URI w przypadku dokumentów, które są przeszukiwane po ich wyodrębnieniu z pliku archiwum (na przykład .zip lub .tar), wygląda następująco:  
*pierwotny\_URI*(?|&)ArchiveEntry=*nazwa\_pozycji*(&ArchiveEntry=*nazwa\_pozycji*)

#### Parametry

*pierwotny\_URI*

Położenie pliku archiwum w źródle danych.

*nazwa\_pozycji*

Nazwa pozycji w pliku archiwum zakodowana jako adres URL.

#### Przykłady

file:///d:/Archive1.zip

file:///d:/Archive1.zip?ArchiveEntry=Folder1/PowerPoint.ppt

file:///d:/Archive1.zip?ArchiveEntry=Folder2/Text.txt

## Przeszukiwacze Content Edition

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Content Edition działający w trybie dostępu do serwera to:

```
vbr://nazwa_serwera/systemowy_id_repozytorium/trwały_id_repozytorium
/id_elementu/id_wersji
/typ_elementu/?[Page=nr_strony&] właściwości_JNDI
```

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Content Edition działający w trybie dostępu bezpośredniego to:

```
vbr://systemowy_id_repozytorium/trwały_id_repozytorium
/id_elementu/id_wersji
/typ_elementu/[?Page=nr_strony]
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_serwera*

Nazwa serwera WebSphere Information Integrator Content Edition.

*systemowy\_id\_repozytorium*

Systemowy identyfikator repozytorium.

*trwały\_id\_repozytorium*

Trwały identyfikator repozytorium.

*id\_elementu*

Identyfikator elementu.

*id\_wersji*

Identyfikator wersji. Jeśli identyfikator wersji jest pusty, ta wartość wskazuje najnowszą wersję dokumentu.

*typ\_elementu*

Typ elementu (CONTENT lub FOLDER).

*nr\_strony*

Numer strony.

*właściwości\_JNDI*

Właściwości JNDI klienta aplikacji J2EE. Istnieją dwa typy właściwości:

**java.naming.factory.initial**

Nazwa klasy serwera aplikacji użytej do utworzenia uchwytu EJB.

**java.naming.provider.url**

Adres URL do usługi nazw serwera aplikacji użytej do żądania uchwytu EJB.

### Przykłady

Documentum:

```
vbr://vbrsrv.ibm.com/Documentum/c06b/094e827780000302//CONTENT/?
java.naming.provider.url=iiop%3A%2F%2Fmyvbr.ibm.com%3A2809&
java.naming.factory.initial=com.ibm.websphere.naming.WsnInitContextFactory
```

FileNet PanagonCS:

```
vbr://vbrsrv.ibm.com/PanagonCS/4a4c/003671066//CONTENT/?Page=1&
java.naming.provider.url=iiop%3A%2F%2Fmyvbr.ibm.com%3A2809&
java.naming.factory.initial=com.ibm.websphere.naming.WsnInitContextFactory
```

## Przeszukiwacze DB2

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz DB2 to:



```
db2://nazwa_bazy_danych/nazwa_tabeli
/nazwa_kolumny1_unikalnego_id/wartość1_unikalnego_id
[/nazwa_kolumny2_unikalnego_id/wartość2_unikalnego_id/...
/nazwa_kolumnyN_unikalnego_id/wartośćN_unikalnego_id]
```

#### Parametry:

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_bazy\_danych*

Wewnętrzna nazwa bazy danych lub alias bazy danych.

*nazwa\_tabeli*

Nazwa tabeli docelowej, w tym nazwa schematu.

*nazwa\_kolumny1\_unikalnego\_id*

Nazwa pierwszej kolumny unikalnego identyfikatora w tabeli.

*wartość1\_unikalnego\_id*

Wartość pierwszej kolumny unikalnego identyfikatora.

*nazwa\_kolumnyN\_unikalnego\_id*

Nazwa *n*-tej kolumny unikalnego identyfikatora w tabeli.

*wartośćN\_unikalnego\_id*

Wartość *n*-tej kolumny unikalnego identyfikatora.

#### Przykłady

Lokalna baza danych wpisana do katalogu:

```
db2://LOCALDB/SCHEMAT1.TABELA1/MODEL/ThinkPadA20
```

Zdalna baza danych nie wpisana do katalogu:

```
db2://mojserwer.mojafirma.com:50001/REMOTEDB/SCHEMAT2.TABELA2/NAZWA/DAWID
```

## Przeszukiwacze DB2 Content Manager

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz DB2 Content Manager to:

```
cm://nazwa_serwera/nazwa_typu_elementu/PID
```

#### Parametry

Kodowanie URL jest stosowane wobec parametru *PID*.

*nazwa\_serwera*

Nazwa serwera bibliotecznego IBM DB2 Content Manager.

*nazwa\_typu\_elementu*

Nazwa typu docelowego elementu.

*PID*

Trwały identyfikator programu DB2 Content Manager.

#### Przykład

```
cm://cmsrvctg/ITEMTYPE1/92+3+ICM8+icmn\lsdb12+ITEMTYPE159+26+A1001001A
03F27B94411D1831718+A03F27B+94411D183171+14+1018
```

## Przeszukiwacze Domino Document Manager

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Domino Document Manager to:

```
dominodoc://nazwa_serwera:nr_portu/id_repliki_bazy_danych
/ścieżka_i_nazwa_bazy_danych
/uniwersalny_id_widoku/uniwersalny_id_dokumentu
/?AttNo=nr_zalącznika&AttName=nazwa_pliku_zalącznika
```

## Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_serwera*

Nazwa serwera Domino Document Manager.

*nr\_portu*

Opcjonalnie: Numer portu serwera Domino Document Manager.

*id\_repliki\_bazy\_danych*

Identyfikator repliki bazy danych.

*ścieżka\_i\_nazwa\_bazy\_danych*

Ścieżka i nazwa pliku bazy danych NSF dokumentów na docelowym serwerze Domino Document Manager.

*uniwersalny\_id\_widoku*

Uniwersalny identyfikator widoku używany przy przeszukiwaniu dokumentów Domino Document Manager.

*uniwersalny\_id\_dokumentu*

Uniwersalny identyfikator dokumentu zdefiniowany w przeszukiwanym dokumencie.

*nr\_zalacznika*

Opcjonalnie: Kolejny numer, zaczynając od zera, dla każdego załącznika.

*nazwa\_pliku\_zalacznika*

Opcjonalnie: Pierwotna nazwa pliku załącznika.

## Przykłady

Dokument Domino Document Manager:

```
dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
```

Załącznik do dokumentu:

```
dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226?AttNo=0&AttName=PlikZalacznika.doc
```

## Przeszukiwacz Exchange Server

Format identyfikatorów dla dokumentów przeszukiwanych przez przeszukiwacz Exchange Server to:

```
exchange://ścieżka_OWA[?useSSL=true]
```

### Parametry

*ścieżka\_OWA*

Ścieżka programu Outlook Web Access (OWA) bez protokołu.

**useSSL=true**

Dodawany, gdy protokół pierwotnej ścieżki OWA to HTTPS.

### Przykłady

Treść dokumentu:

```
exchange://exchangesvr.ibm.com/public/FolderGłówny1/Folder1/Dokument.EML
```

Załącznik do dokumentu:

```
exchange://exchangesvr.ibm.com/public/FolderGłówny1/Folder1/Dokument.EML/PlikZalacznika.doc
```

Włączone dla SSL:

exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/Dokument.EML  
?useSSL=true

## Przeszukiwacze JDBC

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz JDBC to:

```
jdbc://URL_BD/nazwa_tabeli
/nazwa_kolumny1_unikalnego_id/wartość1_unikalnego_id
/[nazwa_kolumny2_unikalnego_id/wartość2_unikalnego_id
/.../nazwa_kolumnyN_unikalnego_id/wartośćN_unikalnego_id]
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*URL\_BD*

Adres URL bazy danych.

*nazwa\_tabeli*

Nazwa tabeli docelowej, w tym nazwa schematu.

*nazwa\_kolumny1\_unikalnego\_id*

Nazwa pierwszej kolumny unikalnego identyfikatora w tabeli.

*wartość1\_unikalnego\_id*

Wartość pierwszej kolumny unikalnego identyfikatora.

*nazwa\_kolumnyN\_unikalnego\_id*

Nazwa *n*-tej kolumny unikalnego identyfikatora w tabeli.

*wartośćN\_unikalnego\_id*

Wartość *n*-tej kolumny unikalnego identyfikatora.

### Przykłady:

Baza danych DB2:

```
jdbc:db2://host01.svl.ibm.com:50000/SAMPLE/DB2INST1.ORG/DEPTNUMB/51
```

Baza danych Oracle:

```
jdbc:oracle:thin:@/host01.svl.ibm.com:1521:ora/SCOTT.EMP/EMPNO/7934
```

Baza danych MS SQL Server 2000:

```
jdbc:microsoft:sqlserver://host01.svl.ibm.com:1433;
DatabaseName=Northwind/dbo.Region/RegionID/100
```

Baza danych MS SQL Server 2005:

```
jdbc:sqlserver://host01.svl.ibm.com:1433;
DatabaseName=Northwind/dbo.Region/RegionID/100
```

## Przeszukiwacze Notes

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Notes to:

```
domino://nazwa_serwera[:nr_portu]/id_repliki_bazy_danych
/ścieżka_i_nazwa_bazy_danych
/[uniwersalny_id_widoku]/uniwersalny_id_dokumentu
[?AttNo=nr_złącznika&AttName=nazwa_pliku_złącznika]
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_serwera*

Nazwa serwera Lotus Notes.

*nr\_portu*

Numer portu serwera Lotus Notes. Numer portu jest opcjonalny.

*id\_repliki\_bazy\_danych*

Identyfikator repliki bazy danych.

*ścieżka\_i\_nazwa\_bazy\_danych*

Ścieżka i nazwa pliku bazy danych NSF na serwerze docelowym Lotus Notes.

*uniwersalny\_id\_widoku*

Uniwersalny identyfikator widoku zdefiniowany w docelowej bazie danych. Ten identyfikator jest określony tylko wtedy, gdy dokument został wybrany z widoku lub folderu. Jeśli widok lub folder do przeszukiwania nie zostanie wskazany (na przykład, jeśli określono, że przeszukiwane mają być wszystkie dokumenty w bazie danych), Uniwersalny identyfikator widoku nie jest określony.

*uniwersalny\_id\_dokumentu*

Uniwersalny identyfikator dokumentu zdefiniowany w dokumencie, który jest przeszukiwany przez przeszukiwacz.

*nr\_zalacznika*

Kolejny numer, zaczynając od zera, dla każdego załącznika. Numer załącznika jest opcjonalny.

*nazwa\_pliku\_zalacznika*

Pierwotna nazwa pliku załącznika. Nazwa pliku załącznika jest opcjonalna.

### **Przykłady**

Dokument wybrany do przeszukania według widoku lub folderu:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf/
8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
```

Dokument, który nie został wybrany do przeszukania według widoku lub folderu:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf//
0205F44FA3F45A9049256DB20042D226
```

Załącznik do dokumentu:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf//
0205F44FA3F45A9049256DB20042D226?AttNo=0&AttName=PlikZalacznika.doc
```

## **Przeszukiwacze QuickPlace**

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz QuickPlace to:

```
quickplace://nazwa_serwera:nr_portu/id_repliki_bazy_danych
/ścieżka_i_nazwa_bazy_danych
/uniwersalny_id_widoku/uniwersalny_id_dokumentu
/?AttNo=nr_zalacznika&AttName=nazwa_pliku_zalacznika
```

### **Parametry**

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_serwera*

Nazwa serwera Lotus QuickPlace.

*nr\_portu*

Opcjonalnie: Numer portu serwera QuickPlace.

*id\_repliki\_bazy\_danych*

Identyfikator repliki bazy danych.

### *ścieżka\_i\_nazwa\_bazy\_danych*

Ścieżka i nazwa pliku bazy danych NSF dokumentów na docelowym serwerze QuickPlace.

### *uniwersalny\_id\_widoku*

Uniwersalny identyfikator widoku używany przy przeszukiwaniu dokumentów QuickPlace.

### *uniwersalny\_id\_dokumentu*

Uniwersalny identyfikator dokumentu zdefiniowany w przeszukiwanym dokumencie.

### *nr\_zalacznika*

Opcjonalnie: Kolejny numer, zaczynając od zera, dla każdego załącznika.

### *nazwa\_pliku\_zalacznika*

Opcjonalnie: Pierwotna nazwa pliku załącznika.

## **Przykłady**

Dokument:

```
quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace%5CPageLibrary4925704300021490.nsf/A7986FD2A9CD47090525670800167225/2B02B1DE3A82B2CE49257043001C2498
```

Załącznik do strony:

```
quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace%5CPageLibrary4925704300021490.nsf/A7986FD2A9CD47090525670800167225/2B02B1DE3A82B2CE49257043001C2498?AttNo=0&AttName==QPCons3.ppt
```

## **Przeszukiwacze listy adresów początkowych**

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz listy adresów początkowych to:

```
seedlist://URL_strony?pageID=id_strony[&useSSL;=true]
```

### **Parametry**

Kodowanie URL ma zastosowanie do wszystkich pól.

#### *URL\_strony*

Adres URL dokumentu (unikalny dla każdego dokumentu).

#### *id\_strony*

Identyfikator obiektu dla dokumentu.

#### **useSSL**

W przypadku protokołu HTTPS identyfikator URI jest uzupełniany członem `&useSSL;=true`. W innych przypadkach opcja `useSSL` jest pomijana.

### **Przykład**

Protokół HTTPS:

```
seedlist://quickrserver.ibm.com:10035/lotus/my poc?uri=dm:bec6090046f1cd52bc5cfc b06e9f4550&verb;=view&pageID;=N1FSZUR1MkJQNjZSMDZQMUMwM1FPNjZCQzY2SUw2SUhPNk1RQ0M2Uk80Nk9PNjVCRUM2UUs2TDFDMA==&useSSL;=true
```

## **Przeszukiwacze systemu plików UNIX**

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz systemu plików UNIX to:

`file:///nazwa_katalogu/nazwa_pliku`

#### **Parametry**

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_katalogu*

Bezwzględna nazwa ścieżki do katalogu.

*nazwa\_pliku*

Nazwa pliku.

#### **Przykład**

`file:///home/user/test.doc`

## **Przeszukiwacze Web Content Management**

Format identyfikatorów URI dokumentów programu WebSphere Content Management przeszukiwanych przez przeszukiwacz Web Content Management jest następujący:

`wcm://URL_strony?pageID=id_strony[&useSSL=true]`

#### **Parametry**

Kodowanie URL ma zastosowanie do wszystkich pól.

*URL\_strony*

Adres URL dokumentu (unikalny dla każdego dokumentu).

*id\_strony*

Identyfikator strony.

#### **useSSL**

W przypadku protokołu HTTPS identyfikator URI jest uzupełniany członem `useSSL=true`. W innych przypadkach opcja `useSSL` jest pomijana.

#### **Przykłady**

Protokół HTTP:

`wcm://wp6server.ibm.com:9081/wps/wcm/myconnect/Web+Content/Site01/SiteArea01/ContentTest01?pageID=6QReDeJ9DI3R0663E03Q06L1E2MR47MHOC3Q862RD6J0863B0GJS86J9E0`

Protokół HTTPS:

`wcm://wp6server.ibm.com:9444/wps/wcm/myconnect/Web+Content/Site01/SiteArea01/ContentTest01?pageID=6QReDeJ9DI3R0663E03Q06L1E2MR47MHOC3Q862RD6J0863B0GJS86J9E0&useSSL=true`

## **Przeszukiwacze WebSphere Portal: WebSphere Portal wersja 5**

Format identyfikatorów URI dokumentów programu WebSphere Portal wersja 5 przeszukiwanych przez przeszukiwacz WebSphere Portal jest następujący:

`wps://URL_strony?portletDefID=ID_def_portletu&portletID=id_portletu&pageID=id_strony[&useSSL=true]`

#### **Parametry**

Kodowanie URL ma zastosowanie do wszystkich pól.

*URL\_strony*

Adres URL dokumentu (unikalny dla każdego dokumentu).

*ID\_def\_portletu*

Identyfikator definicji portletu.

*id\_portletu*

Identyfikator portletu.



*id\_strony*

Identyfikator strony.

**useSSL**

W przypadku protokołu HTTPS identyfikator URI jest uzupełniany członem `useSSL=true`. W innych przypadkach opcja `useSSL` jest pomijana.

#### Przykłady

Treść dokumentu:

```
wps://wpsserver.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+LKnPy1vM0Y_QjzKCN4g3cQbJgQio-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8qHRj/delta/base64xml/L01DU1kvd0NrQUpORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!?portletDefID=3_0_3S&pageID=6_0_6J
```

#### Przykłady

Włączone dla SSL:

```
wps://wpsserver.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+LKnPy1vM0Y_QjzKCN4g3cQbJgQio-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8qHRj/delta/base64xml/L01DU1kvd0NrQUpORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!?portletDefID=7_0_A4&pageID=6_0_6J&useSSL=true
```

## Przeszukiwacze WebSphere Portal: WebSphere Portal wersja 6

Format identyfikatorów URI dokumentów programu WebSphere Portal wersja 6 przeszukiwanych przez przeszukiwacz WebSphere Portal jest następujący:

```
wp6://URL_strony?portletURL=URL_portletu?portletDefID=ID_def_portletu&pageID=id_strony[&useSSL=true]
```

#### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*URL\_strony*

Adres URL dokumentu (unikalny dla każdego dokumentu).

*URL\_portletu*

Unikalny adres URL dokumentu.

*ID\_def\_portletu*

Identyfikator definicji portletu.

*id\_strony*

Identyfikator strony.

**useSSL**

W przypadku protokołu HTTPS identyfikator URI jest uzupełniany członem `useSSL=true`. W innych przypadkach opcja `useSSL` jest pomijana.

#### Przykłady

Protokół HTTP:

```
wp6://wp6server.ibm.com:9081/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGhgYWjjobuhmaGDka0Bu5uHqFRpkaAAAEisaBQ!!?portletUrl=/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BiUZnBwMqsyaBykYGBmS2tcDoDE40MebyRiU-woYGFoaG7oZuhgZGjgbuZh6uXY4ihAYOMGQ7Z0CDBAwDcXPkM&portletDefID=6_24LS10811G1F102A0G6HEJUR10&pageID=3_24LS108110R6B02A0GGU94LN00
```

Protokół HTTPS:

```
wp6://wp6server.ibm.com:9444/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy
8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAzjIPrz6_
Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGh
gYWjjobuhmaGDka0Bu5uHq5RtqaAAA50L41Q!!
?portletUrl=/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59
gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAzjIPrz6_Tzyc1P1C3IjDHQdF
RUBTu-saA!!/d12/d0/Y2BiUZnBwMqsyabykYGBmS2tcDoDE40MebyRiU-woYGFOaG7oZu
hgZGjgbuZh6uXWZihAYOMGQ5Z31BDAAwAk73P2
&portletDefID=6_24LS10811G1F102A0G6HEJMU10
&pageID=3_24LS108110R6B02A0GGU94T410&useSSL=true
```

## Przeszukiwacze systemu plików Windows

Formaty identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz systemu plików Windows to:

*file:///nazwa\_katalogu/nazwa\_pliku*

*file:///nazwa\_folderu\_sieciowego/nazwa\_katalogu/nazwa\_pliku*

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_katalogu*

Bezwzględna nazwa ścieżki do katalogu.

*nazwa\_pliku*

Nazwa pliku.

*nazwa\_folderu\_sieciowego*

Tylko dla dokumentów znajdujących się na serwerach zdalnych, nazwa folderu współużytkowanego w sieci Windows.

### Przykłady

Lokalny system plików:

*file:///d:/katalog/test.doc*

Sieciowy system plików:

*file:///filesvr.ibm.com/katalog/plik.doc*



---

## Administrowanie analizatorem składni

Aby poprawić wydajność pobierania dokumentów, można określić opcje definiujące sposób analizowania i klasyfikowania dokumentów i metadanych przed dodaniem ich do indeksu wyszukiwania korporacyjnego.

Analizę składniową zawartości dokumentów i optymalizację pobierania informacji można skonfigurować przy użyciu następujących operacji:

### **Konfigurowanie opcji analizy dokumentów chińskich, japońskich i koreańskich**

Na potrzeby analizy składni dokumentów w językach chińskim, japońskim i koreańskim można określić opcje podziału na ciągi n-znakowe. Ponadto można usunąć znaki nowego wiersza z dokumentów chińskich i japońskich.

### **Włączanie rodzimego wyszukiwania XML**

Jeśli do kolekcji należą dokumenty XML, można włączyć przeszukiwanie ich przy użyciu rodzimej składni zapytań XML (np. XPath i fragmenty XML). Rodzime wyszukiwanie XML pozwala formułować zapytania w oparciu o związki między różnymi elementami XML.

### **Konfiguracja kategorii**

Dokumenty, których identyfikator URI jest zgodny z określonym wzorcem, lub dokumenty zawierające określone słowa mogą zostać zgrupowane w kategorii. Gdy użytkownicy przeszukują kolekcję, mogą ograniczyć wyniki wyszukiwania do dokumentów należących do określonych kategorii.

### **Konfigurowanie pól wyszukiwania**

Elementy w dokumentach XML można odwzorowywać na pola wyszukiwania w indeksie. Ponadto do pól wyszukiwania można przypisywać elementy metadanych ze źródeł HTML. Tworząc pola wyszukiwania w indeksie wyszukiwania korporacyjnego, daje się użytkownikom możliwość wprowadzania zapytań dotyczących określonych części dokumentów XML i HTML oraz poprawiania precyzji wyników wyszukiwania.

### **Konfigurowanie opcji przetwarzania tekstu**

Jeśli do systemu wyszukiwania korporacyjnego dodano niestandardowe mechanizmy analizy tekstu, można wybrać jeden z nich do użytku przez daną kolekcję. Po powiązaniu mechanizmu analizy z kolekcją można określić opcje odwzorowywania treści sterujące analizą lingwistyczną i adnotowaniem. Można również określić sposób odwzorowywania wyników analizy na indeks wyszukiwania korporacyjnego lub tabele bazy danych JDBC.

### **Odwzorowywanie pól na klasy kryteriów**

Można określić, że dokumenty mające pola zgodne z terminami użytymi w zapytaniu będą sklasyfikowane w wynikach wyszukiwania wyżej niż inne dokumenty zgodne z terminami użytymi w zapytaniu. Podczas odwzorowywania pól na klasy kryteriów należy określić, do jakich pól treści i metadanych mają być stosowane kryteria. Można również skonfigurować wyniki, które są używane do oceny dokumentów przez klasę kryteriów.

### **Pojęcia pokrewne**

 Obsługa lingwistyczna w wyszukiwaniu semantycznym

 Analiza tekstu w wyszukiwaniu korporacyjnym

 Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu

 Aplikacje wyszukiwania semantycznego

 Terminy zapytań wyszukiwania semantycznego

“Obsługa języków i stron kodowych” na stronie 167

“Wykrywanie formatu dokumentów” na stronie 153

---

## Praca z kategoriami

Użycie kategorii pozwala na grupowanie dokumentów o wspólnych charakterystykach, a także wyszukiwanie i pobieranie tylko tych dokumentów, które spełniają kryteria przynależności do określonej grupy.

Jeśli dokumenty są przypisane do kategorii, a używane aplikacje wyszukiwujące obsługują tę opcję, użytkownicy mogą przeszukiwać tylko podzbiór kolekcji, określając nazwę kategorii. Jeśli przeszukiwana jest cała kolekcja, użytkownicy mogą ograniczyć zbiór wyników wyszukiwania i przeglądać tylko dokumenty należące do określonej kategorii, do której należy jeden z dokumentów wynikowych.

Podczas konfigurowania kategorii określa się reguły, które instruuja wyszukiwacz, aby powiązał dokumenty z kategorią. Można pogrupować dokumenty, które mają wspólny wzorzec URI lub zawierają określoną treść (na przykład dokumenty zawierające lub wykluczające określone słowa i frazy).

Do tworzenia kategorii i zarządzania nimi służy konsola administracyjna wyszukiwania korporacyjnego:

- Typ klasyfikowania wybierany jest podczas tworzenia kolekcji. Można nie używać żadnych kategorii lub korzystać z kategorii opartych na regułach.
- W razie potrzeby można zmienić typ klasyfikowania, konfigurując reguły analizowania dla kolekcji. W przypadku zmiany typu klasyfikowania po przeszukaniu i poindeksowaniu dokumentów jakość wyszukiwania jest niższa do momentu ponownego przeszukania wszystkich dokumentów i ponownego zbudowania głównego indeksu.
- Jeśli zostanie wybrana opcja kategorii opartych na regułach, wówczas do administrowania drzewem kategorii, kategoriami i regułami kategorii używa się konsoli administracyjnej. W przypadku zmiany kategorii lub reguł kategorii po przeszukaniu i poindeksowaniu dokumentów jakość wyszukiwania jest niższa do momentu ponownego przeszukania wszystkich dokumentów i ponownego zbudowania głównego indeksu.

## Kategorie oparte na regułach

Użytkownik konfigurując reguły określa dokumenty przypisywane do kategorii w kolekcji wyszukiwania korporacyjnego.

Reguły kategorii można tworzyć dla kolekcji utworzonych w wyszukiwaniu korporacyjnym i dla kategorii opartych na regułach migrowanych z kolekcji produktu IBM WebSphere Portal. Aby skonfigurować reguły klasyfikowania dokumentów, podczas tworzenia kolekcji lub definiowania opcji analizowania dla kolekcji należy wybrać opcję użycia kategorii opartych na regułach.

W analizatorze składni podczas przypisywania dokumentów do jednej lub wielu kategorii użyte zostaną reguły określone przez użytkownika:

- Jeśli dokument jest zgodny z przynajmniej jedną regułą w kategorii, analizator składni przypisze go do tej kategorii.
- Jeśli dokument jest zgodny z przynajmniej jedną regułą w kilku kategoriach, analizator składni przypisze go do wszystkich tych kategorii.
- Jeśli dokument nie jest zgodny z żadną regułą dla kategorii, analizator składni nie przypisze go do tej kategorii. Podczas przeszukiwania kolekcji użytkownicy mogą wyszukać ten dokument i pobrać go, lecz nie znajdą go, jeśli przeszukiwać będą poszczególne kategorie.

Administrator używanego dla kolekcji drzewa kategorii (lub systematyki) decyduje, gdzie w hierarchii kategorii ma zostać dodana nowa kategoria. W drzewie kategorii można także wybrać kategorię, która ma zostać poddana edycji, a następnie dodać lub usunąć reguły klasyfikowania dokumentów albo zmienić treść poszczególnych reguł.

Konfigurując regułę klasyfikowania dokumentów należy określić, czy w wyszukiwaniu korporacyjnym przynależność dokumentu do określonej kategorii jest ustalana na podstawie identyfikatora URI tego dokumentu, czy na podstawie jego treści.

## Reguły wzorca URI

Reguła oparta na identyfikatorze URI jest stosowana względem identyfikatora URI dokumentu. Użytkownik określa częściowy identyfikator URI (wzorzec) i z regułą zgodne są te dokumenty, których identyfikatory URI zawierają ten wzorzec.

Na przykład jeśli określony jest łańcuch reguły `/ka/`, wówczas pierwszy pokazany poniżej identyfikator URI jest zgodny z tą regułą, a drugi identyfikator URI jest niezgodny:

```
file:///korporacja/ka/formularz.doc
http://firma.com/kadry/formularz.htm
```

Ponieważ wszystkie identyfikatory URI są traktowane jak wzorce, system ignoruje wszystkie gwiazdki wprowadzone jako znaki wieloznaczne na początku lub na końcu wzorca. Na przykład wzorce `*/ka/*` i `/ka/` odpowiadają temu samemu zestawowi identyfikatorów URI.

W regułach wzorców URI nie jest rozróżniana wielkość liter. W przypadku identyfikatorów URI zawierających spacje wzorzec URI musi być zgodny z regułami wyszukiwania korporacyjnego dotyczącymi kodowania identyfikatorów URI. W poniższym przykładzie przedstawiono poprawny i niepoprawny sposób określania identyfikatorów URI dla ścieżki systemu plików Windows:

```
Niepoprawny identyfikator URI: file:///c:/program files/
Poprawny identyfikator URI: file:///c:/program+files/
```

## Reguły treści dokumentu

Reguły treści dokumentu należy podawać w takim samym formacie jak zapytania. Jeśli dokument jest poprawny dla określonego zapytania, wówczas jest on zgodny z daną regułą. Podczas konfigurowania reguły należy określić słowa i frazy, które muszą występować lub nie mogą występować w dokumencie, oraz wybrać język reguł tematykacji słów (określenia słów na podstawie rdzenia).

Na przykład w następującej regule określono, że jeśli w dokumencie występuje słowo `ka` lub `kadry`, wówczas dokument jest zgodny z tą regułą:

```
ka "kadry"
```



W innym przykładzie w następującej regule określono, że jeśli w dokumencie występuje słowo "zdrowie", ale nie fraza "świadczenia pracownicze", wówczas dokument jest zgodny z tą regułą:

```
+title:zdrowie -title:"świadczenia pracownicze"
```

Reguły dotyczące treści podlegają tym samym normom językowym, co zapytania tworzone za pomocą interfejsu SI-API (Search and Index API). Jednak w składni reguł dotyczących treści obsługiwany jest tylko podzbiór operacji dostępnych w składni zapytań tworzonych za pomocą interfejsu SI-API. W regułach dotyczących treści dozwolone są tylko następujące operatory zapytania:

- + Wstawienie znaku plus przed terminem oznacza, że dokument musi zawierać dany termin.
- Wstawienie znaku minus przed terminem oznacza, że dokument nie może zawierać danego terminu.
- " " Ujęcie dwóch lub większej liczby terminów w cudzysłów oznacza, że dokument musi zawierać dokładnie taką frazę.

*nazwa\_pola:*

Termin lub frazę należy poprzedzić nazwą pola, aby wskazać, że termin lub fraza muszą (lub nie mogą) wystąpić w określonym polu dokumentu.

Obsługiwane są wszystkie pola treści i metadanych, które w kolekcji zostały skonfigurowane jako pola dostępne do wyszukiwania. Nie są obsługiwane następujące słowa kluczowe pól SI-API i typy pól:

- site:
- url:
- link:
- docID:
- samegroupas:
- pola parametryczne
- znaczniki bezpieczeństwa
- atrybuty (na przykład \$source, \$language, \$doctype itd.)

## Drzewa kategorii

Drzewo kategorii umożliwia wyświetlenie wszystkich - używanych w kolekcji - kategorii opartych na regułach. Drzewo kategorii służy do tworzenia i usuwania kategorii oraz do edytowania reguł przypisywania dokumentów do kategorii.

Drzewo kategorii, zwane także systematyką, jest uporządkowane hierarchicznie. Początkowym punktem tej hierarchii jest kategoria podstawowa; wszystkie pozostałe kategorie są podkategoriami kategorii podstawowej. Można utworzyć dowolną liczbę zagnieżdżonych kategorii i podkategorii w celu udostępnienia użytkownikom odpowiednich opcji wyboru podczas przeglądania i pobierania dokumentów.

Na przykład jeśli dokument jest zgodny z regułami związanymi z kilkoma kategoriami, jest przypisywany do wszystkich tych kategorii. Gdy podczas przeglądania wyników wyszukiwania użytkownicy przeszukują kategorię lub przeglądają dokumenty należące do określonej kategorii, fakt przynależności określonego dokumentu do wielu kategorii zapewnia większe prawdopodobieństwo, że dokument ten zostanie znaleziony.

Administrator drzewa kategorii może określić, czy dokument należy do jednej czy do wielu kategorii, zagnieżdżając nowe kategorie w istniejących kategoriach. Podczas tworzenia

kategorię należy określić, czy ma ona zostać utworzona jako kategoria główna czy jako podkategoria innej kategorii. Drzewo kategorii służy także do usuwania kategorii z kolekcji oraz do modyfikowania reguł przypisywania dokumentów do kategorii. Podczas edycji kategorii można zmienić jej nazwę, dodać lub usunąć reguły klasyfikowania albo zmodyfikować treść poszczególnych reguł.

Podczas administrowania drzewem kategorii warto korzystać z następujących opisów działania funkcji wyszukiwania i przeglądania:

- Jeśli przeszukiwana jest kategoria wysokiego poziomu, dokument spełniający kryteria wyszukiwania jest poszukiwany w całej tej kategorii i we wszystkich jej podkategoriach. Jeśli kategoria nie zawiera żadnych podkategorii, wówczas przeszukiwana jest tylko ta określona kategoria.
- Jeśli podczas przeglądania wyników wyszukiwania użytkownik wybierze opcję przeglądania dokumentów należących do konkretnej kategorii, wówczas wyświetlone zostaną tylko dokumenty należące do tej kategorii. W wynikach wyszukiwania wyświetlane są także nazwy wszystkich podkategorii, aby użytkownik mógł jednocześnie nawigować wśród kategorii i wyświetlać podzbiory dokumentów.

## Wybór typu klasyfikacji

Po wybraniu typu klasyfikacji należy określić metodę, przy użyciu której dokumenty będą kojarzone z kategoriami w kolekcji.

### Zagadnienia wstępne

Aby zmienić typ klasyfikacji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora zmienianej kolekcji.

### Informacje o tym zadaniu

Typ klasyfikacji jest określany w momencie tworzenia kolekcji. W razie potrzeby istnieje możliwość zmiany sposobu klasyfikowania dokumentów w kolekcji. Można użyć kategorii opartych na regułach skonfigurowanych specjalnie na potrzeby tej kolekcji albo całkowicie zrezygnować z używania kategorii.

**Ważne:** Jeśli typ klasyfikacji zostanie zmieniony po przeszukaniu danych i utworzeniu indeksu dla kolekcji, indeks będzie niespójny. Aby zapewnić jak najwyższą dokładność wyników wyszukiwania, należy ponownie przeszukać wszystkie dokumenty znajdujące się w kolekcji i zbudować indeks główny.

### Procedura

Aby wybrać typ klasyfikacji, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Wybór typu klasyfikowania**.
2. Na stronie Wybór typu klasyfikowania wybierz jedną z następujących opcji:

**Brak** Wybierz tę opcję, jeśli dokumenty w tej kolekcji nie mają być klasyfikowane.

#### **Oparty na regułach**

Wybierz tę opcję, jeśli dokumenty mają zostać podzielone na kategorie zgodnie z regułami przygotowanymi specjalnie dla tej kolekcji.

## Konfiguracja kategorii

Dla kolekcji można utworzyć dowolną liczbę kategorii, a każda kategoria może zawierać dowolną liczbę reguł. Reguły określają, które dokumenty będą automatycznie powiązane z daną kategorią.

### Zagadnienia wstępne

Aby skonfigurować kategorie, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą kategorie.

Jako typ klasyfikowania należy wybrać kategorie oparte na regułach.

Na przykład w celu określenia reguł dotyczących powiązania dokumentów z kategoriami należy kliknąć opcję **Pomoc** w trakcie tworzenia lub edycji kategorii.

### Informacje o tym zadaniu

Jeśli używane aplikacje wyszukujące zapewniają obsługę kategorii, użytkownicy mogą wykonywać wyszukiwanie w podzbiorniku kolekcji, podając nazwę kategorii. Użytkownicy mogą także wybrać kategorię w wynikach wyszukiwania i przeglądać wyłącznie dokumenty należące do wybranej kategorii.

**Ważne:** Jeśli po przeszukaniu dokumentów i utworzeniu indeksu dla kolekcji zmienione zostaną kategorie lub reguły kategorii, indeks stanie się niespójny. Aby zapewnić jak najwyższą dokładność wyników wyszukiwania, należy ponownie przeszukać wszystkie dokumenty znajdujące się w kolekcji i zbudować indeks główny.

### Procedura

Aby skonfigurować kategorię, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguracja drzewa kategorii**.
2. Na stronie Drzewo kategorii wybierz w drzewie lokalizację, w której chcesz dodać kategorię, i kliknij opcję **Tworzenie kategorii**.

Jeżeli wybrany został poziom główny, nowa kategoria zostanie utworzona na tym poziomie. Jeśli wybrana zostanie nazwa kategorii, nowa kategoria zostanie utworzona poniżej wybranej kategorii w drzewie kategorii.

Otwarty zostanie kreator umożliwiający określanie reguł dotyczących powiązania dokumentów z nową kategorią:

- a. Na stronie Tworzenie kategorii wprowadź opisową nazwę kategorii i kliknij przycisk **Dalej**.
- b. Na stronie Tworzenie reguł kategorii kliknij przycisk **Dodaj regułę**.
- c. Na stronie Tworzenie reguły kategorii wprowadź unikalną nazwę reguły w polu **Nazwa reguły**. Nazwa ta musi być niepowtarzalna w obrębie wszystkich kategorii w danej kolekcji.
- d. Określ regułę, która ma być stosowana do powiązania dokumentów z daną kategorią, a następnie kliknij przycisk **OK**.
  - Aby użyć identyfikatora URI dokumentu do określenia, czy dokument ten należy do danej kategorii, kliknij opcję **Wzorzec identyfikatora URI**, a następnie podaj wzorzec identyfikatora URI.

Jeśli identyfikator URI będzie zawierał podany tekst, dokument zostanie powiązany z kategorią.

Na przykład: `file:///c:/program+files/finance`

- Aby sprawdzić przez odpytanie treści dokumentu, czy należy on do danej kategorii, kliknij opcję **Treść dokumentu**, wybierz język dokumentu i podaj słowa oraz frazy, które muszą lub nie mogą pojawiać się w treści dokumentu.

Regułę tę można wyrazić w formacie, w jakim zapisywane są zapytania, jednak można w niej używać wyłącznie operatorów włączenia (+), wykluczenia (-), frazy (" ") i nazwy pola (*nazwa\_pola*:). Podział na ciągi n-znakowe nie jest obsługiwany przez reguły treści.

Dokument zawierający podane słowa lub dokument, który ich nie zawiera, jest kojarzony z kategorią.

Na przykład: +finanse -konta +title:"rok fiskalny"

- e. Kliknij przycisk **Zakończ**.

Nowa kategoria jest wyświetlana na stronie Drzewo kategorii razem z innymi kategoriami należącymi do tej kolekcji.

#### Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 120

## Praca z polami wyszukiwania XML

W celu umożliwienia użytkownikom wyszukiwania określonych części dokumentów XML należy odwzorować elementy XML na pola wyszukiwania.

Do odwzorowania elementów XML na pola wyszukiwania należy użyć konsoli administracyjnej wyszukiwania korporacyjnego.

Zazwyczaj indeksowaniem objęte są wszystkie dane tekstowe w dokumencie XML. Przez odwzorowanie elementów XML na pola wyszukiwania można udostępnić informacje strukturalne z dokumentów XML jako podstawę formułowania bardziej precyzyjnych zapytań. Na przykład po odwzorowaniu nazwy elementu XML na nazwę pola dane z elementu XML stają się dostępne dla celów wyszukiwania według nazwy pola oraz są wyświetlane wśród wyników wyszukiwania. Aby zwiększyć dokładność odwzorowania, można określić atrybuty elementu XML (jak również wartości tych atrybutów) jako kryteria, w zależności od których element może się stać polem wyszukiwania.

## Pola wyszukiwania XML

Pola wyszukiwania XML umożliwiają wykonywanie zapytań dotyczących określonych części dokumentów XML.

Dokumenty XML stają się coraz bardziej powszechne, ponieważ zawierają zarówno tekst z częściową strukturą, jak i tekst bez określonej struktury. Struktura XML jest hermetyczna i korzysta z kontekstu zdefiniowanego jawnie przez występujące wokół tekstu elementy XML. Na przykład imię i nazwisko autora może występować w następujący sposób:

```
<author>John Smith</author>
```

W tym kontekście tekst John Smith identyfikuje autora dokumentu XML.

W systemie wyszukiwania korporacyjnego możliwe jest kojarzenie elementów XML z nazwami pól wyszukiwania, określane też odwzorowywaniem. Przy konfigurowaniu opcji analizowania składni dla kolekcji określa się, które elementy XML mają być odwzorowane na które nazwy pól wyszukiwania. Dzięki odwzorowaniu elementów XML na pola wyszukiwania użytkownicy zyskują możliwość wyszukiwania wartości tych elementów przez odwoływanie się w zapytaniach do nazw pól. Zapytania odwołujące się do nazw pól pozwalają uzyskać dokładniejsze wyniki niż zapytania pełnotekstowe przeszukujące całą zawartość dokumentu.

Na przykład, jeśli kolekcja zawiera dokumenty XML, których pola <title> i <author> wybrano jako pola do wyszukiwania w indeksie, użytkownicy będą mogli wprost odwoływać się do tych elementów. Wyszukiwanie wzorca author:Smith pozwala odnaleźć dokumenty mające wartość Smith w elementach odwzorowanych do pola o nazwie author.

Innym przykładem może być element XML o nazwie <summary> (streszczenie), którego zawartość może być użyteczna w wynikach wyszukiwania. Po odwzorowaniu elementu <summary> na pole wyszukiwania i określeniu, że wartość tego elementu ma być uwzględniana w wynikach wyszukiwania, treść tego elementu będzie umieszczana w dokumencie wynikowym.

Odwzorowanie elementu XML na nazwę pola sprawia, że tekst zawarty w tym elemencie staje się dostępny do wyszukiwania przez odwołanie się do wskazanej nazwy pola. Jednak ewentualne atrybuty elementu XML nie są przez to udostępniane do bezpośredniego wyszukiwania. Aby za pomocą zapytania móc wyszukiwać wartości atrybutu, w konsoli administracyjnej wyszukiwania korporacyjnego należy skonfigurować opcje analizowania składni oraz włączyć rodzime wyszukiwanie XML. Po włączeniu wyszukiwania własnego XML struktura dokumentu XML staje się dostępna w indeksie, a w zapytaniu względem dokumentu można określać reguły ograniczające XPath. Na przykład:

```
@xml:xp::'//elementName[@attributeName="attributeValue"]
```

### Pojęcia pokrewne



Termin zapytania wyszukiwania semantycznego

### Zadania pokrewne

“Włączanie obsługi rodzimego wyszukiwania XML” na stronie 151

## Odwzorowanie elementów XML na pola wyszukiwania

Przy odwzorowaniu elementu XML na pole wyszukiwania użytkownik określa, którzy użytkownicy elementów XML mogą wyszukiwać, określając nazwę pola w zapytaniu.

### Zagadnienia wstępne

Aby odwzorować elementy XML na pola wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienie administratora kolekcji, do której należą dane dokumenty XML.

### Ograniczenia

W ramach kolekcji może istnieć wiele odwzorowań pól XML, ale tylko jedno odwzorowanie elementu głównego XML. Element główny dokumentu XML realizuje odwzorowania odpowiednio do definicji.

### Informacje o tym zadaniu

Po utworzeniu odwzorowania pól XML, a także po dodaniu, modyfikacji lub usunięciu pól w istniejącym odwzorowaniu pól XML wszelkie zmiany odnoszą skutek po ponownym uruchomieniu analizatora składni. Nowe i zmodyfikowane odwzorowania są uwzględniane w odniesieniu do nowych danych analizowanych po restarcie analizatora; nie dotyczą one danych wprowadzonych do indeksu już wcześniej. Aby zaktualizować dokumenty już znajdujące się w indeksie, należy powtórnie wykonać przeszukanie i indeksowanie tych dokumentów.

To zadanie korzysta z poniższego przykładowego dokumentu XML w celu przedstawienia sposobu odwzorowania rekordów dotyczących pracowników i umożliwienia użytkownikom bezpośredniego zapytania o niektóre elementy.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<personnel>
 <personnelrecord>
 <phone>5555</phone>
 <email>joe@us.ibm.com</email>
 <jobroles>Manager, architect
 <jobrole>Managing Search Development Group</jobrole>
 <jobrole>Architecting Search Technology</jobrole>
 </jobroles>
 <location>New York</location>
 <section id="expertise">
 <text>Linguistics</text>
 </section>
 </personnelrecord>
</personnel>

```

## Procedura

Aby odwzorować użyte w przykładzie elementy XML na pola wyszukiwania, należy wykonać następujące czynności:

- Przejdź do edycji kolekcji, wybierz stronę Analiza składni, a następnie kliknij opcję **Odwzorowanie elementów XML na pola**.
- Na stronie Odwzorowania pól XML kliknij przycisk **Utwórz odwzorowanie XML**. Zostanie otwarta strona Tworzenie odwzorowania pola XML.
- W polu **Nazwa elementu głównego XML** wpisz nazwę elementu głównego: **personnel**.  
Upewnij się, że określona w tym miejscu nazwa dokładnie odpowiada elementowi głównemu w dokumentach XML, które mają być wyszukiwane. Podczas analizowania i indeksowania dokumentów XML wyszukiwanie korporacyjne wybiera odwzorowanie do użycia zgodnie z nazwą elementu głównego.
- W polu **Nazwa odwzorowania XML** wpisz nazwę tego zestawu reguł odwzorowywania pól XML.  
Po utworzeniu zestawu reguł odwzorowania XML nazwa ta będzie wyświetlana na stronie Odwzorowania pól XML i będzie można ją wybrać w celu dodania, usunięcia lub zmiany reguł odwzorowania.
- Odwzoruj element XML **jobrole** na pole wyszukiwania o nazwie **jobrole**:
  - W polu **Nazwa pola** wpisz **jobrole**.  
**Wskazówka:** Informacje na temat znaków ASCII i metaznaków, których nie można stosować w nazwach pól, można wyświetlić za pośrednictwem ikony **Pomoc**.
  - W polu **Nazwa elementu XML** wpisz **jobrole**.
  - Aby umożliwić użytkownikom tworzenie zapytań dla pola **jobrole**, należy zaznaczyć pole wyboru **Wyszukiwanie przez pola**.
  - Aby dopasowanie było stwierdzane tylko w przypadku tych pól **jobrole**, których cała zawartość jest dokładnie zgodna z kryterium wyszukiwania (tzn. w polu nie ma już żadnych innych wyrazów), zaznacz pole wyboru **Pełne dopasowanie**.
  - Aby umożliwić użytkownikom sortowanie wyników wyszukiwania według wartości w polu **jobrole**, zaznacz pole wyboru **Dopuszczalne sortowanie**.  
Pole jest wyświetlane jako pole z dopuszczalnym sortowaniem, tylko jeśli aplikacja wyszukiwująca obsługuje tę funkcję. W przykładowej aplikacji wyszukiwującej wyszukiwania korporacyjnego lista pól, które można wybrać do sortowania wyników wyszukiwania, nie zawiera pól odwzorowanych z elementów XML.
  - Aby umożliwić użytkownikom wyświetlanie wartości pól **jobrole** w wynikach wyszukiwania, zaznacz pole wyboru **Wyniki wyszukiwania**.
- Odwzoruj element XML **jobroles** na to samo pole wyszukiwania:



- a. Kliknij ikonę **Dodaj pole**, aby dodać pusty wiersz do listy reguł odwzorowania pól.
- b. W polu **Nazwa pola** wpisz jobrole.
- c. W polu **Nazwa elementu XML** wpisz jobroles.

**Wskazówka:** Nazwy elementów XML nie muszą odpowiadać nazwom pól wyszukiwania, a wiele elementów XML może być odwzorowanych na to samo pole wyszukiwania.

- d. Aby umożliwić użytkownikom odwoływanie się do pola jobrole w zapytaniach, wybrać całkowite dopasowanie wartości pola do kryterium zapytania, umożliwić sortowanie wyników względem wartości w tym polu oraz wyświetlać pole w wynikach zapytania, zaznacz odpowiednie pola wyboru.
7. Odwzoruj element XML section mający atrybut o nazwie ID i zawierający wartość expertise do pola wyszukiwania o nazwie expertise:
- a. Kliknij ikonę **Dodaj pole**, aby dodać pusty wiersz do listy reguł odwzorowania pól.
  - b. W polu **Nazwa pola** wpisz expertise.
  - c. W polu **Nazwa elementu XML** wpisz section.
  - d. W polu **Nazwa atrybutu XML** wpisz id.
  - e. W polu **Wartość atrybutu XML** wpisz expertise.
  - f. Aby umożliwić użytkownikom tworzenie zapytań wobec pola expertise, zaznacz pole wyboru **Wyszukiwanie przez pola**.
  - g. Aby dopasowanie było stwierdzane tylko w przypadku tych pól expertise, których cała zawartość jest dokładnie zgodna z kryterium wyszukiwania (tzn. w polu nie ma już żadnych innych wyrazów), zaznacz pole wyboru **Pełne dopasowanie**.
  - h. Aby umożliwić użytkownikom sortowanie wyników wyszukiwania według wartości w polu expertise, jeśli aplikacja wyszukująca obsługuje tę funkcję, zaznacz pole wyboru **Dopuszczalne sortowanie**.
  - i. Aby umożliwić użytkownikom wyświetlanie wartości pól expertise w wynikach wyszukiwania, zaznacz pole wyboru **Wyniki wyszukiwania**.

### Przykłady:

Aby znaleźć w organizacji wszystkie osoby pracujące na produktach wyszukiwania, podaj następujące zapytanie:

jobrole:search

Aby znaleźć w organizacji wszystkie osoby znające się na lingwistyce, podaj następujące zapytanie:

expertise:linguistics

---

## Praca z polami wyszukiwania HTML

W celu umożliwienia użytkownikom wyszukiwania określonych sekcji metadanych w dokumentach HTML należy odwzorować elementy metadanych HTML na pola wyszukiwania.

Do odwzorowania elementów metadanych HTML na pola wyszukiwania należy użyć konsoli administracyjnej wyszukiwania korporacyjnego.

Dzięki odwzorowaniu elementów metadanych HTML na pola wyszukiwania użytkownicy mogą wyszukiwać dokumenty HTML za pomocą dokładniejszych zapytań.

## Pola wyszukiwania HTML

Pola wyszukiwania HTML umożliwiają wykonywanie zapytań dotyczących atrybutów dokumentów HTML.

Elementy metadanych w dokumentach HTML przypominają atrybuty dokumentu, ponieważ również udostępniają informacje o dokumencie, o jego formacie oraz o metodzie dostępu do niego w sieci WWW. Na przykład:

```
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
<meta name="copyright" content="(C) Copyright IBM Corporation 2005" />
<meta name="content.owner" content="(C) Copyright IBM Corporation 2005" />
<meta name="security" content="public" />
<meta name="abstract" content="Opis produktu firmy IBM." />
<meta name="format" content="XHTML" />
```

Wyszukiwanie korporacyjne może powiązać (odwzorować) nazwy elementów metadanych HTML z nazwami pól wyszukiwania. Podczas konfigurowaniu opcji analizowania składni dla kolekcji określa się, jakie elementy metadanych HTML mają być odwzorowane na jakie nazwy pól wyszukiwania. Dzięki odwzorowaniu elementów metadanych HTML na pola wyszukiwania użytkownicy mogą wyszukiwać dokumenty zawierające te elementy, podając w zapytaniach nazwy pól wyszukiwania. Zapytania, w których użyto konkretnych pól, umożliwiają uzyskanie bardziej precyzyjnych wyników wyszukiwania niż zapytania pełnotekstowe, które wyszukują cały dokument.

Na przykład, jeśli kolekcja zawiera dokumenty HTML, a elementy metadanych "copyright" i "abstract" mają być poindeksowane jako pola wyszukiwania, użytkownicy mogą używać tych elementów w zapytaniach. W wyniku zapytania `copyright:IBM` zostaną znalezione dokumenty HTML, które zawierają tekst IBM w elemencie metadanych `copyright`.

Podczas odwzorowywania elementów metadanych HTML na pola wyszukiwania użytkownik określa, czy chce odwzorować wszystkie elementy metadanych HTML, tylko te, które należą do zestawu Dublin Core, czy tylko podane przez niego elementy metadanych HTML. Opis elementów w zestawie elementów metadanych Dublin Core można znaleźć w serwisie WWW Dublin Core Metadata Initiative:

<http://dublincore.org/documents/dcmi-terms/#H2>

Jeśli użytkownik zdecyduje się utworzyć odwzorowania dla wszystkich elementów metadanych HTML lub wszystkich elementów metadanych z zestawu Dublin Core, domyślna nazwa pola wyszukiwania będzie zgodna z nazwą elementu metadanych. Można nadpisać domyślną nazwę pola wyszukiwania i określić inne opcje wyszukiwania konkretnych elementów, dodając je do listy elementów, które będą konfigurowane indywidualnie.

## Odwzorowywanie elementów metadanych HTML na pola wyszukiwania

Podczas odwzorowywania elementów metadanych HTML na pola wyszukiwania określa się elementy metadanych HTML, które użytkownik może wyszukiwać, podając w zapytaniu nazwę pola.

### Zagadnienia wstępne

Aby odwzorować elementy metadanych HTML na pola wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą dokumenty HTML.

### Informacje o tym zadaniu

Po utworzeniu odwzorowania pól HTML, a także po dodaniu, modyfikacji lub usunięciu pól w istniejącym odwzorowaniu pól HTML wszelkie zmiany odnoszą skutek po ponownym uruchomieniu analizatora składni. Nowe i zmodyfikowane odwzorowania są uwzględniane w odniesieniu do nowych danych analizowanych po restarcie analizatora; nie dotyczą one danych wprowadzonych do indeksu już wcześniej. Aby zaktualizować dokumenty już znajdujące się w indeksie, należy powtórnie wykonać przeszukanie i indeksowanie tych dokumentów.

## Procedura

Aby odwzorować elementy metadanych HTML na pola wyszukiwania:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Odwzorowanie metadanych HTML na pola**.
2. Na stronie Odwzorowanie pól HTML określ, które elementy metadanych HTML mają zostać odwzorowane na pola wyszukiwania:
  - Przy odwzorowywaniu wszystkich elementów HTML lub wszystkich elementów należących do zestawu Dublin Core można określić, czy użytkownicy mają mieć możliwość wyszukiwania wartości w polach przez odwołania do nazw pól i czy wartości pól mają być wyświetlane w wynikach wyszukiwania. Ponadto można określić, czy podane przez użytkownika kryteria wyszukiwania muszą dokładnie zgadzać się z całą wartością elementu metadanych (czyli pole nie zawiera innych słów poza podanymi przez użytkownika) oraz czy użytkownicy będą mogli sortować wyniki według tego pola.

System automatycznie użyje nazw elementów metadanych jako nazw pól. Jeśli zachodzi potrzeba wybrania innych nazw pól dla wybranych elementów lub zastosowania odmiennych opcji wyszukiwania, można odwzorować nazwę elementu na wybraną nazwę pola, po czym określić opcje dotyczące tylko tego elementu.

- Aby odwzorować określone elementy metadanych na pola wyszukiwania:
  - a. Kliknij ikonę **Dodaj pole**, aby dodać pusty wiersz do listy reguł odwzorowania pól.
  - b. Wpisz nazwę, która ma być kojarzona z odwzorowywanym elementem metadanych HTML. Użytkownicy będą mogli podawać tę nazwę pola podczas wyszukiwania dokumentów HTML w danej kolekcji.

**Wskazówka:** Informacje na temat znaków ASCII i metaznaków, których nie można stosować w nazwach pól, można wyświetlić za pośrednictwem ikony **Pomoc**.

- c. Wpisz nazwę elementu metadanych, który ma zostać odwzorowany.
- d. Aby umożliwić użytkownikom tworzenie zapytań wobec tego pola, zaznacz pole wyboru **Wyszukiwanie przez pola**.

**Wskazówka:** Jeśli to pole wyboru jest wyczyszczone, zawartości pola nie można wyszukiwać ani przez nazwę pola, ani metodą pełnotekstową. Taka możliwość przydaje się, jeśli wybrano opcję powodującą włączenie wszystkich elementów metadanych HTML lub elementów Dublin Core, ale niektóre pola mają zostać wykluczone z wyszukiwania.

- e. Aby umożliwić użytkownikom tworzenie zapytań względem tego pola i określić, że zaliczenie dokumentu do wyników wymagać będzie całkowitego dopasowania wartości pola (czyli że pole nie może zawierać żadnych innych wyrazów), zaznacz pole wyboru **Pełne dopasowanie**.
- f. Aby umożliwić użytkownikom alfabetyczne sortowanie wyników wyszukiwania (sortowanie łańcuchowe) według tego pola, zaznacz pole wyboru **Z możliwością sortowania**.

Pole jest wyświetlane jako pole z dopuszczalnym sortowaniem, tylko jeśli aplikacja wyszukująca obsługuje tę funkcję. W przykładowej aplikacji wyszukującej wyszukiwania korporacyjnego lista pól, które można wybrać do sortowania wyników wyszukiwania, nie zawiera pól odwzorowanych z elementów metadanych HTML.

- g. Jeśli typem danych tego pola jest DECIMAL, DOUBLE, INTEGER, SHORT, TIME lub TIMESTAMP, a użytkownicy mają mieć możliwość określania zapytań parametrycznych przy prowadzeniu wyszukiwania w tym polu lub możliwość liczbowego sortowania wyników na podstawie wartości tego pola, to zaznacz pole wyboru **Wyszukiwanie parametryczne**.
- h. Aby umożliwić użytkownikom wyświetlanie wartości tego pola w wynikach wyszukiwania, zaznacz pole wyboru **Wyniki wyszukiwania**.

### Przykład:

Użytkownicy mogą tworzyć zapytania odwołujące się do nazw odwzorowanych pól przy wyszukiwaniu dokumentów HTML zawierających określone metadane. Na przykład, jeśli element HTML o nazwie `description` zostałby odwzorowany na pole o nazwie `abstract`, użytkownicy mogliby posługiwać się poniższym wzorem zapytania, chcąc wyszukać dokumenty HTML poświęcone komputerom Thinkpad:

```
abstract:thinkpad
```

---

## Niestandardowe przetwarzanie tekstu

Dzięki zintegrowaniu niestandardowych algorytmów przetwarzania tekstu z kolekcjami wyszukiwania korporacyjnego można poprawić jakość i precyzję wyników wyszukiwania.

Program OmniFind Enterprise Edition obsługuje architekturę IBM Unstructured Information Management Architecture (UIMA), środowisko do tworzenia, wykrywania, składania i wdrażania funkcji analizy tekstu. Twórcy aplikacji mogą tworzyć i testować algorytmy analizy dla przeszukiwanej treści, a następnie tworzyć archiwum mechanizmu przetwarzania (plik `.pear`), które zawiera wszystkie zasoby niezbędne do wykorzystania tego archiwum podczas wyszukiwania korporacyjnego. Aby mieć możliwość przeszukiwania kolekcji za pomocą niestandardowych algorytmów analizy, użytkownik musi dodać do systemu wyszukiwania korporacyjnego archiwum zawierające mechanizm analizy tekstu.

Komponent logiki analizy w mechanizmie analizy tekstu jest nazywany *adnotatorem*. Każdy adnotator wykonuje konkretne zadania analizy lingwistycznej. Mechanizm przetwarzania tekstu może zawierać wiele adnotatorów; może się również składać z kilku mechanizmów analizy tekstu, z których każdy zawiera własne niestandardowe adnotatory.

Informacje tworzone przez adnotatory są nazywane *wynikami analizy*. Wyniki analizy, które odpowiadają wyszukiwanym informacjom, są zapisywane w strukturze danych nazywanej *wspólną strukturą analizy*.

Podczas konfigurowania opcji przetwarzania tekstu dla kolekcji użytkownik wykonuje następujące zadania:

- Wybiera mechanizm analizy tekstu, który ma być używany do tworzenia adnotacji do dokumentów w kolekcji.
- Jeśli kolekcja zawiera dokumenty XML z istotnymi znacznikami i znaczniki te powinny zostać użyte w niestandardowej analizie tekstu, można powiązać pliki odwzorowania z kolekcją i odwzorować dane wyjściowe odwzorowania XML na wspólną strukturę analizy.

Na przykład można odwzorować zawartość elementów `<adresat>` i `<klient>` na adnotacje `Osoba` we wspólnej strukturze analizy. Do tych adnotacji mogą później uzyskać dostęp

niestandardowe adnotatory, wykrywające dodatkowe informacje (na przykład płeć osoby). Można również odwzorować adnotacje osoby w indeksie wyszukiwania korporacyjnego, co umożliwi użytkownikom wyszukiwanie osób bez znajomości oryginalnych elementów XML.

Jeśli użytkownicy mają mieć możliwość określania w zapytaniach oryginalnych elementów XML, nie trzeba definiować żadnych odwzorowań XML. Zamiast tego można skonfigurować opcje analizy składni i umożliwić rodzime wyszukiwanie XML w kolekcji.

- Odwzoruj wspólną strukturę analizy na indeks wyszukiwania korporacyjnego, co umożliwi wyszukiwanie adnotowanych dokumentów metodą semantyczną.

Na przykład w zależności od encji i związków wykrytych przez adnotatory użytkownicy mogą wyszukiwać pojęcia, które występują w tym samym zdaniu (takie jak nazwisko konkretnej osoby i nazwa dowolnego konkurenta), lub słowo kluczowe i pojęcie (takie jak imię Jan i numer telefonu).


- Odwzoruj wspólną strukturę analizy na relacyjną bazę danych. Można odwzorować dane na tabele bazy danych IBM DB2 lub Oracle. Ten typ odwzorowania umożliwia użycie wyników analizy na przykład w aplikacjach bazodanowych służących do eksploracji danych. Pozwala też na wykorzystanie zapytań SQL do wyszukiwania danych znajdujących się poza systemem wyszukiwania korporacyjnego.

### Pojęcia pokrewne

 Integracja niestandardowej analizy tekstu

 Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu

 Przepływ pracy dla integracji analizy niestandardowej

 Algorytmy analizy tekstu

 Aplikacje wyszukiwania semantycznego

 Termin zapytania wyszukiwania semantycznego

## Dodawanie mechanizmów analizy tekstu do systemu

Jeśli użytkownik utworzy niestandardowy mechanizm analizy tekstu, należy go dodać do systemu, aby można było z niego korzystać do wyszukiwania korporacyjnego. Używając mechanizmu analizy w kolekcjach, można analizować dokumenty i opatrywać je adnotacjami oraz zwiększać dokładność wyników wyszukiwania.

### Zagadnienia wstępne

Aby dodać mechanizmy analizy tekstu do systemu, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Informacje o tym zadaniu


Twórcy aplikacji mogą utworzyć archiwum mechanizmu przetwarzania (.pear), które jest zgodne ze strukturą UIMA do analizy tekstu. Archiwum zawiera wszystkie zasoby niezbędne do wyszukiwania w kolekcjach wyszukiwania korporacyjnego. Aby mieć możliwość przeszukiwania kolekcji za pomocą niestandardowych algorytmów analizy, użytkownik musi dodać do systemu wyszukiwania korporacyjnego archiwum zawierające mechanizm analizy tekstu.

Po dodaniu do systemu mechanizmu analizy tekstu można zmienić jego nazwę wyświetlaną i wybrać opcję przeglądania źródła XML. Na podstawie źródła XML można ustalić, jakie informacje są wytwarzane przez dany mechanizm.

Jeśli mechanizm analizy tekstu jest powiązany z kolekcją, nie można go usunąć z systemu.

### Procedura

Aby dodać niestandardowy mechanizm analizy tekstu do systemu wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Analiza składni kliknij opcję **Konfiguruj mechanizmy analizy tekstu**.
4. Na stronie Mechanizmy analizy tekstu kliknij opcję **Dodaj mechanizm analizy tekstu**.
5. Na stronie Dodawanie mechanizmu analizy tekstu wpisz opisową nazwę nowego mechanizmu. W systemie nazwa wyświetlana służy do identyfikowania tego mechanizmu analizy tekstu w konsoli administracyjnej.
6. Określ położenie pliku .pear. Jeśli plik jest mniejszy niż 8 MB, może się on znajdować na komputerze lokalnym i można go odszukać, przeglądając lokalne katalogi. Pliki większe niż 8 MB muszą się znajdować na serwerze indeksowania i należy wpisać jego pełną ścieżkę.
7. Kliknij przycisk **OK**. Mechanizm analizy tekstu będzie wyświetlany na stronie Mechanizmy analizy tekstu.

### Pojęcia pokrewne



Przepływ pracy dla integracji analizy niestandardowej



Integracja niestandardowej analizy tekstu



Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu



Znaczniki XML w analizie i wyszukiwaniu

### Zadania pokrewne



Tworzenie elementów XML w pliku odwzorowania wspólnej struktury analizy

## Powiązanie mechanizmu analizy tekstu z kolekcją

Jeśli z systemem wyszukiwania korporacyjnego są powiązane niestandardowe mechanizmy analizy tekstu, można wybrać jeden z nich do użycia z określoną kolekcją. Użytkownicy, wyszukujący dane w tej kolekcji, będą mogli tworzyć zapytania semantyczne, a przy tym podnieść jakość i dokładność wyników wyszukiwania.

### Zagadnienia wstępne

Aby powiązać mechanizm analizy tekstu z kolekcją, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji.

### Informacje o tym zadaniu

Jeśli mechanizm analizy tekstu już jest powiązany z tą kolekcją, podczas próby powiązania innego mechanizmu zostaną wykonane następujące działania:



- Jeśli zostanie wybrana opcja **Brak analizy niestandardowej**, wszystkie odwzorowania analizy tekstu zdefiniowane wcześniej dla tej kolekcji zostaną zresetowane. Odtąd dla kolekcji będą używane systemowe wartości domyślne.
- Jeśli zostanie wybrana nazwa innego niestandardowego mechanizmu analizy tekstu, wszystkie odwzorowania analizy tekstu zdefiniowane wcześniej dla tej kolekcji zostaną zachowane. Na przykład, jeśli `mechanizm_1` zostanie zmieniony na `mechanizm_2`, to `mechanizm_2` odziedziczy pliki odwzorowań XML, które skonfigurowano dla `mechanizmu_1`.

### Procedura

Aby powiązać mechanizm analizy tekstu z kolekcją, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. Kliknij opcję **Wybierz mechanizm analizy tekstu**. Jeśli żaden niestandardowy mechanizm analizy tekstu nie został dodany do systemu wyszukiwania korporacyjnego lub jeśli dla kolekcji używane są domyślne algorytmy analizy, mechanizm analizy ma nazwę **Domyślnie**.
3. Na stronie Wybór mechanizmu analizy tekstu dla tej kolekcji wybierz nazwę mechanizmu, który ma być używany z tą kolekcją. Jeśli nie są dostępne żadne mechanizmy analizy tekstu lub jeśli wybrano opcję **Brak analizy niestandardowej**, analizator składni zastosuje domyślne reguły analizy tekstu podczas dodawania do dokumentów adnotacji i przygotowywania dokumentów pod kątem indeksu.

### Pojęcia pokrewne



Przepływ pracy dla integracji analizy niestandardowej



Integracja niestandardowej analizy tekstu



Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu

## Odzworowywanie elementów XML na wspólną strukturę analizy

Jeśli kolekcja zawiera dokumenty XML ze znaczącymi znacznikami, elementy opisane tymi znacznikami można odzworować na wspólną strukturę analizy, co umożliwi użytkownikom prowadzenie wyszukiwania semantycznego w relacyjnej bazie danych lub w systemie wyszukiwania korporacyjnego.

### Zagadnienia wstępne

Odzworowanie elementów XML na wspólną strukturę analizy może wykonać tylko osoba mająca rolę administratora wyszukiwania korporacyjnego lub administrator danej kolekcji.

### Ograniczenia

Maksymalna wielkość pliku odzworowania wynosi 8 MB.

### Informacje o tym zadaniu

Chcąc odzworować kilka elementów XML na wspólny typ w celu użycia w wyszukiwaniu semantycznym lub umożliwić dostęp do konkretnych elementów w dokumentach XML za pośrednictwem niestandardowych procesów analizy tekstu, można utworzyć niestandardowe pliki odwzorowań. Pliki te muszą być zgodne z wymogami standardu UIMA w zakresie analizy tekstu.

Dodanie plików odwzorowań do kolekcji korzystającej z mechanizmu niestandardowej analizy tekstu umożliwia odwzorowywanie elementów XML w dokumentach źródłowych na adnotacje we wspólnej strukturze analizy. Adnotacje mogą następnie zostać wykorzystane przez niestandardowy mechanizm analizy tekstu. Wspólną strukturę analizy można odwzorować na indeks i umożliwić użytkownikom tworzenie zapytań względem adnotacji, gdy wyszukują w kolekcji metodą semantyczną.

Na przykład można odwzorować zawartość elementów adresat i klient na adnotacje Osoba we wspólnej strukturze analizy. Do tych adnotacji mogą później uzyskać dostęp niestandardowe adnotatory, wykrywające dodatkowe informacje (na przykład płeć osoby). Można również odwzorować adnotacje osoby w indeksie wyszukiwania korporacyjnego, co umożliwi użytkownikom wyszukiwanie osób bez znajomości oryginalnych elementów XML.

Jeśli użytkownicy mają mieć możliwość określania oryginalnych elementów XML w zapytaniach, nie trzeba konfigurować żadnych plików odwzorowań. Zamiast tego można skonfigurować opcje analizy składni i umożliwić rodzime wyszukiwanie XML w kolekcji.

### Procedura

Aby odwzorować elementy XML na wspólną strukturę analizy:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. W obszarze **Odwzoruj elementy XML na wspólną strukturę analizy** kliknij przycisk **Dodaj odwzorowanie**.
3. na stronie Odwzorowanie elementów XML na wspólną strukturę analizy wpisz opisową nazwę pliku odwzorowania.
4. Określ położenie pliku. Jeśli plik odwzorowania znajduje się w systemie lokalnym, można odszukać go w odpowiednim katalogu za pomocą okna dialogowego. Jeśli plik odwzorowania znajduje się na serwerze indeksowania, należy wpisać pełną ścieżkę do niego.
5. Kliknij przycisk **OK**. Nowy plik odwzorowania zostaje dodany do strony Opcje przetwarzania tekstu.

#### Pojęcia pokrewne


 Przepływ pracy dla integracji analizy niestandardowej

 Integracja niestandardowej analizy tekstu

 Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu

 Znaczniki XML w analizie i wyszukiwaniu

#### Zadania pokrewne

 Tworzenie elementów XML w pliku odwzorowania wspólnej struktury analizy

## Odwzorowywanie wspólnej struktury analizy na indeks

Można określić, które wyniki analizy tekstu ze wspólnej struktury analizy mają zostać odwzorowane na indeks i dostępne dla użytkowników wyszukujących w kolekcji metodą semantyczną.

### Zagadnienia wstępne

Odwzorowanie wspólnej struktury analizy na indeks może wykonać tylko osoba mająca rolę administratora systemu wyszukiwania korporacyjnego albo administrator odpowiedniej kolekcji.

### Ograniczenia

Maksymalna wielkość pliku odwzorowania wynosi 8 MB.

### Informacje o tym zadaniu

Dzięki odwzorowaniu wspólnej struktury analizy na indeks wyszukiwania korporacyjnego użytkownicy zyskują możliwość konstruowania dokładnych semantycznie zapytań i otrzymują wyniki wyszukiwania o lepszej jakości.





Na przykład, w zależności od jednostek i związków wykrytych przez adnotatory, użytkownicy mogą wyszukiwać pojęcia występujące w tym samym zdaniu (takie jak nazwisko danej osoby lub nazwa konkurencyjnej firmy) lub słowa kluczowe i pojęcia (na przykład imię Aleksander i numer telefonu).

### Procedura

Aby odwzorować wspólną strukturę analizy na indeks:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. W obszarze **Odwzoruj wspólną strukturę analizy na indeks** kliknij opcję **Wybierz plik odwzorowania**.
3. Na stronie Wybór pliku odwzorowania dla tej kolekcji wybierz plik odwzorowania, który ma być używany w połączeniu z indeksem wyszukiwania korporacyjnego:
  - Aby dla indeksu wyszukiwania korporacyjnego użyć domyślnych reguł odwzorowania, wybierz opcję **Domyślny**.
  - Aby odwzorować niestandardową wspólną strukturę analizy na indeks, określ położenie pliku odwzorowania. Jeśli plik znajduje się w systemie lokalnym, można go znaleźć, przeglądając system plików. Jeśli plik znajduje się na serwerze indeksowania, należy wpisać pełną ścieżkę do tego pliku.
4. Kliknij przycisk **OK**. Tak określony plik odwzorowania zostaje wyświetlony na stronie Opcje przetwarzania tekstu.

#### Pojęcia pokrewne

-  Przepływ pracy dla integracji analizy niestandardowej
-  Integracja niestandardowej analizy tekstu
-  Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu
-  Odwzorowanie indeksu wyników analizy niestandardowej

#### Zadania pokrewne

-  Tworzenie pliku odwzorowania wspólnej struktury analizy do indeksu

## Odwzorowywanie wspólnej struktury analizy na relacyjną bazę danych

Można określić, które wyniki analizy tekstu ze wspólnej struktury analizy mają być odwzorowywane na relacyjną bazę danych na użytek aplikacji bazodanowych.

## Zagadnienia wstępne

Odwzorowanie wspólnej struktury analizy na relacyjną bazę danych może wykonać tylko osoba mająca rolę administratora systemu wyszukiwania korporacyjnego albo administrator odpowiedniej kolekcji.

## Ograniczenia

Maksymalna wielkość pliku odwzorowania wynosi 8 MB.

## Informacje o tym zadaniu

Dzięki odwzorowaniu wspólnej struktury analizy na tabelę relacyjnej bazy danych zawarte w nich dane stają się dostępne dla aplikacji bazodanowych. Na przykład użytkownicy mogą tworzyć zapytania SQL poza systemem wyszukiwania korporacyjnego w celu wyszukiwania adnotacji dodawanych w wyniku analizy tekstu. Tych informacji można również użyć do dalszego przetwarzania tekstu, np. korzystając z nich w aplikacjach służących do eksploracji danych.

Odwzorowując wspólną strukturę analizy na relacyjną bazę danych, można używać jednego pliku odwzorowania na raz. Załadowanie nowego pliku odwzorowania powoduje zastąpienie pliku używanego dotychczas.

## Procedura

Aby odwzorować wspólną strukturę analizy na relacyjną bazę danych:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. W obszarze **Odwzoruj wspólną strukturę analizy na relacyjną bazę danych** kliknij opcję **Dodaj odwzorowanie**.
3. Na stronie **Odwzorowanie wspólnej struktury analizy na relacyjną bazę danych** wpisz opisową nazwę pliku odwzorowania, który ma służyć do odwzorowania informacji na relacyjną bazę danych.
4. Określ położenie pliku odwzorowania. Jeśli plik znajduje się w systemie lokalnym, można go znaleźć, przeglądając system plików. Jeśli plik znajduje się na serwerze indeksowania, należy wpisać pełną ścieżkę do tego pliku.
5. Kliknij przycisk **OK**. Nazwa wyświetlana pliku odwzorowania jest widoczna na stronie **Opcje przetwarzania tekstu**.

### Pojęcia pokrewne



Przepływ pracy dla integracji analizy niestandardowej



Integracja niestandardowej analizy tekstu



Podstawowe pojęcia używane w odniesieniu do przetwarzania w zakresie analizy tekstu



Odwzorowanie do bazy danych dla wybranych wyników analizy

### Zadania pokrewne



Tworzenie pliku odwzorowania wspólnej struktury analizy do bazy danych

---

## Konfigurowanie wątków dla usługi analizatora składni

Jeśli istnieją wystarczające zasoby pamięci, można zwiększyć liczbę wątków dostępnych dla analizatora składni podczas analizowania dokumentów.

### Zagadnienia wstępne

W przypadku dużej liczby kolekcji można zwiększyć liczbę wątków analizatora składni. Należy się upewnić, że system ma wystarczającą ilość pamięci do obsługi dodatkowych wątków. Analizator składni z jednym wątkiem wymaga 200 MB pamięci. Każdy dodatkowy wątek wymaga kolejnych 50 MB pamięci.

Aby skonfigurować liczbę wątków uruchamianych na potrzeby analizatora składni, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

### Procedura

Aby skonfigurować liczbę wątków analizatora składni:

1. Przejdź do edycji kolekcji, wybierz stronę Analiza składni, a następnie kliknij opcję **Skonfiguruj opcje analizowania**.
2. Określ maksymalną liczbę wątków analizatora składni, które mają być uruchamiane podczas uruchamiania analizatora składni, i kliknij przycisk **OK**.
3. Zrestartuj analizator składni.

---

## Włączanie zaawansowanej analizy terminów złożonych

Jakość wyszukiwania można poprawić, umożliwiając analizatorowi składni użycie zaawansowanych metod analizy dla terminów złożonych. Analiza zaawansowana umożliwia rozbiór terminów złożonych, dzięki czemu każdy z jego członów może być traktowany jak pojedynczy termin.

### Zagadnienia wstępne

Aby określić opcje analizowania terminów złożonych, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

### Informacje o tym zadaniu

Niektóre języki łączą terminy zawierające kilka słów w pojedyncze słowo bez spacji (terminy *złożone*). Funkcja zaawansowanej analizy i rozbioru terminów złożonych jest przydatna podczas wyszukiwania w językach takich jak niemiecki i ma newralgiczne znaczenie w przypadku wyszukiwania w językach takich jak koreański.

Jeśli zaawansowana analiza terminów złożonych zostanie włączona, użytkownicy mogą wyszukiwać terminy bez konieczności używania znaków wieloznacznych w celu odnalezienia złożonych form terminów użytych w zapytaniu. Na przykład wyszukanie terminu **Organ** (organ) może zwrócić dokumenty zawierające termin **Organspender** (dawca organu), ale nie zwróci dokumentów zawierających termin **Organisation** (organizacja). W odróżnieniu od zapytania ze znakiem wieloznacznym **Organ\***, które może zwrócić dowolny łańcuch zawierający ciąg **Organ**, wyszukiwanie dopasowuje tylko pełne słowa podrzędne w ramach większego terminu złożonego.

Jako inny przykład można podać słowo złożone **Mustermann**, które jest dzielone na dwa leksemy (muster i mann) przechowywane w indeksie osobno. Gdy zostanie wprowadzone zapytanie ze znakiem wieloznacznym **Musterma\***, procesy wyszukiwania nie będą w stanie zidentyfikować łańcucha **Musterma** jako przedrostka istniejącego słowa. W rezultacie dokumenty z terminem **Mustermann** nie zostaną znalezione. Jeśli użytkownicy mają mieć możliwość wprowadzania zapytań ze znakami wieloznacznymi w celu znalezienia terminów złożonych, nie należy włączać zaawansowanej analizy terminów złożonych.

Definiowane przez użytkownika terminy słownikowe, takie jak synonimy i słowa-kryteria, są również stosowane do członów terminów złożonych, które są używane jako oddzielne słowa w zapytaniu.

### Procedura

Aby włączyć zaawansowaną analizę terminów złożonych, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Skonfiguruj opcje analizowania**.
2. Zaznacz pole wyboru **Włącz zaawansowaną analizę terminów złożonych** i kliknij przycisk **OK**.

#### Pojęcia pokrewne

“Znaki wieloznaczne w zapytaniach” na stronie 184



Obsługa lingwistyczna w wyszukiwaniu semantycznym



Analiza tekstu w wyszukiwaniu korporacyjnym

---

## Włączanie obsługi rodzimego wyszukiwania XML

Jeśli kolekcja zawiera dokumenty XML, można umożliwić użytkownikom korzystanie z języka znaczników XML podczas wyszukiwania dokumentów, włączając rodzime wyszukiwanie XML dla kolekcji.

### Zagadnienia wstępne

Aby włączyć obsługę wyszukiwania dokumentów XML za pomocą rodzimego wyszukiwania XML, użytkownik musi mieć uprawnienia administracyjne wyszukiwania korporacyjnego lub być administratorem danej kolekcji.

### Informacje o tym zadaniu

Rodzime wyszukiwanie XML, takie jak XPath lub fragmenty XML, może zapewnić dokładniejsze wyniki wyszukiwania dzięki wykorzystaniu języka znaczników XML dokumentów w zapytaniu. Użytkownicy mogą określić warunek, że wyszukiwana fraza musi występować w obrębie określonego elementu lub atrybutu XML.

Na przykład faktury od dostawcy komputerów, które są w formacie XML, mogą zawierać wpisy <zamówienie>, które zawierają elementy <firma> i <typkomputera>. Podczas pobierania faktur, które zawierają zamówienia na notebooki firmy IBM, wyszukiwanie według słowa kluczowego może spowodować pobranie dokumentów, które zawierają notebooki firmy Dell i modele desktop firmy IBM. Dzięki użyciu wyszukiwania XML można określić, że w elemencie <firma> musi wystąpić IBM, w elemencie <typkomputera> musi wystąpić notebook oraz że obydwa elementy muszą wystąpić w tym samym elemencie <zamówienie>. W ten sposób można pobrać faktury, które dotyczą notebooków firmy IBM.



## Procedura

Aby umożliwić użytkownikom przeszukiwanie kolekcji za pomocą rodzimego wyszukiwania XML, należy wykonać poniższe czynności:

1. Przejdź do edycji kolekcji, wybierz stronę Analiza składni, a następnie kliknij opcję **Skonfiguruj opcje analizowania**.
2. Wybierz opcję **Zezwól użytkownikom na wyszukiwanie w dokumentach XML za pomocą rodzimego wyszukiwania XML**.
3. Kliknij przycisk **OK**.

### Pojęcia pokrewne

 Obsługa lingwistyczna w wyszukiwaniu semantycznym

 Analiza tekstu w wyszukiwaniu korporacyjnym

 Aplikacje wyszukiwania semantycznego

 Termin zapytania wyszukiwania semantycznego

“Pola wyszukiwania XML” na stronie 137

---

## Wykrywanie formatu dokumentów

W wyszukiwaniu korporacyjnym do wykrywania typu dokumentu i typu analizatora, jaki ma obsługiwać dany dokument stosowane jest domyślne odwzorowanie rozszerzeń URL i typów MIME.

Edytując plik konfiguracyjny `parserTypes.cfg`, można przesłonić i rozbudować domyślne odwzorowania rozszerzeń URL i typów MIME na typy analizatorów. Plik `parserTypes.cfg` definiuje reguły odwzorowania rozszerzeń plików lub typów MIME na typy analizatorów. Na przykład można odwzorować rozszerzenie plików, takie jak `.content`, i nakazać, aby dokumenty tego typu mają być analizowane przez analizator HTML.

Poszczególne formaty dokumentów mają różne wewnętrzne reprezentacje. System wyszukiwania korporacyjnego do analizowania dokumentów korzysta z filtrów wewnętrznych i zewnętrznych, a wiele dokumentów analizowanych jest za pomocą analizatorów specjalizowanych do obsługi określonego formatu.

Wykrywanie formatu dokumentów i przypisywanie analizatorów wykonywane jest w następujący sposób:

1. Algorytm wykrywania formatu dokumentu sprawdza rozszerzenie adresu URL przetwarzanego dokumentu.
2. System sprawdza typ MIME dokumentu będący częścią metadanych podawanych przez przeszukiwacza.
3. System próbuje powiązać z każdym dokumentem właściwy typ analizatora. W przypadku dokumentów HTML, tekstowych (TXT) i XML, system przypisuje typ analizatora właściwy dla formatu każdego z tych dokumentów.

W przypadku wszystkich innych formatów dokumentów system korzysta z analizatora Stellent. Należy pamiętać, że technologia filtrowania dokumentów Stellent jest własnością firmy Oracle. Odniesienia do technologii Stellent w tej dokumentacji są równoznaczne z odniesieniami do technologii Oracle Outside In Content Access.

Analizator Stellent obsługuje kilkaset różnych formatów dokumentów, ale wyszukiwanie korporacyjne może skorzystać tylko z pewnego podzbioru tych filtrów. Można jednak edytować pliki konfiguracyjne, umożliwiając analizowanie innych typów dokumentów przez analizator Stellent.

**Ważne:** Dodawane przez użytkownika filtry dokumentów nienależące do podzbioru filtrów dokumentów włączonych dla wyszukiwania korporacyjnego w domyślnej konfiguracji systemu nie zostały przetestowane i nie są obsługiwane.

4. Jeśli nie jest dostępny plik `parserTypes.cfg`, to do wykrywania typu dokumentu i przydzielenia mu analizatora stosowane są domyślne odwzorowania. Aby wykryć typ dokumentu, system wykonuje następujące kroki:
  - a. Porównuje rozszerzenie adresu URL z regułami zdefiniowanymi przez klienta w pliku `parserTypes.cfg`.
  - b. Porównuje typ MIME z regułami typów MIME zdefiniowanymi przez klienta w pliku `parserTypes.cfg`.
  - c. Porównuje rozszerzenie adresu URL z domyślnymi regułami wyszukiwania korporacyjnego.
  - d. Porównuje typ MIME z domyślnymi regułami typów MIME wyszukiwania korporacyjnego.

5. Jeśli system nie jest w stanie wykryć formatu dokumentu, to dokument ten jest odrzucany. Pojawić się może komunikat o błędzie informujący, że ten typ dokumentów nie jest obsługiwany.

Jeśli do danego typu dokumentów przydzielony został analizator Stellent, a ten nie jest w stanie rozpoznać formatu dokumentu, to może pojawić się odpowiedni komunikat o błędzie. Taki błąd może się pojawić w sytuacji, gdy:

- Dokument jest uszkodzony.
- Dokument zapisany jest w formacie nieobsługiwanym przez analizatora Stellent. W celu rozwiązania tego problemu należy dopisać odrzucone formaty plików do pliku `stellentTypes.cfg`. Oprócz tego konieczne jest też zaktualizowanie pliku `parserTypes.cfg` poprzez wpisanie, że typ MIME lub rozszerzenie odrzuconego formatu dokumentów powinny zostać powiązane z analizatorem Stellent.

---

## Domyślne typy obsługiwanych dokumentów

W czasie wykrywania formatu dokumentu sprawdzane są tylko określone typy dokumentów.

Następujące formaty dokumentów są rodzimymi typami automatycznie wykrywanymi i analizowanymi przez wbudowane usługi analizatora składni kolekcji:

HTML  
Zwykły tekst  
XML

Dokumenty w następujących formatach są analizowane przez analizator Stellent:

Adobe Portable Document Format (PDF)  
Lotus 1-2-3  
Lotus Freelance Graphics  
Lotus Word Pro  
Just System Ichitaro  
Microsoft Excel (wersje do 2007)  
Microsoft PowerPoint (wersje do 2007)  
Microsoft Visio  
Microsoft Word (wersje do 2007)  
Tekst sformatowany (RTF)  
StarOffice/OpenOffice Calc  
StarOffice/OpenOffice Impress  
StarOffice/OpenOffice Draw  
StarOffice/OpenOffice Writer

Formaty plików Office Open XML w Microsoft Office oraz formaty OpenDocument w OpenOffice są obsługiwane bez konieczności dokonywania zmian w plikach konfiguracyjnych.

Aby analizować inne typy dokumentów, należy zaktualizować pliki konfiguracji (`parserTypes.cfg` i `stellenttypes.cfg`), dodając do nich reguły odwzorowania określonych typów dokumentów na usługi analizatora kolekcji lub filtry Stellent.

**Ograniczenie:** Przetwarzanie tekstu dwukierunkowego w plikach PDF tak, aby był zgodny z logicznym porządkiem czytania tekstu, leży poza możliwościami technologii zastosowanej w przeglądarce Stellent. Analizator składni Stellent nie gwarantuje porządku tekstu wyodrębnianego przy jego użyciu z plików PDF. W przypadku tekstu dwukierunkowego w plikach PDF porządek, w którym tekst jest analizowany, prawdopodobnie nie będzie zgodny z logicznym porządkiem czytania tekstu. To ograniczenie sprawia problem podczas przetwarzania plików PDF napisanych w językach bliskowschodnich, takich jak hebrajski i arabski, które w przeważającej mierze są pisane od prawej do lewej (dwukierunkowo).

---

## Typy dokumentów związane z analizatorami kolekcji i analizatorami Stellent

Chcąc się upewnić, że dokumenty z przestrzeni przeszukiwania są prawidłowo i skutecznie analizowane, można utworzyć pliki konfiguracyjne określające, jakie rodzaje dokumentów mają być analizowane przez analizatora kolekcji, a jakie mają być analizowane przez filtry dokumentów Stellent.

W kolekcji wyszukiwania korporacyjnego dokumenty w większości formatów przetwarzane są przy użyciu wbudowanych analizatorów składni HTML lub XML. Niektóre typy dokumentów z reguły nie podlegają analizie (np. dokumenty w języku Postscript), a inne są obsługiwane przez funkcje analizujące Stellent (np. dokumenty programów Microsoft Word, Microsoft Excel, Microsoft PowerPoint, Lotus Freelance, Lotus 123, PDF, RT i Ichitaro).

Metadane mogą wprowadzać w błąd, w wyniku którego dokumenty HTML i zwykły tekst mogą być omyłkowo przesyłane do analizatora Stellent, a następnie odsyłane do jednego z wbudowanych analizatorów. Taka sytuacja źle wpływa na wydajność pracy systemu. W przypadku innych dokumentów wykrycie ich rodzaju może się okazać niemożliwe, dlatego takie dokumenty są pomijane. Aby tego uniknąć, można utworzyć pliki konfiguracyjne definiujące miejsce i sposób analizowania różnych typów dokumentów.

Powiązanie typów dokumentów z analizatorem kolekcji i analizatorem Stellent wymaga wykonania następujących czynności:

1. Skonfigurowanie typów dokumentów dla analizatora składni kolekcji. W tym kroku należy utworzyć plik konfiguracyjny, który przypisuje typy dokumentów do analizatora składni używanego przez kolekcję. Można utworzyć po jednym pliku konfiguracyjnym na kolekcję.
2. Skonfigurowanie rodzajów dokumentów dla analizatora Stellent. W tym kroku należy utworzyć plik konfiguracyjny, który przypisuje typy dokumentów do filtrów dokumentów Stellent używanych przez kolekcję. Można utworzyć po jednym pliku konfiguracyjnym na kolekcję.
3. Zatrzymanie i ponowne uruchomienie analizatora składni. Aby zmiany odniosły skutek, należy użyć konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania kolekcji, dla której skonfigurowano typy dokumentów, a następnie zatrzymać i ponownie uruchomić analizator składni.

---

## Powiązanie typów dokumentów z analizatorem składni kolekcji

Aby powiązać określone typy dokumentów z analizatorem składni kolekcji, należy utworzyć plik konfiguracyjny o nazwie `parserTypes.cfg`. Konsola administracyjna nie umożliwia wykonania tego zadania.

### Zagadnienia wstępne

Aby wykonać to zadanie, należy się zalogować jako administrator wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Jeśli plik konfiguracyjny nie istnieje, analizator składni kolekcji używa domyślnych reguł usługi analizatora składni. Jeśli plik konfiguracyjny istnieje, reguły w pliku określają:

- Które rozszerzenie URL i który typ MIME będą odwzorowane do którego analizatora.
- W jaki sposób analizować składnię dokumentów, których typ jest nieznan ze względu na niekompletne metadane.

Zawartość pliku `parserTypes.cfg` jest zorganizowana w wiersze, z których każdy zawiera jedną z poniższych reguł:

**EXTENSION** *rozszerzenie analizator składni*

Wszystkie dokumenty, których adres URL kończy się podanym rozszerzeniem, zostaną przetworzone przez podany analizator składni. W rozszerzeniu nie należy wstawiać kropki, a podczas porównywania nie jest rozróżniana wielkość liter.

**CONTENTTYPE** *typ/podtyp analizator składni*

Wszystkie dokumenty, których typ treści odpowiada podanemu typowi/podtypowi, zostaną przetworzone przez podany analizator składni. Jeśli typ treści dokumentu zostanie podany w postaci `t/p`, zgodność występuje wtedy, gdy typ równa się `t` oraz podtyp równa się `p` albo podtypem jest znak wieloznaczny (gwiazdka, `*`).

**UNKNOWN** *analizator składni*

Wszystkie dokumenty, których rozszerzenie i typ treści są nieznane (czyli nie zostały udostępnione przez przeszukiwacz), zostaną przetworzone przez podany analizator składni.

**DEFAULT** *analizator składni*

Wszystkie dokumenty, które nie zostały uwzględnione w innych regułach, zostaną przetworzone przez podany analizator składni.

We wszystkich przypadkach *analizator składni* musi być równy wartości `text`, `html`, `xml`, `stellent` lub `none`, gdzie wartość `none` oznacza, że dokumenty tego typu nie zostaną przeanalizowane.

Jeśli dokument odpowiada więcej niż jednej regule, pierwszeństwo ma reguła, która jest bardziej szczegółowa, bez względu na kolejność występowania reguł:

- Reguła **EXTENSION** jest bardziej szczegółowa niż reguła **CONTENTTYPE**.
- Reguła **CONTENTTYPE**, która obejmuje podtyp, jest bardziej szczegółowa niż ta sama reguła ze znakiem wieloznacznym. Reguła dla typu treści `application/postscript` ma pierwszeństwo przed regułą `application/*`.
- Nie powinno być dwóch reguł dotyczących tego samego rozszerzenia lub typu treści. W takim przypadku kolejność reguł zależy od implementacji.

## Procedura

Aby powiązać typy dokumentów z analizatorem składni kolekcji, należy wykonać poniższe czynności:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. W następujący sposób utwórz plik konfiguracyjny, gdzie *id\_kolekcji* identyfikuje kolekcję, która ma zostać skonfigurowana:  
`ES_NODE_ROOT/master_config/id_kolekcji.parserdriver/parserTypes.cfg`
3. Za pomocą edytora tekstu zmodyfikuj zawartość pliku i określ reguły usługi analizatora składni, a następnie zapisz i zamknij plik.
4. Aby zmiany zostały uwzględnione, użyj konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania analizatora składni kolekcji, a następnie zatrzymaj i ponownie uruchom analizator składni.

## Przykład

W tym przykładzie wbudowany analizator składni HTML przetwarza wszystkie dokumenty mające rozszerzenia `txt`, `htm` lub `html`, mające typ treści zaczynający się członem `text/` oraz

te, dla których rozszerzenie i typ treści nie są znane. Wbudowany analizator składni XML przetwarza wszystkie dokumenty z rozszerzeniem xml lub z typem treści text/xml. Wszystkie inne dokumenty, w szczególności te, których typ treści zaczyna się członem application/, są kierowane do analizatora Stellent.

```
EXTENSION doc stellent
EXTENSION txt html
EXTENSION htm html
EXTENSION html html
EXTENSION xml xml
EXTENSION ps none
CONTENTTYPE text/xml xml
CONTENTTYPE text/* html
CONTENTTYPE application/* stellent
UNKNOWN html
DEFAULT stellent
```

#### **Zadania pokrewne**

“Zmiana reguł wymiany dla niektórych znaczników HTML” na stronie 158

“Powiązanie typów dokumentów z analizatorem składni Stellent” na stronie 160

#### **Odsyłacze pokrewne**

“Domyślne reguły usługi analizatora składni kolekcji”

## **Domyślne reguły usługi analizatora składni kolekcji**

Jeśli nie zostanie utworzony plik konfiguracyjny, który odwzorowuje typy plików i typy treści na analizator składni dla kolekcji, zostaną użyte domyślne reguły analizy składni dokumentów.

Domyślne reguły używane przez analizatora składni kolekcji:

```
EXTENSION pdf stellent
EXTENSION ppt stellent
EXTENSION prz stellent
EXTENSION lwp stellent
EXTENSION doc stellent
EXTENSION rtf stellent
EXTENSION xls stellent
EXTENSION 123 stellent
EXTENSION vsd stellent
EXTENSION vdx stellent
EXTENSION jxw stellent
EXTENSION jsw stellent
EXTENSION jtw stellent
EXTENSION jaw stellent
EXTENSION juw stellent
EXTENSION jbw stellent
EXTENSION jvw stellent
EXTENSION jfw stellent
EXTENSION jtt stellent
EXTENSION jtd stellent
EXTENSION jttc stellent
EXTENSION jtdc stellent
EXTENSION jtdx stellent
EXTENSION ps none
EXTENSION xml xml
EXTENSION txt text
EXTENSION htm html
EXTENSION html html
EXTENSION shtml html
EXTENSION xhtml html
EXTENSION asp html

CONTENTTYPE application/postscript none
CONTENTTYPE application/* stellent
```



```
CONTENTTYPE text/rtf stellent
CONTENTTYPE text/richtext stellent
CONTENTTYPE text/xml xml
CONTENTTYPE text/html html
CONTENTTYPE text/plain text
```

```
UNKNOWN none
DEFAULT none
```

### **Zadania pokrewne**

“Powiązanie typów dokumentów z analizatorem składni kolekcji” na stronie 155

## **Analizowanie dokumentów nieznanego typu**

Jeśli typ dokumentu jest nieznanego (na przykład dokument nie ma rozszerzenia nazwy pliku lub z dokumentem nie jest skojarzony typ MIME), można skonfigurować reguły, które zapobiegną usuwaniu takiego dokumentu przez analizator składni.

### **Informacje o tym zadaniu**

Jeśli analizator składni nie rozpoznaje formatu pliku, to dokument jest poddawany próbie analizy z użyciem domyślnego analizatora składni HTML. Jeśli treść nie jest w formacie HTML, analizator składni odrzuca dokument. Informacje o wszystkich odrzuconych dokumentach są zapisywane w następującym miejscu na serwerze indeksowania:

ES\_NODE\_ROOT/data/id\_kolekcji/dropped\_doc\_logs/dropped\_docs\_pd\_date.log

### **Procedura**

Aby nie doszło do takiej sytuacji, analizator składni należy skonfigurować tak, aby w przypadku nieznanego typu dokumentów używany był analizator składni ASCII:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Utwórz lub poddaj edycji następujący plik konfiguracyjny analizatora składni, gdzie *id\_kolekcji* stanowi identyfikator kolekcji, która ma zostać skonfigurowana:  
ES\_NODE\_ROOT/master\_config/id\_kolekcji.parserdriver/parserTypes.cfg
3. Dodaj następującą regułę do pliku:  
UNKNOWN text
4. Aby zmiany zostały uwzględnione, użyj konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania analizatora składni kolekcji, a następnie zatrzymaj i ponownie uruchom analizator składni.

---

## **Zmiana reguł wymiany dla niektórych znaczników HTML**

Można zmienić reguły wymiany znaczników HTML używane przez analizator składni podczas analizowania dokumentów HTML. Konsola administracyjna nie umożliwia wykonania tego zadania.

### **Zagadnienia wstępne**

Aby wykonać to zadanie, należy się zalogować jako administrator wyszukiwania korporacyjnego.

### **Informacje o tym zadaniu**

Aby udostępnić informacje tekstowe dla indeksu wyszukiwania korporacyjnego, analizator składni HTML wymienia informacje w języku znaczników (znaczniki HTML) na inne znaki, które modelują znaczenie znaczników w sposób stosowany w przeglądarce WWW. Na przykład znacznik akapitu (<p>) wymieniany jest na ogranicznik akapitu w strukturze CAS (Common Analysis Structure) i w indeksie. Reguły wymiany mają wpływ na to, które części informacji tekstowych znajdują się w tym samym akapicie.

Aby znaczenie niektórych znaczników ściślej związać ze standardem HTML, można zaktualizować plik konfiguracyjny analizatora składni dla kolekcji, zmieniając w nim niektóre reguły wymiany.

**Wskazówka:** Inne metody określania sposobu obsługi znaczników HTML w dokumentach HTML można poznać w sekcji <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011251>.

### Procedura

Aby zmienić reguły wymiany znaczników HTML:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Dokonaj edycji pliku `ES_NODE_ROOT/master_config/id_kolekcji/parserdriver/collection.properties`, gdzie `id_kolekcji` stanowi identyfikator kolekcji, która ma zostać skonfigurowana.
3. Ustaw następujący parametr na wartość `true`:  
`trevi.tokenizer.newHtmlTagReplacement=true`

Reguły wymiany następujących znaczników HTML zostaną zmienione na niżej podane wartości. Wszystkie inne reguły wymiany znaczników HTML nadal będą takie, jak reguły domyślne.

```
dfn EMPTY
div NEW LINE
q BLANK
span EMPTY
```

4. Aby zmiany zostały uwzględnione, użyj konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania analizatora składni kolekcji, a następnie zatrzymaj i ponownie uruchom analizator składni.

#### Zadania pokrewne

“Powiązanie typów dokumentów z analizatorem składni kolekcji” na stronie 155

#### Odsyłacze pokrewne

“Domyślne reguły wymiany HTML”

## Domyślne reguły wymiany HTML

Analizator składni HTML interpretuje i zastępuje informacje w języku znaczników w dokumentach HTML w celu dostarczenia informacji tekstowych na potrzeby indeksu wyszukiwania korporacyjnego.

W analizatorze składni HTML dla wyszukiwania korporacyjnego używane są niżej podane reguły wymiany podczas analizowania znaczników HTML. W pierwszej kolumnie widoczna jest nazwa znacznika HTML. W drugiej kolumnie widoczna jest wartość zamiennika.

```
comment BLANK
a EMPTY
applet EMPTY
area EMPTY
```

b	EMPTY
base	EMPTY
big	EMPTY
body	EMPTY
br	NEW LINE
caption	EMPTY
center	EMPTY
del	EMPTY
dfn	PARAGRAPH
dir	PARAGRAPH
div	BLANK
dl	PARAGRAPH
em	EMPTY
form	EMPTY
frame	PARAGRAPH
h1	PARAGRAPH
h2	PARAGRAPH
h3	PARAGRAPH
h4	PARAGRAPH
h5	PARAGRAPH
h6	PARAGRAPH
hr	PARAGRAPH
iframe	EMPTY
img	NEW LINE
li	BLANK
meta	EMPTY
object	EMPTY
ol	BLANK
option	EMPTY
p	PARAGRAPH
q	PARAGRAPH
samp	PARAGRAPH
script	EMPTY
select	EMPTY
spacer	BLANK
span	NEW LINE
strike	EMPTY
strong	EMPTY
style	EMPTY
table	PARAGRAPH
td	NEW LINE
title	PARAGRAPH
tr	NEW LINE
ul	BLANK
xmp	PARAGRAPH

### **Zadania pokrewne**

“Zmiana reguł wymiany dla niektórych znaczników HTML” na stronie 158

---

## **Powiązanie typów dokumentów z analizatorem składni Stellent**

Aby określić, które typy dokumentów mają być analizowane za pomocą filtrów Stellent, należy utworzyć plik konfiguracyjny o nazwie `stellenttypes.cfg`. Konsola administracyjna nie umożliwi wykonania tego zadania.

### **Technologia Oracle Outside In:**

Technologia filtrowania dokumentów Stellent jest własnością firmy Oracle. Odniesienia do technologii Stellent w tej dokumentacji są równoznaczne z odniesieniami do technologii Oracle Outside In Content Access.

### **Ograniczenia**

Produkt OmniFind Enterprise Edition obsługuje filtry Stellent w przypadku następujących typów dokumentów:

- Adobe Portable Document Format (PDF)
- Lotus 1-2-3
- Lotus Freelance Graphics
- Lotus Word Pro
- Ichitaro
- Microsoft Excel
- Microsoft PowerPoint
- Microsoft Visio
- Microsoft Word
- Tekst sformatowany (RTF)

Jeśli w indeksie wyszukiwania korporacyjnego mają być uwzględniane dodatkowe typy dokumentów, których formaty obsługiwane są przez filtr Stellent, można skonfigurować reguły analizowania w pliku konfiguracyjnym `stellenttypes.cfg`.

**Ważne:** Dodawane formaty dokumentów niezgodne z typami dokumentów z wcześniej podanej listy nie zostały przetestowane pod kątem wyszukiwania korporacyjnego i nie są obsługiwane.

Kompletną listę formatów dokumentów Stellent można znaleźć w dokumencie *Outside In Technology: Supported File Formats* (Technologia Outside In: obsługiwane formaty plików) pod adresem [http://www.oracle.com/technology/products/content-management/oit/ds\\_oitFiles.pdf](http://www.oracle.com/technology/products/content-management/oit/ds_oitFiles.pdf).

### Zagadnienia wstępne

Aby wykonać to zadanie, należy się zalogować jako administrator wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Plik konfiguracyjny `stellenttypes.cfg` określa:

- Reguły akceptacji dotyczące typów plików, które mają być analizowane przez analizator składni Stellent. Typ pliku odpowiada jednemu z typów plików rozpoznawanych przez bibliotekę Stellent.
- Reguły rodzime dotyczące typów plików, które mają być odsyłane do analizatora składni kolekcji w celu przeanalizowania z użyciem jednego z wbudowanych analizatorów. Działanie to jest wymagane, ponieważ analizator kolekcji mógłby przekazać dokument do analizatora Stellent z powodu mylnych metadanych.
- Reguły odrzucania dotyczące typów plików, które mają być odrzucane ze względu na brak ich obsługi w ramach wyszukiwania korporacyjnego.

Jeśli plik konfiguracyjny został określony, ale nie istnieje, nie można będzie uruchomić analizatora składni. Jeśli nie określono żadnego pliku konfiguracyjnego w ustawieniu właściwości `OutsideInSupportedTypes` w pliku `stellent.properties`, używane będą domyślne reguły analizy obowiązujące dla analizatorów Stellent.

W pliku konfiguracyjnym są wymienione typy dokumentów i sposoby ich obsługi. Plik jest sformatowany jako sekwencja wierszy, z których każdy reprezentuje jedną regułę według jednego z następujących formatów:

```
accept DEFAULT
accept ALL typ_dokumentu
accept typ_stellent typ_dokumentu
native DEFAULT
native typ_stellent typ_dokumentu
reject typ_stellent
```

Gdzie:

#### *typ\_dokumentu*

Jest wartością używaną dla tokenu zapytania doctype. Dokumenty można wyszukiwać według typu dokumentu. Na przykład użytkownik mógł podać `$doctype::pdf` w celu wyszukania dokumentów PDF.

#### *typ\_stellent*

Jest to jedna z wartości typu filtru z biblioteki Stellent, na przykład FI\_123R1.

### **DEFAULT**

Oznacza, że lista akceptowanych lub rodzimych typów, zależnie od typu reguły, obejmuje wszystkie reguły domyślne. Ta opcja umożliwi rozszerzenie domyślnej konfiguracji bez konieczności jej zastępowania.

**All** Oznacza, że dla podanego tokenu doctype będą akceptowane wszystkie typy, które nie zostały jawnie wymienione.

Reguły w pliku konfiguracyjnym są przetwarzane w następujący sposób:

1. Jeśli istnieje reguła **reject** dla danego typu *typ\_stellent*, to dany dokument nie jest akceptowany.
2. Jeśli istnieje reguła **native** dla danego typu *typ\_stellent* (w tym domyślne reguły analizowania, jeśli określono wartość **native DEFAULT**), to oprócz wartości elementu *typ\_dokumentu* określonego tą regułą, dany dokument jest odsyłany do wbudowanego analizatora składni. Wartość elementu *typ\_dokumentu* musi być równa txt, htm lub xml, co wskazuje odpowiednio na zwykły tekst, HTML lub XML.
3. Jeśli istnieje reguła **accept** dla danego typu *typ\_stellent* (w tym lista domyślna, jeśli określono wartość **accept DEFAULT**), to dany dokument jest akceptowany.
4. W przeciwnym przypadku, jeśli została określona reguła **accept ALL**, dokument zostanie zaakceptowany.
5. W przeciwnym przypadku dokument zostanie odrzucony i nie zostanie przeanalizowany.

Jeśli typ dokumentu zostanie zaakceptowany, zostanie użyta wartość *typ\_dokumentu*, która została określona w zastosowanej regule. Ta wartość jest zwracana do analizatora składni kolekcji wraz z przeanalizowaną treścią.

### **Procedura**

Aby powiązać typy dokumentów z analizatorem składni Stellent, należy wykonać poniższe czynności:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Zmień zawartość pliku `ES_NODE_ROOT/master_config/id_kolekcji.stellent/stellent.properties`, gdzie *id\_kolekcji* identyfikuje kolekcję, która ma zostać skonfigurowana.
3. We właściwości `OutsideInSupportedTypes` podaj bezwzględną ścieżkę do tworzonego pliku konfiguracyjnego.

Na przykład można utworzyć następujący plik konfiguracyjny dla pojedynczej kolekcji i zapisać go wraz z innymi plikami danej kolekcji:

```
ES_NODE_ROOT/master_config/id_kolekcji.stellent/stellenttypes.cfg
```

Innym przykładem może być utworzenie następującego pliku konfiguracyjnego w celu wykorzystania tych samych ustawień dla wszystkich kolekcji oraz zapisanie go z innymi plikami na poziomie systemu. W przypadku wykorzystania tego podejścia należy się upewnić, że ta ścieżka została podana w pliku `stellent.properties` dla każdej kolekcji, co zostało określone w kroku 2 na stronie 162.

```
ES_INSTALL_ROOT/default_config/stellent/stellenttypes.cfg
```

4. Za pomocą edytora tekstu utwórz plik konfiguracyjny i określ reguły analizy Stellent, a następnie zapisz i zamknij plik.
5. Aby zmiany zostały uwzględnione, użyj konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania analizatora składni kolekcji, a następnie zatrzymaj i ponownie uruchom analizator składni.

### Przykłady

Według następującego pliku konfiguracyjnego sesja Stellent, oprócz dokumentów z domyślnej listy obsługiwanych typów dokumentów, akceptuje dokumenty w formacie Microsoft Visio.

```
accept DEFAULT
accept FI_VISI03 visio
accept FI_VISI04 visio
accept FI_VISI05 visio
accept FI_VISI06 visio
```

Według następującego pliku konfiguracyjnego: dokumenty Postscript są akceptowane i podlegają wyszukiwaniu z zastosowaniem typu dokumentu `ps`; dokumenty w formacie X pixmap (XPM) są odsyłane do wbudowanego tekstowego analizatora składni; dokumenty w formacie obrazu PNG są odrzucane; wszystkie pozostałe typy plików są akceptowane i udostępniane do wyszukiwania z zastosowaniem typu dokumentu `other`.

```
accept DEFAULT
accept FI_POSTSCRIPT ps
native FI_XPIXMAP txt
reject FI_PNG
accept ALL other
```

#### Zadania pokrewne

“Powiązanie typów dokumentów z analizatorem składni kolekcji” na stronie 155

#### Odsyłacze pokrewne

“Domyślne reguły analizowania przez analizatory składni Stellent”

## Domyślne reguły analizowania przez analizatory składni Stellent

Jeśli nie zostanie utworzony plik konfiguracyjny odwzorowujący typy plików na filtry dokumentów Stellent, to analizator składni będzie analizować dokumenty według reguł domyślnych.

Analizator składni Stellent akceptuje oraz analizuje dokumenty następujących typów:

```
ACCEPT FI_123R1 123
ACCEPT FI_123R2 123
ACCEPT FI_123R3 123
ACCEPT FI_123R4 123
ACCEPT FI_123R6 123
ACCEPT FI_123R9 123
ACCEPT FI_EXCEL xls
ACCEPT FI_EXCEL2000 xls
ACCEPT FI_EXCEL2002 xls
```



ACCEPT FI_EXCEL2003	xls
ACCEPT FI_EXCEL2007	xlsx
ACCEPT FI_EXCEL3	xls
ACCEPT FI_EXCEL4	xls
ACCEPT FI_EXCEL5	xls
ACCEPT FI_EXCEL97	xls
ACCEPT FI_EXTPOWERPOINT4	ppt
ACCEPT FI_EXTPOWERPOINTMAC4	ppt
ACCEPT FI_FREELANCE	prz
ACCEPT FI_FREELANCE3	prz
ACCEPT FI_ICHITAR03	jxw
ACCEPT FI_ICHITAR04	jsw
ACCEPT FI_ICHITAR08	jtd
ACCEPT FI_PDF	pdf
ACCEPT FI_PDFMACBIN	pdf
ACCEPT FI_POWERPOINT2	ppt
ACCEPT FI_POWERPOINT2000	ppt
ACCEPT FI_POWERPOINT2007	pptx
ACCEPT FI_POWERPOINT3	ppt
ACCEPT FI_POWERPOINT4	ppt
ACCEPT FI_POWERPOINT7	ppt
ACCEPT FI_POWERPOINT9597	ppt
ACCEPT FI_POWERPOINT97	ppt
ACCEPT FI_POWERPOINTMAC3	ppt
ACCEPT FI_POWERPOINTMAC4	ppt
ACCEPT FI_POWERPOINTMACB3	ppt
ACCEPT FI_POWERPOINTMACB4	ppt
ACCEPT FI_RTF	rtf
ACCEPT FI_RTFJ	rtf
ACCEPT FI_STAROFFICEWRITER8	odt
ACCEPT FI_STAROFFICEDRAW8	odg
ACCEPT FI_STAROFFICEIMPRESS8	odp
ACCEPT FI_STAROFFICECALC8	ods
ACCEPT FI_STAROFFICECALC6	sxc
ACCEPT FI_STAROFFICEDRAW6	sxd
ACCEPT FI_STAROFFICEIMPRESS6	sxi
ACCEPT FI_STAROFFICEWRITER6	sxw
ACCEPT FI_STAROFFICECALC52	sdc
ACCEPT FI_STAROFFICEIMPRESS52	sdd
ACCEPT FI_STAROFFICEWRITER52	sdw
ACCEPT FI_VISIO3	vsd
ACCEPT FI_VISIO4	vsd
ACCEPT FI_VISIO5	vsd
ACCEPT FI_VISIO6	vsd
ACCEPT FI_VISIO2003	vsd
ACCEPT FI_WINWORD1	doc
ACCEPT FI_WINWORD1COMPLEX	doc
ACCEPT FI_WINWORD1J	doc
ACCEPT FI_WINWORD2	doc
ACCEPT FI_WINWORD2000	doc
ACCEPT FI_WINWORD2002	doc
ACCEPT FI_WINWORD2003	doc
ACCEPT FI_WINWORD2007	docx
ACCEPT FI_WINWORD5J	doc
ACCEPT FI_WINWORD6	doc
ACCEPT FI_WINWORD7	doc
ACCEPT FI_WINWORD97	doc
ACCEPT FI_WORD4	doc
ACCEPT FI_WORD5	doc
ACCEPT FI_WORD6	doc
ACCEPT FI_WORDPRO	lwp
ACCEPT FI_WORDPRO97	lwp

Analizator składni Stellent zwraca dokumenty następujących typów do analizatora składni kolekcji w celu przetworzenia ich przez jeden z wbudowanych analizatorów składni:

NATIVE FI_7BITTEXT	txt
NATIVE FI_ANSI	txt
NATIVE FI_ANSI8	txt
NATIVE FI_ARABIC_710	txt
NATIVE FI_ARABIC_720	txt
NATIVE FI_ARABIC_WINDOWS	txt
NATIVE FI_ASCII	txt
NATIVE FI_ASCII8	txt
NATIVE FI_CENTRALEU_1250	txt
NATIVE FI_CHINESEBIG5	txt
NATIVE FI_CHINESEGB	txt
NATIVE FI_CYRILLIC1251	txt
NATIVE FI_CYRILLICKO18	txt
NATIVE FI_EBCDIC_1026	txt
NATIVE FI_EBCDIC_273	txt
NATIVE FI_EBCDIC_277	txt
NATIVE FI_EBCDIC_278	txt
NATIVE FI_EBCDIC_280	txt
NATIVE FI_EBCDIC_284	txt
NATIVE FI_EBCDIC_285	txt
NATIVE FI_EBCDIC_297	txt
NATIVE FI_EBCDIC_37	txt
NATIVE FI_EBCDIC_500	txt
NATIVE FI_EBCDIC_870	txt
NATIVE FI_EBCDIC_871	txt
NATIVE FI_HANGEUL	txt
NATIVE FI_HEBREW_E0	txt
NATIVE FI_HEBREW_OLDCODE	txt
NATIVE FI_HEBREW_PC8	txt
NATIVE FI_HEBREW_WINDOWS	txt
NATIVE FI_HTML	htm
NATIVE FI_HTML_ARABIC_ASMO708	htm
NATIVE FI_HTML_ARABIC_DOS	htm
NATIVE FI_HTML_ARABIC_ISO	htm
NATIVE FI_HTML_ARABIC_MAC	htm
NATIVE FI_HTML_ARABIC_WINDOWS	htm
NATIVE FI_HTML_BALTIC_ISO	htm
NATIVE FI_HTML_BALTIC_WINDOWS	htm
NATIVE FI_HTML_CENTRALEUROPEAN_DOS	htm
NATIVE FI_HTML_CENTRALEUROPEAN_ISO	htm
NATIVE FI_HTML_CENTRALEUROPEAN_MAC	htm
NATIVE FI_HTML_CENTRALEUROPEAN_WINDOWS	htm
NATIVE FI_HTML_CHINESEBIG5	htm
NATIVE FI_HTML_CHINESEEUC	htm
NATIVE FI_HTML_CHINESEGB	htm
NATIVE FI_HTML_CHINESESIMPLIFIED_EUC	htm
NATIVE FI_HTML_CHINESESIMPLIFIED_WINDOWS	htm
NATIVE FI_HTML_CHINESETRADITIONAL_WINDOWS	htm
NATIVE FI_HTML_CYRILLIC_DOS	htm
NATIVE FI_HTML_CYRILLIC_ISO	htm
NATIVE FI_HTML_CYRILLIC_KO18R	htm
NATIVE FI_HTML_CYRILLIC_MAC	htm
NATIVE FI_HTML_CYRILLIC_WINDOWS	htm
NATIVE FI_HTML_CYRILLIC1251	htm
NATIVE FI_HTML_CYRILLICKO18	htm
NATIVE FI_HTML_EBCDIC_1026	htm
NATIVE FI_HTML_EBCDIC_273	htm
NATIVE FI_HTML_EBCDIC_277	htm
NATIVE FI_HTML_EBCDIC_278	htm
NATIVE FI_HTML_EBCDIC_280	htm
NATIVE FI_HTML_EBCDIC_284	htm
NATIVE FI_HTML_EBCDIC_285	htm
NATIVE FI_HTML_EBCDIC_297	htm
NATIVE FI_HTML_EBCDIC_37	htm
NATIVE FI_HTML_EBCDIC_500	htm
NATIVE FI_HTML_EBCDIC_870	htm
NATIVE FI_HTML_EBCDIC_871	htm

NATIVE FI_HTML_GREEK_ISO	htm
NATIVE FI_HTML_GREEK_MAC	htm
NATIVE FI_HTML_GREEK_WINDOWS	htm
NATIVE FI_HTML_HEBREW_DOS	htm
NATIVE FI_HTML_HEBREW_ISO_VISUAL	htm
NATIVE FI_HTML_HEBREW_WINDOWS	htm
NATIVE FI_HTML_JAPANESE_MAC	htm
NATIVE FI_HTML_JAPANESE_SHIFTJIS	htm
NATIVE FI_HTML_JAPANESE_EUC	htm
NATIVE FI_HTML_JAPANESE_JIS	htm
NATIVE FI_HTML_JAPANESE_SJIS	htm
NATIVE FI_HTML_KOREAN_JOHAB	htm
NATIVE FI_HTML_KOREAN_WINDOWS	htm
NATIVE FI_HTML_KOREAN_HANGUL	htm
NATIVE FI_HTML_LATIN2	htm
NATIVE FI_HTML_RUSSIAN_DOS	htm
NATIVE FI_HTML_THAI_WINDOWS	htm
NATIVE FI_HTML_TURKISH_DOS	htm
NATIVE FI_HTML_TURKISH_ISO	htm
NATIVE FI_HTML_TURKISH_MAC	htm
NATIVE FI_HTML_TURKISH_WINDOWS	htm
NATIVE FI_HTML_VIETNAMESE_WINDOWS	htm
NATIVE FI_HTML_WESTERNEUROPEAN_ISO	htm
NATIVE FI_HTML_WESTERNEUROPEAN_MAC	htm
NATIVE FI_HTML_WESTERNEUROPEAN_WINDOWS	htm
NATIVE FI_HTML_UNICODE	htm
NATIVE FI_JAPANESE_EUC	txt
NATIVE FI_JAPANESE_JIS	txt
NATIVE FI_LATIN2	txt
NATIVE FI_MAC	txt
NATIVE FI_MAC8	txt
NATIVE FI_PP2KHTML	htm
NATIVE FI_SHIFTJIS	txt
NATIVE FI_UNICODE	txt
NATIVE FI_UTF8	txt
NATIVE FI_W2KHTML	htm
NATIVE FI_WML	xml
NATIVE FI_WML_CHINESEBIG5	xml
NATIVE FI_WML_CHINESE_EUC	xml
NATIVE FI_WML_CHINESEGB	xml
NATIVE FI_WML_CYRILLIC1251	xml
NATIVE FI_WML_CYRILLIC018	xml
NATIVE FI_WML_JAPANESE_EUC	xml
NATIVE FI_WML_JAPANESE_JIS	xml
NATIVE FI_WML_JAPANESE_SJIS	xml
NATIVE FI_WML_KOREAN_HANGUL	xml
NATIVE FI_WML_LATIN2	xml
NATIVE FI_XHTML	htm
NATIVE FI_XL2KHTML	htm
NATIVE FI_XML	xml
NATIVE FI_XML_DOCTYPE_HTML	htm

### Zadania pokrewne

“Powiązanie typów dokumentów z analizatorem składni Stellant” na stronie 160

---

## Obsługa języków i stron kodowych

Przetwarzanie lingwistyczne w wyszukiwaniu korporacyjnym jest obsługiwane inaczej przez analizatory składni i serwery wyszukiwania.

Na potrzeby przetwarzania lingwistycznego analizator składni nie rozróżnia poszczególnych języków i ustawień narodowych. Jeśli jednak użytkownicy wyszukują w kolekcji zawierającej dokumenty w różnych językach, to serwery wyszukiwania pozwalają na ograniczenie wyników wyszukiwania do określonego języka lub ustawienia narodowego.

Na przykład, jeśli metadane dokumentu w języku angielskim definiują ustawienia narodowe dokumentu jako `en_US`, to taki dokument indeksowany jest zarówno jako dokument w języku angielskim (`en`), jak również jako dokument zapisany w amerykańskiej jego odmianie (`en_US`). Ten rodzaj indeksowania pozwala na prawidłowe wyświetlanie informacji istotnych dla konkretnego języka, takich jak liczby, daty i czas. Gdy użytkownicy wyszukują w kolekcji, to dokument może zostać znaleziony niezależnie od tego, czy szukane są dokumenty z ustawieniem `en`, czy ustawieniem `en_US`.

Jeśli dokument indeksowany jest tylko według kodu języka, takiego jak `en`, to indeksowany jest on wyłącznie według kodu języka, a nie według kodu ustawień narodowych. Jeśli użytkownicy wyszukiwać będą w kolekcji dokumentów z ustawieniem `en_US`, to taki dokument nie zostanie znaleziony.

System wyszukiwania korporacyjnego udostępnia obsługę lingwistyczną poniższych języków oraz języków kodowanych dwuznakowo, zgodnie ze standardem ISO 639:

### Języki ze zwykłym tekstem:

- en=angielski
- sq=albański
- az=azerski (łacina)
- bg=bułgarski
- be=białoruski
- ca=kataloński
- hr=chorwacki
- cs=czeski
- da=duński
- nl=holenderski
- et=estoński
- fi=fiński
- fr=francuski
- de=niemiecki
- el=grecki
- hu=węgierski
- is=islandzki
- id=indonezyjski
- in=indonezyjski
- it=włoski
- kk=kazachski
- lv=łotewski
- lt=litewski
- lo=laotański
- mk=macedoński
- ms=malajski
- mt=maltański
- no=norweski
- nb=norweski (bokmal)
- pl=polski
- pt=portugalski
- ro=rumuński

ru=rosyjski  
sr=serbski (cyrylica)  
sh=serbski (łacina)  
sk=słowacki  
sl=słoweński  
es=hiszpański  
sv=szwedzki  
tr=turecki  
uk=ukraiński  
cy=walijski

#### **Języki ideograficzne:**

W przypadku uproszczonego i tradycyjnego języka chińskiego stosowane są rozszerzone kody języków, a nie kody dwuznakowe.

zh-CN=chiński (uproszczony)  
zh-TW=chiński (tradycyjny)  
ja=japoński  
ko=koreański

#### **Języki z tekstem złożonym:**

ar=arabski  
as=assamski  
bn=bengalski  
gu=gujarati  
iw=hebrajski  
he=hebrajski  
hi=hindi  
kn=kannada  
ml=malajalam  
mr=marathi  
or=oriya  
pa=pendźabski  
ta=tamiński  
te=telugu  
th=tajski  
ur=urdu  
vi=wietnamski

System wyszukiwania korporacyjnego może automatycznie wykryć dowolny z tych języków, a także stronę kodową stosowaną w dokumentach zapisanych zwykłym tekstem. Jeśli konieczne jest jawne podanie stosowanego języka lub strony kodowej, to w czasie konfigurowania przeszukiwacza można wyłączyć automatyczne wykrywanie języka i strony kodowej.

---

## **Automatyczne rozpoznawanie języka**

System wyszukiwania korporacyjnego może przetwarzać dokumenty zapisane w praktycznie dowolnym języku.

Jeśli dokument jest zapisany w jednym z poniższych języków, to system będzie w stanie automatycznie rozpoznać ten język. Jeśli język dokumentów jest z góry znany, w czasie konfigurowania przeszukiwacza można określić język dokumentów i wyłączyć tym samym automatyczne rozpoznawanie języka.

arabski  
bułgarski  
czeski  
chiński uproszczony  
chiński tradycyjny  
duński  
holenderski  
angielski  
fiński  
francuski, Kanada

francuski, Francja  
niemiecki, Niemcy  
niemiecki, Szwajcaria  
grecki  
hebrajski  
węgierski  
islandzki  
włoski  
japoński  
koreański  
norweski, bokmal  
polski  
portugalski, Brazylia  
portugalski, Portugalia  
rumuński  
rosyjski  
hiszpański  
szwedzki  
tajski  
turecki

---

## Automatyczne wykrywanie strony kodowej

System wyszukiwania korporacyjnego obsługuje dokumenty zapisane z wykorzystaniem wielu różnych stron kodowych.

W przypadku plików tekstowych system może wykrywać poniższe strony kodowe automatycznie. W przypadku innych formatów dokumentów do wykrywania strony kodowej wykorzystywane są zapisane w dokumencie metadane, takie jak elementy metadanych HTML. Jeśli strona kodowa dokumentów jest z góry znana, to w czasie konfigurowania przeszukiwacza można ją podać, co spowoduje wyłączenie mechanizmu automatycznego wykrywania jej w dokumentach.

### Formaty kodowania Unicode:

UTF-8  
UTF-16BE  
UTF-16LE

### Formaty kodowania wielobajtowego:

Shift-JIS  
ISO-2022-CN  
ISO-2022-JP  
ISO-2022-KR  
GB18030  
EUC-JP  
EUC-KR

### Formaty kodowania jednobajtowego:

ISO-8859-1: duński, holenderski, niemiecki, angielski, francuski, włoski, norweski, portugalski, hiszpański, szwedzki  
ISO-8859-2: czeski, węgierski, polski, rumuński  
ISO-8859-5: rosyjski  
ISO-8859-6: arabski  
ISO-8859-7: grecki  
ISO-8859-8: hebrajski, hebrajski w kolejności widocznej  
ISO-8859-9: turecki  
Windows-1250: czeski, węgierski, polski, rumuński  
Windows-1251: rosyjski  
Windows-1252: duński, holenderski, niemiecki, angielski, francuski, włoski, norweski, portugalski, hiszpański, szwedzki  
Windows-1253: grecki  
Windows-1254: turecki  
Windows-1255: hebrajski  
Windows-1256: arabski  
KOI8-R: rosyjski



Wykrywanie zestawu znaków jest bardzo mało dokładną operacją. Proces wykrywania strony kodowej próbuje określić zestaw znaków najlepiej pasujący do charakterystyk danych bajtowych, ale jest to proces całkowicie statystyczny i nie ma żadnej gwarancji, że jego wyniki będą poprawne.

W celu osiągnięcia lepszych wyników dane wejściowe powinny być zapisane tylko w jednym języku. Poza tym potrzebnych jest przynajmniej kilkaset bajtów zwykłego tekstu w danym języku.

Jeśli system nie znajdzie zgodności między wykrytym kodowaniem, a kodowaniami obsługiwanymi, to stosować będzie domyślną stronę kodową dokumentów danej kolekcji.

---

## Analiza lingwistyczna dokumentów chińskich, japońskich i koreańskich

Aby zwiększyć wydajność pobierania dokumentów w języku chińskim, japońskim i koreańskim, można określić opcje analizy lingwistycznej.

Dla dokumentów w języku chińskim, japońskim i koreańskim można skonfigurować analizator składni, aby używał do analizy leksykalnej podziału na ciągi n-znakowe. Dla dokumentów w języku chińskim i japońskim można ponadto skonfigurować analizator składni w taki sposób, aby ze znaków niedrukowanych usuwał znaki nowego wiersza.

### Podział na ciągi n-znakowe

Podczas tworzenia kolekcji wybierany jest typ analizy leksykalnej, która ma być używana do analizowania składni dokumentów napisanych w językach, w których nie stosuje się białych znaków do oddzielania słów.

Podział według znaków spacji w kodzie Unicode wykorzystuje białe znaki jako separatory wyrazów. Według reguł podziału na ciągi n-znakowe nakładające się sekwencje dowolnej liczby znaków są uznawane za pojedynczy wyraz. W przypadku takich języków, jak japoński, chiński i koreański, w których nie używa się białych znaków jako znaków oddzielających wyrazy, podział na ciągi n-znakowe może dać lepsze wyniki niż podział według białych znaków w kodzie Unicode.

Metodę podziału, która ma zostać użyta do analizowania składni dokumentów, wybiera się podczas tworzenia kolekcji. Po utworzeniu kolekcji można przeglądać te ustawienia wśród opcji analizowania składni, ale nie można ich zmienić.

Informacje na temat sposobu konfigurowania obsługi pełnej analizy n-znakowej i podziału na leksemy w kolekcjach wyszukiwania korporacyjnego, a także informacje na temat sposobu traktowania znaków w kolekcjach skonfigurowanych pod kątem pełnej obsługi n-znakowej można znaleźć w dokumencie <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088>.

### Usuwanie znaków spacji, tabulacji lub nowego wiersza z tekstu

W konfiguracji analizatora składni można zawrzeć opcję powodującą usuwanie z tekstu znaków spacji, tabulacji i nowego wiersza.

#### Zagadnienia wstępne

Aby wykonać to zadanie, należy się zalogować jako administrator wyszukiwania korporacyjnego.

#### Informacje o tym zadaniu

Po włączeniu tej opcji dla kolekcji analizator składni usuwa następujące po sobie znaki spacji, tabulacji i nowego wiersza znajdujące się między literami. Usuwanie tych znaków przydaje się na przykład w przypadku dokumentów w językach, w których do oddzielania wyrazów nie stosuje się tych znaków, na przykład chiński lub japoński.

Konfigurując analizator, można określić, czy usuwanie znaków spacji, tabulacji lub nowego wiersza ma następować tylko między znakami z zestawu dwubajowego (DBCS), czy też zawsze, bez względu na kontekst znaków. Ta druga opcja przydaje się na przykład w tekstach, w których język japoński sąsiaduje z angielskim, a znaki spacji, tabulacji i nowego wiersza mają być usunięte również z fragmentów angielskich.

Analizator składni usuwa następujące znaki:

- Tabulacja (0x09)
- LF lub znak nowego wiersza (0x0A)
- Znak CR lub znak powrotu karetki (0x0D)

Aby zmiany zostały uwzględnione, należy zatrzymać i restartować analizator składni. Aby zastosować zmianę wobec dokumentów już przechowywanych w indeksie, należy powtórnie wykonać przeszukanie dokumentów i ponownie zbudować indeks główny.

### Procedura

Aby usunąć znaki spacji, tabulacji i nowego wiersza z tekstu:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Za pomocą edytora tekstu zmień następujący plik, gdzie *id\_kolekcji* jest identyfikatorem, który został określony (lub przypisany przez system) podczas tworzenia kolekcji:  
ES\_NODE\_ROOT/master\_config/*id\_kolekcji*.parserdriver/collection.properties
3. Określ sposób usuwania znaków spacji, tabulacji i nowego wiersza:
  - Aby usuwać te znaki tylko między znakami dwubajowymi, ustaw wartość właściwości `removeCjNewlineChars` na `true`:  
`removeCjNewlineChars=true`
  - Aby usuwać te znaki w dowolnym miejscu dokumentu, ustaw wartość właściwości `removeCjNewlineChars` na `true` oraz wartość właściwości `removeCjNewlineCharsMode` na `all`:

```
removeCjNewlineChars=true
removeCjNewlineCharsMode=all
```



---

## Administrowanie indeksem

Aby upewnić się, że użytkownicy zawsze mają dostęp do najnowszych informacji, wyszukiwanie korporacyjne tworzy indeks każdej kolekcji i obsługuje go, okresowo aktualizując treść.

Aby udostępnić zgromadzone przez przeszukiwacze dane dla operacji przeszukiwania, należy utworzyć indeksy. Podczas tworzenia kolekcji w wyszukiwaniu korporacyjnym tworzony jest indeks wszystkich danych, które zostały wstępnie przeszukane. Gdy przeszukiwacze przeszukują nowe i zmodyfikowane źródła danych, tworzona jest aktualizacja (tak zwany *indeks przyrostowy*) dla nowej treści. Na koniec aktualizacje należy scalić z indeksem podstawowym. Ten proces jest określany mianem budowania *indeksu głównego*. Za każdym razem, gdy budowany jest indeks przyrostowy lub indeks główny, nowa treść jest kopiowana na serwery wyszukiwania i udostępniana do celów wyszukiwania.

Przeszukiwacze zbierają dane ciągle lub regularnie według harmonogramu. W przypadku częstego aktualizowania indeksów umożliwia się użytkownikom wyszukiwanie najbardziej aktualnych danych. Wciąż aktualizowany indeks należy w końcu odbudować. W miarę rozrastania się indeksu zużywa on więcej zasobów systemowych. Aby zachować optymalną wydajność, należy regularnie budować indeks główny.

O częstotliwości budowania indeksu głównego decydują następujące czynniki:

- zasoby systemu (obszar systemu plików, szybkość procesora, pamięć);
- liczba dokumentów przeznaczonych do przeszukania i ponownego przeszukania;
- rodzaj danych do przeszukiwania;
- częstotliwość zmieniania reguł kategorii (zmiany nie zostają uwzględnione do momentu odbudowania indeksu głównego);
- częstotliwość wymuszonych uruchomień przeszukiwacza (w przeciwieństwie do uruchomień zaplanowanych w harmonogramie);

W przypadku kolekcji zawierających kilka milionów dokumentów, w większości dokumentów WWW, indeks główny należy budować średnio raz dziennie, a aktualizować co godzinę lub dwie.

Aby zapewnić użytkownikom dostęp do indeksu, który jest aktualny i przygotowany dla operacji przeszukiwania, należy wykonać następujące zadania:

- Określenie harmonogramów budowania indeksu.
- Zmodyfikowanie harmonogramu operacji na indeksie.
- Włączenie i wyłączenie harmonogramu operacji na indeksie.
- Skonfigurowanie współbieżnych operacji tworzenia indeksów.

Aby określić istotne dla użytkowników opcje mające wpływ na indeks, można dodatkowo wykonać następujące zadania:

- Skonfigurowanie obsługi znaków wieloznacznych w zapytaniach.
- Skonfigurowanie zasięgów w celu ograniczenia zakresu dokumentów dostępnych do przeglądania przez użytkowników.
- Zwijanie w wynikach wyszukiwania dokumentów pochodzących z tego samego źródła.
- Usunięcie identyfikatorów URI z indeksu.

### Zadania pokrewne

“Monitorowanie aktywności indeksu dla kolekcji” na stronie 315

“Monitorowanie kolejki indeksu wyszukiwania korporacyjnego” na stronie 315

---

## Planowanie operacji budowania indeksu

Operacje budowania indeksu głównego oraz aktualizowania indeksu nowymi danymi mogą być przeprowadzane według harmonogramu definiowanego przez użytkownika.

### Zagadnienia wstępne

Do planowania operacji budowania indeksu wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla danej kolekcji.

### Informacje o tym zadaniu

Aby zapewnić użytkownikom stały dostęp do najbardziej aktualnej postaci danych w przeszukiwanych źródłach, warto zaplanować harmonogram regularnego budowania indeksu. Podczas budowania indeksu głównego odtwarzany jest cały indeks. Procesy indeksowania odczytują wszystkie dane zebrane przez przeszukiwacze i przeanalizowane przez analizator składni. Podczas budowania indeksu przyrostowego do wyszukiwania zostają udostępnione informacje, które przeszukano od czasu ostatniego budowania indeksu głównego.

Domyślnie włączona jest opcja planowania operacji budowania indeksu. Opcja ta nakazuje zaplanowanie zadań budowy indeksu głównego i przyrostowego po uruchomieniu systemu wyszukiwania korporacyjnego. Aby zapobiec uruchamianiu zaplanowanych operacji budowania indeksu, w dowolnym momencie można usunąć zaznaczenie opcji **Włączanie podczas uruchamiania systemu**. Na przykład konieczne może być wyłączenie harmonogramu w celu rozwiązania problemów.

Dla zaoszczędzenia zasobów systemowych i dla poprawy wydajności system automatycznie sprawdza, czy wystąpiły zmiany, które powinny zostać zastosowane w indeksie. Jeśli żadne zmiany nie muszą być zastosowane w indeksie, zaplanowane żądanie budowania jest odrzucane.

### Procedura

Aby utworzyć harmonogram operacji budowania indeksu, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Tworzenie harmonogramu budowania indeksów**.
2. Aby określić, jak często indeks ma być aktualizowany za pomocą nowych danych, określ następujące opcje na stronie Harmonogram budowania indeksu w obszarze **Określ harmonogram budowania indeksu przyrostowego**:
  - a. W obszarze **Początek** określ termin pierwszej operacji budowania indeksu przyrostowego za pomocą pól **Rok**, **Miesiąc**, **Dzień**, **Godzina** i **Minuta**.
  - b. W obszarze **Odstęp czasu aktualizacji** określ częstotliwość budowania indeksu przyrostowego za pomocą pól **dni**, **godzin** i **minut**.

Zazwyczaj indeksy przyrostowe powinny być budowane dosyć często, na przykład co jedną lub dwie godziny. W zależności od tego, jak często zmienia się treść źródłowa, należy podać większy lub mniejszy odstęp czasu. Na przykład może to być proces wykonywany co godzinę (0 dni 1 godzina) albo co 12 godzin (0 dni i 12 godzin).
3. Aby określić częstotliwość całkowitej odbudowy indeksu, określ następujące opcje w obszarze **Określ harmonogram budowania indeksu głównego**:

- a. W obszarze **Początek** określ termin pierwszej operacji budowania indeksu głównego za pomocą pól **Rok, Miesiąc, Dzień, Godzina i Minuta**.
  - b. W obszarze **Odstęp czasu aktualizacji** określ częstotliwość budowania indeksu głównego za pomocą pól **dni, godzin i minut**.  
Zazwyczaj pełne budowanie indeksu należy wykonywać regularnie, np. co 24 godziny. W zależności od tego, jak często zmienia się treść źródłowa, należy podać większy lub mniejszy odstęp czasu. Na przykład może to być proces wykonywany co 12 godzin (0 dni 12 godzin) albo co dwie i pół doby (2 dni i 12 godzin).
4. Kliknij przycisk **OK**.

## Zmiana harmonogramu operacji na indeksie

Można zmienić harmonogram budowania indeksu.

### Zagadnienia wstępne

Aby użytkownik mógł zmodyfikować harmonogram operacji na indeksie, musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub administratora określonej kolekcji.

### Procedura

Aby zmienić harmonogram operacji na indeksie:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i zmień odpowiednie wartości w polach **Miesiąc, Dzień, Rok i Godzina**. Określ, jak często indeks ma być aktualizowany oraz jak często ma być budowany.
2. Kliknij przycisk **Zastosuj**.

## Włączanie i wyłączanie harmonogramów operacji na indeksach

Można włączać i wyłączać harmonogramy budowania indeksu.

### Zagadnienia wstępne

Aby użytkownik mógł włączyć lub wyłączyć harmonogram operacji na indeksie, musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub administratora określonej kolekcji.




### Informacje o tym zadaniu

Istnieje możliwość wyłączenia harmonogramu w celu uniknięcia zaplanowanego procesu tworzenia indeksu. Na przykład może zaistnieć potrzeba wyłączenia harmonogramu w celu uniknięcia utworzenia indeksu w zaplanowanym czasie w celu rozwiązania zaistniałych problemów.

Użytkownik może włączyć lub wyłączyć harmonogram podczas edytowania lub monitorowania kolekcji.

### Procedura

1. Aby podczas edytowania kolekcji włączyć lub wyłączyć harmonogram operacji na indeksie, wykonaj następujące kroki:
  - a. Rozpocznij edycję kolekcji, w której ma nastąpić zmiana.
  - b. Na stronie Indeks zaznacz pole wyboru **Włącz podczas uruchamiania systemu** lub usuń jego zaznaczenie, aby włączyć lub wyłączyć harmonogram aktualizacji indeksu.

- c. Zaznacz pole wyboru **Włącz podczas uruchamiania systemu** lub usuń jego zaznaczenie, aby włączyć lub wyłączyć harmonogram budowania indeksu głównego.
  - d. Kliknij przycisk **Zastosuj**.
2. Aby podczas monitorowania kolekcji włączyć lub wyłączyć harmonogram operacji na indeksie, wykonaj następujące kroki:
  - a. Rozpocznij monitorowanie kolekcji, w której ma nastąpić zmiana.
  - b. Jeśli indeks jest ujęty w harmonogramie, lecz nie ma być budowany w zaplanowanym terminie, na stronie Indeks kliknij ikonę  **Wyłącz harmonogram**. Indeks nie będzie budowany, dopóki harmonogram nie zostanie włączony lub nie zostanie kliknięta ikona  **Uruchom** w celu uruchomienia procesu budowania indeksu.
  - c. Jeśli indeks jest powiązany z harmonogramem, ale harmonogram jego budowania jest wyłączony, kliknij ikonę  **Włącz harmonogram**.

Proces tworzenia indeksu zostanie umieszczony w kolejce z datą i godziną określoną w harmonogramie operacji na indeksie.

## Konfiguracja współbieżnych operacji budowania indeksu

Użytkownik steruje korzystaniem z zasobów indeksowania, określając liczbę kolekcji, które mogą jednocześnie przetwarzać własne żądania budowy indeksu. Jeśli dostępne są wystarczające zasoby systemowe, można poprawić jakość wyszukiwania, określając, że aktualizacje indeksu mają następować w tym samym czasie, co budowanie indeksu głównego.

### Zagadnienia wstępne

Aby określić opcje budowania indeksu dla systemu, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Wyszukiwanie korporacyjne umożliwia współbieżne budowanie wielu indeksów dzięki współużytkowaniu zasobów przez kolekcje, co umożliwia równoległe przetwarzanie żądań budowy indeksu w wielu kolekcjach. Dzięki współużytkowaniu procesów można zagwarantować, że budowanie bardzo dużego indeksu nie zablokuje dostępności innych indeksów, które oczekują na budowanie w kolejce.

Gdy operacja budowania indeksu zostanie zainicjowana na żądanie lub zgodnie z harmonogramem, indeks zostaje umieszczony w kolejce indeksów, gdzie oczekuje na przetworzenie. Ponieważ każda kolekcja ma własny indeks, w kolejce indeksów jednocześnie może być kilka żądań budowy indeksu z różnych kolekcji. Podczas konfigurowania w systemie opcji indeksowania można określić liczbę kolekcji, które mogą równoległe współużytkować zasoby indeksowania i przetwarzać żądania.

Można także określić, że żądania aktualizacji indeksu mają być przetwarzane równocześnie z budową głównego indeksu dla kolekcji. Jeśli ta opcja zostanie włączona, serwery wyszukiwania będą odświeżane danymi o najnowszych dokumentach (za pośrednictwem indeksu przyrostowego) w trakcie bardziej czasochłonnego budowania indeksu głównego. Jednak budowanie indeksu jest procesem w znacznym stopniu obciążającym zasoby systemu. Podczas operacji budowania indeksu wykorzystywana jest duża ilość pamięci systemowej i znaczny obszar dysku. W przypadku włączenia tej opcji i posiadania niewystarczającej ilości miejsca na dysku lub pamięci ogólna wydajność systemu może ulec pogorszeniu.


Zwiększenie limitu równoległych operacji budowania indeksu nie powoduje automatycznego uruchomienia żądań budowy indeksu znajdujących się już w kolejce. Zmiana tej wartości dotyczy tylko nowych żądań budowy indeksu umieszczanych w kolejce po dokonaniu zmiany.



Zmniejszenie limitu równoległych operacji budowania indeksu nie powoduje automatycznego zatrzymania indeksów aktualnie budowanych. Zmiana odnosi skutek po zatrzymaniu aktualnego budowania indeksu, co pozwala rozpocząć budowanie indeksów z kolejki.

### Procedura

Aby określić w systemie opcje budowania indeksu, należy wykonać następujące czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Indeksowanie kliknij opcję **Konfiguruj opcje indeksu**.
4. Na stronie Opcje indeksowania na poziomie systemu wpisz liczbę kolekcji, które mogą współużytkować zasoby systemu i dla których żądania budowy indeksu mogą być przetwarzane równolegle.

Liczba kolekcji współużytkujących zasoby indeksowania nie może przekroczyć liczby kolekcji w systemie wyszukiwania korporacyjnego. Na przykład, jeśli w systemie istnieje pięć kolekcji, należy wprowadzić liczbę nie większą niż pięć.

5. Jeśli dostępna jest wystarczająca ilość zasobów systemowych do równoległego budowania wielu indeksów dla poszczególnych kolekcji, można włączyć opcję budowania indeksów przyrostowych równoległe z indeksem głównym.

**Ograniczenie:** Jeśli ta opcja zostanie wybrana w trakcie budowania indeksu dla dowolnej kolekcji, budowanie indeksu może się nie powieść. Przed włączeniem współbieżnych procesów budowania indeksu przyrostowego i głównego należy wyłączyć wszystkie zaplanowane procesy budowania indeksu i poczekać, aż współbieżnie działające procesy budowania indeksu zatrzymają się (lub zakończą je wcześniej, jeśli jest to stosowne). Po włączeniu tej opcji należy włączyć wcześniej wyłączone zaplanowane procesy budowania indeksu.

---

## Budowanie indeksów tylko w przypadku wykrycia zmian

W przypadku zaplanowanego budowania indeksów budowanie występuje tylko wtedy, gdy zostają wykryte zmiany, które muszą zostać zastosowane w indeksie. Budowanie indeksu można uruchomić za pomocą komendy **startIndexBuild** i wybrać, czy system ma sprawdzać występowanie zmian.

### Zagadnienia wstępne

Aby ręcznie określić, że budowanie indeksu ma występować tylko wtedy, gdy zostają wykryte zmiany, należy zalogować się jako administrator wyszukiwania korporacyjnego.

### Ograniczenia

System nie sprawdzi, czy zaszły zmiany, które mogą mieć wpływ na indeks, dopóki indeks główny nie zostanie zbudowany co najmniej dwukrotnie. Taka implementacja ma na celu przystosowanie do zmian, których nie można wykryć do czasu, aż indeks główny zostanie przynajmniej jeden raz zbudowany, czyli na przykład zmian w definicjach zasięgu lub regułach zwijania wyników wyszukiwania.

### Informacje o tym zadaniu

Procesy budowania indeksu zajmują zasoby systemowe, takie jak procesor, pamięć i pasmo we/wy. Dla zaoszczędzenia zasobów systemowych i dla poprawy wydajności system wyszukiwania korporacyjnego może wykrywać, czy wystąpiły zmiany, które powinny zostać

zastosowane w indeksie, i budować indeks tylko wtedy, gdy jest to konieczne dla zastosowania zmian. System może wykrywać następujące zdarzenia modyfikacji indeksu:

#### Tylko procesy budowania indeksu głównego

- Usunięcie adresów URI według wzorca, w związku z czym dokumenty zostają usunięte z indeksu.
- Usunięcie przeszukiwacza z systemu, w związku z czym dokumenty zostają usunięte z indeksu.

#### Procesy budowania indeksów głównego i przyrostowego

- Zmodyfikowanie definicji zasięgu lub reguł zwijania wyników wyszukiwania.
- Dodanie lub usunięcie dokumentów wprost poprzez ustawienia przeszukiwacza lub funkcję API typu push.
- Zmodyfikowanie współczynników kryterium w celu wpłynięcia na wyniki rankingu statycznego dokumentów w indeksie.
- Zmodyfikowanie reguł określających możliwy sposób użycia znaków wieloznacznych w zapytaniach.

Wszystkie zaplanowane żądania budowania indeksu podlegają wykrywaniu zmian. Budowanie indeksu odbywa się tylko wtedy, gdy istnieją oczekujące zmiany, które muszą być zastosowane w indeksie. Jeśli żadna zmiana nie zostanie wykryta, żądanie budowania indeksu jest odrzucane.

Aby wymusić budowanie indeksu bez względu na wystąpienie zmian, można uruchomić budowanie indeksu za pomocą konsoli administracyjnej wyszukiwania korporacyjnego. Budowanie indeksu można też uruchomić ręcznie za pomocą komendy **startIndexBuild** i określić, czy system ma sprawdzać występowanie zmian, zanim zrealizuje żądanie.

#### Procedura

Aby uruchomić budowanie indeksu z poziomu wiersza komend i zażądać, aby budowanie odbywało się tylko wtedy, gdy zostaną wykryte zmiany:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W konfiguracji z wieloma serwerami należy zalogować się na serwerze indeksowania.
2. Wprowadź następującą komendę:

```
esadmin controller startIndexBuild -opcje
```

Opcje:

#### **-cid** *id\_kolekcji*

Identyfikator kolekcji, dla której ma zostać uruchomione budowanie indeksu.

#### **-buildType** *typ\_budowy*

Określa typ indeksu do zbudowania. Poprawne wartości to **main** (główny) i **delta** (przyrostowy).

#### **-detectChanges**

Opcjonalne. Powoduje, że żądanie budowania indeksu poddawane jest testom wykrywania zmian. Budowanie indeksu odbywa się tylko wtedy, gdy zostaną wykryte zmiany, które muszą zostać zastosowane w indeksie.

W następującym przykładzie żądanie uruchomienia budowania indeksu głównego dla kolekcji **col\_1** realizowane jest tylko wtedy, gdy zostaną wykryte żadne zmiany, które muszą zostać zastosowane w indeksie:

```
esadmin controller startIndexBuild -cid col_1 -buildType main -detectChanges
```

---

## Zatrzymywanie budowania indeksu

Przy użyciu komendy **stopIndex** można zatrzymać budowanie indeksu głównego i przyrostowego nie używając konsoli administracyjnej wyszukiwania korporacyjnego.

### Procedura

Aby zatrzymać budowanie indeksu przy użyciu komendy, a nie konsoli administracyjnej:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W instalacji z wieloma serwerami, zaloguj się na serwerze indeksowania.
2. Wprowadź następującą komendę:

```
esadmin stopIndex -cid id_kolekcji -buildType typGdzie:
```

**-cid *id\_kolekcji***

Określa identyfikator kolekcji, która jest właścicielem indeksu.

***typ*** Określa typ indeksu, którego budowanie ma być zatrzymane. Dozwolone wartości to main (główny) i delta (przyrostowy).

### Przykład

```
esadmin stopIndex -cid col1 -buildType delta
```

---

## Opcje wpływające na podlegający wyszukiwaniu widok indeksu

Po poindeksowaniu dokumentów można określić opcje decydujące o sposobach wyszukiwania dokumentów i przeglądania ich w wynikach wyszukiwania przez użytkowników.

Aby określić opcje wpływające na sposób wyświetlania indeksu dla użytkownika, można wykonać następujące działania:

- Skonfigurować obsługę znaków wieloznacznych w terminach użytych w zapytaniu. W indeks można wbudować obsługę zapytań ze znakami wieloznacznymi albo można określić opcje rozwijające kryteria podczas przetwarzania zapytania.
- Skonfigurowanie zasięgów, aby ograniczyć zakres dokumentów dostępnych do przeszukiwania dla użytkownika. Przy wyszukiwaniu w kolekcji uwzględniane są tylko dokumenty znajdujące się w zdefiniowanym zasięgu, a nie wszystkie dokumenty z indeksu.
- Zwinięcie w wynikach wyszukiwania dokumentów pochodzące z jednego źródła. Dokumenty pasujące do identyfikatora URI lub jego wzorca w indeksie można grupować i w wynikach wyszukiwania wyświetlać tylko dokumenty położone najwyżej. Użytkownik w takiej sytuacji może sam zdecydować o wyświetleniu zwiniętych dokumentów wynikowych.
- Usunięcie identyfikatorów URI z indeksu. Można tymczasowo uniemożliwić użytkownikom wyszukiwanie określonych dokumentów w indeksie.

W przypadku niektórych typów przeszukiwaczy oraz w przypadku kolekcji, które nie umożliwiają obsługi zabezpieczeń, używane jest wykrywanie duplikatów dokumentów dla zapobieżenia sytuacji, w której użytkownikom wyświetlanych będzie wiele identycznych lub prawie identycznych dokumentów w wynikach wyszukiwania.

## Indeksowane opcje wyszukiwania dokumentów

Podczas konfigurowania opcji wyszukiwania w przeszukanych danych albo odwzorowywania metadanych XML i HTML na pola wyszukiwania definiuje się, w jaki sposób dokumenty mają być przeszukiwane i wyświetlane w wynikach wyszukiwania.

Podane opcje wyszukiwania przechowywane są w indeksie razem z dokumentami. Umożliwiają one ograniczenie zakresu zapytań wprowadzanych przez użytkowników oraz ograniczenie danych zwracanych w wynikach wyszukiwania.

#### **Opcje przeszukiwacza:**

W czasie konfigurowania przeszukiwacza do przeszukiwania źródeł danych zawierających pola można podać następujące opcje stanowiące o tym, czy dane pole ma być przeszukiwane, jak ma być przeszukiwane i czy może być zwracane w wynikach wyszukiwania:

- Wyszukiwanie pełnotekstowe
- Wyszukiwanie dla pól
- Pełna zgodność
- Pozwala na sortowanie
- Wyszukiwanie parametryczne
- Wyniki wyszukiwania
- Treść dokumentu

#### **Opcje odwzorowywania pól XML i HTML:**

Jeśli w czasie konfigurowania analizatora składni ustalone zostanie, że elementy XML i elementy metadanych HTML mają być odwzorowywane na pola pozwalające na wyszukiwanie, to należy podać następujące opcje:

- Wyszukiwanie dla pól
- Pełna zgodność
- Pozwala na sortowanie
- Wyniki wyszukiwania

Jeśli konfigurowane są opcje wyszukiwania wyłącznie określonych elementów metadanych HTML, a nie wszystkich elementów lub elementów znajdujących się w zbiorze metadanych Dublin Core, to można też określić, że pola zawierające wartości numeryczne mogą być przeszukiwane za pomocą zapytań sparametryzowanych.

## **Wyszukiwanie pełnotekstowe**

Indeks wyszukiwania korporacyjnego jest indeksem pełnotekstowym zawierającym treści z różnych źródeł danych. Można przeszukiwać treść przy użyciu prostego zapytania sformułowanego w języku naturalnym. Procesy wyszukiwania przeglądają pola i treść dokumentów, wyszukując dokumenty odpowiadające zapytaniu.

Aby umożliwić przeszukiwanie pól za pomocą zapytań pełnotekstowych, w czasie konfigurowania przeszukiwacza należy zaznaczyć pole wyboru **Wyszukiwanie pełnotekstowe**. Aby umożliwić wyszukiwanie w polach tytułu, słów kluczowych i opisu należy zaznaczyć wspomniane przed chwilą pole wyboru oraz pole **Wyszukiwanie w polach**.

#### **Przykład 1:**

Wyszukiwanie pełnotekstowe można zainicjować nawet tak prostym zapytaniem:

łańcuch rowerowy

W celu zaznaczenia słów, które muszą albo nie mogą wystąpić w dokumencie, można użyć specjalnych notacji. Na przykład poprzedzenie słowa znakiem plusa (+) oznacza, że dokument musi zawierać dane słowo, aby spełnić kryteria. Słowo poprzedzone znakiem

minusa (-) powoduje wykluczenie dokumentów zawierających to słowo spośród wyników wyszukiwania. Ujęcie dwóch lub więcej słów w cudzysłów ("") powoduje wyszukanie identycznej frazy.

#### **Przykład 2:**

W przypadku następującego zapytania pełnotekstowego dokument będzie zgodny tylko pod warunkiem, że będzie on zawierać frazę **fantastyka naukowa**, przy tym nie będzie zawierać słowa robot:

+**"fantastyka naukowa"** -robot

### **Wyszukiwanie dla pól**

Wyszukiwanie w polach pozwala na ograniczenie obiektu zapytania do określonych pól i pól metadanych dokumentu. Przykładem może być wyszukiwanie określonych słów tylko w tytułach dokumentów.

Aby umożliwić przeszukiwanie pól według ich nazw, należy zaznaczyć pole wyboru **Wyszukiwanie w polach** w czasie konfigurowania przeszukiwacza albo w czasie konfigurowania odwzorowań pól na elementy XML i HTML. Aby wyszukiwać w polach Title (tytuł), Keywords (słowa kluczowe) i Description (opis), należy zaznaczyć wymienione wyżej pole wyboru oraz pole wyboru **Wyszukiwanie pełnotekstowe**.

#### **Przykład:**

Aby użyć wyszukiwania dla pól w wyszukiwaniu korporacyjnym, należy podać nazwę pola oraz słowo lub frazę, które powinny się znajdować w polu w zapytaniu.

Poniższe zapytanie powoduje wyszukanie dokumentów zawierających słowo **ibm** oraz frazę **wyszukiwanie korporacyjne** w polu tytułu:

title:ibm title:"wyszukiwanie korporacyjne"

### **Pełna zgodność**

Wyszukiwanie z pełną zgodnością może poprawić jakość wyników wyszukiwania przez umożliwienie zadawania bardzo dokładnych zapytań. Wyszukiwanie z pełną zgodnością pozwala na tworzenie zapytań do pól i elementów XML oraz pobieranie tylko tych dokumentów, w których cała wartość pola lub elementu XML jest zgodna z warunkami zapytania. Jeśli wartość pola lub elementu zawiera za mało lub za dużo treści, to dopasowanie nie będzie pełne.

**Wskazówka:** W czasie sprawdzania, czy istnieje pełne dopasowanie, system przekształca wszystkie litery w treści zapytania na małe, usuwa z ciągu znaków nadmiarowe spacje i wykonuje dopasowanie wzorców ze znakami wieloznacznymi. Nie jest jednak wykonywana lematyzacja ani wyszukiwanie synonimów, nie są też usuwane słowa ignorowane. Pełne dopasowanie dla elementu XML wymaga, żeby nazwa elementu, bez elementów zagnieżdżonych, oraz cała wartość elementu dokładnie pasowała do warunków zapytania.

Do wyszukiwania pól z pełnym dopasowaniem można stosować składnię zapytań wyszukiwania korporacyjnego (SI-API) lub składnię zapytań XMLFrag2 (zapytania XPath nie są obsługiwane). Znak równości (=) poprzedzający warunki zapytania oznacza, że wykonane ma zostać wyszukiwanie z pełnym dopasowaniem.

Aby umożliwić w polach wyszukiwanie z pełnym dopasowaniem, należy zaznaczyć pole wyboru **Pełne dopasowanie** podczas konfigurowania przeszukiwacza albo podczas konfigurowania opcji odwzorowania pól na elementy XML i HTML. Jeśli podczas konfigurowania opcji analizy kolekcji użytkownicy otrzymają możliwość wyszukiwania w dokumentach XML za pomocą wyszukiwania rodzimego XML, to można przeszukiwać wszystkie elementy XML zgodnie z warunkami pełnego dopasowania.

### Przykład 1:

Pole o nazwie kolor zawiera wartość ciemny niebieski.

- Poniższe zapytanie z pełnym dopasowaniem jest zgodne, ponieważ zapytanie nie zawiera żadnych innych warunków:

```
kolor:"ciemny niebieski"
```

- Poniższe zapytanie nie spełni warunków dopasowania, ponieważ oprócz słów ciemny niebieski zawiera też słowo bardzo:

```
kolor:"bardzo ciemny niebieski"
```

- Poniższe zapytanie również nie spełni warunków dopasowania, ponieważ pole kolor zawiera też słowo ciemny:

```
kolor:"niebieski"
```

### Przykład 2:

Bez zastosowania pełnego dopasowania poniższe zapytanie XMLFrag2 może zwracać dokumenty zawierające tekst <diagnoza>rak wewnątrzprzewodowy</diagnoza> albo dowolne inne wyrażenia niekoniecznie związane z warunkami zapytania:

```
@xmlf2:.'<diagnoza>rak wewnątrzprzewodowy</diagnoza>'
```

Po zastosowaniu pełnego dopasowania poniższe zapytanie XMLFrag2 zwracać będzie wyłącznie te dokumenty, w których cała zawartość elementu XML pasować będzie do warunków zapytania:

```
@xmlf2:.'<<diagnoza>rak wewnątrzprzewodowy</diagnoza>'
```

## Pozwala na sortowanie

Jeśli źródło danych zawiera pola albo użytkownicy wyszukują w dokumentach XML i HTML, to można pozwolić na sortowanie wyników wyszukiwania według wartości zapisanych w określonym polu. W przykładowej aplikacji wyszukiwania korporacyjnego podawane są nazwy wszystkich pól skonfigurowanych jako pola pozwalające na sortowanie wyników. Użytkownicy mogą wybrać alfabetyczne sortowanie wyników (odpowiednio do sortowania łańcuchowego) przez wybranie jednego z podanych pól, zamiast sortowania wyników według istotności lub daty dokumentu. Można także wybrać, czy dokumenty mają być sortowane w kolejności rosnącej lub malejącej.

Dokumenty znajdujące się w wynikach wyszukiwania, ale niezawierające pola, według którego sortowane są wyniki, umieszczane są na końcu listy wyników. Dokumenty znajdujące się w wynikach wyszukiwania i zawierające pole, według którego sortowane są wyniki, ale poindeksowane jeszcze zanim w konfiguracji włączono możliwość sortowania według tego pola, również zostaną umieszczone na końcu wyników wyszukiwania.

Aby umożliwić użytkownikom alfabetyczne sortowanie wyników według wartości w polu, należy zaznaczyć pole wyboru **Z możliwością sortowania** podczas konfigurowania przeszukiwacza lub podczas konfigurowania opcji odwzorowywania pól dotyczących elementów XML i HTML. Jeśli pole zawiera wartości liczbowe, należy zaznaczyć pole wyboru **Wyszukiwanie parametryczne**, aby określić, że wyniki wyszukiwania mogą być

sortowane liczbowo na podstawie wartości pola.

## Wyszukiwanie parametryczne

Wyszukiwanie parametryczne jest odmianą wyszukiwania dla pól umożliwiającego użycie zapytań porównujących lub obliczających, które operują na polach liczbowych i daty oraz na metadanych. Przykładem może być wyszukanie dokumentów mających określoną wielkość lub utworzonych po określonej dacie. Można również wyszukiwać dokumenty o atrybutach większych, mniejszych lub równych określonym wartościom.

Aby wyniki wyszukiwania były sortowane liczbowo według wartości pola, należy udostępnić pole do wyszukiwania parametrycznego.

Aby umożliwić wyszukiwanie w polu za pomocą zapytania parametrycznego, a tym samym umożliwić liczbowe sortowanie wyników, należy zaznaczyć pole wyboru **Wyszukiwanie parametryczne** podczas konfigurowania przeszukiwacza lub podczas konfigurowania opcji odwzorowywania pól dotyczących konkretnych elementów metadanych HTML.

### Przykład 1:

Poniższe zapytanie pozwala wyszukać elementy kosztujące dokładnie 50 złotych (lub innej waluty przypisanej wartościom w polu **cena**):

```
#cena::=50
```

### Przykład 2:

Poniższe zapytanie pozwala wyszukać dokumenty mające rozmiar pliku powyżej 1024 i nie większy niż 2048:

```
#filesize::>1024<=2048
```

## Wyniki wyszukiwania

Można wyszukać pewne pola i nie pokazać ich w wynikach wyszukiwania. Można też pokazać w wynikach wyszukiwania pole, mimo że nie było ono uwzględnione w zapytaniu. Użytkownik może na przykład wprowadzić zapytanie dotyczące danych finansowych w celu utworzenia odpowiedniego raportu, ale może jednocześnie wyświetlić pensje pracowników w wynikach, które zawierają również nazwiska pracowników.

Aby umożliwić wyświetlanie pola w wynikach wyszukiwania, należy zaznaczyć pole wyboru **Wyniki wyszukiwania** podczas konfigurowania przeszukiwacza albo podczas konfigurowania odwzorowań pól na elementy XML i HTML.

## Treść dokumentu

W przypadku niektórych dokumentów, takich jak dokumenty sieci WWW, za treść dokumentu uznawana jest cała jego zawartość. W przypadku innych typów dokumentów, takich jak dokumenty zawierające pola, można zaznaczyć pola zawierające właściwą treść i nie zawierające metadanych.

Aby zaznaczyć, że dane pole stanowi treść dokumentu, należy w czasie konfigurowania przeszukiwacza zaznaczyć pole wyboru **Treść dokumentu**. Jeśli zaznaczone zostaną pola wyboru **Treść dokumentu** i **Wyszukiwanie pełnotekstowe**, to wartość pola stosowana będzie do wykrywania duplikatów dokumentu, a w wynikach wyszukiwania stanie się częścią dynamicznego podsumowania dokumentu.



## Pojęcia pokrewne



## Wykrywanie duplikatów dokumentów

Wykrywanie duplikatów dokumentów jest techniką służącą zapobieganiu sytuacji, w której wyniki wyszukiwania zawierałyby wielokrotnie dokumenty o identycznej lub prawie identycznej treści.

Jakość wyszukiwania może być gorsza, jeśli w wynikach wyszukiwania znajdzie się wiele egzemplarzy identycznych (lub prawie identycznych) dokumentów. Analiza pod kątem duplikatów dokumentów występuje tylko wtedy, gdy spełnione są obydwa następujące warunki:

- W kolekcji używany jest model rankingu oparty na odsyłaczach. Model ten dotyczy przeszukiwaczy przeszukujących serwisy WWW, takich jak przeszukiwacz sieci WWW lub przeszukiwacz WebSphere Portal.
- Zabezpieczenia kolekcji są wyłączone.

W toku analizy globalnej procesy indeksowania wykrywają duplikaty, skanując treść każdego z dokumentów. Jeśli dwa dokumenty mają identyczną treść, są traktowane jako swoje duplikaty.

Jeśli również metadane dokumentów mają być uwzględniane podczas analizy wykrywającej duplikaty, należy zaznaczyć pole wyboru **Treść dokumentu** podczas konfigurowania przeszukiwaczy dla danej kolekcji i określania opcji przeszukiwania metadanych. W takim przypadku przeszukiwacz będzie przeszukiwać pola metadanych tak jak treść dokumentu i uwzględniać metadane w toku analizowania treści pod kątem duplikatów dokumentów. Podobna analiza będzie występować wtedy, gdy zostaną skonfigurowane opcje analizowania składni dokumentów HTML i XML oraz zostanie zaznaczone pole wyboru **Treść dokumentu**.

W razie określenia, że pole lub pole metadanych stanowi treść dokumentu, treść takiego pola będzie dodawana do dynamicznego podsumowania dokumentu w wynikach wyszukiwania, co może zadecydować o tym, czy dany dokument zostanie w wynikach wyszukiwania wyświetlony. Jeśli w aplikacji wyszukiwającej włączone jest wykrywanie dokumentów będących niemalże duplikatami (właściwość `NearDuplicateDetection` w metodzie `setProperty` ustawiona na `Yes`), dokumenty o podobnych tytułach i podsumowaniach będą ukrywane, gdy użytkownik będzie przeglądać wyniki wyszukiwania. Aby wyświetlić ukryte dokumenty będące prawie duplikatami, użytkownicy mogą kliknąć odpowiedni odsyłacz.

W grupie duplikatów dokumentu jeden dokument jest wzorcem, a pozostałe są jego duplikatami. Wszystkie dokumenty w grupie duplikatów mają identyczną kanoniczną reprezentację treści. W toku indeksowania indeksowana jest treść (leksemy) dokumentu wzorcowego. W przypadku duplikatów dokumentu indeksowane są tylko leksemy metadanych. Gdy dokument wzorcowy zostanie usunięty z indeksu, wzorcem stanie się następny duplikat. Użytkownikom prowadzącym wyszukiwania w kolekcji zwracany jest tylko dokument wzorcowy.

## Pojęcia pokrewne

“Analiza duplikatów dokumentów i bezpieczeństwo kolekcji” na stronie 260

## Znaki wieloznaczne w zapytaniach

Można umożliwić użytkownikom stosowanie znaków wieloznacznych w terminach zapytania oraz wyszukiwanie słów zgodnych z określonym wzorcem.

Termin zapytania ze znakiem wieloznacznym jest to termin, który zawiera znak gwiazdki (\*). Jeśli użytkownik wprowadzi zapytanie, które zawiera znak wieloznacznym, w wynikach wyszukiwania zostaną uwzględnione wszystkie dokumenty w indeksie, które są zgodne z terminem zapytania oraz wszystkie dokumenty w indeksie, które są zgodne ze wzorcem reprezentowanym przez znak wieloznacznym. Na przykład znak wieloznacznym kończący w wyszukiwanym terminie SZU\* może odpowiadać terminom szukaj, szufla i szuler.

Podczas konfigurowania opcji znaków wieloznacznym dla indeksu należy wskazać, czy użytkownicy będą mieli możliwość podawania znaków wieloznacznym w zapytaniach, a jeśli tak, to trzeba też określić sposób ich obsługi:

- Możliwość wyszukiwania słów zgodnym ze wzorcem zawierającym znak wieloznacznym można włączyć dla wszystkich części dokumentu lub ograniczyć tylko pól.
- Obsługę zapytań zawierających znaki wieloznacznym można włączyć dla wszystkich pól lub ograniczyć wyłącznie do określonych pól.
- Można ograniczyć dozwoloną pozycję znaku wieloznacznym do ostatniego znaku terminu użytego w zapytaniu (znak wieloznacznym kończący) lub można zezwolić na występowanie znaku wieloznacznym w każdym miejscu tego terminu. (Nie można stosować znaków wieloznacznym w nazwach pól).
- W zależności od miejsc, w których dozwolone jest występowanie znaku wieloznacznym, można wybrać sposób rozwijania terminów użytych w zapytaniu (terminy zawierające znaki wieloznacznym rozwijane są do wszystkich zgodnym z nimi terminów indeksu). W indeksie mogą być zapisane wszystkie możliwe rozwinięcia terminów lub procesy wyszukiwania mogą rozwijać terminy podczas przetwarzania zapytania.

Wszelkie zmiany w ustawieniach znaków wieloznacznym odnoszą skutek przy następnej operacji budowania indeksu głównego.

## Rozwinięcie indeksu

Aby uwzględnić rozwijanie terminów znajdujących się w indeksie, należy określić, ile początkowych znaków słowa musi być zgodnym ze wzorcem zawierającym znaki wieloznacznym umieszczonym w zapytaniu, żeby nastąpiło uzgodnienie. Wyniki zwracane są tylko przez te terminy, które mają przynajmniej taką liczbę znaków (z wyłączeniem znaku \*). Na przykład jeśli podana zostanie wartość 4, to uzyskanie zgodności możliwe będzie tylko wtedy, gdy termin użyty w zapytaniu będzie zawierać przynajmniej cztery znaki.

Gdy podana zostanie wartość 4, wtedy słowo *technologia* będzie zgodne z terminem *tech\** oraz z terminem *techno\**, ale nie będzie zgodne z terminem *te\**.

Podczas operacji głównego lub przyrostowego budowania indeksu oprócz samych terminów indeksowane są wszystkie możliwe rozszerzenia każdego z terminów w dokumencie. Zaletą tego podejścia jest oszczędność czasu potrzebnego na rozwijanie terminów podczas przetwarzania zapytania. Jednak to podejście powoduje zwiększenie wielkości indeksu, co oznacza konieczność posiadania wystarczających zasobów systemowych do obsługi większego indeksu.

To podejście jest najbardziej użyteczne, gdy wielkość kolekcji jest stosunkowo nieduża lub gdy czas i miejsce potrzebne do zbudowania indeksu są mniej ważne niż czas odpowiedzi na zapytanie. Na przykład to podejście można wybrać w celu przeszukania katalogu lub katalogu pracowników.

Takie podejście dostępne jest tylko w przypadku włączonej obsługi znaków wieloznacznym kończących. Włączenie obsługi znaków wieloznacznym występujących w dowolnym miejscu terminu użytego w zapytaniu powoduje, że nie można wybrać opcji pozwalającej na włączenie rozwinięć terminów do indeksu.

## Rozwinięcie zapytania

Aby rozwijać zapytania i stosować reguły dopasowywania wzorca podczas wprowadzania przez użytkowników zapytań zawierających znaki wieloznaczne, należy podać liczbę zgodnych wariantów terminu zapytania. Jeśli na przykład zostanie podana wartość 50, do 50 wariantów terminu zapytania może zostać zakwalifikowane jako zgodne z terminem zapytania.

Na przykład termin `tech*` będzie pasował do wyrazów `techniczny`, `technika`, `technologia` oraz do około 50 innych wyrazów zaczynających się znakami `tech`.

Jakkolwiek rozwijanie zapytania tylko w nieznaczny sposób wpływa na rozmiar indeksu, może istotnie obniżyć wydajność realizacji zapytania. Procesy wyszukiwania muszą iteracyjnie szukać wszystkich możliwych rozwinięć terminu zapytania ze znakiem wieloznacznym aż do osiągnięcia limitu określonego w ustawieniach znaku wieloznacznego.

To podejście jest najbardziej użyteczne, gdy wielkość kolekcji jest stosunkowo duża i należy zminimalizować miejsce i czas budowy indeksu. Na przykład to podejście można wybrać dla repozytoriów poczty elektronicznej, gdzie indeks musi nadążać za szybko zmieniającymi się dokumentami, a czas odpowiedzi na zapytanie jest mniej ważny.

Takie podejście dostępne jest bez względu na to, czy włączona jest obsługa znaków wieloznacznych kończących czy też obsługa znaków wieloznacznych występujących w dowolnym miejscu terminu użytego w zapytaniu.

## Obsługa znaków wieloznacznych w zapytaniach

Zbiór rozwinięć terminu zapytania ze znakiem wieloznacznym zawiera wszystkie terminy indeksu, które można uzyskać, zastępując znak wieloznaczny arbitralnymi sekwencjami znaków. Zbiór ten jest określony w następujący sposób:

- Jeśli w kolekcji obsługiwane są znaki wieloznaczne mogące wystąpić w dowolnym miejscu terminu użytego w zapytaniu, to każdy termin zapytania zawierający gwiazdkę jest interpretowany jako termin ze znakiem wieloznacznym.
- Zbiór zawiera najwyżej maksymalną liczbę rozwinięć, którą skonfigurował administrator wyszukiwania korporacyjnego. Jeśli indeks zawiera więcej rozwinięć, są one ignorowane. (W wynikach wyszukiwania zawarta jest informacja, czy niektóre rozwinięcia znaków wieloznacznych zostały zignorowane).
- Jeśli obsługa znaków wieloznacznych jest ograniczona do pewnego zbioru pól, to zbiór zawiera tylko te terminy, które występują w jednym z podanych pól. Wystarczy, jeśli termin występuje tylko w jednym z pól w przynajmniej jednym dokumencie w indeksie.
- Jeśli termin użyty w zapytaniu dotyczy określonego pola, to znak wieloznaczny musi wystąpić po specyfikatorze pola (na przykład `nazwa_pola:*sphere`). Nazwa pola nie może zawierać dwukropka (:).
- Jeśli obsługa znaków wieloznacznych jest ograniczona do pewnego zbioru pól, to nazwa pola występująca w zapytaniu, które zawiera znaki wieloznaczne, musi być nazwą jednego z pól określonych w konsoli administracyjnej wyszukiwania korporacyjnego. W przeciwnym razie nie zostaną znalezione żadne rozwinięcia danego terminu.
- Znaki wieloznaczne są obsługiwane tylko w terminach określonych zwykłym tekstem, natomiast nie można ich używać w nazwach elementów XML, nazwach atrybutów ani ich wartościach. Terminy składające się wyłącznie ze znaków wieloznacznych nie są obsługiwane.

## Jak znaki wieloznaczne wpływają na indeks

Obsługa znaków wieloznacznych w kontekście rozszerzania indeksu powoduje zwiększenie rozmiarów indeksu oraz wydłużenie czasu wymaganego na jego zbudowanie.

Rozszerzenie indeksu oznacza indeksowanie każdego terminu razem ze wszystkimi jego przedrostkami. Na przykład razem z terminem **program** indeksowane będą następujące terminy pochodne:

p, pr, pro, prog, progr, progra, program.

Liczba haseł składanych w indeksie rośnie tym samym odpowiednio do średniej długości wyrazu. Kompresja indeksu pozwala zmniejszyć jego rozmiar, chociaż tylko w nieznacznym stopniu. Czas wymagany do zbudowania indeksu rośnie proporcjonalnie do średniej długości wyrazu.

Indeks tworzony na bazie angielskich dokumentów rośnie w przybliżeniu czterokrotnie, ponieważ średnia długość angielskiego słowa wynosi od pięciu do sześciu znaków. Indeks tworzony w n-gramowym podziale na leksemy rośnie w przybliżeniu dwukrotnie, ponieważ każdy n-gram zawiera dwa znaki.

Rozszerzanie indeksu jest zalecane w celu zapewnienia obsługi znaków wieloznacznych w następujących sytuacjach:

- Kolekcja jest dostatecznie mała, przez co miejsce na dysku i czas zużywane w związku z rozszerzaniem indeksu nie powoduje problemów z wydajnością.
- Wszystkie możliwe rozszerzenia znaków wieloznacznych muszą zostać uwzględnione w wynikach wyszukiwania, aby zadośćuczynić oczekiwaniom użytkowników (lub wymogom organizacji).

Wpływ rozszerzenia na obsługę indeksu można ograniczyć, określając minimalną długość przedrostka. Na przykład, jeśli minimalna długość przedrostka zostanie określona jako 3, przedrostki p i pr dla wyrazu **program** nie będą już indeksowane, a indeks angielski rozrośnie się trzykrotnie, a nie czterokrotnie.

Jeśli obsługa znaków wieloznacznych jest realizowana na poziomie zapytań, do indeksu nie są wprowadzane żadne przedrostki. Rozwijanie terminów następuje po wprowadzeniu zapytania, a indeks jest powiększany tylko o niewielką strukturę danych wymaganą do zapewnienia obsługi tego rozszerzenia. Zazwyczaj indeks umożliwiający rozszerzanie zapytań jest o 10% do 20% większy od indeksu pozbawionego tej funkcji, a jego budowanie trwa dłużej o mniej niż 10%. Nałożenie górnego ograniczenia na liczbę rozszerzeń nie wpływa na rozmiar indeksu ani na czas jego budowania.

## **Konfigurowanie opcji dotyczących znaków wieloznacznych w zapytaniach**

Podczas konfigurowania opcji indeksowania dla kolekcji wyszukiwania korporacyjnego można określić, czy użytkownicy mają mieć możliwość stosowania znaków wieloznacznych w terminach używanych w zapytaniach.

### **Zagadnienia wstępne**

Do skonfigurowania opcji znaków wieloznacznych wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należy dany indeks.

### **Informacje o tym zadaniu**

W przypadku konfigurowania opcji dotyczących znaków wieloznacznych wszelkie zmiany odnoszą skutek podczas następnego operacji budowania indeksu głównego.

### **Procedura**

Aby skonfigurować obsługę znaków wieloznacznych w zapytaniach, należy wykonać poniższe czynności:

1. Przejdź do edycji kolekcji, wybierz stronę Indeksowanie, a następnie kliknij opcję **Konfiguruj opcje dotyczące znaków wieloznacznych**.
2. Na stronie Opcje znaków wieloznacznych zaznacz pole wyboru **Obsługa znaków wieloznacznych w zapytaniach**.
3. Opcjonalnie: można określić, że w zapytaniach pełnotekstowych mają być obsługiwane znaki wieloznaczne. Na przykład zapytanie pełnotekstowe `tech*`, które nie dotyczy pola o zadanej nazwie, zwraca wyniki rozwinięte (takie jak technologia lub technika) tylko wtedy, gdy to pole wyboru jest zaznaczone.
4. Określ pola, w których obsługiwane są znaki wieloznaczne:
  - Aby określić, że znaki wieloznaczne nie powinny być uwzględniane w zapytaniach przeszukujących pola, wybierz opcję **Bez pól**.
  - Aby włączyć obsługę znaków wieloznacznych w zapytaniach dla wszystkich pól w dokumencie, wybierz opcję **Wszystkie pola**.
  - Aby ograniczyć obsługę znaków wieloznacznych do niektórych pól, wybierz opcję **Konkretne pola**, a następnie wpisz nazwy pól. Wyniki rozwinięte zwracane są tylko dla podanych pól. Na przykład zapytanie `author:jan*` zwraca wyniki rozwinięte tylko wtedy, gdy dla pola `author` włączona jest obsługa znaków wieloznacznych.
5. Wskaż, czy znak wieloznaczny musi wystąpić na ostatniej pozycji terminu użytego w zapytaniu (znak wieloznaczny kończący) czy też jego pozycja jest dowolna i może on wystąpić w każdym miejscu terminu użytego w zapytaniu.

Po określeniu pozycji i typu znaku wieloznacznego trzeba też podać sposób włączenia obsługi znaków wieloznacznych. Aby uzyskać szczegółowe informacje, kliknij ikonę **Pomoc** w konsoli administracyjnej.

## Zasięgi

Skonfigurowanie zasięgu pozwala przedstawić użytkownikom ograniczony widok kolekcji.

Zasięg jest grupą pokrewnych identyfikatorów URI w indeksie. Skonfigurowanie zasięgu ogranicza zbiór dokumentów wyświetlanych w kolekcji. Przy wyszukiwaniu w kolekcji uwzględniane są tylko dokumenty znajdujące się w zdefiniowanym zasięgu, a nie wszystkie dokumenty z indeksu. Aby można było korzystać z tej funkcji, aplikacje wyszukiwujące muszą zapewniać obsługę wyszukiwania w zakresach.

Tworzenie zasięgu polega na wydzieleniu zakresu identyfikatorów URI w indeksie, które będą dostępne do wyszukiwania dla użytkowników. Ograniczenie zakresu dokumentów dostępnych do wyszukiwania umożliwi lepsze dopasowanie wyników wyszukiwania do oczekiwań użytkowników.

Na przykład można utworzyć jeden zasięg obejmujący identyfikatory URI dla pracowników działu obsługi technicznej oraz drugi zasięg obejmujący identyfikatory URI dla pracowników działu kadr. Jeśli używana aplikacja wyszukiwująca obsługuje zasięgi, pracownicy działu obsługi technicznej i pracownicy działu kadr będą pobierali dokumenty z odpowiadających tym działom zasięgów.

Istnieje możliwość tworzenia dowolnej liczby zasięgów, jednak utworzenie zbyt wielu zasięgów może mieć wpływ na wydajność. Zasięgi należy konfigurować w taki sposób, aby realizacja większości zapytań wymagała filtrowania tylko przez jeden lub dwa zasięgi. Ponieważ zasięgi mogą zawierać pełne identyfikatory URI lub ich zwrócić, jeden dokument może należeć do więcej niż jednego zasięgu.

W czasie konfigurowania zasięgów skuteczne wprowadzenie zmian wymagać może dwukrotnego zbudowania indeksu. Jeśli zasięgi zostaną skonfigurowane jeszcze przed pierwszym budowaniem głównego indeksu, to użytkownicy będą mogli wyszukiwać w kolekcji, ale w wynikach wyszukiwania nie będą pojawiały się dane zasięgu. Dopiero ponowne zbudowanie głównego indeksu sprawi, że wyniki wyszukiwania będą odzwierciedlały identyfikatory URI znajdujące się w zasięgu.

Jeśli zasięgi zostaną skonfigurowane po zbudowaniu głównego indeksu, to wprowadzone zmiany wejdą w życie dopiero po ponownym zbudowaniu tego indeksu.

## Konfiguracja zasięgów

Przy konfigurowaniu zasięgu dla kolekcji wyszukiwania korporacyjnego należy określić identyfikatory URI lub wzorce URI dla zakresu dokumentów w indeksie, który użytkownicy mogą wyszukiwać.

### Zagadnienia wstępne

Aby skonfigurować zasięgi, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą zasięgi.

### Informacje o tym zadaniu

Jeśli aplikacje wyszukujące umożliwiają obsługę zasięgów, podczas wyszukiwania w kolekcji użytkownicy mogą wyszukiwać tylko te dokumenty, które odpowiadają identyfikatorom URI zdefiniowanym w ramach zasięgu.

W czasie konfigurowania zasięgów skuteczne wprowadzenie zmian wymagać może dwukrotnego zbudowania indeksu. W przypadku skonfigurowania zasięgów przed pierwszym budowaniem indeksu użytkownicy będą mogli przeszukiwać kolekcję, ale nie będą widzieć danych zasięgu w wynikach wyszukiwania. Dopiero ponowne zbudowanie głównego indeksu sprawi, że wyniki wyszukiwania będą odzwierciedlały identyfikatory URI znajdujące się w zasięgu.

Jeśli zasięgi zostaną skonfigurowane po zbudowaniu głównego indeksu, to wprowadzone zmiany wejdą w życie dopiero po ponownym zbudowaniu tego indeksu.

### Procedura

Aby skonfigurować zasięg, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Konfiguracja zasięgów**.
2. Na stronie Zasięgi kliknij przycisk **Utwórz zasięg**.
3. Podaj nazwę zasięgu oraz identyfikatory URI i wzorce URI, które definiują granice zasięgu. Można określić również identyfikatory URI i wzorce URI, które mają być wykluczone z zasięgu.
4. Kliknij przycisk **OK**.

Nowy zasięg jest umieszczony na stronie Zasięgi wraz z innymi należącymi do tej kolekcji zasięgami.

### Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 120



## Zwinięte identyfikatory URI

Wyszukiwanie korporacyjne może prezentować wyniki wyszukiwania w ten sposób, że dokumenty pochodzące ze źródeł o takich samych przedrostkach identyfikatorów URI wyświetlane będą w postaci zwiniętej.

Gdy wyniki są zwinięte, pierwszy z nich zazwyczaj jest wyrównany do lewego marginesu. Jeden lub kilka dokumentów o niższych ocenach są zgrupowane i wyświetlane poniżej (wyróżnione wcięciem).

Aby zwinąć w jednej grupie dokumenty wynikowe mające różne przedrostki URI, można skojarzyć przedrostki URI z nazwą tworzonej grupy. Na przykład, jeśli używane są trzy serwery do zarządzania danymi finansowymi, w wynikach wyszukiwania można pogrupować dokumenty ze wszystkich trzech serwerów i zwinąć dokumenty o niższym stopniu zgodności z kryteriami wyszukiwania pod dokumentami o wyższym stopniu zgodności.

Aplikacje wyszukujące mogą zwinąć dokumenty w wynikach wyszukiwania na podstawie przedrostka URI lub nazwy grupy. W przykładowej aplikacji wyszukującej dla wyszukiwania korporacyjnego wyświetlane są dwa najwyżej ocenione dokumenty wynikowe wyszukiwania. Jeśli zwrócone zostaną więcej niż dwa dokumenty o takim samym przedrostku identyfikatora URI (albo dokumenty należące do tej samej grupy identyfikatorów URI), można zaznaczyć opcję wyświetlania takich dokumentów w formie zwiniętej.

Użytkownicy mogą korzystać ze składni zapytań wyszukiwania korporacyjnego (`samegroupas:przedrostek_URI`) i w ten sposób szukać wszystkich dokumentów znajdujących się w tej samej grupie, co podany w zapytaniu przedrostek identyfikatora URI.

## Jak zorganizować przedrostki URI i nazwy grup

Podczas korzystania z konsoli administracyjnej do konfigurowania reguł sterujących zwijaniem wyników wyszukiwania należy określić przedrostki URI dokumentów, które mają być zwijane, oraz skojarzyć przedrostki URI z nazwą grupy (opcjonalnie).

Istotne znaczenie ma kolejność konfigurowanych przedrostków URI. Serwer indeksowania uwzględnia kolejność przedrostków URI podczas obliczania wartości poszczególnych identyfikatorów URI w kolekcji. Dla każdego identyfikatora URI:

1. Serwer indeksowania przegląda kolejno przedrostki URI w regułach dotyczących zwijania wyników wyszukiwania.
2. Kiedy serwer indeksowania znajdzie pierwszy przedrostek URI pasujący do przedrostka dokumentu w indeksie, przypisuje nazwę grupy (lub przedrostek URI, jeśli w regule nie określono nazwy grupy) jako dodatkowy termin wyszukiwania dla tego dokumentu.

Jeśli do przedrostka identyfikatora URI nie da się dopasować dokumentu WWW, to serwer indeksowania w miejsce przedrostka identyfikatora URI zastosuje nazwę hosta z adresu URL. Jeśli do przedrostka identyfikatora URI nie da się dopasować dokumentu NNTP, to serwer indeksowania w miejsce przedrostka identyfikatora URI zastosuje identyfikator pierwszej wiadomości z nagłówka referencji.

Po dodaniu przedrostka URI do listy tych, które mają być zwijane w wynikach wyszukiwania, należy jeszcze ustawić dany przedrostek URI w takiej kolejności, w jakiej serwer indeksowania ma go przeglądać i ewentualnie kojarzyć z dokumentami w indeksie jako dodatkowy termin wyszukiwania:

- Po dodaniu przedrostka URI bez kojarzenia go z nazwą grupy można wybrać indywidualny przedrostek URI i przenieść go w górę lub w dół listy.



- Po dodaniu przedrostka URI i skojarzeniu go z nazwą grupy przesunięcie jednego przedrostka z tej grupy w górę lub w dół listy powoduje przesunięcie całej grupy. Kolejność przedrostków identyfikatorów URI w grupie nie ma żadnego znaczenia. Wybranie pojedynczego przedrostka identyfikatora URI powoduje automatyczne wybranie całej grupy.

## Zwijanie identyfikatorów URI w wynikach wyszukiwania

Można określić opcje grupowania i zwijania dokumentów wynikowych ze źródeł mających ten sam przedrostek URI. Można także określić nazwę grupy, która umożliwi zwijanie razem dokumentów wynikowych o różnych przedrostkach identyfikatora URI.

### Zagadnienia wstępne

Aby podać opcje zwijania wyników wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienie administratora kolekcji dla tej kolekcji.



### Informacje o tym zadaniu

Wprowadzone zmiany dotyczące zwijania wyników wyszukiwania odnoszą skutek dopiero podczas następnej operacji budowania indeksu głównego.

### Procedura

Aby określić opcje zwijania wyników wyszukiwania, należy wykonać następujące czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Zwijanie wyników wyszukiwania**.
2. Na stronie Zwijanie wyników wyszukiwania kliknij opcję **Dodaj przedrostek URI**.
3. Na stronie Dodawanie przedrostka URI dla zwijania wyników wpisz przedrostek URI dla dokumentów, które mają być zwijane w wynikach wyszukiwania. Na przykład:  

```
http://finanse/ROI/
http://serwer1.com/finanse/
db2://LOCALDB/SCHEMA1.TABLE1/
exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/
```
4. Można wpisać opisową nazwę grupy, która ma być powiązana z tym przedrostkiem URI. Aby zwinąć w jedną grupę dokumenty wynikowe z kilku źródeł, należy wpisać tę samą nazwę grupy podczas dodawania każdego przedrostka URI.
5. Kliknij przycisk **OK**.
6. Na stronie Zwijanie wyników wyszukiwania umieść nową regułę w takiej kolejności, w jakiej ma być skanowana przez serwer indeksowania:
  - Jeśli został dodany nowy przedrostek URI, który nie jest powiązany z nazwą grupy, pojawia się on na dole listy. Przenieś go na właściwą pozycję za pomocą klawiszy strzałek.
  - Jeśli z nowym przedrostkiem URI została powiązana nazwa grupy, pojawia się on na końcu zestawu przedrostków URI należących do tej samej grupy. Do przeniesienia całej grupy przedrostków URI na właściwą pozycję należy użyć klawiszy strzałek.
7. Aby zmienić przedrostek identyfikatora URI lub nazwę grupy, wybierz przedrostek URI i kliknij ikonę  **Edytuj**.
8. Aby usunąć przedrostek URI z listy, wybierz przedrostek i kliknij ikonę  **Usuń**.

## Usuwanie z indeksu identyfikatorów URI

Aby zapobiec wyszukiwaniu dokumentów w kolekcji przez użytkowników, należy usunąć z indeksu identyfikatory URI tych dokumentów.

## Zagadnienia wstępne

Aby usunąć identyfikatory URI z indeksu, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

### Informacje o tym zadaniu

Jeśli został określony pełny identyfikator URI, użytkownicy nie zobaczą go w wynikach wyszukiwania. Jeśli jednak użytkownik uruchomi to samo zapytanie, a dokumenty wynikowe dla tego zapytania znajdują się w pamięci podręcznej wyszukiwania, to buforowana strona wynikowa dla usuniętego identyfikatora URI nadal będzie zwracana w wynikach wyszukiwania. Pamięć podręczna wyszukiwania nie jest odświeżana, a identyfikator URI nie jest usuwany z indeksu do czasu następnej operacji budowania indeksu głównego lub przyrostowego.

Jeśli określono wzorzec URI w celu usunięcia wielu adresów jednocześnie, adresy pasujące do podanego wzorca będą widoczne dla użytkowników w wynikach wyszukiwania do czasu wykonania następnej operacji budowania indeksu głównego.

Usunięcie identyfikatora URI z indeksu nie usuwa go z przestrzeni przeszukiwania. Przy następnym przeszukiwaniu dokumentu przez przeszukiwacz identyfikator URI zostanie wbudowany w indeks i będzie dostępny do ponownego wyszukiwania. Aby usunąć identyfikator URI z przestrzeni przeszukiwania, należy zaktualizować reguły przeszukiwania, w celu wykluczenia dokumentu, po czym zatrzymać i zrestartować przeszukiwacz.

### Procedura

Aby usunąć z indeksu identyfikatory URI określonych dokumentów, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Usuwanie identyfikatorów URI z indeksu**.
2. Na stronie Usuwanie identyfikatorów URI z indeksu wpisz identyfikatory URI (lub wzorce URI), które mają zostać usunięte z indeksu.

Na przykład:

```
http://domena.org/kadry/*
db2://zarzadzanieInformacja/ROI*
cm://przedsiębiorstwo/finanse*
```

### Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 120

---

## Administrowanie serwerem wyszukiwania

Opcje określone dla serwerów wyszukiwania to między innymi: używanie obszaru pamięci podręcznej do zwracania wyników wyszukiwania, definiowanie maksymalnej długości wyświetlanych podsumowań dokumentów w wynikach wyszukiwania, powiązanie niestandardowych słowników w celu zwiększenia jakości wyszukiwania, a także zwracanie predefiniowanych identyfikatorów URI w wynikach wyszukiwania, gdy w zapytaniu zostały użyte określone terminy.

Gdy użytkownik wprowadza zapytanie, serwery wyszukiwania znajdują szybko odpowiednie dokumenty, korzystając z indeksu. Serwery wyszukiwania korzystają ze składnicy danych wyszukiwania korporacyjnego (zawierającej przeanalizowane dane podzielone na leksemy) do pobierania metadanych dla odpowiednich dokumentów. Metadane mogą zawierać identyfikatory URI dokumentów, tytuły, opisy, daty, informacje o typach danych itp.

Podczas konfigurowania serwerów wyszukiwania dla kolekcji określa się opcje determinujące sposób przetwarzania zapytań, w tym także opcje, które mogą mieć wpływ na wydajność zapytań:

### **Konfiguracja pamięci podręcznej wyszukiwania**

Aby zoptymalizować wydajność zapytań, można określić, że wyniki wyszukiwania mają być przechowywane w pamięci podręcznej. Można także ustalić wielkość pamięci przeznaczonej na buforowanie wyników przeszukiwania.

### **Konfiguracja maksymalnej długości wyświetlanych podsumowań dokumentów**

Dla większości dokumentów zwracanych w wynikach wyszukiwania wyświetlane jest podsumowanie treści dokumentu pomagające użytkownikom zdecydować, czy jest to dokument, który chcą pobrać. Istnieje możliwość określenia, jak duży obszar wyników wyszukiwania ma być przeznaczony na wyświetlanie podsumowania.

### **Określanie innego języka domyślnego**

Domyślny język wyszukiwania dokumentów w kolekcji jest określany podczas tworzenia kolekcji, jednak w razie potrzeby można wybrać inny język.

### **Powiązanie niestandardowych słowników**

Jeśli twórcy aplikacji utworzyli niestandardowe słowniki synonimów, słów ignorowanych i słów-kryteriów, można określić, które z nich będą używane podczas przeszukiwania kolekcji przez użytkowników.

### **Konfiguracja szybkich odsyłaczy**

Można predefiniować identyfikatory URI, które mają być zwracane po podaniu określonych słów kluczowych i fraz. Gdy użytkownik wprowadzi określone słowo kluczowe lub frazę w zapytaniu, w wynikach wyszukiwania zostanie zwrócony predefiniowany identyfikator URI. Identyfikatory URI związane z szybkimi odsyłaczami są zwracane dodatkowo, oprócz identyfikatorów URI zwracanych przez serwery wyszukiwania w wyniku przeszukania indeksu.

### **Pojęcia pokrewne**

“Ranking dokumentów” na stronie 205

“Niestandardowe słowniki kryteriów” na stronie 209

---

## Pamięci podręczne wyszukiwania

Gdy obciążenie serwerów wyszukiwania jest stosunkowo duże, wydajność procesów wyszukiwania można zwiększyć, używając funkcji buforowania wyników wyszukiwania.

Gdy serwery wyszukiwania przetwarzają zapytania wyszukiwania, najpierw sprawdzana jest zawartość pamięci podręcznej w poszukiwaniu gotowych wyników takiego zapytania. Jeśli serwery wyszukiwania znajdą odpowiednie wyniki zapytania, mogą je szybko zwrócić użytkownikowi. Jeśli serwery wyszukiwania nie znajdą odpowiednich wyników zapytania, zaczną wyszukiwanie w indeksie.

W miarę zapełniania się pamięci podręcznej najstarsze wyniki wyszukiwania oraz wyniki rzadko wpisywanych zapytań są usuwane w celu zwolnienia pamięci dla wyników nowych wyszukiwań.

Konsola administracyjna wyszukiwania korporacyjnego umożliwia włączenie buforowania wyników wyszukiwania i określenie pojemności pamięci podręcznej (liczby wyników zapytań, które mogą być zapisane jednocześnie w pamięci podręcznej).

Aby zmiany wprowadzone do ustawień pamięci podręcznej wyszukiwania zaczęły obowiązywać, należy zrestartować serwery wyszukiwania.

---

## Konfiguracja pamięci podręcznej wyszukiwania

Dla każdej kolekcji można włączyć lub wyłączyć pamięć podręczną wyszukiwania. Istnieje również możliwość określenia opcji sterujących wielkością pamięci podręcznej wyszukiwania.

### Zagadnienia wstępne

Aby skonfigurować pamięć podręczną wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

### Procedura

Aby skonfigurować pamięć podręczną wyszukiwania, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. Na stronie Opcje serwera wyszukiwania zaznacz pole wyboru **Użyj pamięci podręcznej wyszukiwania**.
3. W polu **Maksymalna liczba pozycji w pamięci podręcznej** wpisz maksymalną liczbę wyników zapytania przechowywanych w pamięci podręcznej.
4. Kliknij przycisk **OK**.
5. Aby zmiany zostały uwzględnione, przejdź do monitorowania serwerów wyszukiwania i zrestartuj procesy serwera.

---

## Niestandardowe słowniki synonimów

Aby podnieść jakość wyników wyszukiwania, podczas przeszukiwania kolekcji można włączyć funkcję wyszukiwania synonimów terminów wprowadzonych w zapytaniu.

Po utworzeniu słownika synonimów, dodaniu go do systemu wyszukiwania korporacyjnego i skojarzeniu go z kolekcją użytkownicy mogą wyszukiwać w kolekcjach dokumenty zawierające synonimy określonych terminów z zapytania. Dzięki takiemu poszerzeniu zakresu zapytań zwiększa się prawdopodobieństwo znalezienia wszystkich potrzebnych dokumentów, a nie tylko tych, które zawierają podane sformułowanie w dosłownej postaci. Ponieważ w słowniku synonimów definiuje się, które słowa są dokładnymi odpowiednikami innych, użytkownicy mogą odszukać wszystkie interesujące ich dokumenty bez konieczności samodzielnego wprowadzania każdej odmiany wyrazu oddzielnie.

Na przykład w danej organizacji mogą funkcjonować akronimy i skrócone nazwy działów, sprzętu itp. Dokumenty w niektórych kolekcjach mogą też zawierać słownictwo specyficzne dla danej branży. Przygotowując słownik synonimów, można zapewnić, że zapytania zawierające akronimy (na przykład ACL) zwracać będą również dokumenty omawiające rozwinięcia tych akronimów (na przykład listy ACL, listy kontroli dostępu, access control list itd.).

Język zapytań wyszukiwania korporacyjnego obsługuje też synonimy i pozwala użytkownikom dodawać przed łańcuchem zapytania operator tyldy. Na przykład zapytanie ~WAS może zwrócić dokumenty zawierające pełną nazwę produktu WebSphere Application Server. Twórcy aplikacji mogą także wbudować obsługę synonimów na poziomie właściwości zapytania, co eliminuje konieczność stosowania specjalnej składni.

Słowniki synonimów zawierają odmiany wyrazów. Mają one następujące właściwości:

- Wyrazy nie są związane z jednym językiem, mogą być używane w różnych językach. Dla każdej kolekcji istnieje tylko jeden słownik synonimów.
- Wyrazy nie są odmieniane automatycznie. Wszystkie formy gramatyczne wyrazu należy dodać do listy synonimów. Na przykład liczba pojedyncza i mnoga wyrazu będą stanowiły osobne wpisy w słowniku.

Większość terminów umieszczanych w słowniku stanowi swoje dokładne odpowiedniki semantyczne, co oznacza, że jeśli termin A jest synonimem terminu B, to B jest też synonimem A. Terminy A i B mogą być więc w zapytaniu stosowane zamiennie.

Jednak można także dodawać terminy różne znaczeniowo, na przykład wyrazy o bardziej ogólnym lub bardziej konkretnym znaczeniu. Na przykład w jednej grupie synonimów znajdują się wyrazy budowanie i dom, a w drugiej bank, fundusz i towarzystwo kredytowe.

Im dalsze są związki między terminami, tym większy będzie zestaw wyników wyszukiwania, ale niektóre wyniki mogą nie mieć związku z terminem z zapytania. Interfejs API wyszukiwania i indeksowania udostępnia metody pozwalające wybierać użytkownikom właściwe synonimy przy wprowadzaniu zapytania, a także opcje pokazujące, które kryteria zapytania zostały poszerzone o jakie synonimy.

Aby utworzyć słownik synonimów, ekspert w zakresie obsługi kolekcji musi przygotować listę synonimów w formacie XML albo we współpracy z twórcą aplikacji utworzyć odpowiedni plik XML. Następnie, za pomocą narzędzia wyszukiwania korporacyjnego **essyndictbuilder**, należy przekształcić plik XML do postaci binarnego pliku .dic.

Administrator wyszukiwania korporacyjnego przesyła plik binarny do systemu i przypisuje mu nazwę wyświetlaną. Korzystając z opcji serwera wyszukiwania dla kolekcji, administratorzy kolekcji mogą wybrać słownik synonimów stosowany do wyszukiwania dokumentów w danej kolekcji.

**Ograniczenie:** Niestandardowy słownik synonimów nie może być już modyfikowany po dodaniu go do systemu. Aby zmienić synonimy w słowniku kolekcji, należy:

1. Zmodyfikować odpowiednio źródłowy plik XML.
2. Przekształcić źródło XML na nowy plik słownika.
3. Usunąć poprzedni słownik synonimów z kolekcji, które go używają.
4. Usunąć poprzedni słownik synonimów z systemu.
5. Dodać nowy słownik synonimów do systemu.
6. Powiązać nowy słownik synonimów z kolekcjami, które mają go używać.

Można napisać skrypt obejmujący te kroki, a następnie za pomocą tego skryptu ponownie wdrożyć słownik w systemie wyszukiwania korporacyjnego.

#### Pojęcia pokrewne

 Obsługa synonimów w aplikacjach wyszukiwujących

#### Zadania pokrewne

 Tworzenie pliku XML synonimów

 Tworzenie słownika synonimów

“Ponowne wdrażanie słowników niestandardowych” na stronie 199

## Dodawanie słowników synonimów do systemu

Po utworzeniu niestandardowych słowników synonimów do wyszukiwania w dokumentach kolekcji należy je powiązać z systemem wyszukiwania korporacyjnego. Później można wybrać słownik synonimów, który ma być wykorzystywany podczas wyszukiwania w kolekcji.

#### Zagadnienia wstępne


Aby udostępnić niestandardowe słowniki synonimów do wykorzystywania w zapytaniach wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego.

#### Ograniczenia

Maksymalna wielkość słownika synonimów wynosi 8 MB.

#### Procedura

Aby powiązać synonimy z systemem wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Wyszukiwanie kliknij opcję **Konfigurowanie słowników synonimów**.
4. Na stronie Konfigurowanie słowników synonimów kliknij opcję **Dodaj słownik synonimów**.
5. Na stronie Dodawanie słownika synonimów wpisz unikalną nazwę wyświetlaną dla słownika synonimów i opcjonalnie podaj opis.
6. Określ położenie pliku .dic. Jeśli plik znajduje się w systemie lokalnym, można go znaleźć, przeglądając system plików. Jeśli plik znajduje się na serwerze indeksowania, należy wpisać pełną ścieżkę do tego pliku.
7. Kliknij przycisk **OK**. Niestandardowy słownik synonimów zostanie dodany do systemu wyszukiwania korporacyjnego i stanie się dostępny podczas wyszukiwania w kolekcjach.

## Powiązanie słownika synonimów z kolekcją

Jeśli z systemem wyszukiwania korporacyjnego są powiązane słowniki synonimów, można wybrać jeden z nich do wykorzystania podczas wyszukiwania w kolekcji. Jeśli termin z zapytania odpowiada terminowi ze słownika, wynikowe dokumenty zawierające synonimy również zostaną zwrócone w wynikach wyszukiwania.

#### Zagadnienia wstępne

Aby wybrać słownik synonimów dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienie administratora kolekcji dla tej kolekcji.

### Procedura

Aby powiązać słownik synonimów z kolekcją, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. W polu **Nazwa słownika synonimów** na stronie Opcje serwera wyszukiwania wybierz słownik synonimów, który ma być używany podczas wyszukiwania w kolekcji.  
Lista dostępnych słowników synonimów zawiera wszystkie słowniki synonimów dodane do systemu wyszukiwania korporacyjnego.
3. Kliknij przycisk **OK**.

---

## Niestandardowe słowniki słów ignorowanych

Aby podnieść jakość wyników wyszukiwania, można określić, że pewne słowa mają być automatycznie usuwane z terminów użytych w zapytaniu podczas przetwarzania zapytania.

Słownik słów ignorowanych zawiera często używane terminy specyficzne dla danego przedsiębiorstwa, które nie nadają się z tego względu do zastosowania jako terminy w zapytaniu. Dzięki wyłączeniu tych słów z zapytań można zapewnić, że użytkownicy nie będą zasypywani nieodpowiednimi dokumentami wynikowymi (zostaną zwrócone tylko dokumenty zgodne z innymi terminami użytymi w zapytaniu). Podczas przetwarzania zapytania serwery wyszukiwania usuną słowa ignorowane z zapytań. Do usuwanych słów należą słowa ignorowane ze słownika niestandardowego i słowa ignorowane predefiniowane dla wyszukiwania korporacyjnego (takie jak powszechnie używane przyimki i przedimki).

W wyszukiwaniu korporacyjnym domyślnie wykonywane jest rozpoznawanie słów pomijanych dla danego języka. Ten proces usuwa z zapytania często używane słowa typu **to** i **ten**. Niestandardowy słownik słów ignorowanych należy zdefiniować tylko dla terminów charakterystycznych dla danego przedsiębiorstwa lub domeny.

W czasie przetwarzania zapytania słowa pomijane są usuwane jeszcze przed wygenerowaniem propozycji poprawki pisowni. Jeśli w zapytaniu znajdują się wyłącznie słowa pomijane, w czasie przetwarzania zapytania nie jest usuwane żadne z nich. Aby zapewnić, że zostaną zwrócone jakiegokolwiek wyniki wyszukiwania, mechanizm usuwania słów ignorowanych wyłączany jest w sytuacji, gdy wszystkie słowa w zapytaniu są słowami pomijanymi. Na przykład, jeśli słowem pomijanym będzie słowo **auto**, a użytkownik szukać będzie właśnie słowa **auto**, to w wynikach wyszukiwania znajdą się dokumenty pasujące do szukanego słowa **auto**. Jeśli jednak użytkownik szukać będzie słów **auto volvo**, to w wynikach wyszukiwania znajdą się wyłącznie te dokumenty, które pasować będą do słowa **volvo**.

Aby utworzyć słownik słów pomijanych, ekspert w zakresie kolekcji musi przygotować listę słów pomijanych w formacie XML albo we współpracy z twórcą aplikacji utworzyć odpowiedni plik XML. Następnie przy użyciu narzędzia wyszukiwania korporacyjnego **esstopworddictbuilder** należy przekształcić plik XML do postaci binarnej **.dic**.

Administrator wyszukiwania korporacyjnego przesyła plik binarny do systemu i przypisuje mu nazwę wyświetlaną. Korzystając z opcji serwera wyszukiwania dla kolekcji, administratorzy kolekcji mogą wybrać słownik słów ignorowanych stosowany do wyszukiwania dokumentów w danej kolekcji.



**Ograniczenie:** Niestandardowy słownik słów ignorowanych nie może być już modyfikowany po dodaniu go do systemu. Aby zmodyfikować słowa ignorowane dostępne podczas przetwarzania zapytania, należy:

1. Zmodyfikować odpowiednio źródłowy plik XML.
2. Przekształcić źródło XML na nowy plik słownika.
3. Usunąć poprzedni słownik słów ignorowanych z kolekcji, które go używają.
4. Usunąć poprzedni słownik słów ignorowanych z systemu.
5. Dodać nowy słownik słów ignorowanych do systemu.
6. Powiązać nowy słownik słów ignorowanych z kolekcjami, które mają go używać.

Można napisać skrypt obejmujący te kroki, a następnie za pomocą tego skryptu ponownie wdrożyć słownik w systemie wyszukiwania korporacyjnego.

#### Pojęcia pokrewne



Niestandardowe słowniki słów ignorowanych

#### Zadania pokrewne



Tworzenie pliku słów ignorowanych w formacie XML



Tworzenie słownika słów ignorowanych

“Ponowne wdrażanie słowników niestandardowych” na stronie 199

## Dodawanie do systemu słowników słów ignorowanych

Jeśli tworzone są niestandardowe słowniki słów pomijanych pozwalające na usuwanie określonych słów z zapytań, to trzeba je dodać do systemu wyszukiwania korporacyjnego. Później można wybrać słownik słów ignorowanych, który ma być wykorzystywany podczas wyszukiwania w kolekcji.

#### Zagadnienia wstępne


Aby dodać słowniki słów ignorowanych do systemu, wymagana jest rola administratora wyszukiwania korporacyjnego.

#### Ograniczenia

Maksymalna wielkość słownika słów pomijanych wynosi 8 MB.

#### Procedura

Aby powiązać słowa ignorowane z systemem wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Wyszukiwanie kliknij opcję **Konfiguruj słowniki słów ignorowanych**.
4. Na stronie Konfigurowanie słowników słów ignorowanych kliknij opcję **Dodaj słownik słów ignorowanych**.
5. Na stronie Dodawanie słownika słów ignorowanych wpisz unikalną nazwę wyświetlaną dla słownika.
6. Określ położenie pliku .dic. Jeśli plik znajduje się w systemie lokalnym, można go znaleźć, przeglądając system plików. Jeśli plik znajduje się na serwerze indeksowania, należy wpisać pełną ścieżkę do tego pliku.

7. Kliknij przycisk **OK**. Niestandardowy słownik słów ignorowanych zostanie dodany do systemu wyszukiwania korporacyjnego i stanie się dostępny podczas wyszukiwania w kolekcjach.

## Powiązanie słownika słów ignorowanych z kolekcją

Jeśli z systemem wyszukiwania korporacyjnego są powiązane słowniki słów ignorowanych, można wybrać jeden z nich do wykorzystania podczas wyszukiwania w kolekcji. Jeśli termin użyty w zapytaniu odpowiada terminowi ze słownika, termin ten zostanie usunięty z zapytania przed jego przetworzeniem.

### Zagadnienia wstępne

Aby wybrać słownik słów ignorowanych dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

### Procedura

Aby powiązać słownik słów ignorowanych z kolekcją, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. W polu **Nazwa słownika słów ignorowanych** na stronie Opcje serwera wyszukiwania wybierz słownik słów ignorowanych, który ma być używany podczas wyszukiwania w kolekcji.

Lista dostępnych słowników zawiera wszystkie słowniki słów ignorowanych dodane do systemu wyszukiwania korporacyjnego.

3. Kliknij przycisk **OK**.

---

## Ponowne wdrażanie słowników niestandardowych

Zmian w słowniku synonimów, słów pomijanych lub kryteriów nie można dokonywać za pomocą konsoli administracyjnej wyszukiwania korporacyjnego. Niemniej jednak można do skryptu wprowadzić pewne działania, a następnie za pomocą tego skryptu ponownie wdrożyć słownik.

### Zagadnienia wstępne

Aby ponownie wdrożyć niestandardowy słownik, należy zalogować się jako administrator wyszukiwania korporacyjnego.

### Procedura

Aby ponownie wdrożyć niestandardowy słownik do celów wyszukiwania korporacyjnego:

1. Na serwerze indeksowania otwórz jeden z następujących plików, aby znaleźć nazwę słownika, który ma zostać ponownie wdrożony:
  - Aby znaleźć nazwę słownika synonimów, otwórz plik ES\_NODE\_ROOT/master\_config/SynonymConfiguration.xml.
  - Aby znaleźć nazwę słownika słów pomijanych, otwórz plik ES\_NODE\_ROOT/master\_config/StopWordDictionaryConfiguration.xml.
  - Aby znaleźć nazwę słownika kryteriów, otwórz plik ES\_NODE\_ROOT/master\_config/BoostingWordDictionaryConfiguration.xml.

Na poniższym przykładzie pokazano plik SynonymConfiguration.xml z dwoma słownikami synonimów, które przesłano z nazwami hello i goodbye:

```
% cat $ES_NODE_ROOT/master_config/SynonymConfiguration.xml
```

```
<SynonymConfiguration>
<Synonyms>
<Synonym Name="hello" ID="SynonymId_1">
<Filename>synonym_hello1.dic</Filename>
<Timestamp>1169766691776</Timestamp>
</Synonym>
<Synonym Name="goodbye" ID="SynonymId_2">
<Filename>synonym_goodbye2.dic</Filename>
<Timestamp>1169767224839</Timestamp>
</Synonym>
</Synonyms>
</SynonymConfiguration>
```

Gdy słowniki zostają przesłane, system przypisuje im unikalne identyfikatory pliku. W powyższym przykładzie tymi identyfikatorami są `synonym_hello1.dic` oraz `synonym_goodbye2.dic`. Ścieżka bezwzględna dla tych plików na serwerze indeksowania jest następująca:

- Ścieżka do słownika synonimów to `ES_NODE_ROOT/data/custom_dictionary/synonym_*.dic`.
- Ścieżka do słownika słów pomijanych to `ES_NODE_ROOT/data/custom_dictionary/stopwordDictionary_*.dic`.
- Ścieżka do słownika kryteriów to `ES_NODE_ROOT/data/custom_dictionary/boostingwordDictionary_*.dic`.

W systemach AIX, Linux i Solaris listę dostępnych słowników można uzyskać za pomocą komendy `ls`. Na przykład:

```
% ls -l $ES_NODE_ROOT/data/custom_dictionary/synonym_*.dic
```

```
-rw-rw-r-- 1 esuser users 9 Jan 25 15:11 /home/esuser/node/data/custom_dictionary/synonym_hello1.dic
-rw-rw-r-- 1 esuser users 9 Jan 25 15:11 /home/esuser/node/data/custom_dictionary/synonym_goodbye2.dic
```

2. Po zidentyfikowaniu słownika, który ma zostać zaktualizowany, zatrzymaj system wyszukiwania korporacyjnego.
3. Słownik, który ma zostać zaktualizowany, nadpisz nowym plikiem słownika. Na przykład nadpisz plik `synonym_hello1.dic` lub `synonym_goodbye2.dic`.
4. Jeśli system wyszukiwania korporacyjnego działa na dwóch lub na czterech serwerach, plik słownika nadpisz ręcznie także na tych serwerach wyszukiwania. Pliki słowników znajdują się w tej samej ścieżce, w której znajdują się na serwerze indeksowania (`ES_NODE_ROOT/data/custom_dictionary/`).
5. Uruchom ponownie system wyszukiwania korporacyjnego.

#### Pojęcia pokrewne

“Niestandardowe słowniki synonimów” na stronie 194

“Niestandardowe słowniki słów ignorowanych” na stronie 197

“Niestandardowe słowniki kryteriów” na stronie 209

---

## Podsumowywanie dynamiczne

Dynamiczne podsumowywanie to technika umożliwiająca określenie w dokumencie wynikowym fraz, które najlepiej reprezentują pojęcia wyszukiwane przez użytkownika.

W wyszukiwaniu korporacyjnym technika dynamicznego podsumowywania polega na próbach przechwycenia w dokumentach zdań, w których występuje wiele poszukiwanych wyrazów. Wybieranych jest kilka zdań lub części zdań, które są wyświetlane w wynikach wyszukiwania. Poszukiwane wyrazy są podświetlone w dokumencie HTML z wynikami wyszukiwania.

Podczas konfigurowania opcji serwera wyszukiwania dla kolekcji można określić maksymalną długość podsumowań dokumentów wyświetlanych w wynikach wyszukiwania. Ponieważ podsumowanie obejmuje także znaki podświetlone, bufor zwrócony do aplikacji wyszukującej będzie większy niż określona wartość maksymalna. Jednak wyświetlana długość nie przekroczy określonej wartości maksymalnej. Podsumowanie może też być krótsze, co zależy od danych podsumowania wyodrębnionych z dokumentu źródłowego.

## Dostosowywanie podsumowań dokumentów w konsoli administracyjnej

Można dostosować zakres informacji wyświetlanych w podsumowaniach dokumentów, określając opcje dla serwera wyszukiwania w konsoli administracyjnej wyszukiwania korporacyjnego.

### Zagadnienia wstępne

Aby sterować wyświetlaną długością podsumowań dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora danej kolekcji.

### Informacje o tym zadaniu

Wartość określona dla maksymalnej wyświetlanej długości podsumowań dokumentów jest powiązana z wartością określoną dla liczby zdań, które może zawierać każde podsumowanie. Pierwszeństwo ma wartość, która daje najkrótsze podsumowanie dokumentu.

Jeśli na przykład zostanie określony limit czterech zdań, podsumowanie dokumentu będzie zawierać tylko cztery zdania, nawet jeśli wyświetlana długość dopuszcza większą liczbę znaków niż łączna liczba znaków w tych zdaniach. Kolejny przykład: limit 10 zdań w połączeniu z limitem wyświetlanej długości równym 500 znaków może w efekcie dać podsumowanie dokumentu zawierające mniej niż 10 zdań.

### Procedura

Aby skonfigurować wyświetlaną długość podsumowań dokumentów, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. Na stronie Opcje serwera wyszukiwania określ maksymalną wyświetlaną długość podsumowań dokumentów. Podczas przeglądania wyników wyszukiwania przez użytkowników podsumowania dokumentów nie będą przekraczały określonej wielkości.
3. Określ liczbę zdań, które może zawierać każde podsumowanie dokumentu (podsumowania mogą zawierać do 10 zdań).
4. Kliknij przycisk **OK**.
5. Aby zmiany zostały uwzględnione, przejdź do monitorowania serwerów wyszukiwania i zrestartuj procesy serwera.

## Dostosowywanie podsumowań dokumentów za pomocą edycji właściwości

Każdy dokument wynikowy zwracany w wyniku zapytania wyszukiwania korporacyjnego zawiera podsumowanie. Można zmienić zakres informacji, które zawiera każde podsumowanie, edytując plik właściwości.

### Informacje o tym zadaniu

Opisy wyników wyszukiwania można dostosować, zmieniając wartości następujących właściwości w pliku `ES_NODE_ROOT/master_config/id_kolekcji.runtime.node1/runtime-generic.properties`:

#### **MinWordsPerSentence**

Minimalna liczba słów w każdym ze zdań w podsumowaniu. Do podsumowania włączane są krótsze zdania, jeśli nie ma dostatecznej liczby zdań mających więcej słów niż określa to wartość `MinWordsPerSentence`. Wartością domyślną jest 4.

#### **MaxWordsPerSentence**

Maksymalna liczba słów w każdym ze zdań, które zostanie włączone do podsumowania. Jeśli zdanie ma więcej słów niż określono w tym ograniczeniu, to do podsumowania włączana jest tylko część zdania (część zawierająca terminy użyte w zapytaniu, w liczbie nie przekraczającej wartości `MaxWordsPerSentence`). Reszta tego zdania jest pomijana. Wartością domyślną jest 20.

Zdania włączane do podsumowania wybierane są w oparciu o wewnętrzny algorytm, określający istotność wszystkich zdań zawierających wyszukiwane terminy. Wybór na podstawie istotności dokonywany jest przed odrzuceniem zdań, których długość przekracza maksymalną liczbę słów.

#### **NumberOfReturnedSentences**

Liczba zdań, które tworzą opis dokumentu. Wartością domyślną jest 5.

#### **MaxSentencesPerDocument**

Maksymalna liczba zdań w dokumencie, które będą brane pod uwagę w procesie tworzenia opisu. Wartością domyślną jest 1000.

### **Procedura**

Aby dostosować podsumowania dokumentów w wynikach wyszukiwania, należy wykonać następujące czynności:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerach wyszukiwania. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Korzystając z edytora tekstu zmień następujący plik, gdzie zmienna `id_kolekcji` jest identyfikatorem, który został określony (lub przypisany przez system) podczas tworzenia kolekcji:

```
ES_NODE_ROOT/master_config/id_kolekcji.runtime.node1/
runtime-generic.properties
```

**Wskazówka:** Informację o tym, jaki identyfikator kolekcji jest przypisany do danej nazwy kolekcji, znajdziesz w pliku `ES_NODE_ROOT/master_config/collections.ini`.

3. Zmień właściwości, które należy dostosować, a następnie zapisz i zamknij plik.
4. Zatrzymaj i zrestartuj serwery wyszukiwania, aby zastosować wprowadzone zmiany.

---

## **Praca z szybkimi odsyłaczami**

Szybkie odsyłacze są to dokumenty zwracane w wynikach wyszukiwania, gdy użytkownik wprowadzi zapytanie, w którym występują konkretne słowa i frazy.

Do konfigurowania szybkich odsyłaczy dla kolekcji służy konsola administracyjna wyszukiwania korporacyjnego.

### **Szybkie odsyłacze**

Użycie szybkich odsyłaczy umożliwia udostępnienie użytkownikom odsyłaczy do dokumentów, które wstępnie zaklasyfikowano jako związane z tematem wyszukiwania.

Szybki odsyłacz jest to identyfikator URI umieszczany automatycznie przez mechanizm wyszukiwania korporacyjnego w wynikach wyszukiwania, gdy w zapytaniu występują określone słowa lub frazy. Zazwyczaj identyfikatory URI szybkich odsyłaczy są wyświetlane na początku listy wynikowej w celu zapewnienia użytkownikom łatwego dostępu do dokumentów, które zostały wstępnie określone jako odpowiednie dla danego zapytania.

Szybkie odsyłacze są zwracane dodatkowo oprócz pozostałych wyników wyszukiwania. Procesy wyszukiwania przeszukują indeksy w celu znalezienia dokumentów zgodnych z terminami określonymi w zapytaniu i oprócz identyfikatorów URI szybkich odsyłaczy zwracają identyfikatory URI tych dokumentów.

Podczas konfigurowania szybkiego odsyłacza można dla każdego identyfikatora URI określić opisowy tytuł i podsumowanie, aby ułatwić użytkownikom rozpoznanie dokumentów i podjęcie szybkiej decyzji, czy określony dokument ma zostać pobrany.

Na przykład dla identyfikatora URI <http://www.ibm.com/education/us/> może zostać użyty tytuł taki, jak **Szkolenia firmy IBM w Stanach Zjednoczonych**, a także następujące podsumowanie: **Rozwiązania, produkty i zasoby dla profesjonalistów, instruktorów i studentów w Stanach Zjednoczonych**.

Aby można było używać szybkich odsyłaczy w kolekcjach wyszukiwania korporacyjnego, w aplikacji wyszukującej musi być dostępna opcja wyświetlania szybkich odsyłaczy. W niektórych aplikacjach wyszukujących użytkownicy mogą mieć możliwość włączenia i wyłączenia funkcji zwracania szybkich odsyłaczy podczas wyszukiwania kolekcji.

## Konfiguracja szybkich odsyłaczy

Aby utworzyć szybki odsyłacz do kolekcji wyszukiwania korporacyjnego, należy powiązać identyfikator URI dokumentu ze słowami kluczowymi, które spowodują jej włączenie do wyników wyszukiwania.

### Zagadnienia wstępne

Aby skonfigurować szybkie odsyłacze, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą te odsyłacze.

### Informacje o tym zadaniu

Przykłady określania słów kluczowych i identyfikatorów URI dla szybkich odsyłaczy dostępne są po kliknięciu przycisku **Pomoc** w trakcie tworzenia lub edytowania szybkiego odsyłacza.

Nie ma potrzeby restartowania serwerów wyszukiwania, aby zastosować wprowadzone zmiany.

### Procedura

Aby skonfigurować szybki odsyłacz, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj szybkie odsyłacze**.
2. Na stronie Szybkie odsyłacze kliknij opcję **Utwórz szybki odsyłacz**.
3. Określ słowa kluczowe i frazy, dzięki którym ten szybki odsyłacz będzie zwracany w wynikach wyszukiwania, identyfikator URI dokumentu określonego uprzednio jako odpowiedni dla tego zapytania oraz opcje szybkiego odsyłacza.

W jednej linii można określić jedno lub kilka słów kluczowych bądź jedną frazę (dwa lub więcej słów ujętych w cudzysłów). Słowa kluczowe należy rozdzielać spacją (nie należy oddzielać słów kluczowych za pomocą przecinka). Naciśnij klawisz Enter, aby rozpocząć nowy wiersz.

4. Kliknij przycisk **OK**.

Nowy szybki odsyłacz jest wyświetlany - wraz z innymi należącymi do kolekcji - na stronie Szybkie odsyłacze.

**Odsyłacze pokrewne**

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 120



---

## Ranking dokumentów

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

Serwery wyszukiwania obsługują wzbogaconą składnię zapytań i korzystają z kilku technik zapewniających zwrócenie najlepiej dopasowanych wyników wyszukiwania, takich jak ocenianie wyników na podstawie tekstu i ocena statyczna. Można rozszerzyć domyślny sposób oceny, konfigurując opcje wpływające na ważność dokumentów w wynikach wyszukiwania:

- Można utworzyć niestandardowe słowniki kryteriów, aby wpłynąć na sposób oceny dokumentów zawierających określone słowa-kryteria w wynikach wyszukiwania.
- Można wpływać na wyniki dokumentów zgodnych z określonym wzorcem URI.
- Można wpływać na wyniki dokumentów zawierających pola odwzorowane na klasy kryteriów.

### Pojęcia pokrewne

“Ocena dokumentów na podstawie klas kryteriów” na stronie 213

“Niestandardowe słowniki kryteriów” na stronie 209

---

## Wartościowanie danych tekstowych

Wyszukiwanie korporacyjne w sposób dynamiczny oblicza wynik dla każdego dokumentu, który spełnia warunki zapytania.

Wynik wartościowania tekstu reprezentuje ważność znajdujących się w dokumencie terminów użytych w zapytaniu. Przy obliczaniu wyniku tekstowego poszczególnych dokumentów spełniających warunki zapytania wyszukiwanie korporacyjne uwzględnia wiele czynników, takich jak:

- Czy terminy wyróżniają dokument spośród innych dokumentów. Na przykład, jeśli terminy użyte w zapytaniu obecne są w jednym dokumencie, ale nie w innych, oznacza to, że terminy te są ważne dla danego dokumentu i wyróżniają go spośród innych dokumentów. Terminy użyte w zapytaniu obecne w większości dokumentów w mniejszym stopniu wpływają na wynik wartościowania dokumentu niż te, które obecne są w węższym zestawie dokumentów.
- Liczba wystąpień w dokumencie terminów użytych w zapytaniu. Wynik wartościowania terminu jest obliczany na podstawie każdego wystąpienia danego terminu w dokumencie. Im więcej w dokumencie wystąpień terminów użytych w zapytaniu, tym wyższy jest wynik wartościowania dokumentu. Gdy na przykład wyszukiwany jest termin *thinkpad*, dokument, w którym termin ten występuje wiele razy, otrzyma wyższą ocenę w wynikach wyszukiwania niż inne dokumenty, które zawierają mniej wystąpień.
- W przypadku dokumentów HTML uwzględniane są atrybuty terminu użytego w zapytaniu (czyli kontekst terminu). Wynik wartościowania terminu jest obliczany na podstawie atrybutów danego terminu (takich jak położenie, pogrubienie czcionki, kursywa, zakotwiczenie itp.). Na ogół terminy występujące w tytule dokumentu mają wyższy wynik wartościowania niż terminy występujące w tekście akapitu. Terminy wyróżnione (na przykład przez pogrubienie tekstu) mają wyższy wynik wartościowania niż terminy w formie zwykłego tekstu. Ważność atrybutów można skonfigurować w ustawieniach analizatora składni.
- Waga terminów użytych w zapytaniu. Aby dostosować ważność terminów w dokumencie, można skonfigurować wartości kryteriów dla terminów. W tym przypadku należy

powiązać słownik kryteriów (zawierający terminy oraz ich wartości kryteriów) z kolekcją. Słownik ten jest używany podczas wyszukiwania, a wartości kryteriów terminów w słowniku wpływają na wynik wartościowania dokumentu. Im wyższa wartość kryterium, tym większy wpływ terminu na wynik wartościowania dokumentu.

- Bliskość w dokumencie terminów użytych w zapytaniu. Jeśli terminy użyte w zapytaniu występują w dokumencie blisko siebie, ich bliskość leksykalna jest uwzględniana przy obliczaniu wyniku wartościowania tekstu. Na przykład założmy, że mamy dwa dokumenty. Jeden dokument mówi o parkingu samochodowym w centrum (słowa parking i samochodowy są blisko siebie). Drugi dokument mówi o salonie samochodowym w centrum obok parkingu (słowa parking i samochodowy nie są blisko siebie). W przypadku wyszukiwania frazy **parking samochodowy** bliskość terminów w pierwszym dokumencie sprawi, że znajdzie się on w rankingu wyżej niż drugi dokument.
- Czynniki uwzględnianymi przy określaniu wyniku wartościowania dokumentu są również długość każdego z dokumentów oraz bogactwo słownictwa (na przykład liczba unikalnych słów).

---

## Ocena statyczna

W przypadku niektórych typów dokumentów możliwe jest przypisanie statycznego współczynnika oceniania, który zwiększa jego ważność w wynikach wyszukiwania.

Podczas tworzenia kolekcji określone są opcje **Ważność dokumentu**. Wybrany typ ważności dokumentu decyduje o tym, czy czynnik rankingu statycznego zostaje powiązany z dokumentami w kolekcji. Gdy użytkownik prowadzi wyszukiwanie w kolekcji, w której używany jest ranking statyczny, czynnik rankingu statycznego wpływa na ocenę nadawaną dokumentom zwracanym w wynikach wyszukiwania.

W przypadku treści WWW czynnik rankingu statycznego jest oparty na odsyłaczach. Liczba odsyłaczy do dokumentu z innych dokumentów oraz źródła tych odsyłaczy mogą zwiększyć istotność danego dokumentu w wynikach wyszukiwania.

W przypadku dokumentów zawierających pola daty lub metadane z datą czynnik rankingu statycznego jest oparty na dacie dokumentu. Pole daty dokumentu dostarczane przez przeszukiwacz może być datą ostatniej modyfikacji dokumentu lub datą ostatniego przeszukiwania, zależnie od tego, jak przeszukiwacz został skonfigurowany.

Data dokumentu może podnosić jego istotność. Na przykład nowsze artykuły w grupach dyskusyjnych NNTP mogą być bardziej adekwatne niż artykuły starsze. Jeśli źródło danych zawiera wiele wartości daty, podczas konfigurowania przeszukiwacza można wybrać, która z nich jest najważniejsza dla określenia istotności dokumentów.

Korzystając z oceny statycznej w odniesieniu do kolekcji, należy pamiętać, aby nie mieszać ze sobą w ramach jednej kolekcji źródeł danych stosujących różne typy ocen. Na przykład, jeśli statycznym współczynnikiem oceniania mają być odsyłacze do dokumentu, kolekcja powinna zawierać wyłącznie dokumenty WWW. Gdy źródła o różnych modelach rankingowych występują razem w tej samej kolekcji, ranking dokumentu jest mniej adekwatny, a porządek wyników wyszukiwania może być inny od oczekiwanego.

Należy także zadbać o to, aby dokumenty w kolekcji zawierały pola i wartości umożliwiające stosowanie rankingu statycznego. Wyobraźmy sobie na przykład kolekcję, która jest skonfigurowana tak, aby używany był ranking statyczny oparty na datach dokumentów, natomiast przeszukiwacz w kolekcji jest skonfigurowany tak, aby data dokumentu była ustalana na podstawie konkretnego pola. Jeśli dokument nie zawiera tego pola, ważność dokumentu może nie otrzymać odpowiedniej oceny rankingowej, a porządek wyników wyszukiwania może być inny od oczekiwanego.

## Implikacje rankingu opartego na odsyłaczach

Ranking statyczny, wraz z takimi czynnikami, jak przypisanie wyniku wartościowania do wzorców URI kryteriów, składa się na statyczny wynik wartościowania dokumentu i wpływa na ważność dokumentu. Model rankingu opartego na odsyłaczach jest zwykle stosowany do kolekcji WWW, ponieważ w modelu tym obliczana jest statyczna ocena rankingowa dokumentu oparta na liczbie odsyłaczy do danego dokumentu. Dokument, do którego prowadzą odsyłacze z większej liczby innych dokumentów, zostaje oceniony w rankingu jako istotniejszy.

Dlatego, jeśli ten model zostanie skonfigurowany dla kolekcji innej niż kolekcja WWW lub dla kolekcji mieszanej (czyli zawierającej dokumenty WWW oraz dokumenty innego rodzaju), jakość wyszukiwania może ulec pogorszeniu, ponieważ w przypadku dokumentów innych niż dokumenty WWW koncepcja odsyłaczy nie ma zastosowania.

Gdy włączony jest ranking oparty na odsyłaczach, włączone jest również wykrywanie duplikatów dokumentów. Duplikaty dokumentów mają taką samą statyczną ocenę rankingową jak dokument wzorcowy. Jeśli czynniki kryteriów wzorca URI nie są skonfigurowane dla żadnych dokumentów w grupie duplikatów, to wszystkie zduplikowane dokumenty otrzymują ten sam statyczny wynik wartościowania.

---

## Odtwarzanie domyślnych wartości w rankingu dokumentów statycznych

Jeśli w czasie tworzenia kolekcji skonfigurowane są opcje rankingu dokumentów statycznych, to edytując związane z tą kolekcją pliki `runtime.properties`, można przywrócić wartości domyślne właściwości.

### Zagadnienia wstępne

Aby odtworzyć domyślne wartości rankingu dokumentów kolekcji, trzeba być administratorem wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Aby odtworzyć domyślne wartości rankingu dokumentów kolekcji, należy zaktualizować pliki `runtime.properties` związane z kolekcją i ze wszystkimi serwerami wyszukiwania systemu wyszukiwania korporacyjnego. W konfiguracji z wieloma serwerami plik `runtime.properties` znajduje się na serwerze indeksowania w katalogu `ES_NODE_ROOT/master_config/id_kolekcji.runtime.id_węzła`, gdzie `id_kolekcji` jest identyfikatorem danej kolekcji, a `id_węzła` jest identyfikatorem serwerów wyszukiwania.

Na przykład podczas aktualizowania kolekcji `col1` w systemie wyszukiwania korporacyjnego z wieloma serwerami należy zaktualizować pliki `runtime.properties` związane z tą kolekcją oraz z obydwojema serwerami wyszukiwania (`node3` i `node4`):

```
ES_NODE_ROOT/master_config/col1.runtime.node3/runtime.properties
ES_NODE_ROOT/master_config/col1.runtime.node4/runtime.properties
```

### Procedura

Aby odtworzyć domyślne wartości rankingu dokumentów danej kolekcji:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego.
2. Odszukaj identyfikator kolekcji, dla której chcesz odtworzyć domyślne wartości rankingu dokumentów. Identyfikator kolekcji zapisany jest w pliku `ES_NODE_ROOT/`

master\_config/collections.ini. W celu łatwiejszego przeglądania posortuj zawartość pliku. W poniższym przykładzie identyfikatorem kolekcji jest col1:

```
% sort $ES_NODE_ROOT/master_config/collections.ini | more
collection1.configfile=col1_config.ini
collection1.datadir=/home/esearch/node/data/col1
collection1.description=
collection1.displayname=Collection1
collection1.flags=0
collection1.id=col1
collection1.sectiontype=collection
collection1.type=1
...
```

3. Dokonaj edycji pliku `runtime.properties` związanego z tą kolekcją i wprowadź do niego następujące zmiany:

- a. Usuń następujące właściwości:

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- b. Jeśli w pliku `runtime.properties` znajduje się zapis `trevi.sourcetype=1` oznaczający, że ranking dokumentów układany jest według odsyłaczy, to dokonaj edycji pliku `ES_INSTALL_ROOT/default_config/runtime.1/runtime.properties` i skopiuj poniższe właściwości, a następnie wklej je do pliku `runtime.properties`:

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- c. Jeśli w pliku `runtime.properties` znajduje się zapis `trevi.sourcetype=2` oznaczający, że ranking dokumentów układany jest według dat, to dokonaj edycji pliku `ES_INSTALL_ROOT/default_config/runtime.2/runtime.properties` i skopiuj poniższe właściwości, a następnie wklej je do pliku `runtime.properties`:

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- d. Jeśli w pliku `runtime.properties` znajduje się zapis `trevi.sourcetype=3` oznaczający, że ranking dokumentów kolekcji nie jest układany według czynników statycznych, to dokonaj edycji pliku `ES_INSTALL_ROOT/default_config/runtime.0/runtime.properties` i skopiuj poniższe domyślne właściwości, a następnie wklej je do pliku `runtime.properties`:

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

4. W konfiguracji z wieloma serwerami powtarzaj krok 3, aktualizując plik `runtime.properties` tej samej kolekcji na drugim serwerze wyszukiwania.
5. W konsoli administracyjnej monitoruj stronę Szukaj i ponownie uruchom procesy wyszukiwania dla tej kolekcji.

Jeśli to konieczne, powtarzaj te kroki dla każdej kolekcji, w której chcesz przywrócić domyślne wartości rankingów dokumentów.

---

## Niestandardowe słowniki kryteriów

Aby podnieść jakość wyników wyszukiwania, można wpłynąć na oceny dokumentów w wynikach wyszukiwania przez utworzenie niestandardowego słownika kryteriów.

Jeśli w zapytaniu określono słowo występujące w słowniku kryteriów, ważność dokumentów zawierających to słowo zostanie zwiększona lub zmniejszona zgodnie ze współczynnikiem kryterium skonfigurowanym dla tego słowa w słowniku.

Aby zapewnić zwracanie określonych dokumentów, gdy użytkownik określi w zapytaniu konkretne terminy, można użyć słownika kryteriów. Na przykład założmy, że mamy kolekcję zawierającą wiele dokumentów na temat samochodów. W przypadku takich dokumentów można przyjąć, że istotne są pewne słowa kluczowe odnoszące się do modeli samochodów, nazwy producenta itp. Aby wpłynąć na ocenę wyników wyszukiwania, można przypisać istotność, kojarząc wartość kryterium ze słowami kluczowymi (model, producent itd.) w słowniku kryteriów. Gdy użytkownik, prowadząc wyszukiwania w kolekcji, określi zapytanie zawierające dowolne z tych słów kluczowych, dokumenty o samochodach otrzymają wyższą ocenę w wynikach wyszukiwania niż pozostałe dokumenty.

Współczynniki kryteriów mają wartości z zakresu od -10 do 10. Podczas przetwarzania zapytania serwery wyszukiwania zwiększają ważność dokumentów zawierających słowa z dodatnimi współczynnikami kryteriów i zmniejszają ważność dokumentów zawierających słowa z ujemnymi współczynnikami kryteriów.

Na przykład dokument zgodny z użytymi w zapytaniu terminami o wysokich współczynnikach kryteriów zostanie oceniony wyżej niż w przypadku, gdy współczynnik kryterium nie zostałby zastosowany (współczynnik kryterium jest tylko jednym z czynników, które wpływają na wynik dokumentu).

Tworząc słownik, można przypisać ten sam współczynnik kryterium do dowolnej liczby słów. Słownik może zawierać termin składający się z jednego słowa lub wielu słów. Terminy składające się z wielu słów są dopasowywane jako frazy.

Jeśli słowo podlegające wartości kryterium jest określone w zapytaniu z użyciem operatora OR (na przykład: `jeden | dwa`), dla terminów użytych w zapytaniu obliczana jest średnia ważona. Uzyskany w ten sposób zagregowany wynik jest używany dla wszystkich wystąpień operandów zapytania OR. Różne wyniki nie są liczone dla różnych operandów zapytania OR.

Wpływanie na wyniki przy użyciu słowników kryteriów nie jest obsługiwane w przypadku terminów użytych w zapytaniu w odniesieniu pól. Podczas analizowania terminów użytych w zapytaniu do obliczenia wyniku dokumentu używany jest wyłącznie tekst zapytania, a nie nazwa pola. Aby zastosować współczynniki kryterium do terminów użytych w zapytaniu, które występują w polach, można odwzorować nazwy pól na klasy kryteriów.

Aby utworzyć słownik kryteriów, ekspert w dziedzinie, której dotyczy kolekcja musi utworzyć listę słów-kryteriów w formacie XML lub we współpracy z twórcą aplikacji utworzyć plik XML. W celu konwersji pliku XML do pliku binarnego (.dic) należy użyć narzędzia wyszukiwania korporacyjnego **esboostermdictbuilder**.

Administrator wyszukiwania korporacyjnego przesyła plik binarny do systemu i przypisuje mu nazwę wyświetlaną. Administratorzy kolekcji mogą wybrać słownik kryteriów używany przy wyszukiwaniu dokumentów w kolekcji podczas konfigurowania opcji serwera wyszukiwania dla kolekcji.

**Ograniczenie:** Niestandardowy słownik kryteriów nie może być już modyfikowany po dodaniu go do systemu. Aby zweryfikować listę słów-kryteriów dostępnych do przetwarzania zapytań, należy:

1. Zmodyfikować odpowiednio źródłowy plik XML.
2. Przekształcić źródło XML na nowy plik słownika.
3. Usunąć poprzedni słownik kryteriów z kolekcji, które go używają.
4. Usunąć ten słownik kryteriów z systemu.
5. Dodać do systemu nowy słownik kryteriów.
6. Powiązać nowy słownik kryteriów z kolekcjami, które mają z niego korzystać.

Można napisać skrypt obejmujący te kroki, a następnie za pomocą tego skryptu ponownie wdrożyć słownik w systemie wyszukiwania korporacyjnego.

#### **Pojęcia pokrewne**

“Ranking dokumentów” na stronie 205



Niestandardowe słowniki kryteriów

#### **Zadania pokrewne**



Tworzenie pliku XML słów-kryteriów



Tworzenie słownika kryteriów

“Ponowne wdrażanie słowników niestandardowych” na stronie 199

## **Dodawanie do systemu słowników kryteriów**

Po utworzeniu niestandardowych słowników kryteriów należy je powiązać z systemem wyszukiwania korporacyjnego. Później można wybrać słownik kryteriów, który ma być używany podczas wyszukiwania w kolekcji.

#### **Zagadnienia wstępne**


Do dodawania niestandardowych słowników kryteriów do systemu wymagana jest rola administratora wyszukiwania korporacyjnego.

#### **Ograniczenia**

Maksymalna wielkość słownika kryteriów to 8 MB.

#### **Procedura**

Aby powiązać niestandardowe słowa-kryteria z systemem wyszukiwania korporacyjnego:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby zmienić systemowy widok edycji.
3. Na stronie Wyszukiwanie kliknij opcję **Konfiguruj słowniki kryteriów**.
4. Na stronie Konfigurowanie słowników kryteriów kliknij opcję **Dodaj słownik kryteriów**.
5. Na stronie Dodawanie słownika kryteriów wpisz unikalną nazwę wyświetlaną słownika i opcjonalnie podaj jego opis.
6. Określ położenie pliku .dic. Jeśli plik znajduje się w systemie lokalnym, można go znaleźć, przeglądając system plików. Jeśli plik znajduje się na serwerze indeksowania, należy wpisać pełną ścieżkę do tego pliku.
7. Kliknij przycisk **OK**. Niestandardowy słownik kryteriów zostanie dodany do systemu wyszukiwania korporacyjnego i będzie dostępny na potrzeby wyszukiwania w kolekcjach.



## Powiązanie słownika kryteriów z kolekcją

Jeśli z systemem wyszukiwania korporacyjnego są powiązane słowniki kryteriów, można wybrać jeden z nich do użycia podczas wyszukiwania w kolekcji. Jeśli termin użyty w zapytaniu odpowiada terminowi ze słownika, wówczas ważność dokumentów zawierających ten termin zostanie zwiększona lub zmniejszona zgodnie ze współczynnikiem kryterium przypisanym do danego terminu w słowniku.

### Zagadnienia wstępne

Do wybrania słownika kryteriów dla kolekcji wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora do danej kolekcji.

### Procedura

Aby powiązać słownik kryteriów z kolekcją:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. W polu **Nazwa słownika kryteriów** na stronie Opcje serwera wyszukiwania wybierz słownik kryteriów, który ma być używany podczas wprowadzania przez użytkowników zapytań dla tej kolekcji.  
Lista dostępnych słowników zawiera wszystkie słowniki kryteriów dodane do systemu wyszukiwania korporacyjnego.
3. Kliknij przycisk **OK**.

---

## Ocena dokumentów na podstawie wzorców URI

Ważność dokumentów można zwiększyć lub zmniejszyć, przypisując współczynniki kryteriów do wzorców URI.

Do wszystkich dokumentów dodawanych do indeksu przypisywane są domyślne wyniki oceny statycznej. Domyślny wynik różni się w zależności od tego, czy ocena statyczna została włączona dla kolekcji, a jeśli tak, od typu oceny statycznej (według daty dokumentu lub, dla dokumentów sieci WWW, liczby innych dokumentów zawierających odsyłacze do danego dokumentu).

Użytkownik może wpływać na względną ważność dokumentu, przypisując współczynniki kryteriów do wzorców URI. Współczynnik kryterium jest używany wraz z domyślnym wynikiem oceny statycznej i innymi czynnikami do określenia ostatecznego wyniku statycznego dokumentu.

Istotne znaczenie ma kolejność konfigurowanych wzorców URI. Podczas obliczania wartości każdego dokumentu w kolekcji serwer indeksowania ocenia wzorce URI w kolejności, w jakiej zostały podane. Dla każdego identyfikatora URI:

1. Serwer indeksowania skanuje kolejno poszczególne wzorce URI.
2. Gdy serwer indeksowania znajdzie pierwszy wzorzec URI zgodny z dokumentem w indeksie, zastosuje do dokumentu współczynnik kryterium skonfigurowany dla tego wzorca URI.
3. Jeśli dokumentu nie uda się dopasować do wzorca URI, zostanie użyty domyślny wynik oceny statycznej.

Po skonfigurowaniu współczynnika kryterium dla wzorca URI należy określić kolejność, w jakiej będzie go przeglądał serwer indeksowania.



## Wpływanie na wyniki dokumentów zgodnych z wzorcami URI

Ważność dokumentów zgodnych z wzorcem URI można zwiększyć lub zmniejszyć, stosując współczynnik kryterium do domyślnego wyniku oceny statycznej.

### Zagadnienia wstępne

Aby zmienić ważność dokumentów zgodnych ze wzorcem URI, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora do danej kolekcji.

### Informacje o tym zadaniu

Skonfigurowany współczynnik kryterium jest używany wraz z domyślnym wynikiem oceny statycznej do obliczenia nowego wyniku statycznego dla wszystkich dokumentów zgodnych z określonym wzorcem URI.

Współczynniki kryteriów poprawiają tylko wyniki statyczne i stanowią tylko jeden z elementów wpływających na obliczenia określające ostateczną ocenę dokumentu. Jeśli na przykład istnieje wiele odsyłaczy do danego dokumentu (które dają wysoki wynik początkowy), inny dokument, do którego nie ma żadnych odsyłaczy, zostanie oceniony niżej.



### Procedura

Aby wpłynąć na wyniki dokumentów zgodnych z wzorcem URI:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Wpływaj na wyniki poprzez dopasowywanie wzorców identyfikatorów URI**.
2. Na stronie Oddziaływanie na wyniki poprzez dopasowywanie wzorców identyfikatorów URI kliknij opcję **Dodaj wzorzec URI**.
3. Wpisz wzorzec URI dla dokumentów, których ważność w wynikach wyszukiwania ma zostać zwiększona lub zmniejszona. Na przykład:  
`http://domena.org/kadry/*  
db2://*ROI*  
*/afs/*`
4. Wpisz wartość współczynnika kryterium z zakresu od -10 do 10. Ostateczny wynik statyczny dla wszystkich dokumentów zgodnych ze wzorcem URI zostanie obliczony na podstawie tego współczynnika kryterium.
5. Kliknij przycisk **OK**.
6. Na stronie Oddziaływanie na wyniki poprzez dopasowywanie wzorców identyfikatorów URI ustaw nowy wzorzec URI w takiej kolejności, w jakiej ma go przeglądać serwer indeksowania.

Serwer indeksowania oblicza wyniki oceny statycznej w kolejności, w jakiej wymienione są identyfikatory URI. Aby uzyskać najlepsze rezultaty, bardziej szczegółowe identyfikatory URI należy wymienić jako pierwsze. W poniższym przykładzie podkatalog /forms jest zgodny ze wzorcem identyfikatora URI `http://www.ibm.com/hr/*`. Aby mieć pewność, że wyniki dla dokumentów z podkatalogu /forms są obliczane poprawnie, wzorzec identyfikatora URI dla podkatalogu /forms należy umieścić jako pierwszy:

```
http://www.ibm.com/hr/forms/* 8
http://www.ibm.com/hr/* -2
```

7. Aby zmienić wzorzec identyfikatora URI lub współczynnik kryterium, wybierz wzorzec URI i kliknij opcję  **Edytuj**.
8. Aby usunąć z listy wzorzec URI, wybierz go i kliknij opcję  **Usuń**.
9. Zastosowanie współczynników kryterium w poindeksowanych już wcześniej dokumentach wymaga ich ponownego poindeksowania przez przeszukiwacza.

---

## Ocena dokumentów na podstawie klas kryteriów

Przez odwzorowanie pól na klasy kryteriów można wpłynąć na ocenę dokumentów w wynikach wyszukiwania.

Podczas analizowania dokumentów analizator składni przypisuje *klasy kryteriów* do leksemów dokumentu w zależności od pól, do których należą te leksemy. Te klasy kryteriów są włączane do indeksu i używane podczas wartościowania zapytania do obliczania wyników mających wpływ na ocenę zwracanych dokumentów.

Aby wpłynąć na sposób obliczania wyników, dla klas kryteriów można skonfigurować liczbowe współczynniki kryteriów. Jeśli termin użyty w zapytaniu jest zgodny z leksemem w polu odwzorowanym na klasę kryteriów, wkład tego wystąpienia leksemu wpływa na łączny wynik dokumentu. Wynik jest obliczany przez zastosowanie współczynnika kryterium skonfigurowanego w tej klasie kryteriów.

Można na przykład uwydatnić wyniki pól tytułu. Gdy termin użyty w zapytaniu wystąpi w tytule, takie wystąpienie będzie miało duży wkład w wynik tego dokumentu i pomoże mu w uzyskaniu lepszej oceny w wynikach wyszukiwania.

W celu oddziaływania na oceny dokumentów należy określić współczynniki kryteriów dla klas kryteriów oraz odwzorować pola na klasy kryteriów za pomocą konsoli administracyjnej wyszukiwania korporacyjnego. Na potrzeby wyszukiwania korporacyjnego udostępniono szesnaście predefiniowanych klas kryteriów. Osiem z nich zostało zaprojektowanych do użycia z polami treści, a drugie osiem dla pól metadanych. Wyniki powiązane z domyślnymi klasami kryteriów można edytować. Można również powiązać z klasami kryteriów inne lub dodatkowe pola.

W przypadku zmiany odwzorowań pól konieczne jest ponowne przeszukanie i przeanalizowanie dokumentów, aby zmiany zostały zastosowane do dokumentów, które zostały już wcześniej poindeksowane. Jeśli zmienione zostaną współczynniki określone dla klasy kryteriów, należy przejść do monitorowania serwerów wyszukiwania i zatrzymać, a następnie zrestartować procesy serwera wyszukiwania, aby zmiany te odniosły skutek.

### Wykrywanie zduplikowanych dokumentów i podsumowanie dokumentów

Odwzorowując pole na klasę kryteriów, należy określić, czy służy ono do wykrywania zduplikowanych dokumentów i czy zawartość tego pola może znaleźć się w podsumowaniu dokumentów w wynikach wyszukiwania.

- Jeśli pole jest używane do wykrywania zduplikowanych dokumentów, jest to pole treści i dostępne do wybrania są tylko klasy kryteriów zaprojektowane dla pól treści. Zawartość tego rodzaju pól może być używana w dynamicznych podsumowaniach dokumentów w wynikach wyszukiwania.
- Jeśli pole nie jest używane do wykrywania zduplikowanych dokumentów, jest to pole metadanych i dostępne do wybrania są tylko klasy kryteriów zaprojektowane dla pól metadanych. W tym przypadku za duplikaty uważa się dwa dokumenty, które są dokładnie takie same za wyjątkiem określonego pola, a pole to nie jest używane w dynamicznych podsumowaniach dokumentu.

## Wysoka i niska wartość znajdawalności

Podczas wartościowania zapytania proces wyszukiwania szacuje liczbę wynikowych dokumentów, które zostaną zwrócone. Wartości progowe określają wówczas, czy zapytanie jest uważane za mające znajdawalność niską, wysoką czy też znajdującą się w zakresie pomiędzy nimi:

### Niska wartość znajdawalności

Jeśli szacowana liczba dokumentów wynikowych jest mniejsza niż wartość dolnego progu, zapytanie jest uznawane za zapytanie o niskiej znajdawalności.

### Wysoka wartość znajdawalności

Jeśli szacowana liczba dokumentów wynikowych jest większa niż wartość górnego progu, zapytanie jest uznawane za zapytanie o wysokiej znajdawalności.

### Mieszana wartość znajdawalności

Jeśli szacowana liczba dokumentów znajduje się między wartościami obu progów, wartość znajdawalności tego zapytania jest połączeniem tych dwóch wartości progowych.

Każda klasa kryteriów określa współczynniki kryteriów powiązane podczas przetwarzania zapytań z zapytaniami o niskiej znajdawalności i zapytaniami o wysokiej znajdawalności. Niski współczynnik kryterium wpływa na względną ważność zapytań o niskiej znajdawalności, a wysoki współczynnik kryterium wpływa na względną ważność zapytań o wysokiej znajdawalności. Połączenie tych dwóch współczynników wpływa na względną ważność zapytań o mieszanej wartości znajdawalności.

Wartości współczynników kryteriów kontrolują względną ważność każdego wystąpienia terminu z zapytania w dokumencie. Każde wystąpienie w dokumencie terminu użytego w zapytaniu jest zliczane zgodnie z odpowiednim współczynnikiem kryterium.

Domyślne współczynniki kryteriów można edytować podczas konfigurowania klas kryteriów dla kolekcji. Na przykład można określić współczynniki kryteriów w taki sposób, aby terminy użyte w zapytaniu i występujące w polu tytułu liczyły się pięć razy bardziej niż terminy występujące w zwykłym tekście.

#### Pojęcia pokrewne

“Ranking dokumentów” na stronie 205

## Odzworowywanie pól na klasy kryteriów

Istnieje możliwość wpływania na względną ważność pól przez odzworowywanie nazw pól na klasy kryteriów.

### Zagadnienia wstępne

Do odzworowywania pól na klasy kryteriów wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora do danej kolekcji.



### Informacje o tym zadaniu

System używa współczynnika kryterium do wpływania na ocenę dokumentów zawierających terminy z zapytania w obrębie pól odzworowanych na klasy kryteriów.

Wyszukiwanie korporacyjne rezerwuje niektóre odzworowania pól wewnętrznych i regularnego tekstu, które nie mają innych charakterystyk definiujących. Na klasy kryteriów używane przez pola zarezerwowane można odzworowywać też inne pola, ale nie można edytować ani usuwać pól zarezerwowanych.

## Procedura

Aby odwzorować pola na klasy kryteriów:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij przycisk **Odwzoruj pola na klasy kryteriów**.
2. Na stronie Odwzorowywanie pól na klasy kryteriów kliknij przycisk **Dodaj pole**.
3. Na stronie Dodawanie pola do klasy kryteriów wpisz nazwę pola, które ma zostać odwzorowane na klasę kryteriów.  
Określić można nazwę pola istniejącego w przeszukiwanym źródle lub źródle zewnętrznym, nazwę pola odwzorowanego z elementu XML, nazwę pola odwzorowanego z elementu metadanych HTML lub jedną z predefiniowanych nazw pól.
4. Określ, czy to pole jest używane do wykrywania zduplikowanych dokumentów. Po zaznaczeniu tego pola wyboru lista dostępnych klas kryteriów zawiera klasy odnoszące się do pól treści. Jeśli dokument z tym polem zostanie zwrócony w wynikach wyszukiwania, zawartość pola zostanie wyświetlona w obszarze podsumowania dokumentu.  
Po anulowaniu zaznaczenia tego pola wyboru lista dostępnych klas kryteriów zawiera klasy odnoszące się do pól metadanych. Zawartość pola nie zostanie wyświetlona w obszarze podsumowania dokumentu w wynikach wyszukiwania.
5. Wybierz klasę kryteriów i kliknij przycisk **OK**.  
Dodane pole zostanie wyświetlone na stronie Odwzorowywanie pól na klasy kryteriów. Istnieje możliwość wyboru opcji służącej do edycji klasy kryteriów i konfigurowania innych współczynników kryteriów w celu określenia wyników dokumentów zawierających to pole.
6. Aby zmienić przeznaczenie pola z wykrywania zduplikowanych dokumentów na odwzorowanie na inną klasę kryteriów, kliknij przycisk  **Edytuj**. Nie można edytować pól zarezerwowanych do użycia przez wyszukiwanie korporacyjne.
7. Aby usunąć pole z klasy kryteriów, kliknij przycisk  **Usuń**. Nie można usuwać pól zarezerwowanych do użycia przez wyszukiwanie korporacyjne.
8. Aby zastosować zmiany w dokumentach, które zostały uprzednio poindeksowane, należy jeszcze raz je przeszukać i poindeksować.

### Pojęcia pokrewne

“Ranking dokumentów” na stronie 205

## Konfigurowanie współczynników kryteriów dla klas kryteriów

Współczynniki kryteriów konfigurowane dla klas kryteriów reprezentują oszacowanie odzwierciedlające, jak ważna jest dla danego zapytania obecność wybranych pól w dokumentach wynikowych. Klasy kryteriów z wysokimi współczynnikami kryteriów mogą zwiększać ważność dokumentów wynikowych zawierających pola odwzorowane na daną klasę kryteriów.

### Zagadnienia wstępne


Do konfigurowania współczynników kryteriów dla klas kryteriów wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora do danej kolekcji.

### Informacje o tym zadaniu

System używa współczynników kryteriów skonfigurowanych dla klasy kryteriów, domyślnego wyniku oceny statycznej oraz innych czynników do obliczenia nowego wyniku dla zwracanych dokumentów zawierających pola odwzorowane na tę klasę kryteriów.

## Procedura

Aby skonfigurować współczynniki kryteriów dla klas kryteriów:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij przycisk **Odwzoruj pola na klasy kryteriów**.
2. Na stronie Odwzorowywanie pól na klasy kryteriów kliknij przycisk **Edytuj klasy kryteriów**.
3. Na stronie Klasy kryteriów znajdź klasę kryteriów, którą chcesz zmienić i kliknij przycisk  **Edytuj**.
4. Na stronie Edycja klasy kryteriów określ nowe wartości dla wysokich i niskich współczynników kryteriów. Dla obu współczynników można podać te same wartości.
5. Kliknij przycisk **OK**.
6. Aby zmiany zostały zastosowane, przejdź do monitorowania serwerów wyszukiwania i wybierz ikony służące do zatrzymywania, a następnie restartowania procesów wyszukiwania. Przy wprowadzaniu zapytań przez użytkowników względna ważność dokumentów wynikowych zawierających pola odwzorowane na tę klasę kryteriów będzie określana przez nowe współczynniki kryteriów.

### Pojęcia pokrewne

“Ranking dokumentów” na stronie 205

## Domyślne wartości klas kryteriów

Wyszukiwanie korporacyjne udostępnia 16 klas kryteriów, za pomocą których można wpływać na oceny dokumentów w wynikach wyszukiwania.

Do obliczania wyników dla pól i tekstu pozbawionych dodatkowych definiujących je charakterystyk zarezerwowane zostały do użycia w wyszukiwaniu korporacyjnym następujące pola:

```
es_special_field.regular_text
es_special_field.default_field
es_special_field.default_metadata_field
```

Na klasy kryteriów używane przez pola zarezerwowane można odwzorowywać też inne pola, ale nie można edytować ani usuwać pól zarezerwowanych.

W przypadku wszystkich innych pól istnieje możliwość edytowania współczynników kryteriów używanych przez system do obliczania oceny dokumentu. Można również odwzorować dowolną liczbę pól na dowolną z klas kryteriów, w tym na klasy kryteriów używane przez pola zarezerwowane.

W poniższej tabeli przedstawiono listę nazw klas kryteriów, domyślnych współczynników kryteriów dla zapytań o niskiej wartości znajdawalności, domyślnych współczynników kryteriów dla zapytań o wysokiej wartości znajdawalności, a także nazw pól predefiniowanych odwzorowanych na klasy kryteriów w konfiguracji domyślnej.

Domyślne współczynniki kryteriów różnią się w zależności od metody oceny statycznej wybranej dla danej kolekcji podczas jej tworzenia. Dostępne opcje to brak oceny statycznej, ocena określana na podstawie liczby odsyłaczy do dokumentu (w przypadku źródeł WWW) oraz ocena określana na podstawie daty dokumentu.

Tabela 6. Domyślne wartości klas kryteriów

Domyślne niskie i wysokie współczynniki kryteriów				
Nazwa klasy kryteriów	Brak oceny statycznej	Odsyłacze do dokumentu	Data dokumentu	Odwzorowania pól predefiniowanych
Klasa treści A	Niski: 4 Wysoki: 2	Niski: 6 Wysoki: 1	Niski: 4 Wysoki: 2	es_special_field.regular_text
Klasa treści B	Niski: 5 Wysoki: 4	Niski: 7 Wysoki: 3	Niski: 5 Wysoki: 4	es_special_field.html_emphasized_text  Uwzględnia następujące elementy języka HTML: b, big, caption, dfn, em, h4, h5, h6, strong
Klasa treści C	Niski: 7 Wysoki: 4	Niski: 9 Wysoki: 3	Niski: 7 Wysoki: 4	es_special_field.html_headers  Uwzględnia następujące elementy języka HTML: h1, h2, h3
Klasa treści D	Niski: 2 Wysoki: 5	Niski: 1 Wysoki: 5	Niski: 2 Wysoki: 5	title
Klasa treści E	Niski: 1 Wysoki: 1	Niski: 5 Wysoki: 10	Niski: 1 Wysoki: 1	es_special_field.anchor
Klasa treści F	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.anchor_same_dir
Klasa treści G	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.anchor_same_host
Klasa treści H	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.default_field
Klasa metadanych A	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.default_metadata_field
Klasa metadanych B	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	
Klasa metadanych C	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	
Klasa metadanych D	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	
Klasa metadanych E	Niski: 1 Wysoki: 1	Niski: 5 Wysoki: 1	Niski: 1 Wysoki: 1	keywords
Klasa metadanych F	Niski: 1 Wysoki: 1	Niski: 3 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.urlhost
Klasa metadanych G	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.urlpath
Klasa metadanych H	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	opis

### Pojęcia pokrewne

“Ranking dokumentów” na stronie 205





---

## Aplikacje wyszukujące do wyszukiwania korporacyjnego

Aplikacja wyszukująca umożliwia wyszukiwanie w kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Liczba możliwych do utworzenia aplikacji wyszukujących nie jest ograniczona, a pojedyncza aplikacja wyszukująca pozwala przeszukać dowolną liczbę kolekcji i źródeł zewnętrznych.

### Przykładowa aplikacja wyszukująca

Przykładowa aplikacja wyszukująca zawiera wiele funkcji wyszukiwania i pobierania dostępnych dla wyszukiwania korporacyjnego. Przykładowa aplikacja wyszukiwania jest również działającym przykładem sposobów wykorzystania interfejsu IBM Search and Index API (SI-API) do tworzenia niestandardowych, interaktywnych aplikacji odwzorowujących cele danego przedsiębiorstwa.

Jeśli nie zmieniono właściwości w domyślnym pliku konfiguracyjnym, przykładowa aplikacja wyszukująca pozwala wyszukiwać we wszystkich aktywnych kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Dzięki temu przykładowa aplikacja wyszukująca może służyć do testowania nowych kolekcji i źródeł zewnętrznych przed udostępnieniem ich użytkownikom.

Przykładowa aplikacja wyszukująca jest automatycznie powiązana ze wszystkimi kolekcjami i źródłami zewnętrznymi. W środowisku produkcyjnym o tym, które aplikacje obsługują wyszukiwanie w poszczególnych kolekcjach, decydują administratorzy wyszukiwania korporacyjnego.

### Niestandardowe aplikacje wyszukujące

Aplikacje wyszukujące mogą być uruchamiane jako samodzielne aplikacje WWW w środowisku serwera IBM WebSphere Application Server albo jako portlety w środowisku IBM WebSphere Portal. Korzystając z funkcji API wyszukiwania i indeksowania, można projektować aplikacje wyszukujące, które będą bezproblemowo działały w obu środowiskach, podobnie jak przykładowa aplikacja wyszukująca.

Program Search Application Customizer można zastosować w celu ułatwienia sobie prac nad dostosowaniem aplikacji wyszukującej do własnych potrzeb. Program ten umożliwia dokonywanie wyborów w graficznym interfejsie użytkownika i przeglądanie efektów wszystkich wprowadzanych zmian. Podczas zapisywania wprowadzonych zmian aktualizowany jest plik konfiguracyjny aplikacji wyszukującej.

#### Wskazówka:




Szczegółowe przykłady używania Konfiguratora aplikacji wyszukującej oraz instalowania aplikacji wyszukującej jako portletu w portalu WebSphere Portal można znaleźć w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

#### Pojęcia pokrewne

“Indeksowane opcje wyszukiwania dokumentów” na stronie 179

“Identyfikatory aplikacji wyszukiwania” na stronie 260

“Zabezpieczenia na poziomie dokumentu” na stronie 261

-  Przegląd interfejsu API wyszukiwania i indeksowania
-  Składnia zapytania
-  Określanie właściwości zapytania

---

## Przypisywanie aplikacji wyszukiwujących do kolekcji

Przed rozpoczęciem korzystania z nowej aplikacji wyszukiwującej należy powiązać ją z kolekcjami, w których może ona wyszukiwać dokumenty.

### Zagadnienia wstępne

Aby powiązać aplikacje wyszukiwujące z kolekcjami, w których mogą prowadzić wyszukiwanie, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Procedura

Aby powiązać aplikację wyszukiwującą z jedną lub z większą liczbą kolekcji, należy wykonać poniższe czynności:

1. Kliknij opcję **Zabezpieczenia** na pasku narzędzi konsoli administracyjnej.
2. Na stronie Aplikacje wyszukiwujące kliknij opcję **Konfiguruj aplikacje wyszukiwujące**.
3. Na stronie Konfigurowanie aplikacji wyszukiwujących kliknij opcję **Dodaj aplikację wyszukiwującą**.
4. Wpisz nazwę aplikacji wyszukiwującej.
5. Wybierz kolekcje, w których aplikacja będzie mogła wyszukiwać:
  - Kliknij opcję **Wszystkie kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukiwująca ma mieć dostęp do wszystkich kolekcji dodanych do systemu.
  - Kliknij opcję **Konkretne kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukiwująca ma mieć dostęp tylko do określonych kolekcji.

Po wybraniu tej opcji zostanie wyświetlona lista nazw kolekcji i źródeł zewnętrznych. Zaznacz pole wyboru **Wybierz** dla każdej kolekcji, w której aplikacja będzie mogła wyszukiwać.
6. Kliknij przycisk **OK**.

---

## Funkcje przykładowej aplikacji wyszukiwującej

Przykładowa aplikacja wyszukiwująca przeznaczona do wyszukiwania korporacyjnego zawiera większość funkcji wyszukiwujących, jakie można wbudować w tworzone aplikacje.

Przykładowej aplikacji wyszukiwania można użyć do jednoczesnego wyszukiwania we wszystkich kolekcjach i źródłach zewnętrznych. O ile nie zostaną zmodyfikowane domyślne właściwości aplikacji, można jej użyć do wyszukiwania we wszystkich kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego.

### Funkcje operujące na zapytaniach

Te funkcje umożliwiają:

- Definiowanie prostych zapytań pełnotekstowych.
- Definiowanie bardziej złożonych zapytań w celu poprawienia dokładności wyników wyszukiwania. Na przykład można szukać określonych pól lub elementów XML albo wykorzystać składnię zapytań do szukania dokumentów zawierających konkretne słowa lub frazy albo niezawierających ich.

- Określenie, które kolekcje i źródła zewnętrzne należy przeszukać.
- Przeszukanie konkretnych typów źródeł lub wszystkich typów źródeł.
- Wyszukiwanie dokumentów określonych typów. Na przykład można wyszukać tylko dokumenty programu Microsoft Word lub tylko dokumenty w formacie PDF.
- Określanie języka, w jakim definiowane są warunki zapytania. Można również określić języki dokumentów, które należy wyszukać.
- Wyszukiwanie określonych podzbiorów kolekcji. Przykładem może być aplikacja wyszukująca, która ogranicza zawartość widoku do predefiniowanej grupy dokumentów (zasięg) albo umożliwia wprowadzanie zapytań wyszukujących tylko dokumenty należące do pewnej kategorii.
- Rozszerzanie zapytania o synonimy terminów z zapytania. Jeśli z kolekcją skojarzony jest słownik synonimów, w wynikach zapytania zwrócone zostaną dokumenty zawierające synonimy terminów użytych w zapytaniu.

## Funkcje operujące na wynikach wyszukiwania

Te funkcje umożliwiają:

- Wyświetlenie wyników wyszukiwania zgodnych z terminami użytymi w zapytaniu.
- Określanie liczby dokumentów wynikowych wyświetlanych na każdej stronie, a także przeglądanie tabeli wynikowej w obie strony.
- Ukrywanie i wyświetlanie szczegółowych informacji o dokumentach wynikowych. Na przykład można przejrzeć krótkie opisy dokumentów lub szczegółowe informacje, takie jak nazwy pól w każdym dokumencie wynikowym.
- Zwijanie dokumentów z jednego źródła. Na przykład, jeśli z jednego źródła zwracanych jest 100 dokumentów, dwa najbardziej trafne dokumenty będą w tabeli wynikowej zgrupowane. Pozostałe 98 dokumentów można wyświetlić, wybierając opcję przeglądania większej liczby dokumentów z tego samego źródła.
- Sortowanie dokumentów według ich ważności, daty lub wartości z określonego pola. Przy okazji sortowania według pola lub daty można zaznaczyć, czy wyniki mają być wyświetlane w kolejności rosnącej, czy też w kolejności malejącej.
- Uzyskiwanie propozycji poprawienia pisowni słów, które mogły zostać błędnie zapisane w zapytaniu.
- Wyświetlanie informacji o kategoriach, do których należy dokument wynikowy (jeśli w kolekcji są stosowane kategorie), oraz przeglądanie wyłącznie dokumentów należących do wybranej kategorii.
- Określanie dodatkowych terminów zapytania przeznaczonych do wyszukiwania w ramach wyników wyszukiwania.

## Funkcje pobierania dokumentów

Te funkcje umożliwiają:

- Pobieranie dokumentów przez kliknięcie identyfikatora URI dokumentu i otwarcie go w przeglądarce WWW. Jeśli przeszukiwacz Notes lub Domino Document Manager skonfigurowany jest do korzystania z protokołu DIIOP, to dokumenty przeszukiwane przez te przeszukiwacze mogą być wyświetlane w przeglądarce klienta Lotus Notes zamiast w przeglądarce WWW.  
Jeśli w przeszukiwaczu zostały skonfigurowane zabezpieczenia na poziomie dokumentu, pobierać dokumenty będą mogli tylko użytkownicy mający uprawnienia dostępu do zabezpieczonych treści.
- Pobieranie dokumentów przez kliknięcie szybkich odsyłaczy. Szybki odsyłacz wiąże słowa kluczowe z identyfikatorami URI. Jeśli w zapytaniu użyto określonych słów kluczowych,

powiązane z nimi identyfikatory URI (które predefiniowano jako w wysokim stopniu adekwatne dla tych słów kluczowych) są wyświetlane w górnej części listy wyników wyszukiwania.

---

## Właściwości aplikacji wyszukującej

Można zmienić plik konfiguracyjny aplikacji wyszukującej, definiując opcje dla danego środowiska, zmieniając wygląd aplikacji i określając, które opcje mają być dostępne dla użytkowników po uruchomieniu przez nich tej aplikacji.

Właściwości można także edytować za pomocą programu Search Application Customizer. Dokonując wyborów za pomocą tego programu, można natychmiast zobaczyć efekt wprowadzonych zmian. Kiedy podane opcje kolekcji wyszukiwania i wyświetlania wyników wyszukiwania są zadowalające, można je zapisać, aby zaktualizować plik konfiguracyjny aplikacji wyszukującej.

**Ważne:** Jeśli aplikacja wyszukująca zostanie uruchomiona jako portlet w ramach produktu WebSphere Portal, nie można korzystać z Konfiguratora aplikacji wyszukującej do interaktywnego wprowadzania zmian w aplikacji wyszukującej. Konieczna jest edycja właściwości i konfiguracja instancji portletu z użyciem interfejsu administracyjnego produktu WebSphere Portal.

Plikiem konfiguracyjnym przykładowej aplikacji wyszukującej w przypadku wyszukiwania korporacyjnego jest plik `config.properties`. W tym temacie omawiane są właściwości w tym pliku i opisywane ich wartości domyślne. W przypadku tworzenia plików konfiguracyjnych dla niestandardowych aplikacji wyszukujących właściwości w tych plikach i wartości podane dla tych właściwości mogą różnić się między sobą.

### Parametry środowiska

Istnieje możliwość określania opcji sterujących działaniem aplikacji wyszukującej.

#### **applicationName**

Określa nazwę poprawnej aplikacji wyszukującej. Wartością domyślną jest nazwa `Default`.

Wartość domyślną można zmienić, aby jako domyślna aplikacja używana była inna aplikacja wyszukująca.

**Wskazówka:** Jeśli nazwa aplikacji to `Default`, przykładowej aplikacji wyszukującej można używać do wyszukiwania danych we wszystkich kolekcjach i źródłach zewnętrznych w pojedynczym zapytaniu.

#### **timeout**

Określa liczbę sekund oczekiwania na odpowiedź z serwera wyszukiwania, zanim żądanie wyszukiwania przekroczy limit czasu. Musi to być liczba całkowita (np. 60, a nie 60,5). Jeśli limit czasu nie zostanie określony, wartością domyślną jest 30 sekund.

#### **hostname**

Określa pełną nazwę hosta odpowiadającą serwerowi WWW skonfigurowanemu do obsługi używanej instancji serwera WebSphere Application Server. Wartością domyślną jest `localhost`.

Aby mieć pewność, że aplikacja wyszukująca działa poprawnie, wartość domyślną można zmienić na pełną nazwę hosta, do używania którego skonfigurowany jest

serwer WebSphere Application Server. Na przykład, jeśli nazwą hosta komputera lokalnego jest Komputer, a nazwą hosta serwera WWW jest www.ibm.com, należy określić `www.ibm.com`.

**protocol**

Określa protokół do komunikacji z serwerem WWW: `http` lub `https`. W przypadku pozostawienia pustego pola wartością domyślną jest `http`.

**port**

Określa numer portu serwera WWW skonfigurowanego do obsługi używanej instancji serwera WebSphere Application Server. Wartością domyślną jest `80`, która jest typowa dla protokołu HTTP. Typowym portem używanym przez protokół HTTPS jest `443`.

**trustStore**

Jeśli używany jest protokół HTTPS, określ pełną ścieżkę do pliku kluczy, czyli pliku bazy danych zawierającego klucze publiczne. Informacje te, określane również mianem *zaufanych certyfikatów*, umożliwiają zaufaną komunikację za pomocą protokołu SSL (Secure Sockets Layer). Aby określić ścieżkę w systemie Windows, należy wpisać podwójny ukośnik odwrotny. Na przykład: `x:\\Serwer aplikacji\\webserver.key`

**trustPassword**

Jeśli używany jest protokół HTTPS, określ hasło do podanego pliku kluczy.

**username**

Aplikacja wyszukująca automatycznie ustawia tę wartość na nazwę użytkownika podaną podczas logowania się do tej aplikacji. Nazwę użytkownika podaje się tutaj tylko wtedy, gdy należy przesłonić domyślny sposób uwierzytelniania użytkowników. To pole jest używane tylko wtedy, gdy włączono zabezpieczenia globalne na serwerze WebSphere Application Server.

**password**

Aplikacja wyszukująca automatycznie ustawia tę wartość na hasło podane podczas logowania się do tej aplikacji. Hasło należy podać tylko wtedy, gdy podano nazwę użytkownika. To pole jest używane tylko wtedy, gdy włączono zabezpieczenia globalne na serwerze WebSphere Application Server.

**ssoCookieName**

Określa nazwę informacji cookie, która zawiera łańcuch tokenu pojedynczego logowania się (SSO). Wartością domyślną jest `LtpaToken`.

**proxyHost**

Określa pełną nazwę hosta serwera proxy, jeśli wymagany jest on w celu uzyskania dostępu do serwera wyszukiwania.

**proxyPort**

Określa port dla podanego hosta serwera proxy.

**proxyUser**

Określa nazwę użytkownika, która ma być używana w celu zalogowania się do serwera proxy, jeśli wymaga on podstawowego uwierzytelniania.

**proxyPassword**

Określa hasło dla podanej nazwy użytkownika serwera proxy.

**filter**

Określa klasę, która ma być używana do pobierania dokumentów wymienionych w wynikach wyszukiwania. Domyślna klasa to `com.ibm.es.api.filters.SetDocumentURIFilterFetch`. Wartość tę należy zmienić tylko wówczas, gdy dysponuje się niestandardową klasą, która ma zostać użyta do pobierania dokumentów.

### logging.level

Określa ilość rejestrowanych szczegółów:

**OFF** Nie są rejestrowane żadne komunikaty.

#### **SEVERE**

Komunikaty wskazujące na poważne awarie są rejestrowane. Jest to wartość domyślna.

**INFO** Rejestrowane są komunikaty informacyjne.

**FINE** Komunikaty śledzenia z niewielką ilością szczegółów są rejestrowane. Ta opcja odpowiada poziomowi rejestrowania FINE w klasie Java `java.util.logging.Level`.

**ALL** Rejestrowane są wszystkie komunikaty.

## Ikony typu źródła

Istnieje możliwość dostosowania obrazów reprezentujących typ źródła danych, do którego należy dokument wynikowy. Poniższe ikony typu źródła identyfikujące przeszukiwacze i źródła zewnętrzne obsługiwane po zainstalowaniu produktu OmniFind Enterprise Edition są predefiniowane w pliku `config.properties`.



#### **documentSource.vbr.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Content Edition. Domyślna ikona to `/images/sourceVBR.gif`.



#### **documentSource.db2.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz DB2. Domyślna ikona to `/images/sourceDB2.gif`.



#### **documentSource.cm.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz DB2 Content Manager. Domyślna ikona to `/images/sourceCM.gif`.



#### **documentSource.dominodoc.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Domino Document Manager. Domyślna ikona to `/images/sourceDominoDoc.gif`.



#### **documentSource.exchange.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Exchange Server. Domyślna ikona to `/images/sourceExchange.gif`.



#### **documentSource.database.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz JDBC. Domyślna ikona to `/images/sourceJDBC.gif`.

**documentSource.nntp.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz NNTP. Domyślna ikona to /images/sourceNntp.gif.

**documentSource.notes.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Notes. Domyślna ikona to /images/sourceNotes.gif.

**documentSource.quickplace.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz QuickPlace. Domyślna ikona to /images/sourceWorkplace.gif.

**documentSource.seedlist.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz listy adresów początkowych. Domyślna ikona to /images/sourceSeedlist.gif.

**documentSource.unixfs.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz systemu plików UNIX. Domyślna ikona to /images/sourceUnixFS.gif.

**documentSource.web.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz sieci WWW. Domyślna ikona to /images/sourceWeb.gif.

**documentSource.wcm.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Web Content Management. Domyślna ikona to /images/sourceWorkplace.gif.

**documentSource.wps.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz WebSphere Portal. Domyślna ikona to /images/sourceWPS.gif.

**documentSource.winfs.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz systemu plików Windows. Domyślna ikona to /images/sourceWindowsFS.gif.

**documentSource.ldap.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument należy do zewnętrznego źródła utworzonego dla serwera LDAP. Domyślna ikona to /images/sourceLDAP.gif.

**documentSource.jdbc.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument należy do zewnętrznego źródła utworzonego dla tabeli bazy danych JDBC (Java Database Connectivity). Domyślna ikona to /images/sourceJDBC.gif.



## Ikony przeglądarki klienckiej

Dokumenty wynikowe mogą być wyświetlane w przeglądarce WWW. Dokumenty przeszukane przez wyszukiwacze Notes lub Domino Document Manager skonfigurowane do korzystania z protokołu DIIOP można także wyświetlać w aplikacji przeglądarki klienta Lotus Notes.

Aby umożliwić wyświetlanie dokumentów za pomocą aplikacji przeglądarki klienta, upewnij się, że następująca właściwość ma wartość true:

```
clientViewer.show=true
```

Można dostosować obrazy reprezentujące aplikację przeglądarki klienta. W poniższym przykładzie ikona Lotus Notes oznacza, że dokument można wyświetlić za pomocą aplikacji przeglądarki:

```
client.notes.icon=/images/notes.gif
client.dominodoc.icon=/images/notes.gif
```

W wynikach wyszukiwania ikona i odsyłacz do aplikacji przeglądarki klienta są wyświetlane w następujący sposób:



### Przeglądarka kliencka

## Pola dokumentu

W przypadku typów źródeł danych zawierających pola można określić, które pola będą wyświetlane w dokumentach wynikowych.

### **fields.przedrostek** URI=*lista\_nazw\_pól\_oddzielonych\_znakami\_spacji*

Znaczenie znaku dwukropka (:) w przedrostku URI należy zmienić, poprzedzając go znakiem odwrotnego ukośnika (\). Aby kontynuować listę nazw pól w kolejnym wierszu, poprzedni wiersz należy zakończyć znakiem odwrotnego ukośnika (\). Na przykład:

```
fields.db2\://=databasename tablename
fields.domino\://=databasetitle filename creator
fields.dominodoc\://=librarydbtitle documentdbtitle filename author
fields.exchange\://=from creator
fields.file\://=directory filename
fields.https\://=documentID
fields.http\://=documentID
fields.jdbc\://=databasename tablename
fields.news\://=group from
fields.quickplace\://=placetitle roomtitle creator
fields.seedlist\://=author
fields.vbr\://=itemname repositorytype revisionuser
fields.wcm\://=author owner modifier
fields.web\://=
fields.wp6\://=
fields.wps\://=
```

## Ikony pól

W przypadku typów źródeł danych i dokumentów zawierających pola można dostosować obrazy reprezentujące pola. Wszystkie pola nad podsumowaniem dokumentu zawierają identyfikujący je obraz. W pliku `config.properties` predefiniowane są następujące ikony pól.



### **field.icon.databasetitle**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera tytuł dokumentu. Domyślna ikona to `/images/notesdb.gif`.

**field.icon.databasename**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera nazwę bazy danych, do której należy dokument. Domyślna ikona to `/images/db2.gif`.

**field.icon.tablename**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera nazwę tabeli, do której należy dokument. Domyślna ikona to `/images/table.gif`.

**field.icon.directory**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera nazwę katalogu, do którego należy dokument. Domyślna ikona to `/images/closedFolder.gif`.

**field.icon.filename**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera nazwę pliku dokumentu. Domyślna ikona to `/images/document.gif`.

**field.icon.documentID**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera identyfikator dokumentu. Tego pustego obrazu można użyć na przykład w dokumentach WWW, aby określić obraz dla adresu URL, ale nie wyświetlać go użytkownikowi. Domyślna ikona to `/images/dot.gif`.

**field.icon.group**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera identyfikator dokumentu. Tego pustego obrazu można użyć na przykład w dokumentach WWW, aby określić obraz dla adresu URL, ale nie wyświetlać go użytkownikowi. Domyślna ikona to `/images/document.gif`.

**field.icon.from**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole identyfikuje kogoś, kto wysłał dokument. Domyślna ikona to `/images/author.gif`.

**field.icon.creator**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole identyfikuje twórcę dokumentu. Domyślna ikona to `/images/author.gif`.

**field.icon.author**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole identyfikuje autora dokumentu. Domyślna ikona to `/images/author.gif`.

**field.icon.revisionuser**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole identyfikuje kogoś, kto sprawdził dokument. Domyślna ikona to `/images/author.gif`.

**field.icon.owner**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole identyfikuje właściciela dokumentu. Domyślna ikona to `/images/author.gif`.

**field.icon.modifier**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole identyfikuje kogoś, kto zmodyfikował dokument. Domyślna ikona to `/images/author.gif`.

## Domyślna ikona pola

Dla pól wyświetlanych w wynikach wyszukiwania można określić obraz używany, gdy nie skonfigurowano żadnych ikon pól. W pliku `config.properties` predefiniowana jest następująca domyślna ikona pola.



### **field.defaultIcon**

Określa nazwę i ścieżkę do pliku graficznego będącego domyślną ikoną dla pól w wynikach wyszukiwania. Domyślna ikona to `/images/database.gif`.

## Pola daty

Istnieje możliwość określenia, które pola są polami daty. Pola, których nazwy zostaną określone w tym miejscu, są formatowane w wynikach wyszukiwania tak jak daty. Format daty jest zgodny z ustawieniami narodowymi przeglądarki WWW.

**date.fields=***lista\_nazw\_pól\_oddzielonych\_znakami\_spacji*

Aby kontynuować listę nazw pól w kolejnym wierszu, poprzedni wiersz należy zakończyć znakiem odwrotnego ukośnika (`\`).

Przykład:

```
date.fields=modifieddate createddate
```

## Tytuły dokumentów

Istnieje możliwość określenia alternatywnych tytułów dokumentów przez zastąpienie tekstu tytułu bardziej zrozumiałymi danymi (pozwala to *wyklarować* tytuły). Na przykład zamiast wyświetlania tytułów dokumentów w postaci niewiele mówiącej etykiety Slajd 1, można określić, że ma ona zostać zastąpiona w wynikach wyszukiwania Zamiast niej do zidentyfikowania dokumentu wynikowego można użyć bardziej zrozumiałego pola, takiego jak nazwa pliku.

Istnieje również możliwość określenia alternatywnych tytułów dokumentów przez usunięcie z tytułów dokumentów słów zbędnych (co pozwala *skracać* tytuły). Na przykład, jeśli pewna liczba dokumentów wynikowych rozpoczyna się słowami Microsoft Word -, można poprawić czytelność wyników wyszukiwania, pomijając powtarzający się początkowy tekst.

**titles.clean=***lista\_tytułów\_oddzielonych\_przecinkami*

**titles.truncatePrefix=***lista\_przedrostków\_oddzielonych\_przecinkami*

Pozycje na liście rozdzielanej przecinkami mogą zawierać znaki spacji oraz inne znaki za wyjątkiem przecinka. Aby kontynuować listę w kolejnym wierszu, poprzedni wiersz należy zakończyć znakiem odwrotnego ukośnika (`\`).

Na przykład:

```
titles.clean=Slajd 1, Układ 1, bez tytułu, \
Dokument bez tytułu, Prezentacja PowerPoint, \
(brak tytułu tej strony)
```

```
titles.truncatePrefix=Microsoft Word -, Microsoft PowerPoint -
```

## Wartości domyślne preferencji użytkownika

Istnieje możliwość określenia wartości domyślnych dla strony Preferencje w aplikacji wyszukiwującej. Jeśli użytkownik zmieni te preferencje, nowe wartości będą obowiązywać tylko na czas bieżącej sesji użytkownika. W pliku `config.properties` predefiniowane są następujące preferencje.

**preferences.resultsRange=10**

Określa, że na każdej stronie wyników wyszukiwania może znaleźć się lista 10 dokumentów wynikowych.

**preferences.siteCollapsing=Yes**

Określa, że identyfikatory URI z tego samego źródła mają być zwinięte. Zwijanie serwisu jest dostępne tylko wtedy, gdy wyniki są sortowane według istotności. W przypadku źródeł danych WWW i NNTP identyfikatory URI zgodne z identyfikatorem URI głównego serwisu (na przykład `www.ibm.com`) są zwijane automatycznie. W przypadku innych typów źródeł danych oraz serwisów WWW o głębszych poziomach ścieżek (na przykład `www.ibm.com/hr`) reguły zwijania serwisu muszą zostać skonfigurowane w konsoli administracyjnej wyszukiwania korporacyjnego.

**preferences.spellCorrections=Yes**

Określa, że gdy użytkownik wprowadzi zapytanie zawierające potencjalnie niepoprawnie wpisane słowo, wyświetlane mają być proponowane poprawki dotyczące pisowni. Należy mieć na uwadze, że słowa ignorowane są zawsze usuwane przed utworzeniem propozycji poprawki.

**preferences.extendedHighlighting=No**

Określa, że oprócz pola podsumowania dokumentu terminy użyte w zapytaniu nie będą wyróżniane w polach dodatkowych (takich jak tytuł dokumentu).

## Domyślne kolekcje i źródła zewnętrzne

Można określić kolekcje i źródła zewnętrzne wybrane wstępnie na stronach Preferencje oraz Wyszukiwanie zaawansowane. Użytkownicy mogą edytować zestaw domyślny, aby wyszukiwać w mniejszej liczbie kolekcji i źródeł zewnętrznych niż udostępniona domyślnie. Jeśli zestaw kolekcji i źródeł zewnętrznych zostanie tu ograniczony, użytkownicy mogą wybrać dowolną kolekcję lub źródło zewnętrzne dostępne dla aplikacji wyszukującej podczas modyfikowania swoich preferencji lub opcji wyszukiwania zaawansowanego.

**preferences.defaultCollections=\*****preferences.defaultCollections=lista\_id\_kolekcji\_oddzielonych\_znakami\_spacji**

Określenie znaku gwiazdki (\*) umożliwia wyszukiwanie we wszystkich kolekcjach i źródłach zewnętrznych (kolekcje i źródła zewnętrzne muszą być powiązane z aplikacją wyszukującą w konsoli administracyjnej wyszukiwania korporacyjnego). Jest to domyślne ustawienie w pliku `config.properties`.

Aby ograniczyć elementy, w których użytkownicy mogą wyszukiwać bez modyfikowania swoich preferencji lub opcji wyszukiwania zaawansowanego, należy określić identyfikatory kolekcji dla kolekcji i źródeł zewnętrznych, które mają być domyślnie udostępnione użytkownikom do wyszukiwania.

Na przykład:

```
preferences.defaultCollections=*\npreferences.defaultCollections=id_kolekcji1 id_kolekcji2
```

## Informacje dodatkowe w wynikach wyszukiwania

Ilość informacji podawanych w wynikach wyszukiwania można dostosować. Można również kontrolować możliwość filtrowania wyników wyszukiwania przez użytkowników. Poniższe ustawienia to domyślne ustawienia w pliku `config.properties`.

**refreshButton.show=false**

Określa, czy przycisk **Odśwież** jest wyświetlany na stronie wyszukiwania podstawowego. Przycisk **Odśwież** jest zawsze dostępny w wyszukiwaniu

zaawansowanym. Jeśli opcja ta jest zaznaczona, użytkownicy mogą odświeżyć listę kolekcji i źródeł zewnętrznych dostępnych do wyszukiwania.

W przypadku korzystania z aplikacji Search Application Customizer przycisk **Odśwież** nie jest potrzebny.

Jeśli aplikacja Search Application Customizer nie jest używana, użytkownik może chcieć wyświetlić przycisk **Odśwież** podczas testowania zmian w pliku konfiguracyjnym. Po zapisaniu wprowadzonych zmian można wówczas kliknąć przycisk **Odśwież** w celu sprawdzenia, jak wpłynęły one na aplikację wyszukującą. Jeśli przycisk **Odśwież** nie jest wyświetlony, należy zrestartować aplikację ESSearchServer na serwerze WebSphere Application Server, zanim zmiany zostaną uwzględnione.

Przycisk **Odśwież** jest automatycznie wyświetlany, jeśli nie ma kolekcji lub źródeł zewnętrznych do wyszukiwania (na przykład wtedy, gdy nazwa hosta jest błędnie określona, serwery wyszukiwania nie są uruchomione lub aplikacja ESSearchServer nie jest uruchomiona na serwerze aplikacji WebSphere Application Server). Stanowi to dla użytkownika pomoc przy rozwiązywaniu problemu.

#### **builtQueryString.show=false**

Steruje wyświetlaniem w pełni rozwiniętej składni zapytania w obszarze poprzedzającym listę dokumentów wynikowych. Aby przetworzone właśnie zapytanie wyświetlane było w pełnej postaci, należy ustawić tę opcję na wartość true.

#### **extraQueryData.show=false**

Steruje wyświetlaniem dodatkowych informacji dotyczących zapytania. Aby wyświetlać informacje o ograniczeniach listy ACL, nazwach wyszukiwanych kolekcji i źródeł zewnętrznych oraz języku zapytań, ustaw w tej opcji wartość true.

#### **refineResults.show=true**

Steruje możliwością zawężania przez użytkowników wyników wyszukiwania przez określanie dodatkowych terminów dla zapytania. Po ustawieniu tej opcji na wartość true w dolnej części strony wyników wyszukiwania wyświetlane jest pole zapytania z etykietą **Szukaj w wynikach**.

#### **sorting.show=true**

Steruje wyświetlaniem opcji sortowania wyników wyszukiwania. Aby opcje **Sortuj według** i **Kolejność sortowania**, umożliwiające sortowanie wyników wyszukiwania, nie były wyświetlane, ustaw w tej opcji wartość false.

#### **sourceTypeFilter.show=true**

Steruje wyświetlaniem opcji filtrowania wyników według rodzaju źródła w wynikach wyszukiwania. Aby użytkownicy nie mogli filtrować wyników według typu źródłowego, ustaw w tej opcji wartość false.

Aby użytkownicy nie mogli filtrować wyników według rodzaju dokumentu, usuń wybrane lub wszystkie rodzaje dokumentów w pliku konfiguracyjnym (`documentType.label=rodzaje_dokumentów`).

#### **filter.showOnTwoLines=true**

Określa, czy opcje filtrowania wyników według typu źródła i typu pliku mają być wyświetlane w wynikach wyszukiwania w jednym czy też w dwóch wierszach. Podczas przeglądania wyników wyszukiwania użytkownicy mogą wybrać typ źródłowy i typ pliku, aby wyświetlić tylko dokumenty wynikowe zgodne z wybranymi filtrami.

Aby maksymalnie zwiększyć ilość miejsca dostępnego do wyświetlania wyników wyszukiwania, należy ustawić tę właściwość na wartość false. Aby zwiększyć

czytelność filtrów, zwłaszcza jeśli dostępne filtry są dłuższe niż jeden wiersz, można ustawić tę właściwość na wartość true, dzięki czemu każdy filtr będzie wyświetlany w osobnym wierszu.

**clientViewer.show=true**

Określa, czy do wyświetlania dokumentu wynikowego ma być używana aplikacja przeglądarki klienta Lotus Notes. Aby nie można było używać aplikacji przeglądarki do wyświetlania dokumentów Domino, ustaw w tej opcji wartość false.

**showDetails.show=true**

Steruje wyświetlaniem odsyłaczy Pokaż szczegóły i Ukryj szczegóły w wynikach wyszukiwania. Aby użytkownicy nie mogli wyświetlać dodatkowych szczegółów dotyczących dokumentów wynikowych, ustaw w tej opcji wartość false.

**showDetailsImage.show=true**

Steruje wyświetlaniem szczegółów dotyczących dokumentów wynikowych w oknie. Aby użytkownicy nie mogli wyświetlać dodatkowych szczegółów dotyczących dokumentów wynikowych, umieszczając kursor nad identyfikatorem URI dokumentu, ustaw w tej opcji wartość false.

**numberSearchResultsReturned.show=true**

Steruje wyświetlaniem łącznej liczby wyników wyszukiwania. Aby użytkownicy nie mogli wyświetlać liczby dokumentów zwracanych w wynikach wyszukiwania, ustaw w tej opcji wartość false.

**showMessage.error=true**

Steruje wyświetlaniem komunikatów o błędach. Aby na górze aplikacji wyszukującej nie były wyświetlane komunikaty o błędach, ustaw w tej opcji wartość false.

**showMessage.warning=true**

Steruje wyświetlaniem komunikatów ostrzegawczych. Aby na górze aplikacji wyszukującej nie były wyświetlane komunikaty ostrzegawcze, ustaw w tej opcji wartość false.

**showMessage.info=true**

Steruje wyświetlaniem komunikatów informacyjnych. Aby na górze aplikacji wyszukującej nie były wyświetlane komunikaty informacyjne, ustaw w tej opcji wartość false.

**showMessage.success=true**

Steruje wyświetlaniem komunikatów o powodzeniu. Aby na górze aplikacji wyszukującej nie były wyświetlane komunikaty o pomyślnym zakończeniu działania, ustaw w tej opcji wartość false.

## Niestandardowy baner i logo

Istnieje możliwość dostosowania obrazów wyświetlanych w obszarze banera w górnej części aplikacji wyszukującej. Można na przykład zastąpić domyślne obrazy produktu OmniFind Enterprise Edition obrazami związanymi z marką firmy. Jeśli baner nie ma być wyświetlany, jeden lub oba poniższe wiersze należy przekształcić w komentarz. Właściwość banner.icon identyfikuje grafikę wyświetlaną po lewej stronie obszaru banera. Właściwość banner2.icon identyfikuje grafikę wyświetlaną po prawej stronie obszaru banera.

```
banner.icon=/images/WS_II_0FEdition.gif
banner2.icon=/images/WS_II_mosaic.gif
```

## Niestandardowy obraz tła

Istnieje możliwość dostosowania obrazów wyświetlanych w tle stron aplikacji wyszukującej. Można na przykład zastąpić domyślne obrazy wyszukiwania korporacyjnego obrazami



związanymi z marką firmy. Jeśli obraz tła ma nie być wyświetlany na stronach, jeden lub większą liczbę poniższych wierszy należy przekształcić w komentarz.

```
search.backgroundImage=/images/IIOF_search.gif
preferences.backgroundImage=/images/IIOF_options.gif
advanced.backgroundImage=/images/IIOF_advanced.gif
browse.backgroundImage=/images/IIOF_tree.gif
myProfile.backgroundImage=/images/IIOF_profile.gif
logoff.backgroundImage=/images/IIOF_logout.gif
```

## Odsyłacze

Właściwości w sekcji pliku `config.properties` dotyczącej odsyłaczy umożliwiają wyświetlanie nazw stron aplikacji wyszukującej na każdej stronie jako odsyłaczy, a nie na pasku narzędzi ani na stronach z kartami. Wyświetlanie odsyłaczy jest przydatne, gdy aplikacja wyszukująca działa jako portlet i pożądane jest zminimalizowanie obszaru używanego do wyświetlania aplikacji wyszukującej na stronie portalu.

Jeśli użytkownik preferuje nawigowanie po aplikacji wyszukującej przez wybór opcji na pasku narzędzi i na stronach z kartami, należy przekształcić te wiersze w komentarz.

## Karty wyszukiwania

Właściwości w sekcji pliku `config.properties` dotyczącej zakładek wyszukiwania określają nazwy stron JSP (Java Server Pages) używanych do wyświetlania stron z zakładkami w widoku Wyszukiwania aplikacji wyszukującej (Wyszukiwanie podstawowe, Wyszukiwanie zaawansowane i Drzewo kategorii). Nie należy edytować tych stron, jeśli nie ma się doświadczenia w programowaniu w języku Java i z obsługą stron JSP.

Przykładowe sposoby dostosowania tego obszaru to:

- Kierowanie aplikacji wyszukującej do niestandardowych stron JSP udostępniających inny wygląd stron z kartami.
- Przekształcanie w komentarz pozycji drzewa kategorii. Na przykład jeśli nie zostaną skonfigurowane kategorie dla kolekcji, nie ma potrzeby wyświetlania strony Drzewo kategorii w aplikacji wyszukującej.
- Kopiowanie pozycji dla stron z kartami do sekcji pliku `config.properties` dotyczącej pasków narzędzi i przekształcanie tych wierszy w komentarz. Na przykład może być przydatne wyświetlenie samego paska narzędzi bez wyświetlania stron z kartami.

## Paski narzędzi

Właściwości w sekcji pliku `config.properties` dotyczącej pasków narzędzi określają nazwy stron JSP (Java Server Pages) używanych jako paski narzędzi w aplikacji wyszukującej. Nie należy edytować tych stron, jeśli nie ma się doświadczenia w programowaniu w języku Java i z obsługą stron JSP.

Przykładowe sposoby dostosowania tego obszaru to:

- Kierowanie aplikacji wyszukującej do niestandardowych stron JSP udostępniających inny wygląd paska narzędzi.
- Przekształcanie w komentarz pozycji paska narzędzi dla elementów, które nie mają być wyświetlane. Na przykład na pasku narzędzi może nie być potrzebny odsyłacz do strony Informacje.
- Przeniesienie funkcji wyświetlania strony Wyszukiwanie zaawansowane z sekcji pliku `config.properties` dotyczącej zakładek, aby opcja ta była dostępna tylko na pasku narzędzi.



## Rozszerzone etykiety typu dokumentu

Istnieje możliwość zwiększenia czytelności filtru typu dokumentu przez odwzorowanie rzeczywistych nazw typów dokumentów na terminy, które są zwięźlejsze i bardziej adekwatne. Typy dokumentów dostępne dla aplikacji wyszukujących zdefiniowane są w klasie `AvailableDocumentTypes` interfejsu API wyszukiwania i indeksowania (SI-API). Dla wygody dostępne typy dokumentów zostały również wymienione na końcu pliku `config.properties`.

### **documentType.etykieta=lista\_typów\_dokumentów\_oddzielonych\_znakami\_spacji**

Określa nazwę wyświetlaną w wynikach wyszukiwania w wierszu filtru typu dokumentu oraz listę rzeczywistych typów dokumentów, które mają być wyświetlane po wybraniu tego filtru przez użytkownika.

Na przykład można określić etykietę `html` i odwzorować na tę nazwę rozszerzenia plików i typy MIME różnych dokumentów WWW. Po kliknięciu przez użytkownika pozycji `html` w celu odfiltrowania wyników wyszukiwania wyświetlone zostaną tylko dokumenty z określonymi rozszerzeniami i typami MIME.

W pliku `config.properties` predefiniowane są następujące odwzorowania typów dokumentów:

```
documentType.html=shtml text/html html xhtml htm
documentType.doc=doc application/msword
documentType.ppt=application/mspowerpoint ppt
documentType.xls=xls application/x-excel application/msexcel \
application/x-msexcel application/excel application/vnd.ms-excel
documentType.xml=xml text/xml
documentType.txt=txt text/plain
documentType.pdf=pdf application/pdf
```

Jeśli wartość określona jako etykieta typu dokumentu jest zgodna z nazwą właściwości w pliku `application.properties`, to wyświetlana jest wartość właściwości z pliku `application.properties` a nie wartość, która została tu określona. Na przykład, jeśli jako etykietę filtru typów pliku określono `documentType.unixfs`, to wyświetlaną nazwą filtru typów pliku jest wartość właściwości `unixfs` w pliku `application.properties` (**system plików UNIX**).

## Filtry niestandardowe

Do filtrowania wyświetlanych wyników wyszukiwania można określić zapytania niestandardowe.

### **filterCustom.etykieta=terminy\_dla\_zapytania**

Określa nazwę wyświetlaną w wynikach wyszukiwania w wierszu filtru niestandardowego oraz zapytanie zawężające wyniki wyszukiwania po wybraniu tego filtru przez użytkownika (podczas przeglądania wyników wyszukiwania użytkownicy mogą wybrać filtr niestandardowy, aby wyświetlić tylko dokumenty wynikowe zgodne z predefiniowanym zapytaniem).

W poniższym przykładzie wyniki wyszukiwania są filtrowane w celu wyświetlenia tylko dokumentów należących do bazy danych kadr (`ka`):

```
filterCustom.HR_database_only=databasename::ka
```

Po kliknięciu przez użytkownika pozycji `HR_database_only` w celu odfiltrowania wyników wyszukiwania przetwarzane jest zapytanie `databasename::ka`. Podczas wyświetlenia wyników wyszukiwania wymienione zostaną tylko dokumenty z bazy danych `ka`.

Jeśli wartość określona jako etykieta niestandardowego filtru jest zgodna z nazwą właściwości w pliku `application.properties`, to wyświetlana jest wartość

właściwości z pliku `application.properties` a nie wartość, która została tu określona. Na przykład, jeśli jako etykietę niestandardowego filtra określono `filterCustom.hostData=`, to wyświetlaną nazwą niestandardowego filtra jest wartość właściwości `hostData` w pliku `application.properties` (**Ustawienia serwera**).

W pliku `config.properties` udostępnionych jest kilka przykładowych filtrów niestandardowych przekształconych w komentarz.

## Wykrywanie duplikatów

Przy dodawaniu dokumentów do indeksu wyszukiwania korporacyjnego przeprowadzana jest analiza mająca na celu usunięcie duplikatów, tak aby w wynikach wyszukiwania nie były wielokrotnie wyświetlane te same dokumenty. Aby dalej przefiltrować wyniki wyszukiwania, można określić opcję ukrywania dokumentów, które są prawie identyczne, zapobiegając ich wyświetlaniu w wynikach wyszukiwania.

### **preferences.nearDuplicateDetection=No**

Określa, że dokumenty będące niemalże duplikatami, nie są filtrowane podczas przetwarzania zapytania.

Jeśli określona jest wartość **Yes**, dokumenty o podobnych tytułach i podsumowaniach są ukrywane przed użytkownikami przeglądającymi wyniki wyszukiwania. Komunikat informuje użytkowników, że niektóre dokumenty zostały pominięte, ponieważ są podobne do innych dokumentów w zestawie wynikowym. Aby wyłączyć ukrywanie i wyświetlić wszystkie dokumenty z zestawu wynikowego, użytkownicy mogą kliknąć odpowiedni odsyłacz.

Aby ukrywać dokumenty będące prawie duplikatami, obiekt `Query` interfejsu `SI-API` (Search and Index API) dla aplikacji wyszukującej musi mieć określoną metodę `setProperty` z łańcuchem `NearDuplicateDetection` ustawionym na wartość **Yes** (na przykład `query.setProperty("NearDuplicateDetection", "Yes");`).

## Analiza najlepszych wyników (wykresy słupkowe dla pól metadanych)

Można określić opcje wyświetlania wykresów słupkowych reprezentujących analizę najlepszych wyników. Domyślnie analizie poddawanych jest 500 najlepszych wyników. Każdy wykres odpowiada jednemu polu metadanych, a każdy słupek na wykresie odpowiada wartości pola. Długość słupka wskazuje częstość występowania danej wartości pola. Im dłuższy jest słupek, tym większa jest liczba wystąpień danej wartości pola.

Podczas konfigurowania przestrzeni przeszukiwania dla przeszukiwacza można określić opcje dla pól metadanych. Aby skonfigurować wykres najlepszych wyników dla pola metadanych, należy określić, że dane pole podlega wyszukiwaniu i że może być zwracane w wynikach wyszukiwania.

Właściwości konfigurowane dla wykresu słupkowego mają następujący format, gdzie *liczba* jest numerem jednoznacznie identyfikującym wykres słupkowy w aplikacji wyszukującej, *opcja* jest opcją wykresu słupkowego, a *wartość* jest wartością opcji:

```
topResultsChartsliczba.opcja=wartość
```

### **topResultsCharts*liczba*.titleKey=klucz\_aplikacji**

Określa tytuł wykresu, gdzie *liczba* jest numerem jednoznacznie identyfikującym wykres w aplikacji wyszukującej, a *klucz\_aplikacji* jest etykietą tytułu wykresu. Tą etykietą może być klucz z pliku `application.properties` dla aplikacji wyszukującej lub określona tutaj wartość. W następującym przykładzie etykieta dla tytułu wykresu jest określona przez wartość klucza `topResults.mostRecentDocuments` z pliku

application.properties:  
topResultsCharts3.titleKey=topResults.mostRecentDocuments.

W kolejnym przykładzie tytuł wykresu jest konkretną określoną tutaj wartością:  
topResultsCharts3.titleKey=Organizacje

**topResultsCharts***liczba.enable=true*

Określa, czy dany wykres słupkowy ma być wyświetlany, gdy użytkownicy będą wyświetlać wyniki wyszukiwania. Jeśli określona będzie wartość `false`, wykres nie będzie wyświetlany.

**topResultsCharts***liczba.fieldName=nazwa\_pola*

Określa nazwę pola metadanych, którego wartości mają być analizowane dla danego wykresu. Na przykład `tytulbazydanych`. Dla każdego dodanego wykresu należy określić inną nazwę pola (tego samego pola nie można użyć na więcej niż jednym wykresie).

**topResultsCharts***liczba.maxValues.collapsed=liczba*

Określa liczbę zwiniętych pozycji do wyświetlenia na danym wykresie. Na przykład określenie liczby 5 spowoduje, że wyświetlanych będzie pięć najczęściej występujących wartości w danym polu.

**topResultsCharts***liczba.maxValues.expanded=liczba*

Określa liczbę pozycji wyświetlanych na danym wykresie, gdy widok wykresu jest w pełni rozwinięty. Na przykład określenie wartości 10 spowoduje po rozwinięciu wykresu wyświetlenie nie więcej niż 10 różnych wartości pola.

**topResultsCharts***liczba.fieldValueSeparator=znak*

Określa znak stanowiący ogranicznik wartości w polu przeznaczonym do analizy. Na przykład, jeśli pole zawiera wiele wartości oddzielonych średnikiem (np. `agent;sprzedawca;broker`), to za pomocą tej właściwości można wskazać średnik (;) jako separator wartości pola, dzięki czemu każda wartość będzie mogła zostać dodana do wykresu słupkowego jako oddzielny element. Bez tej opcji cała wartość pola zostanie dodana do wykresu jako jeden element.

**topResultsCharts***liczba.canUserChangeFieldName=true*

Określa, czy użytkownik może wybrać inne pole podczas wyświetlania wyników wyszukiwania i zobaczyć najlepsze wyniki dla tego pola. Jeśli zostanie określona wartość `false`, użytkownicy nie będą mogli wybierać innego pola do analizy podczas wyświetlania wyników wyszukiwania.

Jeśli zostanie określona wartość `true`, wraz z wynikami dla bieżącego pola wyświetlana będzie lista wszystkich pól znalezionych w pierwszych 500 najlepszych wynikach. Gdy użytkownik wybierze pole z listy, etykieta wykresu jest zmieniana na **Dynamiczny wykres pola**, a wyniki wykresu słupkowego dla wybranego pola są wyświetlane do czasu wybrania przez użytkownika innego pola lub zamknięcia przeglądarki. Po następnym uruchomieniu przez użytkownika aplikacji wyszukującej wyświetlane są wyniki dla pierwotnego pola.

**topResultsCharts***liczba.width=liczba*

Określa szerokość wyświetlania wykresu słupkowego w pikslach. Na przykład 300.

**topResultsCharts***liczba.barheight=liczba*

Określa wysokość każdego słupka na wykresie słupkowym w pikslach. Na przykład 10.

**topResultsCharts***liczba.color=#kod\_koloru*

Określa kod szesnastkowy dla podstawowego koloru słupka na wykresie słupkowym. Wartość domyślna to kolor niebieski (`#0309C0`).

**topResultsCharts***liczba.color.gradient=#kod\_koloru*

Określa kolor gradientu słupka na wykresie słupkowym. Wartość domyślna to kolor

turkusowy (#00FFFF). W miarę jak liczba wyników dla konkretnej wartości pola przechodzi od niskiej do wysokiej częstotliwości, kolor słupka zmienia barwę od wartości określonej dla opcji `color` do wartości określonej dla opcji `color.gradient`.

Na przykład: 

#### **topResultsCharts***liczba.sortKey=frequency*

Określa sposób, w jaki mają być sortowane elementy na wykresie słupkowym.

Obsługiwane wartości:

**brak** Elementy na wykresie słupkowym nie są sortowane.

#### **etykieta**

Elementy na wykresie słupkowym są sortowane zgodnie z wartością pola. Jeśli zostaną skonfigurowane etykiety niestandardowe za pomocą właściwości

`topResultsCharts.liczba.listOfLabels.prefixKey=nazwa_pola`, to elementy na wykresie słupkowym będą sortowane również według nazwy etykiety.

#### **częstotliwość**

Elementy na wykresie słupkowym są sortowane zgodnie z liczbą wyników zwróconych dla poszczególnych wartości pola.

#### **topResultsCharts***liczba.sortOrder=descending*

Określa porządek sortowania dla elementów sortowanych według etykiety lub częstotliwości. Obsługiwane wartości:

#### **rosnąco**

Elementy sortowane według etykiety są wymieniane w porządku alfabetycznym od a do z. W przypadku elementów sortowanych według częstotliwości wartości pola występujące częściej znajdują się niżej na liście niż wartości występujące rzadziej.

#### **malejąco**

Elementy sortowane według etykiety są wymieniane w odwrotnym porządku alfabetycznym od z do a. W przypadku elementów sortowanych według częstotliwości wartości pola występujące częściej znajdują się wyżej na liście niż wartości występujące rzadziej.

#### **topResultsCharts***liczba.listOfLabels.prefixKey=nazwa\_pola*

Opcjonalne. Umożliwia określenie informacji, które zawsze mają być wyświetlane dla pewnych wartości pól, przy czym *nazwa\_pola* wskazuje pole, którego wartość ma być analizowana dla danego wykresu. Etykiety, które zostaną określone w tej właściwości, będą zawsze wyświetlane na wykresie słupkowym, nawet wtedy gdy w wynikach wyszukiwania nie będzie wystąpienia danej wartości pola.

Dla pola można skonfigurować dowolną liczbę etykiet. W odniesieniu do każdej etykiety określa się dwie właściwości, które mają następujący format:

*nazwa\_pola*`liczba.value=wartość`

*nazwa\_pola*`liczba.displayValue=wartość_wyświetlana`

gdzie:

*nazwa\_pola*

Jest nazwą pola, dla którego są konfigurowane etykiety.

*liczba* Jest numerem jednoznacznie identyfikującym etykietę.

*wartość*

Jest wartością, która ma być wyświetlana w wynikach analizy.

*wartość\_wyświetlana*

Jest etykietą, która ma być wyświetlana na wykresie słupkowym.

Tą etykietą może być klucz z pliku `application.properties` dla aplikacji wyszukującej lub określona tutaj wartość.

Na przykład:

```
topResultsCharts1.listOfLabels.prefixKey=tytulbazydanych
```

```
tytulbazydanych1.value=Artykuły i dokumenty przedsiębiorstwa JK
tytulbazydanych1.displayValue=Artykuły i dokumenty
```

```
tytulbazydanych2.value=Puste formularze przedsiębiorstwa JK
tytulbazydanych2.displayValue=Puste formularze
```

```
tytulbazydanych3.value=Biuletyny i wytyczne przedsiębiorstwa JK
tytulbazydanych3.displayValue=Biuletyny i wytyczne
```

## Analiza najlepszych wyników (niestandardowy kod HTML)

Opcje wyświetlania wykresów najlepszych wyników można określić przez rozszerzenie interfejsu `API com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` za pomocą niestandardowej klasy Java. W razie zastosowania tego podejścia do wyświetlania wyników można za pomocą kodu HTML sformatować zwrot dowolnych wyników wyszukiwania, a nie tylko zapytań, według których prowadzone są wyszukiwania w polach metadanych. Wyniki mogą być zwracane dla dokumentów w kolekcjach wyszukiwania korporacyjnego lub pochodzić z wyszukiwań w zewnętrznych repozytoriach i serwisach WWW.

Dla aplikacji wyszukującej można skonfigurować dowolną liczbę wykresów. W celu przypisania tytułu do każdego wykresu należy użyć następującej właściwości:

**topResultsCharts*liczba*.titleKey=klucz\_aplikacji**

Określa tytuł wykresu, gdzie *liczba* jest numerem jednoznacznie identyfikującym wykres w aplikacji wyszukującej, a *klucz\_aplikacji* jest etykietą tytułu wykresu. Tą etykietą może być klucz z pliku `application.properties` dla aplikacji wyszukującej lub określona tutaj wartość. W następującym przykładzie etykieta tytułu wykresu jest określona przez wartość klucza `topResults.mostRecentDocuments` z pliku `application.properties`:

```
topResultsCharts3.titleKey=topResults.mostRecentDocuments
```

Każdy element na wykresie odpowiada jednej wartości wyniku wyszukiwania i składa się na niego zestaw właściwości o następującym formacie, gdzie *liczba* jest numerem jednoznacznie identyfikującym wykres, *opcja* jest opcją wykresu, a *wartość* jest wartością opcji:

```
topResultsChartsliczba.opcja=wartość
```

**topResultsCharts*liczba*.enable=true**

Określa, czy dany wykres ma być wyświetlany, gdy użytkownicy będą wyświetlać wyniki wyszukiwania. Jeśli określona będzie wartość `false`, wykres nie będzie wyświetlany.

**topResultsCharts*liczba*.maxValues.collapsed=*liczba***

Określa liczbę zwiniętych pozycji do wyświetlenia na danym wykresie. Na przykład określenie liczby 5 spowoduje, że wyświetlanych będzie pięć najczęściej występujących wyników zgodnych z kryteriami.

**topResultsCharts*liczba*.width=*liczba***

Określa szerokość wyświetlania wykresu, w pikslach. Na przykład 400.

**topResultsCharts*liczba*.dynamicChartClass=*klasa\_niestandardowa***

Określa nazwę niestandardowej klasy Java rozszerzającej interfejs `API com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` i definiującej sposób wyświetlania danych wyjściowych na wykresie. Na przykład:

```
topResultsCharts.3.dynamicChartClass=com.ibm.es.searchui.charts.servlet.
DynamicMostRecentDocuments
```

Następujące właściwości, zawarte w domyślnym pliku `config.properties` dla przykładowej aplikacji wyszukiwującej przeznaczonej do prezentacji przykładów, są używane w przykładowej klasie Java `DogearSearchResults`. Informacje na temat parametrów zapytania, które można uwzględnić w niestandardowej aplikacji wyszukiwującej, można znaleźć w dokumentacji interfejsu API Dogear.

Następujących właściwości nie można ustawić za pomocą Konfiguratora aplikacji wyszukiwującej:

**topResultsCharts***liczba.xsl.fileName=plik\_stylów*

Określa ścieżkę i nazwę arkusza stylów XSL, na podstawie którego mają być formatowane najlepsze wyniki wyświetlane na wykresie. Na przykład `/styles/dogear.xsl`.

**topResultsCharts***liczba.url=url*

Określa adres URL do przeszukiwania.

**topResultsCharts***liczba.url.parameters=ps=liczba*

Określa wielkość strony. Na przykład `ps=3` ogranicza wielkość strony do 3.

Jest to po prostu łańcuch parametru zgodny z definicją w interfejsie API Dogear REST.

## Analiza najlepszych wyników (maksymalna liczba wyników)

**topResult.resultSize=liczba**

Wartość domyślna i maksymalna określająca liczbę wyników, jakie mają być zwracane w toku analizy najlepszych wyników, wynosi 500. Wartość tę można zmniejszyć, ale nie można jej zwiększyć. Na przykład można określić niższą liczbę, jeśli zdarzają się problemy polegające na przekroczeniu limitu czasu żądania podczas analizy najlepszych wyników. Tej właściwości nie można ustawić za pomocą Konfiguratora aplikacji wyszukiwującej.

### Pojęcia pokrewne



Określanie właściwości zapytania



Klasy Java do wyświetlania najczęstszych wyników

## Edycja właściwości przykładowej aplikacji wyszukiwującej

Przykładowa aplikacja wyszukiwująca dla wyszukiwania korporacyjnego umożliwia wyszukiwanie we wszystkich aktywnych kolekcjach i źródłach zewnętrznych w systemie. Można zmodyfikować plik konfiguracyjny, podając opcje środowiska serwera WWW, wskazać inną domyślną aplikację wyszukiwującą lub wybrać opcje wyświetlane po uruchomieniu aplikacji wyszukiwującej.

### Informacje o tym zadaniu

Program instalacyjny instaluje przykładową aplikację wyszukiwującą dla systemu wyszukiwania korporacyjnego na serwerze IBM WebSphere Application Server dla serwerów wyszukiwania korporacyjnego. Aby skonfigurować tę aplikację wyszukiwującą, należy zmodyfikować zawartość pliku konfiguracyjnego `config.properties`, który jest instalowany razem z aplikacją.

Aby wprowadzone zmiany odniosły skutek, należy zatrzymać i ponownie uruchomić aplikację `ESSearchServer` w serwerze WebSphere Application Server.



## Procedura

Aby zmienić właściwości przykładowej aplikacji wyszukującej, należy wykonać następujące czynności:

1. Zaloguj się na serwerze wyszukiwania jako administrator wyszukiwania korporacyjnego.
2. Zmodyfikuj plik `config.properties` przy użyciu standardowego edytora tekstu.  
Plik `config.properties` jest instalowany w następującym miejscu (`ES_INSTALL_ROOT` to katalog instalacyjny programu OmniFind Enterprise Edition na serwerze wyszukiwania):  
`ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war/WEB-INF/config.properties`
3. Zmień właściwości w celu określenia informacji o środowisku serwera WWW i preferencjach wyszukiwania, a następnie zapisz i zamknij plik. Znak `#` w tym pliku oznacza wiersz komentarza.
4. Zatrzymaj i ponownie uruchom aplikację `ESSearchServer`.

Systemy **AIX**,

**Linux lub Solaris**

```
./stopServer.sh ESSearchServer
```

```
./startServer.sh ESSearchServer
```

Systemy **Windows**

```
stopServer ESSearchServer
```

```
startServer ESSearchServer
```

Skrypty te znajdują się w katalogu `WAS_INSTALL_ROOT/AppServer/bin`:

- W przypadku serwera aplikacji WebSphere w wersji 5 domyślna ścieżka instalacji w systemach AIX ma postać `/usr/WebSphere`, w systemach Linux lub Solaris ma postać `/opt/WebSphere`, natomiast w systemach Windows - `C:\Program Files\WebSphere`.
- W przypadku serwera aplikacji WebSphere w wersji 6 domyślna ścieżka instalacji w systemach AIX ma postać `/usr/IBM/WebSphere`, w systemach Linux lub Solaris ma postać `/opt/IBM/WebSphere`, a w systemach Windows - `C:\Program Files\IBM\WebSphere`.

## Pojęcia pokrewne

 Określanie właściwości zapytania

 Klasy Java do wyświetlania najczęstszych wyników

## Zadania pokrewne

“Konfigurowanie serwerów wyszukiwania w celu zaakceptowania tylko bezpiecznych żądań (SSL)” na stronie 247

## Konfigurowanie aplikacji wyszukujących

Program Search Application Customizer jest to graficzny interfejs pozwalający konfigurować aplikacje wyszukujące pod kątem współpracy z systemem wyszukiwania korporacyjnego lub z aplikacjami niestandardowymi.

## Ograniczenia

Search Application Customizer jest aplikacją autonomiczną. Konfiguratora aplikacji wyszukującej nie można uruchomić w ramach portalu WebSphere Portal w celu skonfigurowania aplikacji wyszukujących działających jako portlety. Aby dostosować



aplikacje wyszukujące działające jako portlety, należy przeprowadzić edycję właściwości oraz konfigurowanie instancji portletu za pomocą interfejsu administracyjnego portalu WebSphere Portal.

### Informacje o tym zadaniu

Program Search Application Customizer pozwala wzrokowo oceniać wprowadzane zmiany i modyfikować aplikacje wyszukujące bez edycji pliku konfiguracyjnego. Na przykład można zmienić zawartość baneru i grafikę w tle strony, zmienić układ interfejsu wyszukiwania oraz określić opcje dotyczące pracy z wynikami wyszukiwania.

Podczas modyfikowania ustawień w programie Search Application Customizer można od razu obserwować efekty wprowadzanych zmian. Zapisanie zmian oznacza modyfikację pliku konfiguracyjnego dla aplikacji wyszukującej.

Aby wprowadzone zmiany odniosły skutek, należy zatrzymać i ponownie uruchomić aplikację ESSearchServer w serwerze WebSphere Application Server.

### Wskazówka:

Szczegółowe przykłady używania Konfiguratora aplikacji wyszukującej można znaleźć w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

### Procedura

Aby skonfigurować aplikację wyszukującą:

1. Aby dostosować przykładową aplikację wyszukującą systemu wyszukiwania korporacyjnego, wpisz adres URL programu Search Application Customizer w oknie przeglądarki. Na przykład:

<http://SerwerWyszukiwania.com/ESSearchApplication/palette.do>

*SerwerWyszukiwania.com* to nazwa hosta serwera wyszukiwania.

Jeśli konfiguracja serwera WWW nie przewiduje korzystania z portu 80, wówczas trzeba podać poprawny numer portu. Na przykład:

<http://SerwerWyszukiwania.com:9080/ESSearchApplication/palette.do>

**Wskazówka:** Administrator systemu wyszukiwania korporacyjnego może także otworzyć program Search Application Customizer, wybierając opcję **Search Customizer** w konsoli administracyjnej wyszukiwania korporacyjnego.

2. Aby skonfigurować niestandardową aplikację wyszukującą, wpisz adres URL programu Search Application Customizer, dodając na jego końcu nazwę pliku konfiguracyjnego odpowiedniej aplikacji. Na przykład:

<http://SerwerWyszukiwania.com/ESSearchApplication/palette.do?configFile=/WEB-INF/MojaKonfiguracja.properties>

Jeśli wskazany w ten sposób plik nie istnieje, zostaną wyświetlone wartości z pliku config.properties przykładowej aplikacji wyszukującej.

**Wskazówka:** Plik konfiguracyjny, który ma być używany z aplikacją, można także określić już po uruchomieniu programu Search Application Customizer za pomocą przycisku **Załaduj**, podając nazwę odpowiedniego pliku.

3. Jeśli w konfiguracji serwera WebSphere Application Server włączono globalny mechanizm zabezpieczeń, należy zalogować się, podając poprawny identyfikator i hasło.

4. Wybierz opcje do konfiguracji, takie jak informacje o serwerze wyszukiwania, rodzaje wyświetlanych komunikatów, opcje zapytań i wyników wyszukiwania oraz nazwy obrazów, którymi w wynikach wyszukiwania będą oznaczane różne rodzaje dokumentów. Aby uzyskać pomoc dotyczącą dostępnych ustawień, kliknij opcję **Pomoc dla konfiguratora**. Aby zobaczyć skutki niektórych zmian, na przykład co do prezentacji wyników wyszukiwania, wpisz zapytanie i kliknij przycisk **Szukaj**.
5. Kiedy ustawienia są już zadowalające, kliknij przycisk **Zapisz**, aby zmodyfikować plik konfiguracyjny. Przycisk **Wyczyść** pozwala przywrócić ustawienia wszystkich opcji w programie Search Application Customizer do stanu zgodnego z ostatnim zapisanym plikiem konfiguracyjnym.
6. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerze wyszukiwania, po czym zatrzymaj i uruchom ponownie aplikację ESSearchServer.

#### Systemy AIX, Linux lub Solaris

```
./stopServer.sh ESSearchServer
```

```
./startServer.sh ESSearchServer
```

#### Systemy Windows

```
stopServer ESSearchServer
```

```
startServer ESSearchServer
```

Skrypty te znajdują się w katalogu WAS\_INSTALL\_ROOT/AppServer/bin:

- W przypadku serwera aplikacji WebSphere w wersji 5 domyślna ścieżka instalacji w systemach AIX ma postać /usr/WebSphere, w systemach Linux lub Solaris ma postać /opt/WebSphere, natomiast w systemach Windows - C:\Program Files\WebSphere.
- W przypadku serwera aplikacji WebSphere w wersji 6 domyślna ścieżka instalacji w systemach AIX ma postać /usr/IBM/WebSphere, w systemach Linux lub Solaris ma postać /opt/IBM/WebSphere, a w systemach Windows - C:\Program Files\IBM\WebSphere.

#### Zadania pokrewne

“Konfigurowanie serwerów wyszukiwania w celu zaakceptowania tylko bezpiecznych żądań (SSL)” na stronie 247

## Klonowanie przykładowej aplikacji wyszukiwania

Aby wykorzystać przykładową aplikację wyszukiwania jako model do tworzenia własnej aplikacji wyszukiwania, można skopiować plik config.properties albo wykorzystać program Search Application Customizer.

#### Informacje o tym zadaniu

W celu ułatwienia tworzenia niestandardowych aplikacji wyszukiwania można sklonować opcje konfiguracji zdefiniowane dla przykładowej aplikacji wyszukiwania, a następnie zmienić je, dostosowując do potrzeb własnej aplikacji.

Po sklonowaniu przykładowej aplikacji wyszukiwania, do jej uruchomienia należy podać nazwę pliku konfiguracyjnego. Nazwę pliku konfiguracyjnego należy też podać w celu dostosowania nowej aplikacji wyszukiwania za pomocą programu Search Application Customizer.

Klonując przykładową aplikację wyszukiwania, można szybko stworzyć aplikację wyszukiwania spełniające konkretne zadania lub przeznaczone dla konkretnych odbiorców. Na przykład jedną aplikację wyszukiwania można przygotować dla pracowników działu kadr, a inną - dla pracowników działu sprzedaży.

Aby wprowadzone zmiany odniosły skutek, należy zatrzymać i ponownie uruchomić aplikację ESSearchServer w serwerze WebSphere Application Server.

### Procedura

Aby sklonować przykładową aplikację wyszukiwania:

1. Jeśli chcesz edytować plik konfiguracyjny:
  - a. Skopiuj plik config.properties z przykładowej aplikacji wyszukiwania i zmień jego nazwę.

Plik config.properties jest instalowany w następującym miejscu (ES\_INSTALL\_ROOT to katalog instalacyjny programu OmniFind Enterprise Edition na serwerze wyszukiwania):

```
ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/
ESSearchApplication.war/WEB-INF/config.properties
```

Tworzony plik należy umieścić w podkatalogu WEB-INF.

- b. Dokonaj edycji właściwości, które używane będą w tworzonej aplikacji wyszukiwania, i zapisz plik. Trzeba zmienić przynajmniej właściwość applicationName, podając w niej nazwę nowej aplikacji wyszukiwania.
2. Jeśli chcesz sklonować przykładową aplikację za pomocą programu Search Application Customizer:

- a. Uruchom program Search Application Customizer, dodając do niego nazwę pliku konfiguracyjnego, który chcesz utworzyć. W poniższym przykładzie tworzony jest plik o nazwie myNewFile.properties:

```
http://ESServer.com/ESSearchApplication/palette.do?configFile=/WEB-INF/
myNewFile.properties
```

Taki plik jeszcze nie istnieje, dlatego stosowane są wartości pobierane z pliku konfiguracji domyślnej aplikacji wyszukiwania (config.properties).

**Wskazówka:** Plik konfiguracji aplikacji wyszukiwania można też przygotować, klikając opcję **Załaduj** po uruchomieniu programu Search Application Customizer i podaniu w nim nazwy pliku. Plik tworzony jest po kliknięciu opcji **Zapisz** zapisującej wszystkie zmienione opcje.

- b. Jeśli w konfiguracji serwera WebSphere Application Server włączono globalny mechanizm zabezpieczeń, należy zalogować się, podając poprawny identyfikator i hasło.
  - c. Podaj nazwę swojej aplikacji wyszukiwania oraz opcje jakie mają zostać wykorzystane przy wyszukiwaniu w kolekcjach i kliknij przycisk **Zapisz**. Wszystkie wprowadzone zmiany zapisane zostaną do pliku konfiguracji w podkatalogu WEB-INF.  
Aby uzyskać pomoc dotyczącą dostępnych ustawień, kliknij opcję **Pomoc dla konfiguratora**. Aby zobaczyć skutki niektórych zmian, na przykład co do prezentacji wyników wyszukiwania, wpisz zapytanie i kliknij przycisk **Szukaj**.
3. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerze wyszukiwania, po czym zatrzymaj i uruchom ponownie aplikację ESSearchServer.

Systemy **AIX, Linux lub Solaris**

```
./stopServer.sh ESSearchServer
./startServer.sh ESSearchServer
```

Systemy **Windows**

```
stopServer ESSearchServer
startServer ESSearchServer
```

Skrypty te znajdują się w katalogu WAS\_INSTALL\_ROOT/AppServer/bin:

- W przypadku serwera aplikacji WebSphere w wersji 5 domyślna ścieżka instalacji w systemach AIX ma postać /usr/WebSphere, w systemach Linux lub Solaris ma postać /opt/WebSphere, natomiast w systemach Windows - C:\Program Files\WebSphere.
- W przypadku serwera aplikacji WebSphere w wersji 6 domyślna ścieżka instalacji w systemach AIX ma postać /usr/IBM/WebSphere, w systemach Linux lub Solaris ma postać /opt/IBM/WebSphere, a w systemach Windows - C:\Program Files\IBM\WebSphere.

## Analizowanie najczęstszych wyników

Można pomóc użytkownikom w zawężeniu zestawu wyników wyszukiwania, konfigurując opcje analizowania najczęstszych wyników.

Analiza najczęstszych wyników polega zasadniczo na klasyfikowaniu bieżącego zestawu wyników wyszukiwania na podstawie częstotliwości, z jaką występują analizowane wartości. Użytkownicy mogą filtrować wyniki wyszukiwania przez wybranie jednej z analizowanych wartości, co spowoduje automatyczne dodanie wybranej wartości jako nowego słowa kluczowego do kryteriów wyszukiwania. Wartość analizy najczęstszych wyników polega na tym, że umożliwia użytkownikom precyzowanie wyników wyszukiwania bez konieczności używania złożonej lub zaawansowanej składni wyszukiwania.

W aplikacji wyszukiwania korporacyjnego można za pomocą wykresów słupkowych pokazać, które wartości pól metadanych występują najczęściej w wynikach wyszukiwania. Można również utworzyć niestandardową klasę Java, która posłuży do przedstawienia najczęstszych wyników, w tym wyników z niekorporacyjnych źródeł wyszukiwania, w dowolnym formacie HTML.

### Ograniczenia

Upewnij się, że na komputerze, na którym działa serwer WebSphere Application Server oraz aplikacja wyszukiwująca, są zainstalowane czcionki odpowiednie dla używanego języka. Ten krok jest niezbędny dla zapewnienia, że przy generowaniu wykresu słupkowego zostanie ustawiona czcionka umożliwiająca poprawne przedstawienie znaków zawartych w etykiecie wykresu. Ma to znaczenie krytyczne zwłaszcza w przypadku języków azjatyckich. Jeśli czcionki zostaną zainstalowane po zainstalowaniu serwera WebSphere Application Server, należy zrestartować serwer WebSphere Application Server, aby zmiany odniosły skutek.

Jeśli aplikacja wyszukiwująca działa jako aplikacja autonomiczna, wykresy najczęstszych wyników można skonfigurować przy użyciu Konfiguratora aplikacji wyszukiwującej lub przez edycję pliku konfiguracyjnego aplikacji wyszukiwującej. Jeśli aplikacja wyszukiwująca działa jako portlet w ramach portalu WebSphere Portal, należy dokonać edycji właściwości i skonfigurować instancję portletu za pomocą interfejsu administracyjnego portalu WebSphere Portal. Do konfigurowania opcji analizy najczęstszych wyników nie można użyć Konfiguratora aplikacji wyszukiwującej.

### Informacje o tym zadaniu

Można zastosować graficzną reprezentację najczęstszych wyników, prezentując na wykresie słupkowym najczęściej występujące wartości pól metadanych. Można również za pomocą klasy Java rozszerzyć aplikację wyszukiwującą i prezentować najczęstsze wyniki w innym formacie, na przykład za pomocą kodu HTML można prezentować najczęstsze wyniki w postaci listy nieuporządkowanej.

## Wykresy słupkowe dla pól metadanych

Można określić opcje analizowania pól metadanych i wyświetlania wyników tej analizy na wykresach słupkowych. Wykresy są wyświetlane jako dodatek do wyników zapytania użytkownika. Każdy wykres odpowiada jednemu polu metadanych (takiemu jak wielkość dokumentu, autor, data itd.), a każdy słupek na wykresie odpowiada wartości danego pola.

Długość słupka reprezentuje liczbę dokumentów zawierających konkretną wartość pola w stosunku do pozostałych dokumentów, które zawierają inne wartości w danym polu. Im dłuższy jest słupek, tym większa jest liczba wystąpień danej wartości pola. Użytkownicy mogą doprecyzować wyniki, klikając słupek na wykresie słupkowym. Wartość pola reprezentowana przez wybrany słupek zostaje dodana jako dodatkowy termin zapytania, a nowe wyniki wyszukiwania zostają zawężone przez dodatkowe kryterium wyszukiwania.

Podczas konfigurowania przestrzeni przeszukiwania dla przeszukiwacza można wybrać opcję określania opcji wyszukiwania dotyczących pól metadanych. Na przykład można określić, czy pole metadanych może: podlegać wyszukiwaniu pełnotekstowemu, podlegać wyszukiwaniu na podstawie nazwy pola, być pokazywane w wynikach wyszukiwania, podlegać wyszukiwaniu jako dane parametryczne itd. Aby w wynikach wyszukiwania wyświetlane były wykresy dla pól metadanych, należy skonfigurować opcje pól metadanych dla przeszukiwacza. Należy określić przynajmniej, że pole jest polem podlegającym wyszukiwaniu oraz że może być pokazywane w wynikach wyszukiwania.

Jeśli podczas konfigurowania opcji dla pola metadanych zostanie wybrana opcja **Pelna zgodność**, to gdy użytkownik wybierze słupek z wykresu słupkowego, uruchomione zostanie zapytanie pełnej zgodności. W przypadku wyszukiwania opartego na pełnej zgodności wyniki są zwracane tylko wtedy, gdy termin użyty w zapytaniu jest zgodny z całą wartością pola. Jeśli pole zawiera mniej treści lub dodatkową treść, zgodność nie występuje.

Jeśli podczas konfigurowania opcji dla pola metadanych nie zostanie wybrana opcja **Pelna zgodność**, to gdy użytkownik wybierze słupek z wykresu słupkowego, uruchomione zostanie zapytanie oparte na polu. W tym przypadku w polu musi wystąpić dodatkowy termin użyty w zapytaniu, lecz nie musi być on zgodny z całą wartością pola.

## Wyniki formatowane za pomocą kodu HTML

Interfejs API `com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` można rozszerzyć z zastosowaniem niestandardowej klasy Java. W razie zastosowania tego podejścia do wyświetlania wyników można za pomocą kodu HTML sformatować prezentację dowolnych wyników wyszukiwania, a nie tylko zapytań, według których prowadzone są wyszukiwania w polach metadanych. Wyniki mogą być zwracane dla dokumentów w kolekcjach wyszukiwania korporacyjnego lub pochodzić z wyszukiwań w zewnętrznych repozytoriach i serwisach WWW.

Każdy wykres odpowiada jednej klasie Java, a dana klasa określa, jak najczęstsze wyniki mają być prezentowane na wykresie. Na przykład `com.ibm.es.searchui.charts.servlet.DynamicMostRecentDocuments`, przykładowa klasa udostępniana z przykładowym kodem wyszukiwania korporacyjnego, służy do prezentowania najczęstszych wyników w postaci listy nieuporządkowanej. Dokumenty są sortowane według daty i wyświetlane są wyłącznie tytuły i daty dokumentów.

Kolejna przykładowa klasa udostępniana wraz z wyszukiwaniem korporacyjnym, `com.ibm.es.searchui.charts.servlet.DogearSearchResults`, pokazuje, jak można dostarczyć użytkownikom listę zakładki z programu Lotus Connections Dogear odnoszących się do zapytania użytkownika.



Gdy użytkownik kliknie jeden z wyników wyszukiwania wśród niestandardowo sformatowanych danych wyjściowych, w nowym oknie przeglądarki zostanie wyświetlony odpowiedni dokument.

## Procedura

Niniejsza procedura pokazuje, jak za pomocą Konfiguratora aplikacji wyszukującej określić, że pola metadanych mają być poddawane analizie. Najczęściej występujące wartości pól metadanych są prezentowane graficznie na wykresie słupkowym.

1. Otwórz Konfigurator aplikacji wyszukującej. Jeśli nie jest jeszcze wyświetlony, załaduj plik konfiguracyjny aplikacji wyszukującej, która ma zostać skonfigurowana, a następnie kliknij opcję **Zastosuj**.
2. Przewiń w dół do pozycji **Wykres najczęstszych wyników** i kliknij opcję **Dodaj wykres**.
3. Gdy zostanie wyświetlona lista opcji wykresu, zostaw pole wyboru **Wykres niestandardowy** niezaznaczone (tę opcję zaznacza się tylko wtedy, gdy została utworzona niestandardowa klasa Java do analizowania i zwracania najczęstszych wyników), natomiast zaznacz pole wyboru **Włącz wykres**, aby użytkownikom przeglądającym wyniki wyszukiwania wyświetlany był wykres słupkowy.
4. W polu **Tytuł wykresu** podaj opisową etykietę dla wykresu. Podana tu wartość zastąpi tekst zastępczy **Nowy wiersz wykresu**.
5. Zdecyduj, czy pole wyboru **Włącz dynamiczny wybór pól** ma być zaznaczone. Jeśli ta opcja zostanie włączona, użytkownicy będą mogli wybrać inne pola podczas wyświetlania wyników wyszukiwania i zobaczyć analizę najczęstszych wyników dla danego pola.

Tej opcji można używać, aby umożliwić użytkownikom zawężanie zestawu wyników wyszukiwania. Na przykład użytkownik może wykonać wyszukiwanie w bazie danych w celu znalezienia informacji na temat zatrudnionych kobiet. Po wprowadzeniu zapytania do wyszukiwania w polu określającym płeć pracownika (na przykład plec:K) wartości najczęstszych wyników są wyświetlane jako słupki na wykresie słupkowym. Następnie użytkownik wybiera element **stanowisko** z listy pól istniejących w tej początkowej tabeli wynikowej. Wykres słupkowy wyświetli teraz wyniki dla najczęstszych stanowisk obejmowanych przez kobiety. Na końcu użytkownik wybiera element **projektant** z listy dostępnych pól. Wyniki podają teraz informacje na temat kobiet zatrudnionych w charakterze projektantów.

6. W polu **Nazwa pola metadanych** wpisz nazwę pola metadanych, którego wartości mają być analizowane na potrzeby tego wykresu.
7. Jeśli pole zawiera wiele wartości rozdzielonych znakami ograniczającymi, podaj taki znak w polu **Separator wartości pola**. Na przykład, jeśli pole zawiera dwie wartości, takie jak oddzielone średnikiem imię i nazwisko klienta, można tu podać średnik, aby każda z wartości została oddzielnie dodana do wykresu słupkowego. Jeśli nie zostanie podany znak separatora, cała wartość pola będzie analizowana jako pojedynczy element.
8. Określ opcje wyświetlania wykresu słupkowego, takie jak wielkość i kolor słupków oraz liczba słupków, jaka ma być wyświetlona, gdy wykres jest zwinięty lub rozwinięty.
9. Określ opcje sortowania wyników analizy. Na przykład można sortować według wartości pola lub według częstotliwości, z jaką wartości te występują.
10. Jeśli informacje dla konkretnych wartości pola mają być zawsze wyświetlane, nawet wtedy gdy w wynikach wyszukiwania nie będzie wystąpienia danej wartości pola, kliknij opcję **Dodaj wiersz** w obszarze **Etykiety niestandardowe**. Określ wartość pola, która ma być pokazywana na wykresie słupkowym nawet wtedy, gdy żadne wystąpienie tej wartości nie jest zwracane w wynikach wyszukiwania, oraz określ etykietę dla tego słupka na wykresie słupkowym.

Na przykład, jeśli zawsze ma być widoczne, czy konkurent, na przykład firma JK i Spółka, występuje w wynikach wyszukiwania nawet wtedy, gdy częstotliwość nie kwalifikuje go do pokazywania w najczęstszych wynikach, określ JK i Spółka jako wartość źródłową do analizowania, a następnie podaj opisową etykietę do identyfikowania tego słupka na wykresie słupkowym.

11. Kliknij opcję **Zastosuj** obok tytułu wykresu, aby zastosować opcje określone dla tego wykresu.
12. Jeśli wyniki wyszukiwania już są wyświetlane w obszarze aplikacji wyszukiwującej, analiza najczęstszych wyników zostanie zastosowana do bieżącego wyszukiwania i zostanie wyświetlony nowy wykres. W przeciwnym razie wprowadź zapytanie w celu przetestowania wyświetlania wykresu słupkowego.
13. Jeśli wykres jest satysfakcjonujący, kliknij opcję **Zapisz**, aby zaktualizować plik konfiguracyjny aplikacji wyszukiwującej.

#### Pojęcia pokrewne



Określanie właściwości zapytania



Klasy Java do wyświetlania najczęstszych wyników

---

## Dostęp do aplikacji wyszukiwujących

Dostęp do aplikacji wyszukiwujących odbywa się przez podanie odpowiedniego adresu URL w przeglądarce.

### Zagadnienia wstępne

Aplikację wyszukiwującą należy skonfigurować dla używanego środowiska serwera WWW.

### Informacje o tym zadaniu

Przykładowa aplikacja wyszukiwująca jest instalowana na serwerach wyszukiwania korporacyjnego. Z aplikacji tej można korzystać w udostępnionej postaci do testowania kolekcji i źródeł zewnętrznych przed ich udostępnieniem użytkownikom. Ponadto aplikacja przykładowa może służyć jako model przy tworzeniu własnych aplikacji wyszukiwujących.

### Procedura

Aby uruchomić aplikację wyszukiwującą:

1. Wpisz adres URL aplikacji wyszukiwującej w przeglądarce WWW. Na przykład:

`http://SerwerWyszukiwania.com/ESSearchApplication/`

*SerwerWyszukiwania.com* to nazwa hosta serwera wyszukiwania.

Jeśli konfiguracja serwera WWW nie przewiduje korzystania z portu 80, wówczas trzeba podać poprawny numer portu. Na przykład:

`http://SerwerWyszukiwania.com:9080/ESSearchApplication/`

2. Aby uruchomić niestandardową aplikację wyszukiwującą, należy wpisać adres URL przykładowej aplikacji i dołączyć do niego nazwę pliku konfiguracyjnego aplikacji niestandardowej. Na przykład:

`http://SerwerWyszukiwania.com/ESSearchApplication/search.do?configFile=/WEB-INF/mojaKonfiguracja.properties`

Jeśli wskazany w ten sposób plik nie istnieje, wyświetlona zostanie przykładowa aplikacja systemu wyszukiwania korporacyjnego.



3. Jeśli w konfiguracji serwera WebSphere Application Server włączono globalny mechanizm zabezpieczeń, należy zalogować się w aplikacji, podając poprawny identyfikator i hasło.

Jeśli dowolna z kolekcji dostępnych dla aplikacji wyszukującej ma włączone zabezpieczenia i jeśli zabezpieczone kolekcje zawierają przeszukiwacze skonfigurowane do sprawdzania poprawności referencji użytkownika podczas przetwarzania zapytania, można skonfigurować profil użytkownika. Na stronie *Mój profil* określ referencje umożliwiające dostęp do zabezpieczonych domen. Umożliwia to wyszukiwanie w domenach bez potrzeby logowania się do nich.

Jeśli przeszukiwacz zapewnia obsługę mechanizmu pojedynczego logowania (SSO), wyszukiwanie w domenach zabezpieczonych nie wymaga tworzenia profilu użytkownika.

4. Na stronie *Szukaj* wpisz zapytanie. Wyszukiwanie zostanie przeprowadzone we wszystkich kolekcjach i zewnętrznych źródłach wybranych na stronie *Preferencje*.

---

## Konfigurowanie serwerów wyszukiwania w celu zaakceptowania tylko bezpiecznych żądań (SSL)

W serwerach wyszukiwania można wyłączyć interfejs HTTP i skonfigurować akceptowanie żądań wyszukiwania przesyłanych za pomocą zabezpieczonego protokołu HTTPS.

### Informacje o tym zadaniu

Aby skonfigurować serwery wyszukiwania w celu wyłącznego używania protokołu SSL (Secure Sockets Layer) do przetwarzania żądań, należy wyłączyć interfejs HTTP. Należy również używać tego samego pliku kluczy na obydwu serwerach i na wszystkich komputerach klienckich, takich jak serwer WebSphere Portal, na którym zainstalowany jest portlet wyszukiwania korporacyjnego. Plik kluczy, który jest również nazywany *magazynem kluczy*, zawiera klucze publiczne, które umożliwiają używanie protokołu SSL w komunikacji.

### Procedura

Aby skonfigurować serwery wyszukiwania w celu zaakceptowania tylko bezpiecznych żądań:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W instalacji z wieloma serwerami wyszukiwania wykonaj następujące kroki na każdym serwerze indeksowania:
  - a. Zatrzymaj system wyszukiwania korporacyjnego:

```
esadmin system stopall
```
  - b. Otwórz plik `ES_NODE_ROOT/master_config/nodes.ini`.
  - c. Zmień wartość parametru `id_węzła.searchserverport` (zwykle 80) na wartość numeru portu używaną przez protokół HTTPS (zwykle 443) i zapisz plik. Dla instalacji z wieloma serwerami zaktualizuj obie wartości `id_węzła.searchserverport` (jedna dla każdego serwera wyszukiwania).
2. Aby zaktualizować serwer wyszukiwania, wykonaj następujące czynności: W instalacji z wieloma serwerami wyszukiwania wykonaj następujące kroki na każdym serwerze wyszukiwania:
  - a. Zaloguj się na serwerze wyszukiwania jako administrator wyszukiwania korporacyjnego.
  - b. Otwórz plik `ES_NODE_ROOT/nodeinfo/es.cfg`.
  - c. Zaktualizuj właściwość `TrustStore property`, aby określić pełną ścieżkę do pliku kluczy SSL.
  - d. Zaktualizuj właściwość `HTTPProtocol`, aby określić protokół HTTPS, następnie zapisz plik.

- e. Wydad następującą komendę, gdzie *hasło\_pliku\_kluczy* jest hasłem pliku kluczy. Komenda ta szyfruje wartość hasła i aktualizuje parametr TrustStorePassword w pliku *es.cfg*.

#### Systemy AIX, Linux lub Solaris

`eschangengetrustpw.sh hasło_pliku_kluczy`

#### Windows

`eschangengetrustpw hasło_pliku_kluczy`

- f. Upewnij się, że właściwości *trustStore* i *trustPassword* w pliku *config.properties* dla aplikacji wyszukiwania poprawnie określają pełną ścieżkę i hasło pliku kluczy. Informacje te można sprawdzić i poprawić, otwierając w edytorze plik *config.properties* lub używając Konfiguratora aplikacji wyszukiwanej.
  - g. Jeśli do wyszukiwania korporacyjnego używany jest portlet wyszukiwania, należy upewnić się, że parametry portletu *trustStore* i *trustPassword* określają poprawną pełną ścieżkę i hasło pliku kluczy. Aby sprawdzić lub zmienić te informacje, należy użyć opcji portletu zarządzającego w interfejsie administracyjnym produktu WebSphere Portal.
3. Zrestartuj system wyszukiwania korporacyjnego:
- ```
esadmin system startall
```

Zadania pokrewne

“Edycja właściwości przykładowej aplikacji wyszukiwanej” na stronie 238

“Konfigurowanie aplikacji wyszukiwanych” na stronie 239

“Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 5.1” na stronie 345

“Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 6” na stronie 350

Konfigurowanie serwerów wyszukiwania w celu akceptowania żądań z serwerów proxy

Serwer wyszukiwania można skonfigurować tak, aby akceptował żądania przechodzące przez serwer proxy.

Procedura

Aby umożliwić wysyłanie żądań do serwerów wyszukiwania poprzez serwer proxy:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego. W instalacji z wieloma serwerami, zaloguj się na serwerze indeksowania.
`esadmin system stopall`
2. Otwórz plik `ES_NODE_ROOT/master_config/nodes.ini`.
 - a. Zmień właściwość `ProxyServer` tak, aby wskazywała pełną nazwę hosta serwera proxy.
 - b. Zmień właściwość `ProxyServerPort` tak, aby wskazywała numer portu serwera proxy.
 - c. Opcjonalne: Jeśli serwer proxy wymaga uwierzytelniania wszystkich żądań, zmień właściwość `ProxyServerUserName` tak, aby wskazywała poprawną nazwę użytkownika serwera proxy.
 - d. Opcjonalne: Jeśli serwer proxy wymaga uwierzytelniania wszystkich żądań, wprowadź następującą komendę, gdzie *hasło_serwera_proxy* jest hasłem użytkownika serwera proxy. Komenda ta szyfruje wartość hasła i aktualizuje wartość parametru `ProxyServerUserPassword` w pliku *es.cfg*.

AIX, Linux i Solaris

`eschangeproxypw.sh haslo_serwera_proxy`

Windows

`eschangeproxypw haslo_serwera_proxy`

3. Upewnij się, że właściwości `proxyHost` i `proxyPort` w pliku konfiguracyjnym (pliku z rozszerzeniem `.properties`) aplikacji wyszukującej określają poprawną pełną nazwę hosta i numer portu serwera proxy.

Jeśli serwer proxy wymaga uwierzytelniania, upewnij się, że właściwości `proxyUser` i `proxyPassword` określają poprawną nazwę użytkownika i hasło dla serwera proxy.

Informacje te można sprawdzić i poprawić, otwierając w edytorze plik konfiguracyjny lub używając Konfiguratora aplikacji wyszukującej.

4. Jeśli do wyszukiwania korporacyjnego używany jest portlet wyszukiwania, upewnij się, że parametry portletu `proxyHost` i `proxyPort` określają poprawną pełną nazwę hosta i numer portu serwera proxy.

Jeśli serwer proxy wymaga uwierzytelniania, upewnij się, że właściwości `proxyUser` i `proxyPassword` określają poprawną nazwę użytkownika i hasło dla serwera proxy. Aby sprawdzić lub zmienić te informacje, należy użyć opcji portletu zarządzającego w interfejsie administracyjnym produktu WebSphere Portal.

5. Zrestartuj system wyszukiwania korporacyjnego:

`esadmin system startall`

Obsługa źródeł zewnętrznych

Źródło zewnętrzne jest źródłem danych udostępnianym do wyszukiwania przy użyciu aplikacji wyszukiwania korporacyjnego bez konieczności przeszukiwania, analizowania i indeksowania dokumentów w tym źródle danych.

Jako źródła zewnętrzne na potrzeby wyszukiwania mogą służyć następujące typy źródeł danych:

- Bazy danych obsługujące protokół JDBC (Java Database Connectivity). Obsługiwane są tylko bazy danych IBM DB2, Oracle, Microsoft SQL Server 2000 i Microsoft SQL Server 2005. Dla każdej tabeli udostępnionej do wyszukiwania tworzone jest osobne źródło zewnętrzne.

Ograniczenie: Obsługa baz danych SQL Server 2005 jest ograniczona do tabel, które nie zawierają danych typu Variant. Sterownik JDBC dla serwera SQL Server 2005 nie jest obsługiwany w systemach AIX.

- Serwery LDAP (Lightweight Directory Access Protocol). Dla każdego serwera LDAP tworzone jest jedno źródło zewnętrzne.

Po skonfigurowaniu informacji na temat źródła zewnętrznego konieczne jest powiązanie go z co najmniej jedną aplikacją wyszukującą. Użytkownicy mogą następnie wyszukiwać dane w tym źródle zewnętrznym jednocześnie z wprowadzaniem zapytań dla kolekcji utworzonych przez przeszukanie, przeanalizowanie i poindeksowanie danych na potrzeby wyszukiwania korporacyjnego.

Pojęcia pokrewne



Programy stowarzyszające interfejsu API wyszukiwania i indeksowania

Dodawanie do systemu źródeł zewnętrznych

Podczas dodawania zewnętrznego źródła do systemu wyszukiwania korporacyjnego należy określić typ dodawanego źródła. Kreator pomaga w określeniu informacji na temat źródła danych i sposobu wyszukiwania w nim danych.

Zagadnienia wstępne

Do dodawania źródeł zewnętrznych do systemu wymagana jest rola administratora wyszukiwania korporacyjnego.

Ograniczenia

Aby możliwe było przeszukiwanie bazy danych Oracle jako źródła zewnętrznego, na serwerach wyszukiwania korporacyjnego musi być zainstalowany program kliencki Oracle.

Sterownik JDBC dla serwera Microsoft SQL Server 2005 nie jest obsługiwany w systemach AIX.

Informacje o tym zadaniu

Dodanie do systemu informacji o źródle zewnętrznym pozwala użytkownikom wprowadzać zapytania dla tego źródła za pomocą aplikacji wyszukiwania korporacyjnego. Włączyć można wyszukiwanie na serwerach LDAP (Lightweight Directory Access Protocol) i w tabelach baz danych JDBC (Java database connectivity).

Podczas konfigurowania serwera LDAP kreator pomaga w określeniu informacji umożliwiających nawiązanie przez system połączenia z tym serwerem oraz w określeniu opcji dotyczących sposobów wyszukiwania na tym serwerze.

Podczas konfigurowania bazy danych JDBC kreator pomaga w określeniu informacji umożliwiających nawiązanie przez system połączenia z tą bazą danych, w wybraniu tabel, które mają zostać udostępnione do wyszukiwania, a także w określeniu opcji dotyczących sposobów wyszukiwania danych w tych tabelach. Dla każdej tabeli dodawanej do systemu tworzone jest źródło danych podlegające odrębnym operacjom wyszukiwania.

Informacje dotyczące sterowników dla serwera SQL Server 2000 znajdują się pod adresem <http://www.microsoft.com/downloads/details.aspx?familyid=07287B11-0502-461A-B138-2AA54BFDC03A&displaylang=en>. Informacje dotyczące sterowników dla serwera SQL Server 2005 znajdują się pod adresem <http://www.microsoft.com/downloads/details.aspx?familyid=e22bc83b-32ff-4474-a44a-22b6ae2c4e17&displaylang=en>.

Procedura

Aby dodać do systemu źródło zewnętrzne:

1. Aby włączyć bazy danych JDBC do systemu wyszukiwania korporacyjnego, przed dodaniem źródła zewnętrznego wykonaj następujące czynności. Czynność ta umożliwia systemowi znalezienie właściwych sterowników JDBC i należy ją wykonać jeden raz.
 - a. Zaloguj się na serwerze przeszukiwacza jako administrator wyszukiwania korporacyjnego.
 - b. Dokonaj edycji pliku `ES_INSTALL_ROOT/configurations/interfaces/discovery__interface.ini` i określ zmienne środowiskowe `CLASSPATH` oraz `LD_LIBPATH`, aby zawierały ścieżkę klasy do sterowników JDBC oraz ścieżkę do plików bibliotek.
 - c. Dokonaj edycji pliku `ES_INSTALL_ROOT/configurations/interfaces/customcommunication__interface.ini` i określ zmienne środowiskowe `CLASSPATH` oraz `LD_LIBPATH`, aby zawierały ścieżkę klasy do sterowników JDBC oraz ścieżkę do plików bibliotek.
 - d. Opcjonalnie: aby używać sterownika JDBC Oracle do obsługi lokalnych lub wpisanych do katalogu baz danych, dodaj ścieżkę do bibliotek do zmiennej środowiskowej `LD_LIBPATH` (na przykład `LD_LIBPATH=../home/oracle/OraHome1/lib32`) i dokonaj edycji pliku `escrset.sh`, aby określał ścieżkę do bibliotek, a następnie wyeksportuj zmienną zawierającą katalog instalacyjny Oracle. Na przykład:

```
ORACLE_HOME=/home/oracle/OraHome1
export ORACLE_HOME
```
 - e. Zrestartuj system wyszukiwania korporacyjnego razem z powszechną warstwą komunikacji (CCL):

Systemy AIX, Linux lub Solaris

```
esadmin stop
stopccl.sh
startccl.sh
esadmin start
```

Wiersz komend systemu Windows

```
esadmin stop
stopccl
startccl
esadmin start
```

Narzędzie administracyjne usług systemu Windows

Aby uruchomić warstwę CCL w tle:

- 1) Wpisz `esadmin stop`.
 - 2) Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
 - 3) Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Zatrzymaj**. Po zatrzymaniu usługi kliknij przycisk **Uruchom**.
 - 4) Wpisz `esadmin start`.
2. Kliknij opcję **Źródła zewnętrzne**, aby otworzyć widok Źródła zewnętrzne.
 3. Kliknij opcję **Dodaj źródło zewnętrzne**.
 4. Wybierz typ źródła zewnętrznego, które ma zostać dodane (serwer LDAP lub baza danych JDBC).
 5. Kliknij przycisk **Dalej**, aby rozpocząć konfigurowanie źródła zewnętrznego.

Otwarty zostanie kreator odpowiedni dla typu tworzonego źródła. Postępuj zgodnie z instrukcjami wyświetlanymi w kreatorze, aby skonfigurować źródło zewnętrzne. Kliknij przycisk **Pomoc** na dowolnej stronie kreatora, aby dowiedzieć się więcej na temat opcji dostępnych do określenia.

Poniższe domyślne nazwy i położenia sterownika JDBC mogą okazać się pomocne podczas konfigurowania informacji na temat połączeń dla baz danych DB2 i Oracle:

DB2: Legacy JDBC Driver

Nazwa sterownika: `COM.ibm.db2.jdbc.app.DB2Driver`

Przykładowe położenie: `główny_kat_inst_db2/java/db2java.zip`

DB2: Universal JDBC Driver

Nazwa sterownika: `com.ibm.db2.jcc.DB2Driver`

Przykładowe położenia:

`główny_katalog_instalacyjny_db2/java/db2jcc.jar`

`główny_katalog_instalacyjny_db2/java/db2jcc_license_cu.jar`

Oracle Nazwa sterownika: `oracle.jdbc.driver.OracleDriver`

Przykładowe położenie: `główny_kat_inst_oracle/jdbc/lib/ojdbc14.jar`

Microsoft SQL Server 2000

Nazwa sterownika: `com.microsoft.jdbc.sqlserver.SQLServerDriver`

Przykładowe położenia:

`katalog_główny_mssql_jdbc/lib/mssqlserver.jar`

`katalog_główny_mssql_jdbc/lib/msbase.jar`

`katalog_główny_mssql_jdbc/lib/msutil.jar`

Microsoft SQL Server 2005

Nazwa sterownika: `com.microsoft.sqlserver.jdbc.SQLServerDriver`

Przykładowe położenie: `kat_inst/sqljdbc_1.0/język/sqljdbc.jar`

Na przykład: `kat_inst/sqljdbc_1.0/enu/sqljdbc.jar`

6. Po określeniu opcji dotyczących wyszukiwania w źródle zewnętrznym kliknij przycisk **Zakończ**.

Nowe źródło zewnętrzne zostanie wymienione w widoku Źródła zewnętrzne wraz z innymi źródłami zewnętrznymi dodanymi do systemu.

Pojęcia pokrewne

 Programy stwarzające interfejsu API wyszukiwania i indeksowania

Powiązanie aplikacji wyszukujących ze źródłami zewnętrznymi

Zanim możliwe będzie wyszukiwanie w źródle zewnętrznym należy powiązać z nim co najmniej jedną aplikację wyszukującą.

Zagadnienia wstępne

Do powiązania aplikacji wyszukujących ze źródłami zewnętrznymi, w których mogą one wykonywać wyszukiwanie, wymagana jest rola administratora wyszukiwania korporacyjnego.

Procedura

Aby powiązać aplikację wyszukującą z jednym lub większą liczbą źródeł zewnętrznych:

1. Kliknij opcję **Zabezpieczenia** na pasku narzędzi konsoli administracyjnej.
2. Na stronie Aplikacje wyszukujące kliknij opcję **Konfiguruj aplikacje wyszukujące**.
3. Na stronie Konfigurowanie aplikacji wyszukujących kliknij opcję **Dodaj aplikację wyszukującą**.
4. Wpisz nazwę aplikacji wyszukującej.
5. Wybierz źródła zewnętrzne, w których aplikacja będzie mogła wyszukiwać:
 - Kliknij opcję **Wszystkie kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukująca ma mieć dostęp do wszystkich źródeł zewnętrznych dodanych do systemu.
 - Kliknij opcję **Konkretne kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukująca ma mieć dostęp tylko do określonych źródeł zewnętrznych.

Po wybraniu tej opcji zostanie wyświetlona lista nazw kolekcji i źródeł zewnętrznych. Należy zaznaczyć pole wyboru **Wybierz** dla każdego źródła zewnętrznego, w którym aplikacja będzie mogła wyszukiwać.
6. Kliknij przycisk **OK**.

Pojęcia pokrewne



Programy stowarzyszające interfejsu API wyszukiwania i indeksowania

Zabezpieczenia wyszukiwania korporacyjnego

Mechanizmy zabezpieczeń w wyszukiwaniu korporacyjnym pozwalają chronić źródła przed nieautoryzowanym wyszukiwaniem oraz ograniczać grupę użytkowników, którzy będą mogli korzystać z funkcji administracyjnych.

Wyszukiwanie korporacyjne umożliwia wyszukiwanie w szerokim zakresie źródeł danych. Wyszukiwanie korporacyjne oferuje kilkupoziomowy mechanizm koordynowania i egzekwowania zabezpieczeń pozwalający zagwarantować, że dostęp do danych mają wyłącznie użytkownicy do tego uprawnieni oraz że tylko uprawnieni użytkownicy mają dostęp do konsoli administracyjnej.

Wskazówka:

Szczegółowe przykłady konfigurowania zabezpieczeń dla wyszukiwania korporacyjnego można znaleźć w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition). Scenariusze pokazują, jak włączać zabezpieczenia globalne na serwerze WebSphere Application Server z repozytorium LDAP, konfigurować komponent do zarządzania tożsamościami oraz konfigurować rozmaite przeszukiwacze dla zapewnienia egzekwowania zabezpieczeń na poziomie dokumentu.

Serwer WWW

Pierwszym poziomem zabezpieczeń jest serwer WWW. Jeżeli na serwerze aplikacji WebSphere włączone zostaną zabezpieczenia globalne, to możliwe jest przydzielanie użytkownikom ról administracyjnych oraz uwierzytelnianie użytkowników administrujących całym systemem. Po zalogowaniu się do konsoli administracyjnej użytkownikowi udostępniane są tylko te funkcje i kolekcje, którymi może on administrować.

Aplikacje wyszukiwania mogą też korzystać z mechanizmów zabezpieczeń serwera aplikacji WebSphere do uwierzytelniania użytkowników przeszukujących kolekcje.

Zabezpieczenia na poziomie kolekcji

Podczas tworzenia kolekcji można włączyć zabezpieczenia na poziomie kolekcji. Ustawienia tego nie można zmienić po utworzeniu kolekcji. Jeśli nie włączono zabezpieczeń na poziomie kolekcji, nie można później ustawiać opcji zabezpieczeń na poziomie dokumentów.

Jeśli zabezpieczenia na poziomie kolekcji są włączone, to:

- Procesy analizy globalnej wyszukiwania korporacyjnego stosują inne reguły do indeksowania podwójnych dokumentów;
- Możliwe jest konfigurowanie opcji wymuszających zabezpieczenia na poziomie dokumentów, na przykład łączenie tokenów zabezpieczeń z przeszukiwanymi dokumentami, wprowadzenie wymogu sprawdzania poprawności bieżących referencji w czasie przetwarzania zapytań oraz zaznaczanie, czy w dokumentach indeksowany ma być tekst zakotwiczenia.
- Odwzorowując aplikacje wyszukiwania (a nie poszczególnych użytkowników) na kolekcje i przeszukiwane przez nie źródła zewnętrzne, można wymusić stosowanie zabezpieczeń. Następnie dopuszcza się użytkowników lub odmawia im dostępu do aplikacji wyszukiwania, stosując przy tym standardowe mechanizmy kontroli dostępu.

Jest to kompromis pomiędzy włączeniem zabezpieczeń kolekcji a jakością wyszukiwania. Włączenie zabezpieczeń kolekcji zmniejsza ilość informacji indeksowanej dla każdego dokumentu. Efektem ubocznym jest to, że w wyniku niektórych zapytań zwracana będzie mniejsza liczba trafień.

Zabezpieczenia na poziomie dokumentu

Podczas konfigurowania przeszukiwaczy dla kolekcji można uaktywnić zabezpieczenia na poziomie dokumentów. Na przykład można określić opcje nakazujące powiązać tokeny zabezpieczeń z danymi zbieranymi przez przeszukiwacze. Aplikacja wyszukująca może wtedy skorzystać tokenów przechowywanych w dokumentach znajdujących się w indeksie i w ten sposób wymusić kontrolę dostępu oraz sprawić, że dostęp do danych będą mieli wyłącznie użytkownicy z odpowiednimi referencjami.

W przypadku pewnych źródeł danych opcje można skonfigurować tak, aby w czasie przetwarzania zapytania kontrola dostępu realizowana była przez sprawdzanie poprawności referencji logowania użytkownika. Ta dodatkowa warstwa zabezpieczająca sprawia, że uprawnienia użytkowników sprawdzane są w czasie rzeczywistym względem własnego źródła danych. Możliwość ta stanowi ochronę przed sytuacjami, w których referencje użytkownika zmieniają się po zindeksowaniu dokumentu i jego tokenów zabezpieczeń.

W fazie analizy globalnej odpowiedzialnej za przetwarzanie tekstu zakotwiczenia standardowo tekst znajdujący się w jednym dokumencie (dokument źródłowy) kojarzony jest z innym dokumentem (dokumentem docelowym), w którym wcale nie musi on występować. W czasie konfigurowania przeszukiwacza sieciowego można nakazać mu wykluczać z indeksu tekst zakotwiczenia, jeśli tekst ten jest odsyłaczem do dokumentu, którego przeszukiwacz nie może przeszukać.

Zabezpieczenie kolekcji wychodzi poza mechanizmy uwierzytelniania i kontroli dostępu używane przez wyszukiwanie korporacyjne do zabezpieczania indeksowanych treści. Istnieją także zabezpieczenia przed przechwytywaniem danych przez nieupoważnionych użytkowników w trakcie ich przesyłania. Na przykład do komunikowania się z serwerem indeksowania i z aplikacją wyszukującą serwery wyszukiwania wykorzystują takie protokoły, jak SSL, SSH i HTTPS.

Dodatkowym zabezpieczeniem jest szyfrowanie. Na przykład hasło administratora wyszukiwania korporacyjnego określone podczas instalowania produktu jest przechowywane w formie zaszyfrowanej. Hasła zapisane przez użytkowników w ich profilach są także zaszyfrowane.

Dla dodatkowego bezpieczeństwa należy się upewnić, że komponenty sprzętowe serwera są należycie izolowane i chronione przed ingerencją nieupoważnionych osób. Instalując firewall, można zabezpieczyć serwery wyszukiwania korporacyjnego przed ingerencją za pośrednictwem innego segmentu sieci. Należy się też upewnić, że na serwerach wyszukiwania korporacyjnego nie ma żadnych otwartych portów. System należy skonfigurować w taki sposób, aby nasłuchiwanie nadchodzących żądań odbywało się wyłącznie na portach w sposób jawny przypisanych aplikacjom i operacjom wyszukiwania korporacyjnego.

Bezpieczeństwo instalacji

Program instalacyjny programu OmniFind Enterprise Edition tworzy środowisko wymuszające odpowiednie zabezpieczenia dotyczące administrowania kolekcjami lub przeszukiwania ich przez mechanizm wyszukiwania korporacyjnego.

Identyfikator administratora wyszukiwania korporacyjnego

W procesie instalacji instalator wyświetla zachętę do podania identyfikatora użytkownika oraz hasła, które będą używane jako dane administratora wyszukiwania korporacyjnego. System korzysta z wprowadzonych referencji do:

- uwierzytelniania administratora wyszukiwania korporacyjnego w czasie wykonywania różnych zadań administracyjnych;
- przygotowania wewnętrznej bazy danych wyszukiwania korporacyjnego;
- uruchomienia wszystkich sesji lub procesów wyszukiwania korporacyjnego.

Identyfikator użytkownika podawany w czasie instalacji musi być prawidłowym identyfikatorem użytkownika w systemie operacyjnym z uprawnieniami do administrowania. Program instalacyjny przechowuje te referencje w postaci zaszyfrowanej w pliku właściwości każdego serwera wyszukiwania korporacyjnego.

Szyfrowanie

W celu ochrony istotnych danych, komunikaty z danymi uwierzytelniania przesyłane przez system wyszukiwania korporacyjnego są szyfrowane. Proces ten powoduje pewien nakład pracy wynikający z konieczności zaszyfrowania identyfikatorów i haseł uwierzytelniania. Wszystkie przechowywane w systemie hasła (w plikach konfiguracyjnych, bazie danych wyszukiwania korporacyjnego itd.) również zapisywane są w postaci zaszyfrowanej.

Bezpieczeństwo serwera aplikacji WebSphere

Jeśli na serwerze wyszukiwania nie został wcześniej zainstalowany serwer aplikacji WebSphere, to program instalacyjny w sposób cichy zainstaluje ten serwer z wyłączonymi funkcjami globalnych zabezpieczeń. Jeśli później w serwerze aplikacji WebSphere zostaną włączone globalne zabezpieczenia, to serwer ten będzie odpowiedzialny za uwierzytelnianie administratora wyszukiwania korporacyjnego.

Jeśli włączone zostaną zabezpieczenia globalne, to trzeba też podać identyfikator administratora wyszukiwania korporacyjnego oraz jego hasło, które w czasie instalowania serwera aplikacji WebSphere zostały dodane do rejestru użytkowników, takiego jak katalog LDAP (Lightweight Directory Access Protocol).

Jeśli globalne zabezpieczenia zostaną włączone już po zainstalowaniu serwera OmniFind Enterprise Edition, to konieczne jest zaktualizowanie wartości konfiguracji i uruchomienie polecenia **eschangewaspw**, przez co referencje serwera aplikacji WebSphere zostaną zaszyfrowane i zapisane w pliku właściwości wyszukiwania korporacyjnego.

Uwierzytelnianie a kontrola dostępu

W ramach ochrony treści przed nieautoryzowanymi użytkownikami oraz kontroli dostępu do funkcji administracyjnych wyszukiwanie korporacyjne obsługuje mechanizmy uwierzytelniania i autoryzacji użytkowników (kontrola dostępu)

Uwierzytelnianie

Uwierzytelnianie jest procesem, w którym system weryfikuje tożsamość użytkownika próbującego uzyskać dostęp do systemu. Kontrola dostępu najczęściej opiera się na tożsamości użytkownika żądającego dostępu do danego zasobu, dlatego uwierzytelnianie jest niezwykle istotnym elementem skutecznych zabezpieczeń.

Uwierzytelnianie użytkowników wyszukiwania korporacyjnego implementowane jest za pomocą referencji, które muszą składać się przynajmniej z identyfikatora i hasła użytkownika.

Do uwierzytelniania użytkowników korzystających z konsoli administratora wyszukiwanie korporacyjne stosuje funkcje uwierzytelniania udostępniane przez serwer aplikacji WebSphere.

Autoryzacja (kontrola dostępu)

Autoryzacją może być dowolny mechanizm, poprzez który system udziela lub odbiera użytkownikowi prawo dostępu do pewnych danych lub prawo wykonania określonego działania. Często użytkownik musi zalogować się do systemu za pośrednictwem pewnego rodzaju uwierzytelniania. Mechanizmy kontroli dostępu porównują tożsamość użytkownika z listami kontroli dostępu, określając w ten sposób operacje, jakie użytkownik może lub jakich nie może wykonać. Kontrola dostępu obejmuje:

- uprawnienia do plików, takie jak prawo do tworzenia, odczytywania, edycji lub usuwania pliku;
- uprawnienia programów, takie jak prawo do uruchomienia programu;
- uprawnienia do danych, takie jak prawo do odczytywania lub aktualizowania informacji znajdujących się w bazie danych.

Role administracyjne

Wyszukiwanie korporacyjne wykorzystuje role jako metodę kontroli dostępu do różnych funkcji konsoli administracyjnej.

Jeśli zainstalowany jest program OmniFind Enterprise Edition (OmniFind Enterprise Edition), to instalator skonfiguruje identyfikator użytkownika oraz hasło na potrzeby administratora wyszukiwania korporacyjnego. Korzystając z konsoli administracyjnej po raz pierwszy, należy się zalogować jako użytkownik o skonfigurowanym identyfikatorze i hasle. Jeśli w serwerze aplikacji WebSphere nie zostaną włączone zabezpieczenia globalne, skonfigurowany identyfikator użytkownika będzie jedynym identyfikatorem pozwalającym na dostęp do administracyjnej konsoli wyszukiwania.

Jeśli globalny mechanizm zabezpieczeń serwera WebSphere Application Server zostanie włączony, można zarejestrować dodatkowych użytkowników jako użytkowników z administracyjnymi uprawnieniami wyszukiwania korporacyjnego. Przypisując użytkownikom do ról, można ograniczyć dostęp do wybranych kolekcji i kontrolować funkcje, z których każdy z użytkowników z uprawnieniami administracyjnymi będzie mógł korzystać. Identyfikatory użytkowników przypisywane do ról administracyjnych w wyszukiwaniu korporacyjnym muszą istnieć w rejestrze użytkowników serwera WebSphere Application Server.

Podczas logowania użytkownika z uprawnieniami administracyjnymi wyszukiwanie korporacyjne uwierzytelnia jego identyfikator. W konsoli będą dostępne wyłącznie te kolekcje i funkcje, którymi dany użytkownik ma prawo administrować.

Użytkowników można przypisywać do następujących ról administracyjnych:

Administrator wyszukiwania korporacyjnego

Ci użytkownicy tworzą kolekcje i mają prawo do administrowania wszystkimi aspektami funkcjonowania systemu wyszukiwania korporacyjnego. W trakcie instalacji produktu OmniFind Enterprise Edition określone są identyfikator i hasło użytkownika dla pierwszego użytkownika wyszukiwania korporacyjnego z uprawnieniami administracyjnymi. Po pierwszym zalogowaniu się użytkownik ten będzie mógł przypisać rolę administratora wyszukiwania korporacyjnego innym użytkownikom.

Administrator kolekcji

Ci użytkownicy mogą edytować i monitorować operacje na kolekcjach, do których administrowania mają uprawnienia, oraz sterować tymi operacjami. Nie mogą oni tworzyć kolekcji. Administratorzy kolekcji mogą monitorować i obsługiwać działania systemowe, tylko jeśli takie uprawnienie zostanie im nadane przez administratora wyszukiwania korporacyjnego.

Operator

Ci użytkownicy mogą monitorować operacje na kolekcjach, do których administrowania mają uprawnienia, i sterować nimi. Mogą oni na przykład uruchamiać i zatrzymywać działania kolekcji, ale nie mogą tworzyć ani edytować kolekcji. Operator może monitorować i obsługiwać działania systemowe, tylko jeśli takie uprawnienie zostanie mu nadane przez administratora wyszukiwania korporacyjnego.

Monitor

Ci użytkownicy mogą monitorować operacje na kolekcjach, do których administrowania mają uprawnienia. Nie mogą oni sterować operacjami (takimi jak uruchamianie i zatrzymywanie serwerów), tworzyć kolekcji ani ich edytować. Użytkownik monitorujący może obserwować, ale nie obsługiwać, działania systemowe, tylko jeśli takie uprawnienie zostanie mu nadane przez administratora wyszukiwania korporacyjnego.

Konfiguracja użytkowników z uprawnieniami administracyjnymi

Konfigurując role administracyjne można ograniczyć dostęp do kolekcji i sterować dostępem użytkowników z uprawnieniami administracyjnymi do poszczególnych funkcji.

Zagadnienia wstępne

Przed przypisaniem użytkownika do roli administracyjnej należy się upewnić, że na serwerze WebSphere Application Server jest włączone zabezpieczenie. Należy również upewnić się, że w rejestrze użytkowników serwera WebSphere Application Server istnieje identyfikator użytkownika.

Aby konfigurować użytkowników administracyjnych, wymagana jest rola administratora wyszukiwania korporacyjnego.

Procedura

Aby przypisać użytkowników do ról administracyjnych, należy wykonać poniższe czynności:

1. Kliknij przycisk **Zabezpieczenia**, aby otworzyć widok Zabezpieczenia.
2. Na stronie Role administracyjne kliknij przycisk **Dodaj użytkownika**.
3. Wpisz identyfikator użytkownika, który ma zostać zarejestrowany, i wybierz odpowiednią rolę administracyjną.
4. Jeśli użytkownik nie jest rejestrowany jako administrator wyszukiwania korporacyjnego, określ, czy może on uzyskiwać dostęp do stron z paska narzędzi **System**.

Na przykład można umożliwić niektórym operatorom lub administratorom kolekcji monitorowanie systemowych plików dziennika.

5. Jeśli użytkownik nie jest rejestrowany jako administrator wyszukiwania korporacyjnego, wybierz kolekcje i źródła zewnętrzne, którymi będzie mógł administrować.

Można zaznaczyć pola wyboru dla pojedynczych kolekcji i źródeł zewnętrznych lub umożliwić użytkownikowi administrowanie wszystkimi kolekcjami i źródłami zewnętrznymi.

Zabezpieczenia na poziomie kolekcji

W celu zapewnienia zabezpieczeń na poziomie kolekcji należy skonfigurować opcje dotyczące indeksowania treści oraz opcje umożliwiające aplikacjom wyszukującym realizację wyszukiwania w określonych kolekcjach.

Podczas tworzenia kolekcji można wybrać opcję powodującą włączenie zabezpieczeń kolekcji. Wybranie tej opcji umożliwia późniejsze skonfigurowanie mechanizmów zabezpieczeń na poziomie dokumentów. Kiedy włączone są zabezpieczenia kolekcji, globalne procesy analizy w wyszukiwaniu korporacyjnym stosują także odmienne reguły przy indeksowaniu duplikatów dokumentów.

Po utworzeniu aplikacji wyszukiwania jej identyfikator pozwala na określenie kolekcji i zewnętrznych źródeł przeszukiwanych przez tę aplikację, a także użytkowników z uprawnieniem dostępu do tej aplikacji.

Analiza duplikatów dokumentów i bezpieczeństwo kolekcji

W przypadku włączonego mechanizmu zabezpieczeń kolekcji globalne procesy analizowania nie identyfikują duplikatów dokumentów w kolekcji.

Proces indeksowania w czasie globalnej analizy oznacza dokumenty będące swoimi duplikatami (lub będące niemal duplikatami). Następnie wszystkie te dokumenty zostają powiązane z pojedynczą kanoniczną reprezentacją treści. Zezwalając na identyfikowanie duplikatów dokumentów, można się upewnić, że wyniki wyszukiwania nie zawierają wielu dokumentów o identycznej (lub niemal identycznej) treści.

Jeśli przy tworzeniu kolekcji zostanie włączony mechanizm zabezpieczeń, duplikaty dokumentów nie będą identyfikowane i nie będą powiązane z ogólną reprezentacją kanoniczną. Każdy dokument będzie indeksowany niezależnie od innych. Dzięki temu reguły zabezpieczeń są przetwarzane dla każdego dokumentu, przez co użytkownicy przeszukują wyłącznie te dokumenty, w których znacznik bezpieczeństwa jest zgodny z ich referencjami. Dwa dokumenty mogą mieć prawie identyczną treść, ale bezpieczeństwo może być zapewniane przy zastosowaniu różnych list kontroli dostępu.

Na przykład założmy, że w przypadku dwóch duplikatów dokumentów, dokument_A i dokument_B, użytkownik ma prawa dostępu tylko do dokumentu dokument_B. Jeśli dokument_B zostanie wyeliminowany przez mechanizm wykrywania duplikatów, to użytkownik nie zobaczy tego dokumentu w wynikach wyszukiwania ze względu na ograniczenia dostępu, jakimi objęty jest dokument_A.

Wyłączenie analizy duplikatów dokumentów może zwiększyć bezpieczeństwo dokumentów w kolekcji, jednak może się to odbyć kosztem pogorszenia jakości wyszukiwania, gdy użytkownicy w wynikach wyszukiwania otrzymywać będą wiele kopii tego samego dokumentu.

Identyfikatory aplikacji wyszukiwania

Możliwość przeszukiwania różnych kolekcji zależy od odwzorowania aplikacji wyszukujących na kolekcje i zewnętrzne źródła, które mogą być przeszukiwane. Aplikacja o nazwie Default umożliwia wykorzystanie przykładowej aplikacji wyszukiwania do wyszukania we wszystkich kolekcjach i źródłach zewnętrznych.

Każda aplikacja wyszukiwania musi przekazać interfejsowi API wyszukiwania korporacyjnego prawidłową nazwę aplikacji (APPID). Aplikacja wyszukiwania może wyszukiwać wyłącznie w tych kolekcjach i źródłach zewnętrznych, które powiązane są z jej identyfikatorem APPID.

Aby aplikacja wyszukująca mogła uzyskać dostęp do kolekcji lub źródła zewnętrznego, administrator wyszukiwania korporacyjnego musi powiązać tę aplikację z określonymi kolekcjami i źródłami, w których może ona wyszukiwać. Aplikacja wyszukująca może przeszukiwać wszystkie kolekcje i źródła zewnętrzne w systemie wyszukiwania korporacyjnego lub tylko określone przez użytkownika.

Z przykładową aplikacją wyszukiwania (ESSearchApplication) związany jest plik właściwości, w którym podana jest nazwa tej aplikacji. Domyślne położenie tego pliku właściwości to `ES_INSTALL_ROOT\installedApps\ESSearchApplication.ear\ESSearchApplication.war\WEB-INF\config.properties`.

Początkowo aplikacja ta ma nazwę Default. Jeśli nazwa ta zostanie zmieniona, zmieni się też lista kolekcji i zewnętrznych źródeł, w których wyszukiwać może aplikacja `ESSearchApplication`.

Aby ustalić, jacy użytkownicy mają dostęp do określonych kolekcji, należy powiązać użytkowników (lub grupy użytkowników) z aplikacją kliencką, stosując przy tym standardowe opcje sterujące serwerem aplikacji WebSphere. W podobny sposób opcji tych używa się do blokowania dostępu do określonych adresów URL. Na przykład można ograniczyć dostęp do adresu URL uruchamiającego aplikację wyszukującą.

Więcej informacji na temat identyfikatorów aplikacji wyszukującej oraz metod wbudowywania mechanizmów zabezpieczeń w niestandardowe aplikacje wyszukujące zawiera opis funkcji API wyszukiwania i indeksowania dla wyszukiwania korporacyjnego.

Pojęcia pokrewne



Przegląd interfejsu API wyszukiwania i indeksowania



Zabezpieczenia interfejsu API wyszukiwania i indeksowania

Zabezpieczenia na poziomie dokumentu

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentu. Kontrola dostępu na poziomie dokumentów zapewnia, że wyniki wyszukiwania zawierają tylko te dokumenty, które użytkownik wysyłający dane zapytanie może przeglądać.

System wyszukiwania korporacyjnego obsługuje wiele sposobów konfigurowania reguł bezpieczeństwa na poziomie dokumentów:

- Dokumenty, przed dodaniem ich do indeksu, mogą być wstępnie filtrowane oraz wiązane z tokenami zabezpieczeń.
- W przypadku niektórych typów danych wyniki wyszukiwania mogą być filtrowane na późniejszym etapie w celu skontrolowania referencji użytkownika względem bieżących danych kontroli dostępu. Komponent zarządzania tożsamościami może zaszyfrować wszystkie referencje, jakie użytkownicy muszą stosować w celu uzyskania dostępu do różnych repozytoriów, i zapisać je w profilu użytkownika. Jeśli zasoby przeznaczone do wyszukania chronione są przez produkt udostępniający zabezpieczenia wymagające pojedynczego zalogowania się (SSO), to komponent zarządzania tożsamościami może kontrolować dostęp do dokumentów bez konieczności tworzenia profili przez użytkowników.
- W przypadku większości typów przeszukiwaczy do powiązania znaczników bezpieczeństwa z dokumentami w indeksie można zastosować specjalną klasę (wtyczkę) języka Java.

- W dokumentach przeszukiwanych przez przeszukiwacza sieci WWW tekst zakotwiczenia powiązany z odsyłaczami do zabronionych dokumentów może zostać wykluczony z indeksu.

Pojęcia pokrewne

 Zabezpieczenia aplikacji

Wstępne i końcowe filtrowanie wyników wyszukiwania

Istnieją dwa różne rozwiązania filtrowania dokumentów zapewniające, że wyniki wyszukiwania zawierać będą wyłącznie te dokumenty, do przeglądania których użytkownik ma uprawnienia.

- Pierwsze rozwiązanie polega na skopiowaniu list kontroli dostępu (ACL) dokumentu do indeksu jeszcze w czasie przeszukiwania i nakazaniu mechanizmowi wyszukiwania porównywania referencji użytkowników ze skopiowanymi listami ACL dokumentów. Wstępne filtrowanie dokumentów i określanie, które dokumenty są dodawane do indeksu, zwiększa wydajność wyszukiwania. Mimo to bardzo trudne jest modelowanie wszystkich reguł zabezpieczeń z różnych źródeł indeksów i zaimplementowanie jednolitej logiki porównań. Rozwiązanie to nie będzie też reagowało w podobnym stopniu na jakiegokolwiek zmiany wprowadzane w źródłowych listach ACL.
- Drugie rozwiązanie polega na późniejszym filtrowaniu dokumentów ze zbioru wyników poprzez odpytywanie źródeł zaplecza o dane zabezpieczeń. W rozwiązaniu tym właśnie źródła zaplecza stają się ostatecznymi arbitrami odnośnie dokumentów prezentowanych użytkownikowi, dzięki czemu zbiór wyników zawsze jest zgodny z bieżącą kontrolą dostępu. Niestety, rozwiązanie to powoduje zmniejszenie wydajności wyszukiwania, ponieważ wymaga ona nawiązania połączeń z wszystkimi źródłami zaplecza. Jeżeli jakieś źródło nie jest dostępne, to ze zbioru wyników odfiltrowane muszą zostać odsyłacze do dokumentów oraz dokumenty, do przeglądania których użytkownik nie otrzymał uprawnień.

Ważne: W konfiguracji z wieloma serwerami filtrowanie na późniejszym etapie niektórych rodzajów źródeł wykonywane jest przez serwer przeszukiwacza. Jeżeli serwer przeszukiwacza jest wyłączony w celu konserwacji, użytkownicy tworzący zapytania do kolekcji wyszukiwania korporacyjnego nie otrzymają żadnych wyników. Co więcej, wyniki nie są zwracane również w przypadku, gdy nie są dostępne serwery zaplecza wymagane do kontroli dostępu do dokumentów.

W wyszukiwaniu korporacyjnym obsługa mechanizmu kontroli dostępu używa kombinacji dwóch przedstawionych wyżej rozwiązań. Taka konstrukcja pozwala na uzyskanie optymalnej wydajności przy jednoczesnym zachowaniu reguł zabezpieczeń pochodzących z repozytorium dokumentów. Dzięki przechowywaniu w indeksie danych kontroli dostępu wysokiego poziomu system może dostarczyć pośredni (prawdopodobnie zmniejszony) zbiór wyników, który może być następnie przefiltrowany w ramach bieżących ustawień kontroli dostępu. W takim systemie stosowane jest założenie, że użytkownik z dostępem do danego repozytorium ma duże szanse na uzyskanie dostępu do określonego dokumentu z tego repozytorium.

Dane kontroli dostępu przechowywane w indeksie mogą różnić się w zależności od rodzaju przeszukiwacza. Na przykład przeszukiwacz Notes może przechowywać dane kontroli dostępu na poziomie baz danych i serwerów, natomiast przeszukiwacz QuickPlace może przechowywać dane kontroli dostępu dla serwerów, miejsc i pokoi.

Wszystkie typy danych w systemie wyszukiwania korporacyjnego pozwalają w czasie przeszukiwania na indeksowanie własnych list kontroli dostępu. Niektóre typy źródeł danych pozwalają również na późniejsze filtrowanie zbioru wyników i kontrolowanie bieżących

referencji użytkowników (ten typ obsługi udostępniany jest przez własne mechanizmy zabezpieczeń albo przez komponent zarządzania tożsamościami wyszukiwania korporacyjnego).

Taki dwukierunkowy model zabezpieczeń spełnia następujące zadania:

- Wyodrębnianie własnych informacji z list kontroli dostępu w czasie przeszukiwania.
- Przechowywanie informacji z list ACL i z baz danych w indeksie.
- Tworzenie kontekstu zabezpieczeń użytkownika w momencie jego zalogowania lub w momencie inicjowania sesji. W tym zadaniu trzeba brać pod uwagę różne identyfikatory, jakich użytkownik musi używać, uzyskując dostęp do poszczególnych źródeł zaplecza.
- Wyszukiwanie zgodne z kontekstem zabezpieczeń użytkownika i tworzenie przejściowego zbioru wyników zawierającego tylko te dokumenty, do których użytkownik ma dostęp na poziomie repozytoriów.
- Końcowe filtrowanie pośredniego zbioru wyników wykonywane przez konsultacje ze źródłami zaplecza, które dołączały dokumenty do zbioru wyników dla bieżących, własnych list ACL.

Sprawdzanie poprawności na podstawie zapisanych znaczników bezpieczeństwa

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentów przez zapisywanie danych zabezpieczeń w indeksie.

Domyślnie każdy dokument ma przypisany publiczny znacznik, który powoduje udostępnienie dokumentu wszystkim użytkownikom. Jeśli zabezpieczenia kolekcji zostały włączone, token publiczny może zostać zastąpiony wartością przekazywaną przez administratora lub wydobytą z pola przeszukiwanego dokumentu. Podczas konfigurowania przeszukiwacza można określić, że ma on korzystać z tokenów zabezpieczeń, aby umożliwić dostęp do przeszukiwanych dokumentów tylko wybranym użytkownikom.

Administrator kolekcji w czasie konfigurowania przeszukiwacza może ustalić opcje zabezpieczeń dla poszczególnych tabel, systemów plików itd. Oznacza to, że dla różnych źródeł danych z przestrzeni przeszukiwania można ustalić różne reguły zabezpieczeń. Administrator może:

- określić, że dokumenty są publiczne (wszyscy użytkownicy mają do nich dostęp);
- każdemu dokumentowi przypisać zdefiniowane przez użytkownika znaczniki bezpieczeństwa;
- z przeszukiwanych danych wydobyć znaczniki bezpieczeństwa i przypisać je do każdego dokumentu.

Znaczniki bezpieczeństwa (z wyjątkiem domyślnych znaczników publicznych) są w całości definiowane przez użytkownika. Znacznik bezpieczeństwa może reprezentować identyfikator użytkownika lub grupy, rolę użytkownika albo dowolną inną wartość, która będzie miała znaczenie w danym źródle danych.

Na przykład administrator może określić, że pole `hrDeptName` ma być stosowane do kontroli dostępu do dokumentów przeszukiwanych przez przeszukiwacza Notes. Administrator może też zdecydować, że w sytuacji, gdy w dokumencie takie pole nie istnieje, kontrolę dostępu do dokumentów mają zapewniać dwa znaczniki zdefiniowane przez użytkownika: `hrgroup1` i `hrgroup2`.

Znaczniki bezpieczeństwa udostępniane są przeszukiwaczowi poprzez jego plik konfiguracji. Dla każdego dokumentu przeszukiwacz udostępnia wartość znacznika zabezpieczeń w postaci

metadanych. Komponent indeksowania odczytuje znaczniki bezpieczeństwa i wykorzystuje je do przesyłania informacji na temat dokumentu znajdującego się indeksie. Jeśli administrator źródła danych zaktualizuje listę kontroli dostępu, to nowe informacje kontroli dostępu będą stosowane po kolejnym tworzeniu głównego indeksu lub indeksu przyrostowego.

Do określania wartości tokenów zabezpieczeń można stosować niestandardowe reguły biznesowe zakodowane w klasie Języka Java. Podczas konfigurowania właściwości przeszukiwacza należy określić nazwę wtyczki, której przeszukiwacz ma używać podczas przeszukiwania dokumentów. Tokeny zabezpieczeń dodawane przez wtyczkę są przechowywane w indeksie i mogą być używane do kontrolowania dostępu do dokumentów.

W jaki sposób aplikacje wyszukiwania korzystają ze znaczników bezpieczeństwa

Zadaniem klienckiej aplikacji wyszukiwania jest dostarczenie w czasie wyszukiwania znaczników bezpieczeństwa, tak aby dokumenty mogły zostać odpowiednio przefiltrowane. Jeśli znacznik bezpieczeństwa nie zostanie dostarczony, to w czasie przetwarzania wyszukiwania automatycznie stosowany jest domyślny znacznik publiczny.

Przykładowa aplikacja wyszukiwania korporacyjnego demonstruje jeden ze sposobów implementowania bezpieczeństwa na poziomie dokumentów. W tym przykładzie zakłada się, że administrator przypisał wartość znacznika bezpieczeństwa do grupy dokumentów (co oznacza, że znacznik bezpieczeństwa nie będzie wydobywany z przeszukiwanego dokumentu). Aplikacje wyszukiwania korzystają z identyfikatora logowania użytkownika w celu określenia, do których dokumentów użytkownik ten będzie miał dostęp. Aplikacja wyszukiwania korzysta z identyfikatora grupy, do której należy użytkownik, a nie z identyfikatora tego użytkownika. Dzięki wykorzystaniu identyfikatora grupy jako znacznika bezpieczeństwa można dodawać użytkowników do grupy i usuwać ich z niej bez konieczności jednoczesnego odbudowania indeksu.

Znacznik bezpieczeństwa przypisany przez administratora do zbioru dokumentów reprezentuje identyfikator grupy użytkowników systemu operacyjnego. W przestrzeni przeszukiwania poszczególnym dokumentom przypisywane są różne identyfikatory grup. Na przykład:

Dokument1-5: Znacznik bezpieczeństwa = Grupa1
Dokument6-10: Znacznik bezpieczeństwa = Grupa2

Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań

Jeśli podczas tworzenia kolekcji włączono opcje zabezpieczeń, w niektórych typach domen możliwe jest sprawdzanie poprawności bieżących referencji użytkownika podczas wprowadzania zapytania.

Przed wysłaniem odpowiedzi na zapytanie serwery wyszukiwania nawiązują kontakt z rodzimymi repozytoriami w celu zweryfikowania bieżących uprawnień użytkownika, a następnie usuwają z wyników wyszukiwania wszystkie dokumenty, których użytkownik nie może przeglądać.

Podczas konfigurowania poniższych typów przeszukiwaczy można włączyć opcję sprawdzania referencji użytkownika przez porównanie ich z bieżącymi ustawieniami kontroli dostępu obsługiwanych przez rodzime repozytorium. Po przeszukaniu i poindeksowaniu dokumentów do sprawdzania tożsamości użytkowników próbujących przeszukać kolekcję stosowany jest komponent zarządzania tożsamościami wyszukiwania korporacyjnego.

- przeszukiwacz Content Edition (tylko repozytoria Documentum, FileNet Panagon Content Services, Hummingbird DM, Portal Document Manager oraz SharePoint)

- przeszukiwacz DB2 Content Manager
- przeszukiwacz Domino Document Manager
- przeszukiwacz Notes
- przeszukiwacz QuickPlace
- przeszukiwacz systemu plików Windows

W przypadku poniższych rodzajów przeszukiwaczy bieżące referencje użytkowników mogą być sprawdzane wtedy, gdy do przeszukiwania kolekcji wyszukiwania korporacyjnego użytkownicy korzystają z portletu wyszukiwania portalu WebSphere.

- przeszukiwacz Web Content Management
- przeszukiwacz WebSphere Portal

Pojęcia pokrewne

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino” na stronie 282

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows” na stronie 285

Zadania pokrewne

“Konfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika” na stronie 282

Zarządzanie tożsamościami na potrzeby wyszukiwania korporacyjnego

Zarządzanie wieloma referencjami użytkowników jest typowym problemem niemal każdego przedsiębiorstwa. System wyszukiwania korporacyjnego rozwiązuje ten problem przez udostępnienie opcjonalnego komponentu zarządzania tożsamościami.

Wskazówka:

Szczegółowe przykłady konfigurowania komponentu do zarządzania tożsamościami na potrzeby wyszukiwania korporacyjnego można znaleźć w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

W przedsiębiorstwie informacje mogą być przechowywane na wiele różnych sposobów. Mogą być rozproszone po całym przedsiębiorstwie i zarządzane za pomocą oprogramowania najlepiej przystosowanego do tego zadania. Na przykład użytkownicy korporacyjni mogą korzystać z relacyjnych baz danych za pośrednictwem aplikacji SQL albo uzyskiwać dostęp do ważnych dla nich dokumentów za pośrednictwem systemu zarządzania dokumentami.

Kontrola dostępu do ważnych informacji znajdujących się w tych repozytoriach najczęściej realizowana jest przez oprogramowanie zarządzające. Użytkownicy identyfikowani są w systemie za pomocą kombinacji identyfikatora użytkownika oraz hasła. Po przeprowadzeniu systemowego uwierzytelniania oprogramowanie zarządzające decyduje o tym, które dokumenty może przeglądać i modyfikować dany użytkownik, posilując się przy tym uprawnieniami dostępu zdefiniowanymi dla tego użytkownika.

Często zdarza się tak, że użytkownicy mają zdefiniowane różne identyfikatory i hasła dla poszczególnych repozytoriów. Użytkownicy muszą podać referencje, aby móc przeglądać dokumenty znajdujące się w kolekcjach wyszukiwania korporacyjnego wymagających sprawdzenia poprawności bieżących referencji, podobnie jak powinni zidentyfikować się przy dostępie do głównego repozytorium firmowego. Użytkownicy korzystający z wielu tożsamości muszą podać odpowiednie referencje dla każdej z tych tożsamości.

Jeśli w konsoli administracyjnej podane zostanie, że do zarządzania tożsamościami używany ma być system wyszukiwania korporacyjnego, to w serwerach wyszukiwania można zastosować następujące rozwiązania pozwalające na sprawdzenie poprawności bieżących referencji użytkownika w czasie przetwarzania jego zapytań:

- W aplikacji wyszukiwania użytkownik może zostać poproszony o zarejestrowanie w profilu użytkownika referencji wymaganych do uzyskania dostępu do poszczególnych domen. Profil użytkownika jest zaszyfrowywany i zapisywany w bezpiecznej składowi danych. Pozwala on użytkownikowi na wyszukiwanie dokumentów w zabezpieczonych domenach. Jeśli referencje dla domeny wymagającej sprawdzenia poprawności bieżących referencji nie zostaną zdefiniowane, to dokumenty z tej domeny będą wykluczane z wyników wyszukiwania.
- Jeśli dokumenty z kolekcji przeszukiwane są przez przeszukiwacza obsługującego zabezpieczenia z pojedynczym logowaniem (SSO), a do kontroli dostępu do dokumentów stosowane mają być zabezpieczenia SSO, to system wykorzysta ten rodzaj zabezpieczeń do uwierzytelniania użytkowników na czas trwania sesji wyszukiwania. W czasie przeszukiwania zabezpieczonych domen użytkownik nie musi tworzyć swojego profilu ani podawać swojego identyfikatora i hasła;

Gdy użytkownicy przeszukują kolekcje wymagające w czasie przekazywania zapytania sprawdzenia poprawności bieżących referencji, to system może przyznać dostęp do dokumentów lub odmówić go na podstawie profilu użytkownika albo za pomocą metod zabezpieczeń SSO.

Odczytywanie informacji o grupach użytkownika

W celu sprawdzenia poprawności referencji użytkownika komponent zarządzania tożsamościami musi uzyskać informacje o grupach użytkownika dla każdej z jego tożsamości i dodać te informacje do ciągu znaków kontekstu zabezpieczeń użytkownika (User Security Context - USC). Informacje o grupach stosowane są do filtrowania wyników zgodnie z danymi kontroli dostępu przechowywanymi w indeksie wyszukiwania korporacyjnego albo zgodnie z danymi uwierzytelniania SSO. Komponent zarządzania tożsamościami wykonuje swoje zadanie, używając znaczników SSO albo referencji użytkowników do nawiązania połączenia z systemem zaplecza i zażądania listy grup, których członkiem jest dany użytkownik.

Podczas konfigurowania opcji zarządzania tożsamościami w konsoli administracyjnej można też określić, jak często odświeżane mają być informacje o grupach użytkowników. Nowe dane o grupach można pobierać po każdym zalogowaniu użytkownika do aplikacji wyszukiwania, ale równie dobrze można robić to cyklicznie, na przykład co trzy dni.

Bezpieczeństwo bez komponentu zarządzania tożsamościami

Nie wszystkie firmy chcą, żeby zarządzaniem tożsamościami społeczności ich użytkowników zajmował się komponent zarządzania tożsamościami. Po wyłączeniu w konsoli administracyjnej wyszukiwania korporacyjnego komponentu zarządzania tożsamościami zadanie utworzenia ciągu znaków kontekstu zabezpieczeń użytkownika spada na aplikację wyszukiwania. Wygenerowany ciąg znaków USC służy do przygotowania ograniczeń listy kontroli dostępu dla każdego zapytania. Na przykład:

```
Query q = factory.createQuery("IBM");
q.setACLConstraints("Kontekst zabezpieczeń użytkownika w formacie XML");
```

Wskazówka: W ramach ułatwienia tworzenia własnych funkcji zarządzania tożsamościami w interfejsie Search and Index API (SI-API) udostępnione zostało rozszerzenie pozwalające na programowe sterowanie bazą danych zarządzania tożsamościami. Interfejs ten pozwala na

wygenerowanie kontekstu zabezpieczeń użytkownika za pomocą obiektów języka Java, dzięki czemu ciąg znaków XML tworzony jest całkowicie automatycznie.

Ciąg znaków zapytania XML musi mieć poniższą formę, w której ... oznacza pełny ciąg znaków XML:

```
@SecurityContext::'...'
```

Oto format ciągu znaków XML:

```
<identities id="nazwa_logowania_uzytkownika">
  <ssoToken>wartosc_znacznika</ssoToken>
  <identity id="domena_zabezpieczen">
    <type>Notes</type>
    <username>nazwa_uzytkownika_domeny</username>
    <password encrypt="no">haslo_uzytkownika_domeny</password>
    <groups>
      <group id="g1" />
      <group id="g2" />
    </groups>
    <properties>
      <property name="nazwa_wlasciwosci">wartosc_wlasciwosci</property>
      ...
    </properties>
  </identity>
  ...
</identities>
```

identities

Wartość atrybutu id jest identyfikatorem użytkownika pozwalającym mu na zalogowanie się do systemu.

ssoToken

Opcjonalnie: określa znacznik LTPA (Lightweight Third-Party Authentication), który tworzony jest dla użytkownika na czas trwania sesji przeglądarki. Ten parametr używany jest tylko w przypadku, gdy docelowa domena pozwala na stosowanie zabezpieczeń SSO, a przeszukiwacz skonfigurowany jest do stosowania zabezpieczeń tego typu.

identity

Zawiera referencje użytkownika dla określonego źródła danych. Wartość atrybutu id oznacza domenę zawierającą informację o referencjach użytkownika (w przypadku domeny Domino będzie to nazwa domeny Domino).

type

Określa typ danych w odniesieniu do typu przeszukiwacza (Notes, DB2, Exchange Server itd.).

username

Określa nazwę użytkownika, jakiej należy użyć w domenie wyszukiwania.

password

Określa hasło dla podanej nazwy użytkownika. Atrybut `encrypt` musi mieć wartość `no` (wyszukiwanie korporacyjne nie udostępnia metod szyfrowania poza komponentem zarządzania tożsamościami).

groups

Podaje nazwy grup, których członkiem jest użytkownik. Dla każdej nazwy grupy stosowany jest osobny element `group`.

properties

Definiuje listę właściwości dla danego połączenia, takich jak identyfikator administratora i zaszyfrowane hasło użyte podczas tworzenia przeszukiwacza albo informacja o włączeniu mechanizmów SSO dla źródła.

nazwa_właściwości

Nazwa właściwości.

wartość_właściwości

Wartość właściwości.

Sprawdzanie poprawności użytkownika z zastosowaniem profilu użytkownika

Aplikacje wyszukiwania mogą żądać od użytkowników zarejestrowania w profilu użytkownika referencji, które potrzebne są do uzyskania dostępu do różnych domen.

Aby móc przeszukiwać domenę wymagającą sprawdzenia referencji podczas wprowadzania zapytania, użytkownicy muszą udostępnić aplikacji wyszukującej referencje, za pomocą których logują się do tej domeny. Dzięki zastosowaniu komponentu zarządzania tożsamościami użytkownicy mogą przechowywać w swoich profilach referencje dla dowolnej liczby domen. Referencje są zaszyfrowane i bezpiecznie zapisane w systemie wyszukiwania korporacyjnego.

Jeśli referencje dla domeny wymagającej sprawdzenia poprawności bieżących referencji nie zostaną zdefiniowane, to dokumenty z tej domeny będą wykluczane z wyników wyszukiwania.

Użytkownicy mogą tworzyć profile i rejestrować w nich swoje referencje, kiedy korzystają z aplikacji wyszukującej. W przykładowej aplikacji wyszukiwania możliwość ta udostępniana jest w ramach opcji **Mój profil**. Każda inna aplikacja wyszukiwania może rozwiązywać tę sprawę w inny sposób.

Kolekcje mogą zawierać dokumenty pochodzące z najróżniejszych typów źródeł. Na przykład kolekcja może zawierać dokumenty przeszukane w systemie plików Windows oraz kilku bazach danych Lotus Notes. Komponent zarządzania tożsamościami rozróżnia wszystkie rodzaje źródeł i żąda podania referencji tylko wtedy, gdy są one potrzebne do uzyskania dostępu do wymagających tego domen.

Domyślnie wszystkie referencje pozwalają na wyszukiwanie, dlatego użytkownik musi podać identyfikator i hasło właściwe dla zabezpieczonych domen. Jeśli użytkownik zapomniał swojego identyfikatora lub hasła związanego z określoną domeną, to domenę tę można wykluczyć z wyszukiwania przez wyczyszczenie pola wyboru. Wykluczenie domeny powoduje, że pochodzące z niej zabezpieczone dokumenty nie będą pojawiały się w wynikach wyszukiwania.

Po utworzeniu profilu użytkownik może zacząć wprowadzać żądania wyszukiwania. Komponent zarządzania tożsamościami będzie mógł utworzyć ciąg znaków kontekstu zabezpieczeń użytkownika, który używany będzie przy kolejnych żądaniach. Jeśli użytkownik nie będzie korzystał z komponentu zarządzania tożsamościami, to przy próbach wysłania zapytania do domen wymagających sprawdzenia poprawności bieżących referencji aplikacja wyszukiwania będzie musiała utworzyć ciąg znaków kontekstu zabezpieczeń użytkownika.

Przy podjętej przez użytkownika następnej próbie wyszukiwania kolekcji wyszukiwania korporacyjnego komponent zarządzania tożsamościami powtórzy proces sprawdzania poprawności referencji, ale tym razem będzie mógł znaleźć profil użytkownika. Jeśli nie zostaną wprowadzone żadne zmiany, użytkownik zostanie automatycznie kierowany do miejsca pozwalającego na wprowadzanie żądań i nie będzie proszony o utworzenie profilu.

Jeśli komponent zarządzania tożsamościami wykryje zmiany w dowolnej referencji użytkownika, to przy próbie dostępu do aplikacji wyszukiwania zostanie wyświetlona strona

profilu. Może się tak zdarzyć na przykład wtedy, gdy zmienione zostanie hasło dostępu do dowolnej domeny włączonej do wyszukiwania albo do kolekcji dodana zostanie domena wymagająca uwierzytelniania użytkowników.

Użytkownicy mogą zignorować zalecenie zaktualizowania profilu, ale w ten sposób godzą się na wykluczenie pewnych dokumentów z wyników wyszukiwania.

W przykładowej aplikacji wyszukiwania dostarczanej razem z wyszukiwaniem korporacyjnym użytkownicy mogą aktualizować swój profil, wybierając z paska narzędzi opcję **Mój profil**.

Sprawdzanie poprawności użytkownika z zastosowaniem zabezpieczeń SSO

Jeśli dokumenty w kolekcji zostały przeszukane przez przeszukiwacza obsługującego zabezpieczenia z pojedynczym logowaniem (SSO), to w czasie konfigurowania opcji zarządzania tożsamościami można zaznaczyć, że do kontroli dostępu do dokumentów stosowane mają być zabezpieczenia SSO.

Włączanie zabezpieczeń SSO

Uwierzytelnianie z pojedynczym logowaniem pozwala na jednokrotne uwierzytelnienie użytkownika i udzielenie mu dostępu do wielu zasobów bez konieczności każdorazowego podawania referencji. W systemie wyszukiwania korporacyjnego uwierzytelnianie SSO bardzo ułatwia zarządzanie wieloma nazwami i hasłami użytkowników, które muszą być podawane przy dostępie do dokumentów z zabezpieczonych kolekcji.

Serwer aplikacji IBM WebSphere i serwer Lotus Domino obsługują typ zabezpieczeń SSO znany pod nazwą Lightweight Third-Party Authentication (LTPA). Gdy użytkownik próbuje uzyskać dostęp jednego z tych produktów, proszony jest o podanie nazwy użytkownika i hasła. Podana nazwa użytkownika i hasło sprawdzane są w zapisach repozytorium LDAP współużytkowanego przez oba produkty. Po zakończeniu uwierzytelniania użytkownika tworzona jest informacja cookie na temat sesji, w której zapisany jest znacznik LTPA. Użytkownik może wtedy uzyskać dostęp do zasobów na dowolnym serwerze o takiej samej konfiguracji uwierzytelniania bez konieczności ponownego wprowadzania referencji. Znacznik istnieje tak długo, jak długo ważna jest sesja przeglądarki.

Aby włączyć obsługę zabezpieczeń SSO w kolekcjach wyszukiwania korporacyjnego:

- Upewnij się, że na serwerach wyszukiwania korporacyjnego włączone są globalne zabezpieczenia serwera aplikacji WebSphere oraz rejestr LDAP. Rejestrem LDAP może być dowolny produkt LDAP obsługiwany przez serwer aplikacji WebSphere.
- Upewnij się, że mechanizm uwierzytelniania WebSphere jest skonfigurowany do użycia aktywnego mechanizmu uwierzytelniania LTPA. W czasie konfigurowania mechanizmu LTPA podaj prawidłową, choć elastyczną nazwę domeny, na przykład firma.serwer.com.
- Upewnij się, że klucz LTPA został wyeksportowany przez serwer aplikacji WebSphere i zaimportowany do pozostałych produktów w tej samej domenie, które również mają obsługiwać mechanizm LTPA.

Po sprawdzeniu za pomocą przeglądarki, że podane wyżej ustawienia zabezpieczeń działają prawidłowo, można skorzystać z konsoli administracyjnej wyszukiwania korporacyjnego i skonfigurować w niej przeszukiwacze do obsługi uwierzytelniania SSO.

Mechanizm SSO i zarządzanie tożsamościami

Gdy użytkownicy przeszukują kolekcje wymagające sprawdzenia poprawności bieżących referencji, to system może przyznać lub odmówić dostępu do dokumentów na podstawie zabezpieczeń SSO. Użytkownicy nie są proszeni o podanie referencji podczas przeszukiwania

źródeł obsługujących uwierzytelnianie SSO. Komponent zarządzania tożsamościami używany jest wtedy, gdy spełnione zostaną wszystkie poniższe warunki:

- Mechanizm SSO jest włączony w serwerze aplikacji WebSphere oraz w docelowych domenach.
- Zabezpieczenia są włączone w co najmniej jednej kolekcji, którą może przeszukiwać aplikacja wyszukująca.
- W konsoli administracyjnej wyszukiwania korporacyjnego włączone są opcje nakazujące użycie komponentu zarządzania tożsamościami oraz zabezpieczeń SSO.
- Przy konfigurowaniu poniższych rodzajów przeszukiwaczy wybrane zostały opcje nakazujące użycie zabezpieczeń SSO i wymuszające zabezpieczenia na poziomie dokumentów (na przykład indeksowanie ustawień kontroli dostępu lub sprawdzanie poprawności referencji w czasie przetwarzania zapytań):
 - Content Edition (dostępny tylko dla repozytoriów Portal Document Manager)
 - Domino Document Manager (dostępny tylko dla przeszukiwaczy korzystających z protokołu DIIOP)
 - Notes (dostępny tylko dla przeszukiwaczy korzystających z protokołu DIIOP)
 - QuickPlace (dostępny tylko dla przeszukiwaczy korzystających z protokołu DIIOP)

Bezpieczeństwo portletu wyszukiwania

Gdy użytkownicy za pomocą portletu wyszukiwania dla systemu wyszukiwania korporacyjnego prowadzą wyszukiwania w kolekcjach, korzystając z portalu WebSphere Portal, bezpieczeństwo jest również zapewniane dokumentom przeszukiwanym przez przeszukiwacze listy adresów początkowych, Web Content Management oraz WebSphere Portal.

Bezpieczne wyszukiwanie jest realizowane w przypadku tych typów źródeł tylko wtedy, gdy używany jest portlet, a nie aplikacja wyszukująca działająca poza portalem WebSphere Portal. Wewnątrz portalu WebSphere Portal referencje użytkownika są uzyskiwane za pośrednictwem portletu wyszukiwania. Gdy użytkownik już zaloguje się w portalu WebSphere Portal, wszystkie żądania wyszukiwania będą obejmować dane bezpieczeństwa użytkownika (nazwę użytkownika, przynależność do grup itd.). Ponieważ te informacje są zawsze dostępne, mechanizm SSO oparty na elementach protokołu LTPA nie jest wymagany.

Jeśli jednak do ochrony serwisów i dokumentów na serwerze WebSphere Portal używany jest inny produkt, podczas konfigurowania przeszukiwacza muszą zostać określone opcje SSO. Na przykład, jeśli używany jest taki produkt, jak IBM Tivoli Access Manager WebSEAL lub CA SiteMinder SSO Agent for PeopleSoft, należy określić referencje, które umożliwią przeszukiwaczowi dostęp do dokumentów na serwerze za pośrednictwem pojedynczego logowania. W tym przypadku zostaje włączony mechanizm SSO na potrzeby dostępu przeszukiwacza do zabezpieczonej treści, a nie zostaje włączony mechanizm SSO na potrzeby bezpiecznego wyszukiwania.

Konfigurowanie zarządzania tożsamościami

Komponent zarządzania tożsamościami wyszukiwania korporacyjnego można wykorzystać do określenia sposobu sprawdzania poprawności referencji w czasie przetwarzania zapytań.

Zagadnienia wstępne

Aby konfigurować opcje zarządzania tożsamościami, wymagana jest rola administratora wyszukiwania korporacyjnego.

Informacje o tym zadaniu

Gdy użytkownicy przeszukują kolekcje wymagające sprawdzenia poprawności bieżących referencji w czasie obsługi zapytań, komponent zarządzania tożsamościami wyszukiwania korporacyjnego może skorzystać z profilu użytkownika lub z metod zabezpieczania przez pojedyncze logowanie i w ten sposób umożliwić dostęp do dokumentów lub odmówić go.

Procedura

Aby skonfigurować zarządzanie tożsamościami:

1. Kliknij przycisk **Zabezpieczenia**, aby otworzyć widok Zabezpieczenia.
2. Na stronie Aplikacje wyszukujące kliknij opcję **Konfiguruj zarządzanie tożsamościami**.
3. Na stronie Konfiguracja zarządzania tożsamościami zaznacz pole wyboru pozwalające na wykorzystanie komponentu zarządzania tożsamościami wyszukiwania korporacyjnego do sterowania sposobem sprawdzania poprawności referencji użytkowników w czasie przetwarzania ich zapytań. Jeśli to pole wyboru nie będzie zaznaczone, aplikacja wyszukująca będzie musiała podać łańcuch kontekstu bezpieczeństwa użytkownika (USC), gdy użytkownicy będą stosować zapytania do domen wymagających sprawdzania poprawności bieżących referencji podczas przetwarzania zapytań.
4. Określ, jak często komponent do zarządzania tożsamościami ma wyodrębniać referencje użytkownika z rekordów grupy w rejestrze użytkowników serwera WebSphere Application Server. Dane referencji można odświeżać za każdym razem, gdy użytkownik próbuje uzyskać dostęp do aplikacji wyszukiwania lub po upłygnięciu określonej liczby dni.
5. W przypadku przeszukiwaczy obsługujących uwierzytelnianie z pojedynczym logowaniem należy zaznaczyć, czy komponent zarządzania tożsamościami do sprawdzania poprawności użytkowników ma używać znaczników bezpieczeństwa, czy też profili użytkowników. Można też wybrać pole wyboru nakazujące stosowanie uwierzytelniania z pojedynczym logowaniem we wszystkich typach przeszukiwaczy albo wybierać pola wyboru poszczególnych typów przeszukiwaczy.

Ważne: Komponent zarządzania tożsamościami tylko wtedy wykorzystuje mechanizmy bezpieczeństwa z pojedynczym logowaniem, gdy są one prawidłowo skonfigurowane na serwerze aplikacji WebSphere i w docelowych domenach.

Analiza tekstu zakotwiczenia

Jeśli włączony jest mechanizm zabezpieczeń kolekcji, w procesach analizy globalnej stosowane są specjalne reguły dotyczące indeksowania tekstu zakotwiczenia w dokumentach przeszukiwanych przez przeszukiwacze sieci WWW. Jeśli nie włączono mechanizmu zabezpieczeń kolekcji, podczas konfigurowania indywidualnych przeszukiwaczy sieci WWW można określić, czy ma być indeksowany tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów.

Tekst zakotwiczenia jest informacją w obrębie odsyłacza hipertekstowego opisującą stronę, do której ten odsyłacz prowadzi. Na przykład w poniższym odsyłaczu tekst **Składnia zapytania** jest tekstem zakotwiczenia w odsyłaczu wskazującym na stronę `syntax.htm`:

```
<a href="../doc/syntax.htm">Składnia zapytania</a>
```

Zazwyczaj przeszukiwacz sieci WWW korzysta z odsyłaczy w dokumentach w celu przeszukania dodatkowych dokumentów i dopisuje te dowiązane strony do indeksu. W trakcie analizy globalnej procesy indeksowania kojarzą tekst zakotwiczenia nie tylko z dokumentem, w którym jest on osadzony (dokumentem źródłowym), lecz także z dokumentem docelowym. W podanym wyżej przykładzie tekst zakotwiczenia **Składnia zapytania** powiązany jest ze stroną docelową `syntax.htm` oraz stroną źródłową która zawiera kotwicę. Takie powiązanie pozwala na pobieranie dokumentu docelowego przez zapytania określające tekst znajdujący

się w dokumencie źródłowym. Z drugiej strony, jeśli użytkownicy będą mogli przeglądać dokument docelowy, ale dostęp do dokumentu źródłowego będzie zabroniony, to takie powiązanie może stanowić lukę w zabezpieczeniach.

Jeśli podczas tworzenia kolekcji zostaną włączone zabezpieczenia kolekcji, przetwarzanie tekstu zakotwiczenia jest wyłączone. Tekst zakotwiczenia nie jest już indeksowany z dokumentem, chyba że występuje w jawnej postaci w treści lub w metadanych dokumentu. Takie zabezpieczenia sprawiają, że użytkownicy nie mogą korzystać z informacji w dokumentach, do których nie mają dostępu. W wynikach wyszukiwania pojawią się wyłącznie te dokumenty, których treść lub metadane będą zgodne z treścią zapytania.

Włączenie zabezpieczeń kolekcji umożliwia zwiększenie bezpieczeństwa dokumentów WWW, umożliwiając użytkownikom wyszukiwania wyłącznie dokumentów z tokenami zabezpieczeń pasującymi do ich referencji. Jednak w przypadku rezygnacji z przetwarzania tekstu zakotwiczenia może dojść do sytuacji, w której wyniki wyszukiwania mogą nie zawierać wszystkich dokumentów potencjalnie odpowiadających warunkom zapytania.

Jeśli zabezpieczenia kolekcji nie zostaną włączone, w czasie konfigurowania zaawansowanych właściwości przeszukiwacza sieci WWW można zaznaczyć, czy indeksowane mają być teksty zakotwiczenia w odsyłaczach do zabronionych dokumentów.

Indeksowanie tekstu zakotwiczenia w odsyłaczach do dokumentów zabronionych

Jeśli dokument zawiera odsyłacze do dokumentów, do których przeszukiwacz sieci WWW nie ma dostępu, podczas konfigurowania przeszukiwacza sieci WWW można określić, czy tekst zakotwiczenia odsyłaczy ma być zachowywany w indeksie.

Zagadnienia wstępne

Opcje indeksowania tekstu zakotwiczeń może skonfigurować uczestnik roli administratorów wyszukiwania korporacyjnego oraz administrator kolekcji dla konfigurowanego przeszukiwacza WWW.


Informacje o tym zadaniu

Dyrektywy zawarte w pliku `robots.txt` lub w metadanych dokumentów WWW mogą uniemożliwić przeszukiwaczowi sieci WWW dostęp do dokumentów serwisu. Jeśli dokument, do którego przeszukiwacz sieci WWW ma dostęp, zawiera odsyłacze do zabronionych dokumentów, można zdecydować o sposobie obsługi tekstu zakotwiczeń dla tych odsyłaczy.

Podczas konfigurowania przeszukiwaczy sieci WWW można określić, czy ma być indeksowany tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów. Ze względów bezpieczeństwa zalecane jest włączenie opcji blokującej indeksowanie tekstu zakotwiczenia w odsyłaczach do zabronionych dokumentów. Jednak w wyniku pominięcia tekstu zakotwiczeń przy indeksowaniu wyniki wyszukiwania mogą nie zawierać niektórych dokumentów potencjalnie spełniających warunki zapytania.

Procedura

Aby włączyć lub wyłączyć indeksowanie tekstu zakotwiczenia w odsyłaczach do zabronionych dokumentów:

1. Przeprowadź edycję kolekcji i na stronie Przeszukiwanie znajdź przeszukiwacza sieci, którego chcesz skonfigurować. Następnie kliknij przycisk  **Właściwości przeszukiwacza**.
2. Kliknij opcję **Edytuj zaawansowane właściwości przeszukiwacza sieci WWW**.

3. Aby indeksować tekst zakotwiczenia we wszystkich dokumentach objętych działaniem danego przeszukiwacza, zaznacz pole wyboru **Indeksuj tekst zakotwiczenia w odsyłaczach do dokumentów zabronionych**. Użytkownicy mogą uzyskać informacje o stronach niedostępnych dla przeszukiwacza sieci WWW dzięki wyszukiwaniu tekstu zakotwiczenia w odsyłaczach prowadzących do tych stron.

Aby wykluczyć z indeksowania tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów, usuń zaznaczenie tego pola wyboru. Użytkownicy nie będą mieli możliwości uzyskania informacji o stronach niedostępnych dla przeszukiwacza sieci WWW. Tekst zakotwiczenia zostanie wykluczony z indeksu razem z zabronionymi dokumentami.

4. Kliknij przycisk **OK**, a następnie - na stronie Właściwości przeszukiwacza sieci WWW - kliknij ponownie przycisk **OK**.
5. Aby zmiany zostały uwzględnione, zatrzymaj i ponownie uruchom przeszukiwacz.

Zastosowanie tych zmian w poindeksowanych już wcześniej dokumentach wymaga ich ponownego poindeksowania przez przeszukiwacza. Jeśli w poprzednim przeszukiwaniu do indeksu trafiły informacje o zabronionych dokumentach, zostaną one usunięte.

Włączanie zabezpieczeń w wyszukiwaniu korporacyjnym

Jeśli planowane jest włączenie zabezpieczeń, w czasie gdy użytkownicy administrują systemem wyszukiwania korporacyjnego lub korzystają z niego, to należy skonfigurować zabezpieczenia globalne serwera aplikacji WebSphere. Informacje dotyczące bezpieczeństwa trzeba też skonfigurować w plikach konfiguracji wyszukiwania korporacyjnego, konsoli administracyjnej oraz aplikacjach wyszukiwania.

Wskazówka:

Szczegółowe przykłady sposobu włączania zabezpieczeń globalnych na serwerze WebSphere Application Server z repozytorium LDAP, w tym przykłady konfigurowania rozmaitych przeszukiwaczy dla zapewnienia egzekwowania zabezpieczeń na poziomie dokumentu, można znaleźć w dokumentacji technicznej (redbook) firmy IBM zatytułowanej IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios (Scenariusze konfiguracji oraz implementacji oprogramowania IBM OmniFind Enterprise Edition).

Procedura

Aby włączyć zabezpieczenia w systemie wyszukiwania korporacyjnego, należy:

1. Zdecydować, jaki rodzaj rejestru użytkowników stosowany będzie do ich uwierzytelniania. Na przykład wielu administratorów serwerów aplikacji WebSphere decyduje się na zastosowanie rejestru użytkowników LDAP (Lightweight Directory Access Protocol).
2. W konsoli administracyjnej wyszukiwania korporacyjnego wybrać opcję **Bezpieczeństwo** i przypisać przynajmniej jednego użytkownika z rejestru użytkowników serwera aplikacji WebSphere do roli **Administratorsa wyszukiwania korporacyjnego**.

Ewentualnie można też podać identyfikator administratora wyszukiwania korporacyjnego, który definiowany jest w czasie instalowania systemu OmniFind Enterprise Edition do rejestru użytkowników serwera aplikacji WebSphere.

Ważne: Po włączeniu zabezpieczeń globalnych dostęp do konsoli administracyjnej i możliwości administrowania wyszukiwaniem korporacyjnym mają wyłącznie te identyfikatory użytkowników, którym przypisana została rola administratora wyszukiwania korporacyjnego.

3. Aby włączyć zabezpieczenia globalne i skonfigurować rejestr użytkowników, należy zastosować procedury opisane w dokumencie “Konfigurowanie zabezpieczeń globalnych oraz rejestru użytkowników protokołu LDAP w serwerze aplikacji WebSphere”.
4. Jeśli zabezpieczenia globalne włączone zostaną po zainstalowaniu produktu OmniFind Enterprise Edition, należy zaktualizować system wyszukiwania korporacyjnego, dodając do niego informacje o identyfikatorze oraz hasle użytkownika serwera WebSphere Application Server. Aby dodać te informacje, należy skorzystać z komendy **eschangewaspw**:
 - Jeśli system OmniFind Enterprise Edition został zainstalowany na jednym tylko serwerze, należy zastosować procedurę opisaną w dokumencie “Włączanie zabezpieczeń systemu wyszukiwania korporacyjnego z jednym serwerem” na stronie 276.
 - Jeśli system OmniFind Enterprise Edition został zainstalowany na wielu serwerach, należy zastosować procedurę opisaną w dokumencie “Włączanie zabezpieczeń systemu wyszukiwania korporacyjnego z wieloma serwerami” na stronie 277.
5. Wykonaj zadania właściwe dla typów dokumentów, jakie mają być przeszukiwane i wyszukiwane. Więcej informacji znaleźć można w dokumencie “Wymagania konfiguracyjne przeszukiwacza w zakresie bezpieczeństwa” na stronie 278.

Konfigurowanie zabezpieczeń globalnych oraz rejestru użytkowników protokołu LDAP w serwerze aplikacji WebSphere

Aby włączyć zabezpieczenia w programie OmniFind Enterprise Edition, należy najpierw włączyć globalne zabezpieczenia serwera aplikacji WebSphere.

Informacje o tym zadaniu

Częścią procesu włączania zabezpieczeń globalnych jest konfiguracja rejestru użytkowników pozwalającego na uwierzytelnianie ich identyfikatorów. W tym zadaniu definiowany jest sposób konfigurowania rejestru użytkowników protokołu LDAP (Lightweight Directory Access Protocol) przy jednoczesnym włączaniu zabezpieczeń globalnych. Co prawda serwer aplikacji WebSphere obsługuje też inne rodzaje rejestrów użytkowników, ale do uwierzytelniania użytkowników wyszukiwania korporacyjnego nie można wykorzystać rejestru lokalnego systemu operacyjnego. Jeśli wykorzystany miałby być rejestr systemu operacyjnego, to na serwerach wyszukiwania korporacyjnego istnieć musiałyby konta wszystkich użytkowników z całego przedsiębiorstwa.

To zadanie opiera się na serwerze aplikacji WebSphere w wersji 6. Jeśli zastosowana będzie inna wersja tego serwera, to pewne różnice mogą wystąpić w domyślnych ścieżkach oraz etykietach interfejsu użytkownika. Zadanie to tworzy też rejestr LDAP za pomocą serwera IBM Tivoli Directory Server. Jeśli zastosowany zostanie inny rodzaj rejestru lub rejestr niestandardowy, konieczne będzie wprowadzenie informacji właściwych dla tego rejestru.

W tym zadaniu podsumowywane są kroki, jakie trzeba wykonać w celu skonfigurowania zabezpieczeń globalnych dla systemu wyszukiwania korporacyjnego. Szczegółowe informacje uzyskać można w Centrum informacyjnym *WebSphere Application Server w wersji 6.0.x* dostępnym pod adresem: <http://publib.boulder.ibm.com/infocenter/wasinfo/v6r0/index.jsp>

Procedura

Aby włączyć zabezpieczenia serwera aplikacji WebSphere:

1. W celu otwarcia konsoli administracyjnej serwera aplikacji WebSphere wprowadź na serwerze wyszukiwania korporacyjnego poniższy adres URL, w którym *localhost* oznacza komputer lokalny, lub nazwę serwera, na przykład *omnifind.search.xyz.com*.
`http://localhost:9060/ibm/console`
2. Kliknij opcję **Bezpieczeństwo**, a następnie opcję **Zabezpieczenia globalne**.
3. Skonfiguruj serwer WebSphere tak, żeby korzystał z rejestru LDAP:
 - a. W sekcji **Rejestry użytkowników** kliknij pozycję **LDAP**.
 - b. Podaj identyfikator użytkownika oraz hasło stosowane do uruchamiania serwera aplikacji.
 - c. Jako typ rejestru wybierz IBM Tivoli Directory Server.
 - d. Określ nazwę hosta serwera LDAP w postaci adresu IP lub w postaci nazwy domenowej. Domyślny numer portu to 389.
 - e. Podaj podstawową nazwę wyróżniającą (DN - Distinguished Name) będącą początkowym punktem przeszukiwania rejestru, na przykład `ou=sales,o=ibm,c=us`.
 - f. Niektóre serwery LDAP nie obsługują powiązań anonimowych w czasie wyszukiwania w rejestrze, dlatego należy podać nazwę wyróżniającą dla serwera aplikacji, taką jak `cn=searchuser,o=ibm,c=us`, a następnie wprowadzić hasło serwera aplikacji. Serwer aplikacji korzysta z podanej nazwy wyróżniającej i hasła do tworzenia powiązań z rejestrem.
 - g. Aby komunikacja pomiędzy serwerami WebSphere i LDAP była szyfrowana protokołem SSL (Secure Sockets Layer), zaznacz pole wyboru **Włącz SSL**.
 - h. Kliknij przycisk **Zastosuj**, a następnie kliknij przycisk **OK**.
4. W sekcji **Właściwości ogólne** zaznacz pola wyboru **Włącz zabezpieczenia globalne** i **Wymuś zabezpieczenia Java 2**.
5. W celu włączenia mechanizmu aktywnego uwierzytelniania wybierz opcję Prosty mechanizm uwierzytelniania serwera WebSphere (SWAM).
6. Jako aktywny rejestr użytkowników wybierz rejestr LDAP (Lightweight Directory Access Protocol).
7. Kliknij przycisk **OK**.
8. Kliknij odsyłacz **Zapisz** znajdujący się u góry strony. Po pojawieniu się zachęty do zapisania zmian kliknij przycisk **Zapisz**.
9. Kliknij przycisk **Wyloguj** znajdujący się na pasku narzędzi.
10. Zatrzymaj i ponownie uruchom aplikację ESSearchServer.

Systemy **AIX, Linux lub Solaris**

```
./stopServer.sh ESSearchServer
./startServer.sh ESSearchServer
```

Systemy **Windows**

```
stopServer ESSearchServer
startServer ESSearchServer
```

Skrypty te znajdują się w katalogu `WAS_INSTALL_ROOT/AppServer/bin`:

- W przypadku serwera aplikacji WebSphere w wersji 5 domyślna ścieżka instalacji w systemach AIX ma postać `/usr/WebSphere`, w systemach Linux lub Solaris ma postać `/opt/WebSphere`, natomiast w systemach Windows - `C:\Program Files\WebSphere`.

- W przypadku serwera aplikacji WebSphere w wersji 6 domyślna ścieżka instalacji w systemach AIX ma postać /usr/IBM/WebSphere, w systemach Linux lub Solaris ma postać /opt/IBM/WebSphere, a w systemach Windows - C:\Program Files\IBM\WebSphere.
11. Ponownie uruchom konsolę administracyjną serwera aplikacji WebSphere.
 12. Teraz serwer uruchamia się w trybie bezpiecznym, dlatego w celu zalogowania się do konsoli wpisać trzeba identyfikator użytkownika oraz hasło podane w czasie konfigurowania rejestru użytkowników LDAP (zobacz krok3b na stronie 275).

Włączanie zabezpieczeń systemu wyszukiwania korporacyjnego z jednym serwerem

Jeśli globalne zabezpieczenia serwera aplikacji WebSphere zostaną włączone po zainstalowaniu serwera OmniFind Enterprise Edition, należy użyć komendy **eschangewaspw** i w ten sposób zaktualizować plik konfiguracji wyszukiwania korporacyjnego **es.cfg**, dopisując do niego hasło użytkownika serwera aplikacji WebSphere.

Zagadnienia wstępne

Upewnij się, że w pliku `config.properties` aplikacji `ESSearchApplication` podana jest nazwa użytkownika i hasło serwera aplikacji WebSphere. Domyślnie plik ten znajduje się w katalogu `ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war/WEB-INF`.

Informacje o tym zadaniu

Komenda **eschangewaspw** szyfruje hasło przed zapisaniem go w pliku `es.cfg`.

Procedura

Aby włączyć zabezpieczenia globalne wyszukiwania korporacyjnego w systemie z pojedynczym serwerem:

1. Na serwerze wyszukiwania korporacyjnego zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego:
`esadmin system stopall`
2. Upewnij się, że w pliku `ES_NODE_ROOT/nodeinfo/es.cfg` pozycja `WASUser` zawiera poprawną nazwę użytkownika serwera aplikacji WebSphere.
3. Uruchom poniższy skrypt, gdzie `haslo_WAS` oznacza hasło właściwe dla nazwy użytkownika serwera aplikacji WebSphere podawanej w pliku `ES_NODE_ROOT/nodeinfo/es.cfg` (zobacz krok 2).

Systemy AIX, Linux lub Solaris

```
eschangewaspw.sh haslo_WAS
```

Windows

```
eschangewaspw haslo_WAS
```

4. W systemie Windows wybierz kolejno opcje **Panel sterowania** → **Narzędzia administracyjne** → **Usługi** i dodaj taki sam identyfikator i hasło użytkownika WebSphere Application Server do usług WebSphere Application Server i `ESSearchServer`.
5. Zrestartuj system wyszukiwania korporacyjnego:
`esadmin system startall`

Włączanie zabezpieczeń systemu wyszukiwania korporacyjnego z wieloma serwerami

Jeśli globalne zabezpieczenia serwera aplikacji WebSphere zostaną włączone po zainstalowaniu serwera OmniFind Enterprise Edition, należy użyć komendy **eschangewaspw** i w ten sposób zaktualizować plik konfiguracji wyszukiwania korporacyjnego **es.cfg**, dopisując do niego hasło użytkownika serwera aplikacji WebSphere.

Zagadnienia wstępne

Upewnij się, że w pliku `config.properties` aplikacji `ESSearchApplication` podana jest nazwa użytkownika i hasło serwera aplikacji WebSphere. Domyślnie plik ten znajduje się w katalogu `ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war/WEB-INF` serwera wyszukiwania.

Informacje o tym zadaniu

Komenda **eschangewaspw** szyfruje hasło przed zapisaniem go w pliku `es.cfg`.

Procedura

Aby w istniejącym, wieloserwerowym systemie wyszukiwania korporacyjnego włączyć zabezpieczenia globalne, należy:

- Wykonaj poniższe kroki na serwerze indeksowania wyszukiwania korporacyjnego:
 - Zaloguj się jako administrator wyszukiwania korporacyjnego i zatrzymaj system wyszukiwania korporacyjnego:
`esadmin system stopall`
 - Upewnij się, że w pliku `ES_NODE_ROOT/nodeinfo/es.cfg` pozycja `WASUser` zawiera poprawną nazwę użytkownika serwera aplikacji WebSphere.
 - Uruchom poniższy skrypt, gdzie `haslo_WAS` oznacza hasło właściwe dla nazwy użytkownika serwera aplikacji WebSphere podawanej w pliku `ES_NODE_ROOT/nodeinfo/es.cfg` (zobacz krok 1b).

AIX, Linux i Solaris

```
eschangewaspw.sh haslo_WAS
```

Windows

```
eschangewaspw haslo_WAS
```

- Wykonaj poniższe kroki na drugim serwerze wyszukiwania (w przypadku konfiguracji dwuserwerowej), lub na serwerze przeszukiwacza i obu serwerach wyszukiwania (w przypadku konfiguracji czteroserwerowej):
 - Zaloguj się jako administrator wyszukiwania korporacyjnego.
 - Uruchom poniższy skrypt, gdzie `haslo_WAS` oznacza hasło właściwe dla nazwy użytkownika serwera aplikacji WebSphere podawanej w pliku `ES_NODE_ROOT/nodeinfo/es.cfg` (zobacz krok 1b).

Systemy AIX, Linux lub Solaris

```
eschangewaspw.sh haslo_WAS
```

Wiersz komend systemu Windows

```
eschangewaspw haslo_WAS
```

- W systemie Windows wybierz kolejno opcje **Panel sterowania** → **Narzędzia administracyjne** → **Usługi** i dodaj taki sam identyfikator i hasło użytkownika WebSphere Application Server do usług WebSphere Application Server i `ESSearchServer`.

- Na serwerze indeksowania wyszukiwania korporacyjnego zrestartuj system wyszukiwania korporacyjnego:
esadmin system startall

Wymagania konfiguracyjne przeszukiwacza w zakresie bezpieczeństwa

Przeszukiwacz musi mieć uprawnienia dostępu do danych bezpieczeństwa, aby w ten sposób zbierać informacje pozwalające na wymuszenie zabezpieczeń na poziomie dokumentów. W przypadku niektórych typów danych konfiguracja bezpiecznego środowiska wymaga podjęcia dodatkowych kroków.

Tabela 7. Wymagania konfiguracyjne przeszukiwacza w zakresie bezpieczeństwa

Przeszukiwacze Content Edition

Przed utworzeniem przeszukiwacza w celu uzyskania dostępu do repozytoriów w trybie bezpośrednim należy skonfigurować system WebSphere Information Integrator Content Edition, tak aby działał on w trybie bezpośrednim, oraz skonfigurować konektor z serwerem przeszukiwacza.

Przed utworzeniem przeszukiwacza w celu uzyskania dostępu do repozytoriów w trybie serwera należy uruchomić skrypt konfiguracyjny serwera przeszukiwacza (escrvbr.sh w systemach AIX, Linux lub Solaris albo escrvbr.vbs w systemach Windows).

Podczas konfigurowania przeszukiwacza podać należy identyfikator użytkownika oraz hasło umożliwiające przeszukiwaczowi dostęp do każdego przeznaczonego do przeszukania repozytorium. W miarę potrzeby dla każdego repozytorium w przestrzeni przeszukiwania można podać różne identyfikatory użytkowników i hasła

Tematy pokrewne:

- “Dostęp do repozytoriów Content Edition w trybie bezpośrednim” na stronie 43
- “Dostęp do repozytoriów WebSphere II Content Edition w trybie serwera” na stronie 44
- “Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu WebSphere II Content Edition” na stronie 45
- “Konfigurowanie serwera przeszukiwacza w systemie Windows dla programu WebSphere II Content Edition” na stronie 46

Przeszukiwacze DB2

Przed utworzeniem przeszukiwacza należy uruchomić skrypt konfiguracyjny serwera przeszukiwacza (escrdb2.sh w systemach AIX, Linux lub Solaris albo escrdb2.vbs w systemach Windows).

Podczas konfigurowania przeszukiwacza przeszukującego zdalne, niewpisane do katalogu bazy danych, należy podać identyfikator użytkownika oraz hasło umożliwiające przeszukiwaczowi dostęp do każdej przeznaczonej do przeszukania bazy danych znajdującej się na docelowym serwerze baz danych. Jeśli to konieczne, dla każdej bazy danych z przestrzeni przeszukiwania podać można różne identyfikatory użytkowników i hasła.

Tematy pokrewne:

- “Konfigurowanie serwera przeszukiwania w systemie UNIX do przeszukiwania źródeł DB2” na stronie 49
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł DB2” na stronie 50

Przeszukiwacze DB2 Content Manager

Przed utworzeniem przeszukiwacza należy uruchomić skrypt konfiguracyjny serwera przeszukiwacza (escrcm.sh w systemach AIX, Linux lub Solaris albo escrcm.vbs w systemach Windows).

Podczas konfigurowania przeszukiwacza podać należy identyfikator użytkownika oraz hasło umożliwiające przeszukiwaczowi dostęp do każdego przeznaczonego do przeszukania serwera. Jeśli to konieczne, dla każdego serwera w przestrzeni przeszukiwania podać można różne identyfikatory użytkowników i hasła.

Tematy pokrewne:

- “Konfigurowanie serwera przeszukiwaczy w systemie UNIX do współpracy z przeszukiwaczami DB2 Content Manager” na stronie 57
- “Konfigurowanie serwera przeszukiwaczy w systemie Windows do współpracy z przeszukiwaczami DB2 Content Manager” na stronie 59

Tabela 7. Wymagania konfiguracyjne przeszukiwacza w zakresie bezpieczeństwa (kontynuacja)

Przeszukiwacze Domino Document Manager, Notes, i QuickPlace

Aby przeszukać serwery Lotus Domino, które korzystają z protokołu zdalnych wywołań procedur (NRPC):

- W systemie AIX upewnij się, że na serwerze przeszukiwacza jest zainstalowany i dostępny moduł I/O Completion Port.
- Przed utworzeniem przeszukiwacza uruchom skrypt konfiguracyjny serwer przeszukiwacza (`escrnote.sh` w systemach AIX, Linux lub Solaris albo `escrnote.vbs` w systemach Windows).
- Na serwerze przeszukiwacza wyszukiwania korporacyjnego musi być zainstalowany serwer Domino, który dodatkowo musi być elementem przeznaczonej do przeszukiwania domeny Domino.
- Aby możliwe było sprawdzanie poprawności bieżących referencji użytkownika w czasie wysyłania przez niego żądania, przeznaczony do przeszukiwania serwer Domino musi być skonfigurowany jako zaufany serwer Lotus Domino.
- Podczas konfigurowania przeszukiwacza podaj ścieżkę do pliku identyfikatora użytkownika usługi Lotus Notes, który ma uprawnienia dostępu do serwera (na przykład `c:\Program Files\lotus\notes\data\name.id` lub `/local/notesdata/name.id`), oraz hasło właściwe dla tego pliku identyfikatora.

Aby przeszukać serwery Lotus Domino korzystające z protokołu Domino Internet Inter-ORB (DIIOP):

- W systemie AIX upewnij się, że na serwerze przeszukiwacza jest zainstalowany i dostępny moduł I/O Completion Port.
- Skonfiguruj serwer przeszukiwacza, tak aby mógł korzystać z tego protokołu.
- Podczas konfigurowania przeszukiwacza podaj pełny identyfikator użytkownika Lotus Notes z uprawnieniami dostępu do serwera (na przykład `Nazwa użytkownika/Dowolne miasto/Moja firma`) oraz hasło właściwe dla tego identyfikatora.

W celu przeszukiwania serwerów QuickPlace trzeba tak skonfigurować serwer QuickPlace, aby obsługiwał on mechanizm zabezpieczeń Local User lub Directory Assistance, w zależności od tego, który z nich ma być stosowany.

Tematy pokrewne:

- “Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 81
- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 76
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 78
- “Konfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika” na stronie 282
- “Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 80
- “Konfigurowanie serwera QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego” na stronie 284
- “Konfigurowanie opcji Directory Assistance na serwerze QuickPlace” na stronie 284

Przeszukiwacze Exchange Server

Tabela 7. Wymagania konfiguracyjne przeszukiwacza w zakresie bezpieczeństwa (kontynuacja)

<p>Podczas konfigurowania przeszukiwacza należy podać pełny identyfikator użytkownika z uprawnieniami dostępu do publicznych folderów na przeznaczonym do przeszukiwania serwerze Exchange Server oraz hasło właściwe dla tego identyfikatora.</p> <p>Aby przeszukiwacz mógł w czasie przeszukiwania danych korzystać z protokołu SSL i zarządzania kluczami serwera Exchange Server, należy też podać pełną ścieżkę do pliku kluczy oraz hasło umożliwiające przeszukiwaczowi dostęp do tego pliku. Plik kluczy musi istnieć na serwerze przeszukiwacza wyszukiwania korporacyjnego.</p>	<p>Temat pokrewny:</p> <ul style="list-style-type: none">• “Sprawdzanie dostępu do zabezpieczonych dokumentów serwera Exchange” na stronie 281
Przeszukiwacze JDBC	
<p>Podczas konfigurowania przeszukiwacza można podać identyfikator użytkownika oraz hasło pozwalające na przeszukiwanie tabel w docelowej bazie danych. Jeśli to konieczne, dla każdej bazy danych z przestrzeni przeszukiwania podać można różne identyfikatory użytkowników i hasła.</p>	
Przeszukiwacze NNTP	
<p>Przeznaczone do przeszukiwania serwery NNTP muszą umożliwiać serwerowi przeszukiwacza odczytywanie danych.</p>	
Przeszukiwacze systemu plików UNIX	
<p>Przeznaczone do przeszukiwania podkatalogi w systemach AIX, Linux i Solaris muszą umożliwiać serwerowi przeszukiwacza odczytywanie danych.</p>	
Przeszukiwacze sieci WWW	
<p>Przeszukiwacz sieci WWW stosuje się do protokołu wykluczania robotów. Jeśli na serwerze WWW w głównym katalogu znajduje się plik robots.txt, to przeszukiwacz analizuje zawartość tego pliku i tylko wtedy przeszukuje znajdujące się na tym serwerze serwisy WWW, gdy otrzyma na to zezwolenie. Więcej informacji na temat tego protokołu znaleźć można na stronie http://www.robotstxt.org/wc/exclusion.html.</p> <p>Podczas konfigurowania przeszukiwacza sieci WWW należy:</p> <ul style="list-style-type: none">• Podać nazwę agenta użytkownika, jaką stosować ma przeszukiwacz. Reguły zapisane w plikach robots.txt znajdujących się na przeznaczonych do przeszukiwania serwerach mogą zezwalać na dostęp lub odmawiać go na podstawie tej właśnie nazwy.• Opcjonalnie: jeśli w serwerze WWW stosowane jest podstawowe uwierzytelnianie protokołu HTTP, można podać referencje pozwalające przeszukiwaczowi sieci WWW na dostęp do stron zabezpieczonych hasłem.• Opcjonalnie: jeśli w serwerze WWW do ograniczenia dostępu stosowane są formularze HTML, można podać referencje pozwalające przeszukiwaczowi sieci WWW na dostęp do stron zabezpieczonych hasłem.	<p>Tematy pokrewne:</p> <ul style="list-style-type: none">• “Serwisy WWW zabezpieczone przez podstawowe uwierzytelnianie HTTP” na stronie 101• “Serwisy WWW zabezpieczone przez uwierzytelnianie oparte na formularzu” na stronie 102
Przeszukiwacze listy adresów początkowych, Web Content Management i WebSphere Portal	

Tabela 7. Wymagania konfiguracyjne przeszukiwacza w zakresie bezpieczeństwa (kontynuacja)

Przed utworzeniem przeszukiwacza konieczne jest uruchomienie skryptu, który zintegruje wyszukiwanie korporacyjne z serwerem WebSphere Portal. Dla poszczególnych wersji serwera WebSphere Portal dostępne są różne wersje skryptu.

Temat pokrewny:

- “Skrypty konfiguracyjne do integrowania wyszukiwania korporacyjnego z programem WebSphere Portal” na stronie 344

Podczas konfigurowania przeszukiwacza należy podać pełną nazwę wyróżniającą, która pozwoli mu pobrać strony z przeszukiwanego serwera, na przykład `uid=admin,cn=RegularEmployees,ou=Software Group,o=IBM,c=US`. Nie można też zapomnieć o podaniu hasła właściwego dla tej nazwy wyróżniającej. Nazwa wyróżniająca musi być zgodna z nazwą wyróżniającą zapisaną w konfiguracji portalu WebSphere Portal.

Trzeba się też upewnić, że podana nazwa wyróżniająca zdefiniowana jest w komponencie PAC (Portal Access Control) portalu WebSphere Portal. Przeszukiwacz korzysta z komponentu PAC w celu uzyskania danych kontroli dostępu do przeszukiwanych dokumentów.

Przeszukiwacze systemu plików Windows

Przeznaczone do przeszukania podkatalogi muszą umożliwić serwerowi przeszukiwacza odczytywanie danych. Podczas konfigurowania przeszukiwacza do przeszukiwania zdalnych systemów plików należy podać identyfikator użytkownika z uprawnieniami dostępu do zdalnych danych oraz hasło właściwe dla tego identyfikatora.

Tematy pokrewne:

- “Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows” na stronie 285
- “Bezpieczne wyszukiwanie w zaufanych domenach systemu Windows” na stronie 287

Aby możliwe było sprawdzanie poprawności bieżących referencji użytkownika w czasie wysyłania przez niego żądania, należy się upewnić, że konta w domenie są prawidłowo skonfigurowane. Wymagania dotyczące kont domen dla plików, które zostały przeszukane na komputerze lokalnym, różnią się od wymagań dla plików, które zostały przeszukane na zdalnym serwerze Windows.

Sprawdzanie dostępu do zabezpieczonych dokumentów serwera Exchange

Aby użyć przeszukiwacza Exchange Server do przeszukiwania dokumentów chronionych przez firewall, należy sprawdzić, czy serwer przeszukiwacza może uzyskiwać dostęp do serwera folderu publicznego Microsoft Exchange Server.

Informacje o tym zadaniu

Jeśli serwer przeszukiwacza nie może uzyskać dostępu do zabezpieczonego serwera Exchange Server, z serwera zostanie przesłany kod HTTP 501 (Nie zaimplementowano). Można również zobaczyć komunikaty wskazujące na odebranie nieoczekiwanej odpowiedzi HTTP.

Procedura

Aby upewnić się, że serwer przeszukiwacza ma dostęp do dokumentów za firewallem, należy wykonać poniższe czynności:

1. Uruchom przeglądarkę WWW na serwerze przeszukiwacza.

- Przejdź do adresu URL serwera folderu publicznego Exchange Server, który ma zostać przeszukany. Na przykład: <http://exchange.TwojaFirma.com/public/>
- Sprawdź, że możesz otworzyć stronę serwera Exchange.
Jeśli nie możesz uzyskać dostępu do serwera Exchange Server, skontaktuj się z administratorem serwera w swojej organizacji.

Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino

Jeśli przeznaczony do przeszukania serwer Domino korzysta z protokołu zdalnych wywołań NRPC (Notes remote procedure call), to przeszukiwacz musi zostać skonfigurowany tak, aby można było wymusić stosowanie reguł dostępu na poziomie dokumentów.

Aby na serwerze Domino korzystającym z protokołu NRPC wymusić zabezpieczenia dokumentów na poziomie dokumentów, należy zainstalować serwer Domino na serwerze przeszukiwacza. Ten serwer Domino musi należeć do domeny Domino użytkownika. Aby zainstalować i skonfigurować serwer Domino, należy postępować zgodnie z instrukcjami znajdującymi się w dokumentacji produktu Lotus Domino.

Aby serwery wyszukiwania mogły sprawdzać, czy użytkownik przeszukujący zabezpieczoną kolekcję ma uprawnienie do przeglądania dokumentów pasujących do kryteriów wyszukiwania, należy też wykonać następujące zadania:

- “Konfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika”.
- “Konfigurowanie zabezpieczeń globalnych oraz rejestru użytkowników protokołu LDAP w serwerze aplikacji WebSphere” na stronie 274.

Pojęcia pokrewne

“Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań” na stronie 264

“Przeszukiwacze Notes” na stronie 73

Konfigurowanie zaufanych serwerów Lotus Domino do sprawdzania poprawności referencji użytkownika

Aby wymusić zabezpieczenie dokumentów przeszukiwanych przez przeszukiwacza Notes korzystającego z protokołu zdalnych wywołań procedur Notes (NRPC), przeznaczone do przeszukania serwery Domino muszą zostać skonfigurowane jako zaufane serwery Lotus Domino.

Zagadnienia wstępne

Poniższa procedura jest wymagana, jeśli planuje się wymuszenie zabezpieczeń na poziomie dokumentu podczas przeszukiwania zdalnych baz danych. W przypadku lokalnych baz danych na serwerze przeszukiwacza procedura ta nie jest konieczna.

Aby można było skonfigurować serwery zaufane, na przeszukiwaczu musi być zainstalowany serwer Domino. Ten serwer Domino musi należeć do domeny Domino użytkownika.

Informacje o tym zadaniu

Podczas konfigurowania opcji zabezpieczeń na poziomie dokumentów dla przeszukiwacza Notes należy określić, czy kontrola dostępu ma być wymuszana poprzez sprawdzanie poprawności bieżących referencji użytkownika w momencie wysyłania przez niego zapytania. Aby możliwe było wymuszenie zabezpieczeń tego typu, serwery Domino przeznaczone do przeszukiwania muszą być zaufanymi serwerami Lotus Domino.

Gdy użytkownicy przeszukują domeny wymagające sprawdzenia ich bieżących referencji, serwer zaufany umożliwia identyfikatorowi serwera Domino przełączenie kontekstu na bieżący identyfikator użytkownika. Baza danych Domino zostaje otwarta w taki sposób, jakby otworzył ją bieżący użytkownik, i stosowane są wszystkie informacje z listy kontroli dostępu do bazy danych, które odnoszą się do tego użytkownika.

Możliwość przełączenia kontekstu w ten sposób jest zwykle dostępna wyłącznie dla baz danych zapisanych w katalogu `data` lokalnego serwera Domino. Począwszy od oprogramowania Lotus Domino w wersji 6.5.1 możliwość taką zapewnia serwer zaufany. W celu skonfigurowania serwera zaufanego administrator Domino określa serwery Domino, które mają mieć możliwość wykonywania operacji wymagających szczególnej uwagi z punktu widzenia bezpieczeństwa, takich jak działanie w imieniu innego użytkownika podczas uzyskiwania dostępu do baz danych ze zdalnego komputera.

Procedura

Aby skonfigurować serwer zaufany, należy wykonać następujące czynności na wszystkich serwerach Domino przeszukiwanych za pomocą przeszukiwacza Notes:

1. Na serwerze Domino otwórz klienta Lotus Domino Administrator za pomocą pliku identyfikacyjnego administratora domeny Domino.
2. Kliknij opcję **Plik**, a następnie wybierz opcję **Otwórz serwer**.
3. Wpisz nazwę serwera Domino, dla którego chcesz aktywować możliwości serwera zaufanego.
4. Wybierz kartę **Konfiguracja**.
5. Rozwiń obiekt **Serwer**, wybierz dokument **Bieżący serwer** i kliknij przycisk **Edytuj serwer**.
6. Wybierz kartę **Bezpieczeństwo**, przewiń na sam dół dokumentu, odszukaj pozycję **Serwery zaufane** i kliknij ikonę strzałki w dół.
7. Podaj jedną z następujących opcji:

LokalneSerweryDomeny

Wybierz tę opcję, jeśli wszystkie serwery w domenie Domino mają być traktowane jako serwery zaufane.

nazwa_serwera

Podaj nazwę serwera Domino, który ma mieć możliwość przeszukiwania i wyszukiwania jako serwer zaufany.

Jeśli serwer Domino, który ma być przeszukiwany, znajduje się w innej domenie Domino, należy podać jego nazwę lub wybrać grupę **InneSerweryDomeny**. Należy też przestrzegać procedur Domino dotyczących wzajemnej certyfikacji pliku identyfikatorów serwera wyszukiwania korporacyjnego Domino z innymi domenami Domino. Informacje na temat tych procedur można znaleźć w dokumentacji serwera Domino.

8. Kliknij przycisk **Zapisz i zamknij**, aby zapisać zmiany.
9. Zatrzymaj i zrestartuj zdalne serwery Domino, które skonfigurowano jako serwery zaufane.

Pojęcia pokrewne

“Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań” na stronie 264

“Przeszukiwacze Notes” na stronie 73

Konfigurowanie serwera QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego

Jeśli planujesz skonfigurować przeszukiwacz QuickPlace do korzystania z opcji użytkownika lokalnego w celu wdrożenia zabezpieczeń, przed utworzeniem przeszukiwacza należy skonfigurować katalog Domino na serwerze Lotus QuickPlace.

Informacje o tym zadaniu

Podczas konfigurowania przeszukiwacza QuickPlace można wybrać tryb zabezpieczeń używany przez przeszukiwacz do wymuszania zabezpieczeń na poziomie dokumentu. W przypadku wybrania trybu użytkownika lokalnego należy się upewnić, że wszystkie lokalne identyfikatory użytkownika i grupy lokalne zostały zarejestrowane w katalogu Domino (hierarchii katalogu Domino musi odpowiadać hierarchii serwera QuickPlace).

Należy się również upewnić, że identyfikator użytkownika i hasło określone dla używanego przeszukiwacza zostały zarejestrowane w katalogu Domino i mają uprawnienia do odczytu przeszukiwanej bazy danych.

Do korzystania z serwera QuickPlace potrzebna jest tylko nazwa użytkownika. Jednak do przeszukania źródeł QuickPlace wymagany jest w pełni rozwinięty identyfikator użytkownika. Rozwinięty identyfikator użytkownika ma następujący format:

nazwa użytkownika/nazwa lokalizacji/QP/nazwa domeny

Ta procedura umożliwi określenia w pełni rozwiniętej wersji identyfikatora użytkownika, upewnienie się, że dany identyfikator użytkownika ma uprawnienia do odczytu z bazy danych QuickPlace, oraz dodanie identyfikatora użytkownika do katalogu Domino. Katalog Domino musi zawierać identyfikator użytkownika, który będzie używany do przeszukiwania baz danych QuickPlace, oraz wszystkich lokalnych użytkowników i grupy programu QuickPlace (hierarchia katalogu Domino musi odpowiadać hierarchii serwera QuickPlace).

Procedura

Aby skonfigurować serwer QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego, należy wykonać poniższe czynności:

1. Potwierdź uprawnienia identyfikatora użytkownika:
 - a. Otwórz dokument Server na serwerze QuickPlace.
 - b. Otwórz stronę plików, a następnie otwórz listę kontroli dostępu (ACL) dla bazy danych, którą należy przeszukać.
 - c. Potwierdź, że lokalny identyfikator użytkownika, który zostanie uwzględniony w konfiguracji przeszukiwacza, istnieje na liście ACL oraz że ten identyfikator użytkownika ma uprawnienia do odczytu bazy danych.

W kroku 2 należy określić w pełni rozwiniętą formę tego identyfikatora użytkownika.
2. Dodaj nowego użytkownika do katalogu Domino:
 - a. Otwórz dokument Server na serwerze QuickPlace.
 - b. Na stronie osób i grup w elemencie drzewa osób dodaj w pełni rozwinięty identyfikator użytkownika potwierdzony w kroku 1.
 - c. W polu **Hasło dostępu do Internetu** podaj hasło dla tego identyfikatora użytkownika.

Konfigurowanie opcji Directory Assistance na serwerze QuickPlace

Jeśli planujesz skonfigurować przeszukiwacz QuickPlace do korzystania z katalogu LDAP w celu wdrożenia zabezpieczeń, przed skonfigurowaniem przeszukiwacza należy utworzyć bazę danych Directory Assistance na serwerze Lotus QuickPlace.

Ograniczenia

Serwer QuickPlace, który ma być przeszukiwany, musi obsługiwać zadania protokołów DIIOP i HTTP.

Procedura

Aby skonfigurować opcję LDAP Directory Assistance na serwerze QuickPlace, należy wykonać poniższe czynności:

1. Utwórz bazę danych Directory Assistance:
 - a. Otwórz dokument Server na serwerze QuickPlace.
 - b. Utwórz bazę danych za pomocą szablonu **Directory Assistance(6)**. Ten szablon znajduje się na serwerze.
 - c. Kliknij opcję **Dodaj bazę danych Directory Assistance**, aby utworzyć dokument w bazie danych.
 - d. Otwórz kartę Podstawowe i w polu **TypDomeny** wybierz opcję **LDAP**.
 - e. Otwórz kartę Konteksty nazw i upewnij się, że pole wyboru **Zaufane dla referencji** jest zaznaczone.
 - f. Otwórz kartę LDAP i podaj informacje o serwerze LDAP.
 - g. Zapisz i zamknij dokument Server.
2. Skonfiguruj serwer QuickPlace do korzystania z bazy danych Directory Assistance:
 - a. Otwórz dokument Server na serwerze QuickPlace.
 - b. Otwórz kartę Podstawowe i w polu **Nazwa bazy danych Directory Assistance** podaj nazwę bazy danych utworzonej w kroku 1.
 - c. Zapisz i zamknij dokument Server.

Serwer QuickPlace może teraz używać serwera LDAP jako dodatkowego katalogu Domino.

Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows

Aby włączyć opcję sprawdzania bieżących referencji podczas wyszukiwania przez użytkownika dokumentów przeszukiwanych przez przeszukiwacz systemu plików Windows, należy skonfigurować konto domeny zarówno na serwerze przeszukiwacza, jak i na serwerze Microsoft Windows.

Podczas konfigurowania przeszukiwacza systemu plików Windows należy określić, czy przeszukiwane mają być podkatalogi na komputerze lokalnym, czy podkatalogi na komputerze zdalnym. Jeśli zabezpieczenia kolekcji są włączone, można też określić opcje kontroli dostępu do dokumentów w przeszukiwanych podkatalogach.

W przypadku wymuszenia kontroli dostępu polegającej na sprawdzaniu bieżących referencji użytkownika podczas wprowadzania przez niego zapytania należy także zadbać o poprawną konfigurację kont domen. Wymagania dotyczące kont domen dla plików, które zostały przeszukane na komputerze lokalnym, różnią się od wymagań dla plików, które zostały przeszukane na zdalnym serwerze Windows.

Ważne: Jeśli oba poniższe warunki zostaną spełnione, w czasie przetwarzania zapytania referencje użytkowników nie mogą być sprawdzane.

- poddawany przeszukiwaniu serwer Windows nie jest członkiem domeny;
- poddawany przeszukiwaniu katalog jest katalogiem zdalnym, mającym postać `\\nazwa_serwera\nazwa_hosta`.

Sprawdzanie poprawności za pomocą lokalnych danych dotyczących kontroli dostępu

Sprawdzając bieżące referencje użytkownika, system korzysta zarówno z informacji o lokalnym koncie użytkownika, jak i z informacji o koncie domeny (jeśli komputer należy do domeny Windows). Aby możliwe było sprawdzenie referencji podczas przetwarzania zapytania, wśród informacji na temat zabezpieczeń dla dokumentów, które mają być wyszukiwane, muszą się znajdować obie nazwy użytkownika.

Konta lokalne

W przypadku konta lokalnego nazwa użytkownika ma następujący format:

NAZWA KOMPUTERA\NAZWA UŻYTKOWNIKA

Aby się zalogować, użytkownicy podają tylko swoje nazwy, ale do właściwego określenia uprawnień użytkownika systemu Windows wykorzystywana jest pełna nazwa. Na przykład jeśli nazwą lokalnego konta użytkownika jest abcuser, pełną nazwą konta może być WINSERVER1\abcuser.

Kiedy użytkownicy korzystają z aplikacji wyszukującej i konfiguruje profile używane do wyszukiwania zabezpieczonych dokumentów w systemie lokalnym, muszą podawać nazwę użytkownika, której używają do logowania się w systemie Windows (np. abcuser).

Konta domen

W przypadku konta domeny nazwa użytkownika ma następujący format:

NAZWA DOMENY\NAZWA UŻYTKOWNIKA

Aby się zalogować, użytkownicy podają te informacje w następującym formacie:

NAZWA UŻYTKOWNIKA@NAZWA DOMENY

Na przykład jeśli podczas konfigurowania praw dostępu użytkownika do pliku wybrana zostanie domena WIN1\abcuser, konto jest wyświetlane jako abcuser@win1.firma.com.

Kiedy użytkownicy korzystają z aplikacji wyszukującej i konfiguruje profile umożliwiające im wyszukiwanie dokumentów w zabezpieczonej domenie, muszą podawać nazwę użytkownika, której używają do logowania się w systemie Windows (np. abcuser@win1.firma.com).

Aby możliwe było wymuszenie sprawdzania bieżących referencji na komputerach lokalnych, konta użytkowników używane przez serwer przeszukiwacza muszą mieć następujące uprawnienia użytkowników Windows. Do określania uprawnień użytkowników służą narzędzia administracyjne systemu Windows: **Narzędzia administracyjne** → **Zasady zabezpieczeń lokalnych** → **Zasady lokalne** → **Przypisywanie praw użytkownika lokalnego**.

- Identyfikator użytkownika, którym posługuje się serwer przeszukiwacza, musi mieć ustawione prawo **Działanie jako element systemu operacyjnego**. Prawo to skonfigurowane jest dla użytkownika administracyjnego wyszukiwania korporacyjnego na serwerze przeszukiwacza podczas instalowania programu OmniFind Enterprise Edition.
- Użytkownicy muszą mieć uprawnienie **Logowanie lokalne**.

Sprawdzanie poprawności za pomocą danych dotyczących kontroli dostępu do domeny zdalnej

W systemie operacyjnym Windows każdy katalog o nazwie rozpoczynającej się od łańcucha *\\nazwa_serwera* uważany jest za katalog zdalny. Na przykład:

\\software\utilities\IBM

Aby uzyskać dostęp do katalogu zdalnego, użytkownicy muszą podawać swoje nazwy w następującym formacie:

NAZWA UŻYTKOWNIKA@NAZWA DOMENY

Kiedy użytkownicy korzystają z aplikacji wyszukującej i konfiguruje profile umożliwiające im wyszukiwanie zabezpieczonych dokumentów w systemie zdalnym, muszą podawać nazwę użytkownika, której używają w celu uzyskania dostępu do zdalnego systemu Windows (np. abcuser@win1.firma.com).

Aby możliwe było wymuszenie sprawdzania bieżących referencji na komputerach zdalnych, konta użytkowników muszą mieć następujące uprawnienia użytkowników systemu Windows. Do określania uprawnień użytkowników służą narzędzia administracyjne systemu Windows: **Narzędzia administracyjne** → **Zasady zabezpieczeń domeny**.

- Serwer przeszukiwacza i serwer Windows, który ma być przeszukiwany, muszą należeć do tej samej domeny.
- Identyfikator użytkownika, którym posługuje się serwer przeszukiwacza, musi mieć ustawione prawo **Działanie jako element systemu operacyjnego**. Prawo to konfigurowane jest dla użytkownika administracyjnego wyszukiwania korporacyjnego na serwerze przeszukiwacza podczas instalowania programu OmniFind Enterprise Edition.
- Użytkownicy muszą mieć uprawnienie **Logowanie w trybie wsadowym**.

Pojęcia pokrewne

“Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań” na stronie 264

“Przeszukiwacze systemu plików Windows” na stronie 114

“Bezpieczne wyszukiwanie w zaufanych domenach systemu Windows”

Bezpieczne wyszukiwanie w zaufanych domenach systemu Windows

Dla zapewnienia realizacji zabezpieczeń na poziomie dokumentu w odniesieniu do zdalnych systemów plików Windows system wyszukiwania korporacyjnego umożliwia weryfikację listy kontroli dostępu (ACL) w zaufanych domenach.

Konfigurowanie przeszukiwacza

Aby skonfigurować przeszukiwacz systemu plików Windows do celów obsługi zaufanych domen, należy określić opcje w nowym pliku konfiguracyjnym. Konsola administracyjna wyszukiwania korporacyjnego nie umożliwia konfigurowania takiej możliwości.

1. Aby zapewnić obsługę zabezpieczeń na poziomie dokumentu w zaufanych domenach Windows, dokonaj edycji następującego pliku:

```
ES_NODE_ROOT/master_config/id_sesji/winfscrawler_ext.xml
```

Wskazówka: W celu ustalenia identyfikatora sesji przeszukiwacza systemu plików Windows, który ma zostać skonfigurowany, można przeprowadzić monitorowanie przeszukiwacza za pomocą konsoli administracyjnej wyszukiwania korporacyjnego lub użyć komendy **esadmin report collections**.

2. Określ nazwę domeny Windows oraz nazwę NETBIOS usługi Active Directory. Na przykład:

```
<ExtendedProperties>
  <SetAttribute XPath="/Crawler/DataSources/Server/Target"
    Name="Domain">jk.enterprises.com
  </SetAttribute>
```



```
<SetAttribute XPath="/Crawler/DataSources/Server/Target"  
  Name="NetBIOSDomain">JKE1  
</SetAttribute>  
</ExtendedProperties>
```

3. Zatrzymaj i ponownie uruchom przeszukiwacz, aby zmiany zostały uwzględnione.

Ograniczenia

- Dokumenty nie mogą obejmować list ACL z wielu domen Windows. Użytkownicy i grupy domen muszą należeć do jednej domeny Windows na kolekcję.
- Aby obsługiwana była weryfikacja dostępu do zdalnego systemu plików, serwery Windows muszą działać w tej samej domenie Windows lub w zaufanych domenach Windows.
- Przeszukiwacz systemu plików Windows odczytuje nazwę NETBIOS usługi Active Directory powiązaną z serwerem Windows przeznaczonym do przeszukiwania i na podstawie nazwy NETBIOS filtruje listę ACL plików. Usługa Active Directory, z którą łączy się serwer przeszukiwacza, jest w relacji zaufania do innej usługi Active Directory, która definiuje konta użytkowników i konta grup.
- Konto użytkownika określone dla przeszukiwacza, aby za jego pomocą uzyskiwał dostęp do zdalnego serwera Windows, musi należeć do domeny Windows, w której mają być egzekwowane i weryfikowane prawa dostępu.
- System operacyjny Windows pozwala na nawiązywanie połączenia z folderami sieciowymi na jednym serwerze plików tylko za pomocą jednego konta. Za pomocą pozostałych kont nie można w tym samym czasie nawiązać połączenia z tym samym serwerem plików. Dlatego nie można skonfigurować różnych kont dla różnych przeszukiwaczy, aby przeszukiwały ten sam serwer Windows, nawet wtedy gdy przeszukiwacze znajdują się w różnych kolekcjach.

Pojęcia pokrewne

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows” na stronie 285

Wyłączanie zabezpieczeń w wyszukiwaniu korporacyjnym

Serwer aplikacji WebSphere pozwala na wyłączenie zabezpieczeń w aplikacji wyszukiwania korporacyjnego. Jeżeli wcześniej skonfigurowane zostały reguły zabezpieczeń na poziomie dokumentów, można określić, że reguły te mają być ignorowane. Ustawienia zabezpieczeń wpływają również na sposób wyświetlania zwiniętych wyników wyszukiwania.

Wyłączanie zabezpieczenia dla aplikacji korporacyjnej na serwerze WebSphere Application Server

Aby ustalić, które działania wyszukiwania korporacyjnego wymagają uwierzytelniania użytkownika, można wyłączyć globalne zabezpieczenia poszczególnych aplikacji korporacyjnych na serwerze aplikacji WebSphere.

Informacje o tym zadaniu

Program instalacyjny OmniFind Enterprise Edition instaluje trzy aplikacje korporacyjne serwera WebSphere Application Server:

- Aplikacja ESAdmin zawiera interfejs konsoli administracyjnej wyszukiwania korporacyjnego.
- Aplikacja ESSearchApplication zawiera interfejs przykładowej aplikacji wyszukującej.
- Aplikacja ESSearchServer udostępnia wszystkie mechanizmy zdalnej komunikacji związane z implementacją interfejsu SI-API wyszukiwania korporacyjnego i pozwala interfejsom SI-API na komunikowanie się z serwerem wyszukiwania.

Domyślnie wszystkie trzy aplikacje korporacyjne obsługują globalne zabezpieczenie programu WebSphere Application Server. Gdy te aplikacje wykryją, że zostało włączone zabezpieczenie globalne, rozpoczną uwierzytelnianie wszystkich otrzymywanych żądań.

Pewne organizacje mogą chcieć włączyć lub wyłączyć zabezpieczenia określonych aplikacji korporacyjnych. Na przykład można uwierzytelniać użytkowników uzyskujących dostęp do konsoli administracyjnej wyszukiwania korporacyjnego, ale pomijać uwierzytelnianie użytkowników korzystających z interfejsów SI-API przykładowej aplikacji wyszukiwania.

Procedura

Aby wyłączyć zabezpieczenia aplikacji korporacyjnej:

1. Na serwerze wyszukiwania uruchom WebSphere Application ServerKonsolę administracyjną.
Konsolę administracyjną można otworzyć na kilka sposobów:
 - Użyj menu **Start** systemu Windows i wybierz w nim odpowiedni program.
 - W przypadku serwera WebSphere Application Server w wersji 5 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/admin`, gdzie *nazwa_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9090.
 - W przypadku serwera WebSphere Application Server w wersji 6 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/ibm/console`, gdzie *nazwa_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9060.
2. Po wyświetleniu monitu o podanie identyfikatora użytkownika i hasła należy podać identyfikator administratora i hasło, które zostały określone podczas włączania zabezpieczenia globalnego w programie WebSphere Application Server.
3. Po zalogowaniu się do konsoli administracyjnej kliknij opcję **Aplikacje**, a następnie opcję **Aplikacje korporacyjne**.
4. Zaznacz pole wyboru obok nazwy aplikacji korporacyjnej, dla której ma zostać wyłączone zabezpieczenie.
5. Przejdź niżej i kliknij odsyłacz **Przypisz użytkownikom/grupom role dotyczące zabezpieczenia**.
6. Znajdź rolę **WszyscyUwierzytelniani** i zaznacz pole wyboru w kolumnie **Wszyscy?**.
7. Kliknij przycisk **OK**.
8. Kliknij odsyłacz **Zapisz**, aby zapisać zmiany.
9. W przypadku korzystania z programu WebSphere Network Deployment zaznacz pole wyboru **Synchronizuj zmiany z węzłami**.
10. Kliknij przycisk **Zapisz**.
11. Zatrzymaj i ponownie uruchom aplikację ESSearchServer.

```
Systemy AIX,  
Linux lub Solaris  
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

```
Systemy Windows  
stopServer ESSearchServer  
startServer ESSearchServer
```

Skrypty te znajdują się w katalogu `WAS_INSTALL_ROOT/AppServer/bin`:

- W przypadku serwera aplikacji WebSphere w wersji 5 domyślna ścieżka instalacji w systemach AIX ma postać /usr/WebSphere, w systemach Linux lub Solaris ma postać /opt/WebSphere, natomiast w systemach Windows - C:\Program Files\WebSphere.
- W przypadku serwera aplikacji WebSphere w wersji 6 domyślna ścieżka instalacji w systemach AIX ma postać /usr/IBM/WebSphere, w systemach Linux lub Solaris ma postać /opt/IBM/WebSphere, a w systemach Windows - C:\Program Files\IBM\WebSphere.

Wyłączanie zabezpieczeń na poziomie dokumentu

Użytkownikom można umożliwić wyszukiwanie w kolekcji niezależnie od praw dostępu powiązanych z dokumentami w indeksie. W przypadku przeszukiwaczy obsługujących sprawdzanie bieżących referencji można pozwolić użytkownikom na wyszukiwanie w kolekcji także bez sprawdzania bieżących praw dostępu podczas przetwarzania zapytania.

Zagadnienia wstępne

Włączanie i wyłączanie zabezpieczeń wszystkich dokumentów w kolekcji na poziomie dokumentów wymaga roli administratora wyszukiwania korporacyjnego lub uprawnień administratora danej kolekcji.

Ograniczenia

Opcje zabezpieczeń na poziomie dokumentu można określać tylko wtedy, gdy podczas tworzenia kolekcji zostały włączone zabezpieczenia.

Informacje o tym zadaniu

Przeszukiwacze można skonfigurować tak, aby w czasie przeszukiwania łączyły tokeny zabezpieczeń z dokumentami. Aplikacje wyszukujące mogą używać tych - przechowywanych w indeksie - tokenów do wymuszania kontroli dostępu podczas wyszukiwania w kolekcji przez użytkowników. W przypadku niektórych przeszukiwaczy można też określić, że podczas wprowadzania zapytania należy sprawdzać bieżące prawa dostępu związane z dokumentami znajdującymi się w ich własnych repozytoriach.

Aby usunąć ograniczenia związane z bezpieczeństwem, można określić, że serwery wyszukiwania mają ignorować wszystkie tokeny zabezpieczeń przekazywane z zapytaniem. Można także umożliwić użytkownikom wyszukiwanie dokumentów bez porównywania ich referencji z bieżącymi prawami dostępu.

Do testowania nowej kolekcji lub rozwiązywania problemów z aplikacją wyszukującą przydatne może być czasowe wyłączenie zabezpieczeń na poziomie dokumentu.

Procedura

Aby wyłączyć kontrolę dostępu na poziomie dokumentu:

1. Przeprowadź edycję kolekcji, wybierz stronę Ogólne i kliknij przycisk **Włącz lub wyłącz zabezpieczenia na poziomie dokumentów**.
2. Jeśli nie chcesz, żeby tokeny zabezpieczeń, jakie przeszukiwacze powiązały z dokumentami, były używane do obsługi zapytań użytkowników dotyczących tej kolekcji, na stronie Zabezpieczenia wszystkich dokumentów na poziomie dokumentów zaznacz pole wyboru **Ignoruj w indeksie kontrolę dostępu na poziomie dokumentów**.

Przeszukiwacze będą nadal dodawały tokeny zabezpieczeń do dokumentów, ale serwery wyszukiwania zignorują te tokeny i umożliwią użytkownikom wyszukiwanie we wcześniej zabezpieczonych dokumentach.

3. Jeśli podczas wprowadzania przez użytkowników zapytań bieżące prawa dostępu powiązane z dokumentami w ich rodzimych repozytoriach nie mają być sprawdzane, zaznacz pole wyboru **Nie sprawdzaj poprawności bieżących referencji podczas przetwarzania zapytania**. To pole wyboru jest dostępne tylko dla dokumentów przeszukanych przez przeszukiwacze, które obsługują tę możliwość.

Jeśli to pole wyboru zostanie zaznaczone, inne opcje zabezpieczeń na poziomie dokumentu pozostają aktywne. Na przykład jeśli podczas konfigurowania przeszukiwacza została określona opcja zapisywania praw dostępu w indeksie, to prawa te są stosowane, o ile nie zostanie zaznaczone także pole wyboru **Ignoruj w indeksie prawa dostępu na poziomie dokumentu**.

Wyłączanie zabezpieczeń w zwiniętych wynikach wyszukiwania

Jeśli włączone są zabezpieczenia kolekcji, to wyniki wyszukiwania z tego samego serwera nie mogą zostać zwinięte, chyba że podane zostanie, że referencje użytkowników nie mają być kontrolowane podczas przetwarzania zapytań.

Zagadnienia wstępne

Aby dla wszystkich dokumentów w kolekcji włączyć lub wyłączyć sprawdzanie poprawności bieżących referencji, trzeba być członkiem administracyjnej roli wyszukiwania korporacyjnego albo być administratorem danej kolekcji.

Ograniczenia

Opcje zabezpieczeń na poziomie dokumentu można określać tylko wtedy, gdy podczas tworzenia kolekcji zostały włączone zabezpieczenia.

Informacje o tym zadaniu

Podczas sprawdzania poprawności bieżących referencji źródło każdego dokumentu jest kontrolowane i prawdopodobnie przekazywane do sprawdzenia poprawności. Chcąc zwinąć w wynikach wyszukiwania dokumenty o jednakowym przedrostku identyfikatora URI albo dokumenty należące do wcześniej skonfigurowanej grupy zwiniętych identyfikatorów URI, należy wyłączyć sprawdzanie poprawności bieżących referencji.

Procedura

Aby wyłączyć sprawdzanie poprawności bieżących referencji i umożliwić tym samym zwijanie dokumentów w wynikach wyszukiwania:

1. Przeprowadź edycję kolekcji, wybierz stronę Ogólne i kliknij przycisk **Włącz lub wyłącz zabezpieczenia na poziomie dokumentów**.
2. Na stronie Zabezpieczenia wszystkich dokumentów na poziomie dokumentów zaznacz pole wyboru **Nie sprawdzaj poprawności bieżących referencji w czasie przetwarzania zapytań**.
3. Monitoruj kolekcję, wybierz stronę Szukaj, zatrzymaj i uruchom ponownie procesy serwera wyszukiwania.

W momencie odpytywania kolekcji przez użytkowników w wynikach wyszukiwania zwijane będą dokumenty o takim samym przedrostku identyfikatora URI lub należące do serwisów przygotowanych do zwijania. W przykładowej aplikacji wyszukiwania użytkownicy mogą przeglądać zwinięte wyniki, klikając odsyłacz **Więcej wyników z tego samego źródła**.

Uruchamianie i zatrzymywanie systemu wyszukiwania korporacyjnego

Po utworzeniu kolekcji konieczne jest uruchomienie serwerów przeszukiwania, analizy i indeksowania danych (serwery wyszukiwania uruchamiane są automatycznie). Po wprowadzeniu zmian do kolekcji należy zatrzymać i ponownie uruchomić te serwery.

Większość serwerów wyszukiwania korporacyjnego może działać w sposób nieprzerwany lub zgodnie z określonym harmonogramem. Na przykład można ustalić harmonogram budowania indeksu głównego i przyrostowego. Po uruchomieniu systemu wyszukiwania korporacyjnego zatrzymanie i ponowne uruchomienie procesów serwera konieczne jest tylko wtedy, gdy wprowadzone zostaną zmiany w ustawieniach konfiguracji (na przykład aktualizacja kategorii lub zwiększenie wielkości pamięci podręcznej wyszukiwania).

Jeśli zostaną wprowadzone zmiany w treści kolekcji albo w regułach zbierania danych przez przeszukiwacze pobierające dane ze źródeł danych w przedsiębiorstwie, zmiany te wejdą w życie po zatrzymaniu i ponownym uruchomieniu przeszukiwaczy. Jeśli reguły przeszukiwania nie zostaną zmienione, przeszukiwacz sieci WWW działać będzie bez przerwy, a pozostałe przeszukiwacze stosować się będą do przydzielonych im harmonogramów.

Aby zwiększyć dostępność serwerów wyszukiwania wtedy, gdy serwer indeksowania i konsola administracyjna są niedostępne, można określić komendy uruchamiające serwery wyszukiwania w trybie autonomicznym.

Uruchamianie systemu wyszukiwania korporacyjnego

Aby umożliwić użytkownikom wyszukiwanie w kolekcjach, należy uruchomić procesy systemowe, a następnie serwery, które przeszukują, analizują i indeksują kolekcje (serwery wyszukiwania zostają uruchomione automatycznie).

Zagadnienia wstępne

Należy skonfigurować źródła danych, które mają być przeszukiwane, i podać opcje określające sposób analizowania, indeksowania i wyszukiwania danych. Na przykład, jeśli użytkownicy mają mieć możliwość wyświetlania w wynikach wyszukiwania szczegółów kategorii, przed uruchomieniem analizatora składni należy skonfigurować kategorie.

Aby uruchomić serwery wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji. Uruchomić serwery może również operator z uprawnieniami do administrowania daną kolekcją.

Serwery wyszukiwania korporacyjnego dla kolekcji należy uruchomić w poprawnej kolejności. Na przykład indeksowanie danych może się odbyć dopiero po uprzednim uruchomieniu przeszukiwacza i przeprowadzeniu przeszukania tych danych.

Ograniczenia

Aby móc uruchomić system wyszukiwania korporacyjnego, należy użyć konta użytkownika, które może zostać lokalnie uwierzytelnione. Przy próbach uruchomienia systemu za pomocą konta typu Andrew File System (AFS) operacja kończy się błędem.

Procedura

Aby uruchomić system wyszukiwania korporacyjnego:

1. Jeśli system wyszukiwania korporacyjnego funkcjonuje w konfiguracji dwu- lub czteroserwerowej, zaloguj się jako administrator wyszukiwania korporacyjnego i uruchom wspólną warstwę komunikacyjną (CCL) na każdym serwerze:

AIX, Linux i Solaris



```
startccl.sh -bg
```

Wiersz komend systemu Windows

```
startccl
```

Narzędzie administracyjne usług systemu Windows


Aby uruchomić warstwę CCL w tle:

- a. Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
 - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
2. Uruchom komponenty systemu wyszukiwania korporacyjnego:
 - a. Zaloguj się jako administrator wyszukiwania korporacyjnego na dowolnym serwerze wyszukiwania.
 - b. Uruchom wszystkie komponenty systemowe: `esadmin system startall`
Ta komenda powoduje uruchomienie następujących procesów i aplikacji:
 - Serwer WWW (w konfiguracji z wieloma serwerami serwer WWW zostaje uruchomiony na obu serwerach wyszukiwania)
 - Aplikacje ESSearchServer i ESAdmin w środowisku WebSphere Application Server (w konfiguracji z wieloma serwerami aplikacje zostają uruchomione na obu serwerach wyszukiwania)
 - Sesja ESAdmin na serwerze indeksowania
 - Warstwa CCL na komputerze, na którym uruchomiono komendę, jeśli warstwa CCL nie została uruchomiona już wcześniej
 - Sieciowy serwer bazy danych wyszukiwania korporacyjnego
 - Centrum informacyjne wyszukiwania korporacyjnego (w konfiguracji z wieloma serwerami Centrum informacyjne zostaje uruchomione na obu serwerach wyszukiwania)
 3. Uruchom konsolę administracyjną wyszukiwania korporacyjnego i zaloguj się jako administrator wyszukiwania korporacyjnego. Jeśli korzystasz z roli administracyjnej, możesz zalogować się jako administrator kolekcji lub jako operator mający uprawnienia do uruchamianej kolekcji.
 4. W widoku Kolekcje odzyskaj kolekcję, którą chcesz administrować, i kliknij ikonę  **Monitoruj**.
 5. Na stronie Przeszukiwanie kliknij ikonę  **Uruchom** dla każdego przeszukiwacza, który ma zostać uruchomiony.
 - Przeszukiwacz WWW natychmiast po uruchomieniu przystępuje do przeszukiwania danych. Przeszukiwacze tego typu działają w sposób ciągły, przeszukując cyklicznie dokumenty WWW.
 - Jeśli uruchamiany jest jeden z pozostałych typów przeszukiwaczy, zostanie uruchomiona sesja przeszukiwacza. Przeszukiwacz rozpocznie przeszukiwanie w zaplanowanym terminie. Jeśli przeszukiwacz nie figuruje w harmonogramie lub jeśli trzeba wcześniej uruchomić przeszukiwacz, przejdź do monitorowania przeszukiwacza i kliknij ikonę uruchomienia dla każdego źródła danych, które ma zostać przeszukane.



Po uruchomieniu przeszukiwacza można zezwolić mu na nieprzerwane działanie. Jeśli utworzono harmonogram dla przeszukiwacza, zostanie on ponownie uruchomiony w zaplanowanym dniu i o zaplanowanej godzinie.

6. Po przeszukiwaniu danych otwórz stronę Analiza składni i kliknij ikonę  **Uruchom**, aby uruchomić analizator składni.

Można zezwolić analizatorowi składni na nieprzerwane działanie. Zwykle nie ma potrzeby zatrzymywania analizatora, chyba że wprowadzone zostały zmiany w sposobie analizowania danych (takie jak aktualizacja kategorii lub odwzorowań pól XML).

7. Opcjonalnie: aby wymusić rozpoczęcie procesów indeksowania w wybranym momencie, zamiast czekać na termin ustalony harmonogramem, otwórz stronę Indeks i w obszarze **Główne** kliknij ikonę  **Uruchom**.

Można zezwolić procesom indeksowania na nieprzerwane działanie. Budowanie indeksu będzie inicjowane w zaplanowanych terminach.

Wskazówka: Serwery wyszukiwania są uruchamiane automatycznie i można im pozwolić działać w trybie ciągłym. Zwykle nie ma potrzeby zatrzymywania serwerów wyszukiwania, chyba że zostały wprowadzone zmiany w ustawieniach pamięci podręcznej wyszukiwania lub ustawieniach podsumowania dokumentów. Aby ponownie uruchomić serwery wyszukiwania, otwórz stronę Szukaj, kliknij ikonę  **Zatrzymaj**, a następnie ikonę  **Uruchom**.

Aby zwiększyć dostępność serwerów wyszukiwania wtedy, gdy konsola administracyjna jest niedostępna, można określić komendy uruchamiające serwery wyszukiwania w trybie autonomicznym. Jeśli serwer indeksowania nie jest uruchomiony, konsola administracyjna jest niedostępna.

Pojęcia pokrewne

“Role administracyjne” na stronie 258

Zadania pokrewne

“Logowanie się do konsoli administracyjnej” na stronie 19

“Wybieranie komponentów do uruchamiania lub zatrzymywania” na stronie 297

“Administrowanie serwerami wyszukiwania pracującymi w trybie autonomicznym” na stronie 298

Zatrzymywanie systemu wyszukiwania korporacyjnego

Może wystąpić konieczność zatrzymania i zrestartowania serwera wyszukiwania korporacyjnego, jeśli dokonano zmian w jego konfiguracji lub w celu rozwiązania problemów.

Zagadnienia wstępne


Aby zatrzymać serwery wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji. Zatrzymać serwery może również operator z uprawnieniami do administrowania daną kolekcją.

Informacje o tym zadaniu

Serwery wyszukiwania korporacyjnego można zatrzymywać niezależnie od siebie. Na przykład zatrzymanie i zrestartowanie przeszukiwacza w celu uaktywnienia zmian dokonanych w jego konfiguracji nie wymaga zatrzymania i zrestartowania serwerów wyszukiwania.

Procedura

Aby zatrzymać serwery wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz administrować, i kliknij ikonę  **Monitoruj**.

2. Na stronie Przeszukiwanie znajdź przeszukiwacz, którym chcesz administrować, i zatrzymaj lub wstrzymaj go.


Jeśli zmieniona została przestrzeń przeszukiwania lub właściwości przeszukiwacza, aby zmiany zostały uwzględnione, należy zatrzymać i zrestartować przeszukiwacz. Jeśli zmodyfikowano przestrzeń przeszukiwania i zmiany mają także być uwzględnione dla dokumentów, które już zostały poindeksowane, należy wykonać ponowne przeszukiwanie tych dokumentów.


Wskazówka: Może zostać wyświetlony komunikat o przekroczeniu limitu czasu przez żadaną operację, mimo że proces nadal działa w tle. Aby dowiedzieć się, czy zadanie zostało ukończone, kliknij przycisk **Odśwież** w konsoli administracyjnej (nie klikaj przycisku **Odśwież** w przeglądarce WWW). Proces jest zakończony, gdy ikona statusu przeszukiwacza wskazuje, że został on zatrzymany.

3. Na stronie Analiza składni kliknij opcję  **Zatrzymaj**, aby zatrzymać analizator składni.

Jeśli zmienione zostały reguły analizowania danych, aby zmiany zostały uwzględnione, należy zatrzymać i zrestartować analizator składni. Zmiany będą stosowane tylko w odniesieniu do nowo przeszukanych dokumentów. Aby zastosować zmiany wobec dokumentów, które już się znajdują w indeksie, należy uruchomić pełne przeszukiwanie, obejmujące wszystkie dokumenty, co pozwala poddać je ponownej analizie składni i indeksowaniu.

4. Na stronie Indeks kliknij ikonę  **Zatrzymaj**, aby zatrzymać indeks będący w trakcie budowania.

Można również zatrzymać budowanie indeksu podczas monitorowania kolejki indeksów. W tym celu wybierz na pasku narzędzi ikonę **System**, otwórz stronę Indeks i kliknij ikonę  **Zatrzymaj** odpowiadającą indeksowi, którego budowanie ma zostać zatrzymane.

5. Na stronie Wyszukiwanie kliknij opcję  **Zatrzymaj**, aby zatrzymać serwery wyszukiwania. Zwykle nie ma potrzeby zatrzymywania i restartowania serwerów wyszukiwania, chyba że zmienione zostały ustawienia pamięci podręcznej wyszukiwania lub ustawienia podsumowania dokumentu.

6. Aby zatrzymać system wyszukiwania korporacyjnego zamiast poszczególnych serwerów:

- a. Zaloguj się jako administrator wyszukiwania korporacyjnego na dowolnym serwerze wyszukiwania.

- b. Zatrzymaj wszystkie komponenty systemu: `esadmin system stopall`

Ta komenda powoduje, że zatrzymane zostaną następujące procesy i aplikacje:

- Serwer WWW (w konfiguracji z wieloma serwerami serwer WWW zostaje zatrzymany na obu serwerach wyszukiwania)
- Aplikacje ESSearchServer i ESAdmin w środowisku WebSphere Application Server (w konfiguracji z wieloma serwerami aplikacje zostają zatrzymane na obu serwerach wyszukiwania)
- Sesja ESAdmin na serwerze indeksowania
- Wspólna warstwa komunikacyjna (CCL) wyszukiwania korporacyjnego na komputerze, na którym uruchomiono komendę
- Sieciowy serwer bazy danych wyszukiwania korporacyjnego
- Centrum informacyjne wyszukiwania korporacyjnego (w konfiguracji z wieloma serwerami Centrum informacyjne zostaje zatrzymane na obu serwerach wyszukiwania)

Zadania pokrewne

“Logowanie się do konsoli administracyjnej” na stronie 19

“Wybieranie komponentów do uruchamiania lub zatrzymywania”

“Administrowanie serwerami wyszukiwania pracującymi w trybie autonomicznym” na stronie 298

Wybieranie komponentów do uruchamiania lub zatrzymywania

Komponenty przeznaczone do uruchomienia lub zatrzymania wybiera się za pomocą komend **esadmin system startall** i **esadmin system stopall**.

Informacje o tym zadaniu

Plik `ES_INSTALL_ROOT/default_config/AutoRunComponents.properties` zawiera listę komponentów wyszukiwania korporacyjnego, które mogą być uruchamiane lub zatrzymywane za pomocą komend **esadmin system startall** i **esadmin system stopall**. Domyślnie wszystkie wymienione komponenty są uruchamiane i zatrzymywane przez te komendy.

Aby zapobiec uruchamianiu lub zatrzymywaniu określonych komponentów, należy zmienić zawartość pliku właściwości.

Procedura

Aby wybrać komponenty przeznaczone do uruchomienia lub zatrzymania wraz z uruchamianiem lub zatrzymywaniem systemu wyszukiwania korporacyjnego:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerze, na którym planujesz wydać komendę **esadmin system startall** lub **esadmin system stopall**.
2. Zmień zawartość pliku `ES_INSTALL_ROOT/default_config/AutoRunComponents.properties`.
3. Aby zapobiec uruchamianiu komponentu, dodaj pole dla komponentu o nazwie `Component.startable.id_komponentu=false`, gdzie `id_komponentu` jest identyfikatorem komponentu, który ma nie być uruchamiany.
4. Aby zapobiec zatrzymywaniu komponentu, dodaj pole dla komponentu o nazwie `Component.stopable.id_komponentu=false`, gdzie `id_komponentu` jest identyfikatorem komponentu, który ma nie być zatrzymywany.
5. Zapisz i zamknij plik.

Przy następnym użyciu komend **esadmin system startall** lub **esadmin system stopall** zmieniony komponent nie będzie uruchamiany lub zatrzymywany, zgodnie z dokonanymi zmianami w pliku właściwości.

Przykład: W tym przykładzie serwer HTTP będzie uruchamiany za pomocą komendy **esadmin system startall** (ustawienie domyślne), nie będzie jednak zatrzymywany za pomocą komendy **esadmin system stopall** (zgodnie z wyróżnionym wierszem w poniższym przykładzie):

```
#####  
# Szczegóły komponentu nr 3.  
#####  
Component.name.3=IBM HTTP server  
Component.impl.class.3=com.ibm.es.control.util.component.impl.HTTPControlImpl  
Component.nodes.3=search  
# Domyślnie można uruchamiać wszystkie komponenty  
Component.stopable.3=false
```

Zadania pokrewne

“Uruchamianie systemu wyszukiwania korporacyjnego” na stronie 293

“Zatrzymywanie systemu wyszukiwania korporacyjnego” na stronie 295

Administrowanie serwerami wyszukiwania pracującymi w trybie autonomicznym

Aby zapewnić wysoką dostępność serwerów wyszukiwania można uruchomić je dla oddzielnych kolekcji, nawet jeśli serwer indeksowania nie jest uruchomiony.

Ograniczenia

Aby zatrzymać i uruchomić serwery wyszukiwania, należy być administratorem wyszukiwania korporacyjnego.

Możliwość uruchomienia i zatrzymania serwerów wyszukiwania pracujących w trybie autonomicznym nie jest dostępna w konsoli administracyjnej wyszukiwania korporacyjnego. Jeśli serwer indeksowania nie jest uruchomiony, konsola administracyjna jest niedostępna.

Przed uruchomieniem i zatrzymaniem serwerów wyszukiwania w trybie autonomicznym należy upewnić się, że serwery przeszukiwacza, analizatora składni, indeksowania i wyszukiwania zostały uruchomione dla danej kolekcji przynajmniej raz. Jest to niezbędne, aby zapewnić synchronizację wymaganych plików na serwerach wyszukiwania.

Jeśli włączone są zabezpieczenia na poziomie dokumentu, należy również upewnić się, że uruchomiony jest serwer przeszukiwacza. Jest to niezbędne do zapewnienia wymuszania zabezpieczeń na poziomie dokumentu. Jeśli serwer przeszukiwacza nie jest uruchomiony, w wynikach wyszukiwania będą umieszczane tylko te dokumenty, dla których nie jest wymagane uwierzytelnianie.

Informacje o tym zadaniu

Jeśli serwer indeksowania nie jest dostępny, można zapewnić, że użytkownicy nadal będą mogli wyszukiwać informacji, uruchamiając serwery wyszukiwania w trybie autonomicznym. Komendy uruchamiające i zatrzymujące serwery wyszukiwania można wydać w dowolnym serwerze wyszukiwania korporacyjnego w instalacji wieloserwerowej. Komendy będą próbować zatrzymać bądź uruchomić serwery dla podanej kolekcji na wszystkich dostępnych serwerach wyszukiwania.

Procedura

Aby uruchomić lub zatrzymać serwery wyszukiwania w trybie autonomicznym:

1. Aby uruchomić serwery wyszukiwania dla kolekcji, gdy serwer indeksowania nie jest uruchomiony:
 - a. W serwerze przeszukiwacza i serwerach wyszukiwania zaloguj się jako administrator wyszukiwania korporacyjnego i uruchom usługę CCL (common communication layer):

AIX, Linux i Solaris

```
startccl.sh -bg
```

Wiersz komend systemu Windows

```
startccl
```

Narzędzie administracyjne usług systemu Windows

Aby uruchomić warstwę CCL w tle:

- 1) Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
- 2) Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.

- b. Na serwerach wyszukiwania uruchom skrypt **startServer**, znajdujący się w katalogu *WAS_INSTALL_ROOT/AppServer/bin*, aby uruchomić aplikację ESSearchServer na serwerze WebSphere Application Server:

Systemy AIX, Linux lub Solaris

```
./startServer.sh ESSearchServer
```

Windows

```
startServer ESSearchServer
```

- c. Wprowadź poniższą komendę, podając jako *id_kolekcji* identyfikator kolekcji, do której należą serwery wyszukiwania, które chcesz uruchomić:

```
esadmin startSearch -cid id_kolekcji
```
2. Aby zatrzymać serwery wyszukiwania dla kolekcji, gdy serwer indeksowania nie jest uruchomiony:
 - a. Zaloguj się jako administrator wyszukiwania korporacyjnego na dowolnym serwerze wyszukiwania.
 - b. Wprowadź poniższą komendę, podając jako *id_kolekcji* identyfikator kolekcji, do której należą serwery wyszukiwania, które chcesz zatrzymać:

```
esadmin stopSearch -cid id_kolekcji
```

Zadania pokrewne

“Uruchamianie systemu wyszukiwania korporacyjnego” na stronie 293

“Zatrzymywanie systemu wyszukiwania korporacyjnego” na stronie 295

Monitorowanie działań związanych z wyszukiwaniem korporacyjnym

Monitorując działania systemu i działania związane z kolekcjami, można wyświetlić informacje o statusie różnych procesów, obserwować zmienne w celu wykrycia potencjalnych problemów lub dopasować ustawienia konfiguracyjne w celu poprawienia wydajności.

Konsola administracyjna wyszukiwania korporacyjnego umożliwia monitorowanie systemu i dostosowanie operacji do własnych potrzeb. Można także wyświetlić szczegółowe informacje statystyczne dotyczące każdego podstawowego działania (przeszukiwanie, analizowanie, indeksowanie i wyszukiwanie). Statystyka obejmuje średnie czasy odpowiedzi i informacje o postępie, na przykład liczbę przeszukanych lub poindeksowanych dokumentów w trakcie sesji.

Klikając odpowiednie ikony, użytkownik może zatrzymać lub uruchomić większość działań. Istnieje możliwość wstrzymania operacji, zmodyfikowania związanych z nią ustawień konfiguracyjnych lub rozwiązania zaistniałych problemów, a następnie zrestartowania procesu przetwarzania w odpowiednim momencie, gdy możliwe będzie kontynuowanie operacji.

Zadania pokrewne

“Uruchamianie systemu wyszukiwania korporacyjnego” na stronie 293

“Zatrzymywanie systemu wyszukiwania korporacyjnego” na stronie 295

Szacowanie liczby dokumentów w kolekcji

Podczas tworzenia lub edycji kolekcji wyszukiwania korporacyjnego należy podać szacunkową liczbę dokumentów, które mają być w niej przechowywane. Na podstawie tej wartości system szacuje zapotrzebowanie kolekcji na pamięć i miejsce na dysku, nie stwarzając przy tym żadnych ograniczeń co do wielkości kolekcji.

Zagadnienia wstępne

Aby zmienić szacowaną wielkość kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienie administratora danej kolekcji.

Informacje o tym zadaniu

Gdy kolekcja osiągnie szacowaną wielkość, system nie przestanie dodawać dokumentów do indeksu. Jeśli skonfigurowano alerty dla kolekcji i włączono opcję powiadamiania, gdy liczba dokumentów w indeksie przekroczy limit, domyślny limit jest zgodny z wartością podaną jako szacowana liczba dokumentów w kolekcji. System monitoruje liczbę dokumentów z uwzględnieniem tego oszacowania oraz procentowego progu dla powiadomienia. Gdy liczba dokumentów w kolekcji zbliży się do zdefiniowanego limitu, wysyłane jest powiadomienie pocztą elektroniczną.

Procedura

Aby podać oszacowanie potencjalnej wielkości kolekcji, należy wykonać następujące czynności:

1. Przejdź do edycji kolekcji, wybierz stronę Ogólne, a następnie kliknij opcję **Konfiguruj opcje ogólne**.

2. W polu **Szacowana liczba dokumentów** wpisz przewidywaną maksymalną wielkość kolekcji. Domyślna wartość to 1 000 000 dokumentów.

Monitorowanie kolekcji


Użytkownik może wyświetlić ogólne informacje na temat statusu poszczególnych komponentów kolekcji albo wybrać opcje dotyczące wyświetlania szczegółowych informacji o wybranych komponentach i identyfikatorach URI.

Zagadnienia wstępne


Do monitorowania kolekcji uprawnieni są wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego. Aby edytować właściwości przeszukiwacza, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora kolekcji dla tej kolekcji.

Procedura

Aby monitorować kolekcję, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**. Wyświetlone zostaną informacje na temat bieżącego statusu każdego z komponentów kolekcji.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Ogólne jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby wyświetlić widok umożliwiający monitorowanie kolekcji.

2. Aby wyświetlić szczegółowe informacje o identyfikatorze URI, kliknij ikonę  **Szczegóły URI**.

Na przykład można wyświetlić informacje o tym, czy dany identyfikator URI znajduje się w indeksie lub czy indeks, w którym identyfikator URI się znajduje, został skopiowany na serwery wyszukiwania.

3. Aby monitorować wybrany komponent i wyświetlić szczegółową statystykę na temat jego działania, kliknij ikonę **Status**.

Wyświetlanie szczegółów dotyczących identyfikatora URI

Użytkownik może wyświetlić szczegółowe informacje na temat identyfikatora URI. Wyświetlić można aktualne i historyczne informacje odnośnie przeszukiwania, indeksowania i wyszukiwania dokumentu reprezentowanego przez dany identyfikator URI.

Zagadnienia wstępne

Przed wprowadzeniem żądania wyświetlenia raportu o identyfikatorze URI lub wysłania raportu na wskazany adres e-mail należy się upewnić, że komponent, który ma być źródłem pochodzenia informacji, jest aktywny. Na przykład, aby wyświetlić informacje szczegółowe na temat przeszukiwania, indeksowania i wyszukiwania dokumentu, należy się upewnić, że uruchomiony jest przeszukiwacz sieci WWW, serwer indeksowania i serwer wyszukiwania. Aby możliwe było śledzenie usuniętego dokumentu, muszą być skonfigurowane opcje rejestrowania dotyczące funkcji śledzenia dokumentów.

Informacje o tym zadaniu


Gromadzenie informacji na temat identyfikatora URI jest procesem czasochłonnym. Można wybrać opcję umożliwiającą wyświetlenie informacji objętych żądaniem, a następnie poczekać na ich wyświetlenie. Metodą bardziej efektywną jest wysłanie raportu na podany przez użytkownika adres e-mail.

Przed wysłaniem raportu należy się upewnić, że dla wyszukiwania korporacyjnego zostały skonfigurowane informacje o serwerze poczty. Informacje te należy podać podczas konfigurowania opcji poczty elektronicznej na stronie Rejestrowanie w widoku System.


Serwer indeksowania i serwer wyszukiwania mogą udostępnić informacje na temat wszystkich identyfikatorów URI (na przykład, czy identyfikator URI znajduje się w indeksie oraz czy został skopiowany na serwer wyszukiwania). Aby wyświetlić informacje na temat przeszukiwania dokumentu, należy podać identyfikator URI dokumentu, który został przeszukany przez przeszukiwacz sieci WWW.

Procedura

Aby wyświetlić szczegóły na temat identyfikatora URI, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Ogólne jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby wyświetlić widok umożliwiający monitorowanie kolekcji.

2. Kliknij ikonę  **Szczegóły URI**.
3. Na stronie Szczegóły dotyczące identyfikatorów URI wpisz identyfikator URI, którego mają dotyczyć wyświetlane informacje.
4. Zaznacz pola wyboru odpowiadające rodzajowi informacji, które mają być wyświetlone:

Szczegóły przeszukiwacza (dostępne tylko dla przeszukiwaczy sieci WWW)

Zaznacz to pole wyboru, aby wyświetlić informacje o operacji przeszukiwania dokumentu przez przeszukiwacz sieci WWW, a także o jego aktualnym statusie w przestrzeni przeszukiwania.

Szczegóły indeksu

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został poindeksowany i skopiowany na serwer wyszukiwania.

Szczegóły wyszukiwania

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, w jaki sposób dokument może być wyszukiwany i czy jest on w ogóle dostępny do wyszukiwania.

Dokumenty usunięte przez analizator składni

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty z systemu wyszukiwania korporacyjnego podczas analizowania, a jeśli tak, to jaka była tego przyczyna.

Dokumenty usunięte z indeksu

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty z systemu wyszukiwania korporacyjnego podczas indeksowania lub analizowania, a jeśli tak, to jaka była tego przyczyna.

5. Aby zaczekać na wyświetlenie raportu, kliknij opcję **Wyświetl raport**.
6. Aby wysłać raport na podany adres e-mail w celu przejrzania go w późniejszym terminie, kliknij opcję **Wyślij raport**.
 - a. Na stronie Wysłanie szczegółowego raportu o identyfikatorach URI, w polu **Adres poczty elektronicznej do wysłania powiadomienia** wpisz adres poczty elektronicznej, na który ma zostać wysłany raport.

b. Kliknij przycisk **Wyślij raport**.

Pojęcia pokrewne

“Śledzenie dokumentu” na stronie 319

Zadania pokrewne

“Wyświetlanie raportów o usuniętych dokumentach” na stronie 320

Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 120

Monitorowanie przeszukiwaczy


Użytkownik może wyświetlić ogólne informacje na temat statusu poszczególnych przeszukiwaczy w kolekcji albo wybrać opcje dotyczące wyświetlania szczegółowych informacji o działaniu przeszukiwacza.


Zagadnienia wstępne


Jeśli uprawnienia przyznane użytkownikowi ograniczone są do monitorowania kolekcji, ma on możliwość przeglądania statystyki przeszukiwacza, ale nie może wpływać na jego działanie (uruchamiać lub zatrzymywać przeszukiwacz).

Procedura


Aby monitorować przeszukiwacz, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**.
2. Otwórz stronę Przeszukiwanie.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Przeszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby zmienić widok na umożliwiający monitorowanie przeszukiwaczy.

3. Jeśli przeszukiwacz jest uruchomiony lub wstrzymany i trzeba wyświetlić szczegółowe informacje na jego temat, kliknij ikonę  **Szczegóły**. Rodzaj wyświetlanej statystyki zależy od typu przeszukiwacza.

Jeśli uprawnienia przypisane użytkownikowi umożliwiają administrowanie procesami dotyczącymi kolekcji, może on uruchamiać, zatrzymywać i wstrzymywać przeszukiwacz podczas wyświetlania szczegółów na temat jego działania. Jeśli dla przeszukiwacza został utworzony harmonogram, można także włączyć lub wyłączyć ten harmonogram przeszukiwania.

4. Jeśli przeszukiwacz jest zatrzymany lub wstrzymany i trzeba rozpocząć sesję przeszukiwacza, kliknij ikonę  **Uruchom** lub **Wznów**.

W przypadku przeszukiwaczy sieci WWW:



Jeśli przeszukiwacz był zatrzymany, rozpoczyna na nowo pracę i przeszukuje całą przestrzeń przeszukiwania. Jeśli przeszukiwacz był wstrzymany, wznawia przeszukiwanie od początku elementu, na którym został wstrzymany.

Aby wymusić natychmiastowe rozpoczęcie przez przeszukiwacz pełnego przeszukiwania, kliknij ikonę **Szczegóły**, a następnie ikonę **Uruchom ponownie pełne przeszukiwanie**. Przeszukiwacz rozpocznie przeszukiwanie całej przestrzeni przeszukiwania, w tym stron, które nie uległy zmianie od czasu ich ostatniego przeszukiwania. Ponowne przeszukiwanie wszystkich dokumentów może być potrzebne na przykład wtedy, gdy zmianie uległy reguły analizowania dokumentów i trzeba zastosować te reguły do dokumentów, które zostały poindeksowane wcześniej.

W przypadku przeszukiwaczy wszystkich innych typów:

Jeśli przeszukiwacz był zatrzymany, rozpocznie przeszukiwanie w zaplanowanym dniu i o zaplanowanej godzinie. Podczas pierwszego przeszukiwania źródła danych przez przeszukiwacz zostaje wykonane pełne przeszukiwanie. Gdy zaplanowane przeszukiwanie jest powtarzane, przeszukiwacz przeszukuje albo wszystkie aktualizacje źródła danych (dodane, usunięte i zmodyfikowane dokumenty), albo tylko dodane i zmodyfikowane dokumenty. Typ przeszukiwania można skonfigurować w harmonogramie przeszukiwacza.

Jeśli przeszukiwacz nie figuruje w harmonogramie lub jeśli zachodzi potrzeba wcześniejszego uruchomienia przeszukiwacza, kliknij ikonę **Szczegóły**. Następnie w obszarze szczegółów przestrzeni przeszukiwania kliknij ikonę typu przeszukiwania, które chcesz uruchomić: pełne przeszukiwanie, wszystkie aktualizacje lub tylko nowe i zmodyfikowane dokumenty. Należy kliknąć odpowiednią ikonę uruchamiającą dla każdego źródła danych, które ma zostać przeszukane (takiego jak serwer, baza danych lub podfolder).

5. Jeśli przeszukiwacz jest uruchomiony i trzeba go zatrzymać, kliknij ikonę  **Zatrzymaj** lub  **Wstrzymaj**. Przeszukiwacz zatrzymuje przeszukiwanie do momentu ponownego uruchomienia lub wznowienia jego działania.

Jeśli przeszukiwacz był wstrzymany, wznowia przeszukiwanie od początku elementu, na którym został wstrzymany. Na przykład przeszukiwacz DB2 wznowia przeszukiwanie od pierwszego wiersza w tabeli, która była przeszukiwana w momencie wstrzymania jego pracy.

Wyświetlanie szczegółów na temat działalności przeszukiwacza sieci WWW


Zapoznając się ze szczegółami na temat działań przeszukiwacza sieci WWW, można ocenić jego ogólną wydajność i odpowiednio dopasować właściwości przeszukiwacza oraz definicje przestrzeni przeszukiwania.

Zagadnienia wstępne


Do monitorowania działań przeszukiwacza uprawnieni są wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego. Aby uruchamiać lub zatrzymywać przeszukiwacz, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora danej kolekcji.

Procedura

Aby wyświetlić szczegółowe informacje na temat przeszukiwacza sieci WWW, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję zawierającą Przeszukiwacz sieci WWW, który chcesz monitorować, i kliknij ikonę  **Monitoruj**.
2. Otwórz stronę Przeszukiwanie.

Wskazówka: W przypadku edycji kolekcji, jeśli strona Przeszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby zmienić widok na umożliwiający monitorowanie przeszukiwaczy.

3. Jeśli przeszukiwacz sieci WWW przeznaczony do monitorowania jest uruchomiony lub zatrzymany, kliknij ikonę  **Szczegóły**.

4. Na stronie szczegółowych informacji o przeszukiwaczu sieci WWW przejrzyj lub wybierz poniższe opcje, aby wyświetlić szczegółową statystykę o aktualnej i dotychczasowej aktywności przeszukiwacza.
 - Kliknij opcję **Szczegóły wątku**, aby zobacz, ile obecnie wątków aktywnie przeszukuje serwisy WWW, a ile znajduje się w stanie nieaktywnym.
 - Kliknij opcję **Aktywne serwisy**, aby wyświetlić informacje o serwisach WWW, które przeszukiwacz aktualnie przeszukuje.
 - Kliknij opcję **Ostatnio przeszukane adresy URL**. Powoduje to wyświetlenie informacji o tym, co przeszukiwacz ostatnio przeszukał. Jeśli elementy listy nie ulegają zmianom przy odświeżaniu widoku, oznacza to, że przeszukiwanie nie jest aktualnie prowadzone.
 - Kliknij opcję **Historia przeszukiwacza**, aby wyświetlić raporty o dotychczasowej aktywności przeszukiwacza.
 - W obszarze **Status URL** wpisz adres URL, o którym mają zostać wyświetlone informacje.
 - a. Kliknij ikonę **Szczegóły adresu URL**, aby wyświetlić informacje o statusie adresu URL. Wyświetlenie szczegółów adresu URL jest możliwe tylko w przypadku adresów uprzednio objętych przeszukiwaniem.
 - b. Kliknij ikonę **Szczegóły serwisu**, aby określić informacje na temat serwisu WWW, do którego adres URL należy i które mają się znaleźć w raporcie. Szczegóły na temat serwisu można wyświetlić bez względu na to, czy serwis był już wcześniej przeszukiwany.

Na przykład opcja ta umożliwia sprawdzenie, czy wybrany adres URL znajduje się w przestrzeni przeszukiwania, czy został przeszukany czy tylko wykryty, kiedy powinien zostać przeszukany ponownie oraz kiedy przeprowadzono ostatnią próbę przeszukania danego serwisu WWW. Istnieje także możliwość wyświetlenia zawartości pliku robots.txt danego serwisu, co może pomóc w wyjaśnieniu, dlaczego serwis nie jest objęty przeszukiwaniem.

Szczegółowe informacje o wątkach przeszukiwacza sieci WWW

Przeszukiwacz sieci WWW można monitorować w celu sprawdzenia liczby wątków aktywnie przeszukujących serwisy WWW oraz liczby wątków, które pozostają nieaktywne.

Przy przeglądaniu szczegółów dotyczących przeszukiwacza sieci WWW podczas monitorowania kolekcji możliwe jest wyświetlenie statusu wątków przeszukiwacza. Najbardziej prawdopodobne statusy to:

Oczekiwanie

Wskazuje, że wątek nie ma adresu URL do przeszukania. Taka sytuacja może wystąpić, gdy wątek skończył przeszukiwanie, a przeszukiwacz nie może wystarczająco szybko znaleźć więcej adresów URL do przeszukania. Na przykład, jeśli właściwość przeszukiwacza decydująca o tym, jak długo przeszukiwacz musi odczekać, zanim będzie mógł pobrać inną stronę z tego samego serwisu, ma zbyt dużą wartość, może to uniemożliwić wystarczająco szybkie dostarczanie adresów URL.

Pobieranie

Wskazuje, że wątek pobiera stronę z serwisu WWW.

Zakończone

Wskazuje, że wątek przesyła do przeszukiwacza przeszukane strony, ale nie jest jeszcze gotowy do przeszukania następnego adresu URL.

Zawieszone

Wskazuje, że przeszukiwacz został wstrzymany.

W idealnej sytuacji wszystkie wątki pobierają strony przez cały czas. Jeśli jednak wątki często są w stanie zakończenia, mogą wystąpić problemy z przepustowością bazy danych.

Jeśli wątki często znajdują się w stanie oczekiwania, należy sprawdzić, jaka wartość została określona we właściwościach przeszukiwacza w polu **Maksymalna liczba aktywnych hostów**. Jeśli wartość ta jest mała, w przestrzeni przeszukiwania może nie być wystarczającej liczby serwisów do utrzymywania wszystkich wątków w stanie zajętości lub może brakować adresów URL dostępnych do przeszukania. Do sytuacji mogących powodować obniżenie aktywności należą błędy wyszukiwania na serwerze DNS oraz błędy wyszukiwania przez roboty.

Aktywne serwisy przeszukiwacza sieci WWW

Istnieje możliwość monitorowania przeszukiwacza sieci WWW w celu uzyskania informacji o serwisach WWW, aktywnie przeszukiwanych przez przeszukiwacz.

Przy przeglądaniu szczegółów dotyczących przeszukiwacza sieci WWW podczas monitorowania kolekcji możliwe jest wyświetlenie danych statystycznych dotyczących aktywnych serwisów. Dane statystyczne przedstawiają:

- liczbę adresów URL, które przeszukiwacz pobiera ze swojej wewnętrznej bazy danych do pamięci do przeszukiwania w danej chwili;
- liczbę adresów URL, które przeszukiwacz próbował przeszukać do tej pory;
- ile czasu pozostało do dezaktywowania serwisu i usunięcia go z pamięci dla tej iteracji przeszukiwacza;
- jak długo serwis znajduje się już w pamięci.

Te informacje zmieniają się co chwila, w miarę jak przeszukiwacz kolejno realizuje skonfigurowane dla niego reguły przeszukiwania. W idealnej sytuacji liczba aktywowanych adresów URL jest bliska wartości skonfigurowanej w polu **Maksymalna liczba aktywnych hostów** we właściwościach pamięci przeszukiwacza.

Jeśli liczba aktywowanych adresów URL jest bliska zeru, przeszukiwacz nie znajduje właściwych adresów URL. Warunki, które mogą powodować tak niską aktywność, obejmują niepowodzenia wyszukiwania na serwerze DNS, zagadnienia związane z połączeniem z siecią, błędy bazy danych oraz problemy związane z definicją przestrzeni przeszukiwania. Na przykład:

- Jeśli od dłuższego czasu w pamięci znajduje się wiele serwisów, a przeszukano tylko kilka adresów URL, należy sprawdzić, czy nie występują problemy z nawiązaniem połączenia z siecią.
- Jeśli na liście jest niewystarczająca liczba serwisów, należy sprawdzić, czy nie występują problemy związane z definicją przestrzeni przeszukiwania lub z wyszukiwaniem DNS.
- Jeśli przeszukiwanie serwisów odbywa się z należyłą szybkością, a mimo to serwisy są usuwane z pamięci z wieloma nieprzeszukanymi adresami URL, należy zmodyfikować właściwości pamięci przeszukiwacza i dopasować odpowiednio wartość limitu czasu w polu **Czas, przez jaki każdy host może pozostać aktywny**, aby serwisy mogły dłużej przebywać w pamięci.

Szybkość przeszukiwania przeszukiwacza sieci WWW

Działanie przeszukiwacza sieci WWW można monitorować, aby uzyskać informacje dotyczące szybkości, z jaką przeszukiwacz pobiera strony z serwisów WWW.

Przy przeglądaniu szczegółów dotyczących przeszukiwacza sieci WWW podczas monitorowania kolekcji możliwe jest wyświetlenie statystyki dotyczącej szybkości, z jaką

przeszukiwacz przeszukuje dane (szybkość przeszukiwania). Można również wyświetlić statystykę dotyczącą liczby adresów URL przeszukanych przez wyszukiwacz WWW od momentu rozpoczęcia bieżącej sesji.

Szybkość przeszukiwania to liczba stron przeszukiwanych w ciągu sekundy. Ta liczba jest skorelowana z różnymi właściwościami wyszukiwacza sieci WWW, które użytkownik może konfigurować:

- liczba wątków wyszukiwacza
- liczba aktywnych serwisów
- czas, który wyszukiwacz musi odczekać przed pobraniem innej strony z tego samego serwera WWW

Jeśli wyszukiwacz ma jeden aktywny serwis na wątek wyszukiwacza i musi odczekać dwie sekundy, zanim będzie mógł pobrać inną stronę z tego samego serwera WWW, to wyszukiwacz nie może przeszukiwać szybciej niż jedna strona na wątek na dwie sekundy. Na przykład, jeśli wyszukiwacz używa domyślnej liczby wątków (200), to może przeszukiwać z szybkością 100 stron na sekundę dla 200 wątków.

Jeśli jest dwa razy więcej aktywnych serwisów niż wątków wyszukiwacza, a wyszukiwacz musi odczekać dwie sekundy, zanim będzie mógł pobrać inną stronę z tego samego serwera WWW, to wyszukiwacz może uzyskać jedną stronę na wątek w ciągu sekundy. Jednak czynnikami ograniczającymi są wówczas szybkość pobierania danych w sieci oraz przepustowość bazy danych. Wysoka wydajność wyszukiwacza wskazywana jest wtedy, gdy szybkość przeszukiwania jest równa liczbie wątków wyszukiwacza, aktywnych serwisów oraz czasowi oczekiwania wyszukiwacza.

Innym czynnikiem, który można przejrzeć podczas monitorowania wydajności wyszukiwacza sieci WWW, jest liczba adresów URL, które wyszukiwacz przeszukał od momentu uruchomienia bieżącej sesji wyszukiwacza. Podzielenie tej liczby przez całkowity czas działania wyszukiwacza daje wynik w postaci średniej długookresowej przepustowości. Jeśli ta liczba się nie zwiększa, wyszukiwacz kończy działanie lub nie może go kontynuować. Na przykład błędy związane z nawiązaniem połączenia z siecią, błędy bazy danych i niepowodzenia wyszukiwania na serwerze DNS mogą uniemożliwić kontynuowanie pracy wyszukiwacza.

Tworzenie raportów wyszukiwacza sieci WWW

Zapoznając się z raportami dotyczącymi dotychczasowych działań wyszukiwacza sieci WWW, można ocenić jego ogólną wydajność i odpowiednio dopasować właściwości wyszukiwacza oraz definicje przestrzeni wyszukiwania.

Zagadnienia wstępne

Jeśli uprawnienia przyznane użytkownikowi ograniczone są do monitorowania kolekcji, ma on możliwość przeglądania statystyki wyszukiwacza lub tworzenia raportów o działaniu wyszukiwacza, ale nie może wpływać na jego działanie (uruchamiać lub zatrzymywać wyszukiwacz).

Informacje o tym zadaniu

Różne typy raportów zapewniają dostęp do informacji o różnych aspektach działania wyszukiwacza sieci WWW. W przypadku niektórych typów raportów informacje są zwracane, jak tylko zostaną one pobrane z wewnętrznej bazy danych wyszukiwacza. Tworzenie raportu o serwisie i raportów o kodach statusu HTTP wymaga pewnego czasu. Podczas tworzenia raportów tego typu można podać adres e-mail, na który ma zostać


przesłany raport; wówczas nie trzeba czekać, aż wyniki zostaną zwrócone na konsolę administracyjną wyszukiwania korporacyjnego.


Aby uzyskać informacje o właściwej interpretacji danych statystycznych zawartych w raportach, można użyć przycisku **Pomoc** w trakcie monitorowania przeszukiwacza sieci WWW i tworzenia raportu.

Procedura

Aby utworzyć raporty o przeszukiwaczu sieci WWW, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję zawierającą Przeszukiwacz sieci WWW, który chcesz monitorować, i kliknij ikonę  **Monitoruj**.
2. Otwórz stronę Przeszukiwanie.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Przeszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby zmienić widok na umożliwiający monitorowanie przeszukiwaczy.

3. Jeśli przeszukiwacz sieci WWW, którego mają dotyczyć tworzone raporty, jest uruchomiony lub zatrzymany, kliknij ikonę  **Szczegóły**.
4. Na stronie szczegółowych informacji o przeszukiwaczu sieci WWW wybierz opcję określającą typ raportu, który ma być utworzony.
 - W obszarze **Podsumowanie statusu przeszukiwacza** kliknij opcję **Historia przeszukiwacza**, aby wygenerować raporty na temat przeszukiwacza i wszystkich wykrytych lub przeszukanych przez niego serwisów.
 - W obszarze **Status URL** określ adres URL serwisu, dla którego ma zostać utworzony raport, a następnie kliknij ikonę **Szczegóły serwisu**.
5. W przypadku obu raportów (historii przeszukiwacza i o serwisie) można zaznaczyć pola wyboru odpowiadające charakterystynom, które powinny być ujęte w raporcie, po czym kliknąć **Wyświetl raport**.

Dla tych typów danych statystycznych przeszukiwacz zwraca raport na konsolę administracyjną natychmiast po pobraniu informacji z wewnętrznej bazy danych.

6. W przypadku tworzenia raportu o historii przeszukiwacza można określić opcje dotyczące raportu o serwisie, a następnie kliknąć przycisk **Uruchom raport**.

Raport zawierający wybrane dane statystyczne zostaje wygenerowany i zapisany w pliku o wskazanej nazwie (nazwa pliku musi być bezwzględna). Zakończenie generowania raportu może być sygnalizowane wysłaniem wiadomości e-mail.

7. W przypadku tworzeniu raportu o historii przeszukiwacza można określić opcje dotyczące raportu o kodach statusu HTTP, a następnie kliknąć przycisk **Uruchom raport**.

Raport zawiera informacje na temat liczby kodów statusu HTTP rozprawdzonych dla każdego serwisu. Raport zostaje zapisany w pliku o podanej nazwie (nazwa pliku musi być bezwzględna). Zakończenie generowania raportu może być sygnalizowane wysłaniem wiadomości e-mail.

Za pomocą tego raportu można się przekonać, które serwisy zwracają dużą liczbę kodów statusu typu 4xx (wskazujących na problem ze znalezieniem strony), typu 5xx (wskazujących na problem z serwerem), typu 6xx (dotyczących problemów z połączeniem) itp.

Raport ten jest najbardziej użyteczny w przypadku przeszukiwaczy, które przez pewien czas są aktywne (na przykład przeszukiwacz działający całymi tygodniami). W ten sposób można zidentyfikować zaginione serwisy, serwisy nowo powstałe, serwisy o dużej liczbie adresów URL (co może wskazywać na nadmiarowe przeszukiwanie bazy danych Lotus Notes) oraz serwisy z rekurencyjnym systemem plików obsługiwane przez serwer HTTP.

Jeśli serwisy generujące dużą liczbę kodów statusu HTTP nie są uwzględniane w indeksie, można poprawić wydajność przeszukiwacza, usuwając te serwisy z przestrzeni przeszukiwania.

Kody statusu HTTP zwracane do przeszukiwacza sieci WWW

Podczas monitorowania przeszukiwacza WWW można wyświetlać informacje o kodach statusu HTTP, które przeszukiwacz otrzymuje z przeszukiwanych stron.

Podsumowanie tabeli

Podczas monitorowania historii przeszukiwacza WWW lub statusu określonego adresu URL można wyświetlać informacje o kodach statusu HTTP zwróconych do przeszukiwacza. Jest to przydatne do zarządzania przestrzenią przeszukiwania oraz optymalizowania wydajności przeszukiwacza. Na przykład, jeśli przeszukiwacz odbiera dużą liczbę kodów statusu HTTP dla pewnego adresu URL, a kody te wskazują na niemożliwość przeszukiwania stron w odpowiedniej lokalizacji, można poprawić wydajność systemu, usuwając ten adres URL z przestrzeni przeszukiwania.

W poniższej tabeli wymieniono kody statusu HTTP wraz ze sposobem ich interpretacji przez przeszukiwacz. Wartości od 100 do 505 są standardowymi kodami statusu HTTP (więcej informacji znajduje się pod adresem <http://www.w3.org/Protocols/rfc2616/rfc2616.html>). Pozostałe kody HTTP są właściwe tylko dla systemu wyszukiwania korporacyjnego i przeszukiwacza WWW.

Tabela 8. Kody statusu HTTP przeszukiwacza WWW

Kod	Opis	Kod	Opis	Kod	Opis	Kod	Opis
NULL	Nie przeszukany	400	Nieprawidłowe żądanie	500	Błąd wewnętrzny serwera	693	Niepowodzenie instrukcji Select (URLFetcher)
100	Kontynuacja	401	Brak uprawnień	501	Nie zaimplementowano	694	Błąd zapisu (URLFetcher)
101	Przełączanie protokołów	402	Wymagana opłata	502	Nieprawidłowy gateway	695	Niekompletny nagłówek bloku (URLFetcher)
200	Powodzenie	403	Zabronione	503	Usługa niedostępna	699	Nieoczekiwany błąd (URLFetcher)
201	Utworzone	404	Nie znaleziono	504	Przekroczenie limitu czasu gatewaya	700	Błąd analizatora składni (brak zakończenia nagłówka)
202	Zaakceptowane	405	Niedozwolona metoda	505	Nieobsługiwana wersja HTTP	710	Błąd analizatora składni (nagłówek)
203	Niemiarodajna informacja	406	Nie akceptowalne	611	Błąd odczytu	720	Błąd analizatora składni (brak kodu HTTP)
204	Brak treści	407	Wymagane uwierzytelnienie serwera proxy	612	Błąd połączenia	730	Błąd analizatora składni (treść)
205	Zerowanie treści	408	Przekroczenie limitu czasu żądania	613	Przekroczenie limitu czasu odczytu	740 lub 4044	Wykluczone przez plik robots.txt

Tabela 8. Kody statusu HTTP przeszukiwacza WWW (kontynuacja)

Kod	Opis	Kod	Opis	Kod	Opis	Kod	Opis
206	Częściowa treść	409	Konflikt	614	Uzgadnianie SSL zakończyło się niepowodzeniem	741	Roboty tymczasowo niedostępne
300	Wielokrotne wybory	410	Niedostępne	615	Inny błąd odczytu	760	Wykluczone w definicji przestrzeni przeszukiwania
301	Przeniesione na stałe	411	Wymagana długość	616	Anomalia FBA	761	Niedozwolone w lokalnej przestrzeni wyszukiwania, dozwolone w globalnej
302	Znalezione	412	Warunek wstępny nie został spełniony	617	Błąd kodowania	770	Nieprawidłowy protokół lub niestandardowy port systemu
303	Patrz inny	413	Zbyt duża jednostka żądania	618	Przekierowanie bez docelowego URL	780	Wykluczone przez wykluczenia typów plików
304	Nie zmodyfikowane	414	Identyfikator URI w żądaniu jest za długi	680	Niepowodzenie wyszukiwania na serwerze DNS	786	Niepoprawny adres URL
305	Użyj serwera proxy	415	Nieobsługiwany typ nośnika	690	Zniekształcony adres URL	2004	Brak znacznika META indeksu
306	(Nieużywane)	416	Zakres w żądaniu nie spełnia wymagań	691	Utrata połączenia (URLFetcher)	3020	Przekierowanie programowe
307	Tymczasowe przekierowanie	417	Oczekiwania nie zostały spełnione	692	Przekroczenie limitu czasu zapisu (URLFetcher)		

Uwagi do tabeli

Kody statusu 4xx

Kod 400 (błędne żądanie) występuje rzadko. Zgodnie ze standardem kodów statusu HTTP kody 4xx mają sygnalizować błąd po stronie klienta (przeszukiwacza). Jednak zazwyczaj problem dotyczy serwera lub adresu URL, który przeszukiwacz otrzymał jako odsyłacz. Na przykład niektóre serwery WWW nie tolerują adresów URL przechodzących z katalogu głównego serwisu (takich jak `http://xyz.ibm.com/../../sales`). Inne serwery WWW nie mają problemu z przechodzeniem do katalogu macierzystego i ignorują operator katalogu macierzystego (`..`), gdy przeszukiwacz już jest w katalogu głównym.

Niektóre serwery traktują żądanie dostępu do katalogu głównego serwisu jako błąd, a niektóre przestarzałe odsyłacze mogą żądać operacji, które nie są już rozpoznawane ani implementowane. Gdy serwer aplikacji otrzyma żądanie strony, która już nie jest udostępniana, zgłasza wyjątek, w reakcji na który serwer WWW zwraca kod statusu 400, ponieważ żądanie zostaje tym samym uznane za nieprawidłowe.

- 615** Wskazuje, że serwer przeszukiwacza pobierający dane z serwisów WWW napotkał nieoczekiwany wyjątek. Duża liczba kodów tego typu może wskazywać na problem z przeszukiwaczem.

Kody statusu 61x

Kody statusu 61x z wyjątkiem kodu 615 wskazują na problemy spotykane podczas przeszukiwania, na przykład przekroczenie limitu czasu. Następujące kody statusu mogą wymagać podjęcia działań korygujących:

611, 612 i 613

Przyczyną tych kodów może być powolna praca serwerów lub słaba wydajność sieci.

611 Wskazuje na błąd podczas pobierania dokumentu przez przeszukiwacz.

612 Wskazuje na błąd w czasie, gdy przeszukiwacz próbował nawiązać połączenie z serwerem WWW.

613 Wskazuje, że błąd wystąpił podczas pobierania dokumentu przez przeszukiwacz.

- 614** Wskazuje, że przeszukiwacz nie może przeszukiwać bezpiecznych serwisów (HTTPS). Jeśli użytkownik jest przekonany, że te serwisy powinny być dostępne, należy sprawdzić prawidłowość instalacji certyfikatów na serwerze przeszukiwacza i na docelowym serwerze WWW. Na przykład, jeśli serwis jest certyfikowany przez autoryzowane ośrodki certyfikacji (CA), można dodać nowe autoryzowane ośrodki certyfikacji do zaufanych certyfikatów używanych przez przeszukiwacz.

Można również sprawdzić, w jaki sposób samopodpisane certyfikaty są skonfigurowane w przeszukiwanych serwisach. Przeszukiwacz jest skonfigurowany do akceptowania samopodpisanych certyfikatów. Niektóre serwisy tworzą samopodpisany certyfikat dla głównego adresu URL (na przykład <http://sales.ibm.com/>), a następnie próbują go używać w poddomenach (na przykład <http://internal.sales.ibm.com/>). Przeszukiwacz nie może akceptować używanych w ten sposób certyfikatów. Przeszukiwacz akceptuje samopodpisane certyfikaty tylko w nazwie domeny będącej podmiotem certyfikatu (sales.ibm.com), a osoba podpisująca certyfikat sprawdza zgodność nazwy domeny strony, do której zażądano dostępu.

- 616** Wskazuje, że mimo ponownego uwierzytelnienia na pobieranej stronie nadal jest wyświetlany formularz logowania (FBA).

Jeśli informacje zawarte w pliku konfiguracyjnym FBA (formularz logowania oraz dane uwierzytelniające, takie jak nazwa i hasło użytkownika) nie wystarczają do uwierzytelnienia przeszukiwacza, wszystkim stronom, do których dostęp wymaga uwierzytelnienia przez formularz, przypisywany jest kod statusu 616. Administrator powinien sprawdzić, dlaczego konfiguracja FBA nie działa w sposób należyty.

- 617** Wskazuje, że nie można utworzyć łańcucha w oparciu o bajtową treść dokumentu, ponieważ łańcuch kodowania (zestaw znaków) jest nieprawidłowy lub dokument zawiera niepoprawne bajty.

- 618** Wskazuje na nieprawidłowy adres URL przekierowania w przypadku, gdy przeszukiwacz otrzymuje niżej wymienione kody statusu HTTP. Istnieje możliwość, że położenie nagłówka odpowiedzi HTTP nie jest poprawne.

301 Przeniesione na stałe
302 Znalezione

- 680** Wskazuje, że przeszukiwacz nie mógł uzyskać w przestrzeni przeszukiwania

adresów IP hostów, być może z powodu problemów z dostępem do sieci. Ten typ błędu oznacza, że przeszukiwacz nie może przeszukiwać całych serwisów, a nie tylko niektórych adresów URL. Duża liczba kodów tego typu powoduje znaczne ograniczenie przepustowości.

Kody statusu 69x

Kody statusu od 690 do 699 nigdy nie są zapisywane w trwałej bazie danych przeszukiwacza. Kody te sygnalizują warunki niezwiązane z faktycznymi wynikami pobierania danych ze zdalnego hosta, lecz dotyczące tymczasowych stanów wewnętrznych przeszukiwacza, na przykład zamknięcie jednego komponentu w czasie, gdy inny komponent oczekuje na wynik jego działania lub wysłał do niego wynik własnego działania. Takie kody statusu można spotkać w niektórych dziennikach, ale nigdy w trwałej bazie danych, toteż nie powinny stanowić kryterium wyboru.

Kody statusu 7xx

Kody 7xx w większości podlegają regułom przestrzeni przeszukiwania:

710 - 730

Wskazują, że problemy uniemożliwiły przeszukiwaczowi pełne pobranie danych lub że przeszukiwacz napotkał w serwisie niepoprawne dane HTML. W razie napotkania dużej liczby kodów statusu tego typu należy zwrócić się o pomoc do przedstawiciela wsparcia klienta dla danego systemu wyszukiwania korporacyjnego.

740 lub 4044

Wskazują, że treść pliku nie może zostać poindeksowana, ponieważ dokument jest wykluczony przez ograniczenia zawarte w pliku robots.txt serwisu.

740 Wskazuje, że w indeksie mogą być uwzględnione zakotwiczone odsyłacze wskazujące na wykluczony dokument.

4044 Wskazuje, że zakotwiczone odsyłacze w dokumentach wskazujące na wykluczony dokument są wykluczone także z indeksu.

741 Wskazuje, że plik robots.txt serwisu pozwala na przeszukiwanie, ale pobranie danych zakończyło się niepowodzeniem. Jeśli w sposób powtarzalny nie można przeszukać tego adresu URL, adres ten jest usuwany z przestrzeni przeszukiwania. W razie napotkania dużej liczby kodów statusu tego typu należy sprawdzić, czy serwis docelowy nie jest tymczasowo lub trwale niedostępny. Jeśli serwis docelowy nie jest już dostępny, należy go usunąć z przestrzeni przeszukiwania.

Pozostałe kody 7xx są spotykane głównie w sytuacjach, gdy przestrzeń przeszukiwania została zmodyfikowana już po pewnym okresie pracy przeszukiwacza. Kody statusu tego rodzaju zazwyczaj nie sygnalizują problemów wymagających podjęcia działań korygujących.

3020 Wskazuje, że dokument o kodzie statusu 200 zawiera nagłówek lokalizacji odnoszący agenta użytkownika do innego adresu URL.

Monitorowanie analizatora składni

Monitorowanie analizatora składni służy do wyświetlania informacji o dokumentach, które są analizowane, przed dodaniem ich do indeksu wyszukiwania korporacyjnego. Istniejące opcje umożliwiają przeglądanie danych statystycznych i administrowanie działaniem analizatora składni.

Zagadnienia wstępne

Jeśli rola administracyjna użytkownika jest ograniczona do monitorowania kolekcji, może on wyświetlać status analizatora składni, ale nie może uruchamiać ani zatrzymywać analizatora.

Informacje o tym zadaniu


Podczas monitorowania szczegółowych informacji o analizatorze składni jest wyświetlany obraz stanu działania analizatora, który zawiera dane statystyczne dotyczące działania analizatora składni w określonym momencie. Informacje statystyczne obejmują: liczbę dokumentów przeszukanych, analizowanych lub oczekujących na analizę i liczbę przeanalizowanych dokumentów oczekujących na zapisanie w indeksie.


Kiedy analizator składni jest aktywny, komunikaty zawierają dodatkowe informacje o jego stanie. Na przykład:




- Analizator składni może być w trakcie analizowania dokumentów.
- Analizator składni może być bezczynny. Analizator oczekuje w stanie uśpienia, aż pojawi się więcej dokumentów do analizy. W przypadku błędów analizator składni oczekuje na restart. Analizator składni jest restartowany automatycznie, gdy nie są dostępne usługi analizatora (automatyczny restart następuje na przykład wtedy, gdy nie można nawiązać połączenia z usługą analizatora lub gdy wszystkie wirtualne maszyny języka Java analizatora są zajęte innymi kolekcjami).
- Analizator składni może być wstrzymany (na przykład wtedy, gdy oczekuje na zakończenie budowania indeksu).

Procedura

Aby monitorować analizator składni dla kolekcji, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**.
2. Otwórz stronę Analiza składni.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Analiza składni jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

3. Jeśli analizator składni jest uruchomiony i trzeba wyświetlić szczegółowe informacje o statusie analizy składni, kliknij ikonę  **Szczegóły**.
Jeśli rola administracyjna użytkownika zezwala na administrowanie procesami kolekcji, może on uruchamiać i zatrzymywać analizator składni podczas wyświetlania szczegółowych informacji o jego działaniu.
4. Jeśli analizator składni jest zatrzymany i trzeba go uruchomić, kliknij ikonę  **Uruchom**.
Jeśli najpierw jest tworzona kolekcja, analizator składni należy uruchamiać dopiero po tym, jak przeszukiwacz rozpocznie przeszukiwanie danych. W ten sposób można dostarczyć analizatorowi składni dane do analizy i klasyfikowania. Jeśli reguły analizy nie są zmieniane, analizator może działać bez przerwy.
5. Jeśli analizator składni jest uruchomiony i trzeba go zatrzymać, kliknij ikonę  **Zatrzymaj**.

Po zmianie reguł analizowania należy zatrzymać i zrestartować analizator składni. Na przykład jeśli zostanie zmieniona konfiguracja analizatora składni, zanim zmiany zaczną obowiązywać, należy go zatrzymać i zrestartować.

Monitorowanie aktywności indeksu dla kolekcji


Monitorowanie operacji na indeksie związanym z kolekcją pozwala wyświetlić postęp operacji budowania indeksu, włączyć lub wyłączyć harmonogram operacji na indeksie bądź uruchomić lub zatrzymać operacje indeksowania.

Zagadnienia wstępne






Do monitorowania operacji na indeksach uprawnieni są wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego. Aby uruchomić lub zatrzymać budowanie indeksu albo włączyć lub wyłączyć harmonogram indeksu, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora kolekcji dla tej kolekcji.

Procedura

Aby monitorować operacje na indeksie kolekcji, należy wykonać poniższe czynności:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**.
2. Otwórz stronę Indeks.

Wskazówka: Jeśli wykonywana jest edycja kolekcji i strona Indeksowanie jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

3. Jeśli indeks jest powiązany z harmonogramem, lecz operacja budowania nie ma być wykonana w zaplanowanym terminie, kliknij ikonę  **Wyłącz harmonogram**. Indeks nie będzie budowany aż do momentu włączenia harmonogramu lub uruchomienia procesu budowania indeksu.
4. Jeśli indeks jest powiązany z harmonogramem, ale harmonogram jego budowania jest wyłączony, kliknij ikonę  **Włącz harmonogram**. Indeks zostanie umieszczony w kolejce oczekiwania na budowanie zgodnie z terminem określonym w harmonogramie operacji na indeksie.
5. Jeśli indeks jest zatrzymany i trzeba go uruchomić, kliknij ikonę  **Uruchom**.
Zazwyczaj indeksowanie wykonywane jest regularnie zgodnie z harmonogramem. W przypadku zatrzymania indeksu w trakcie budowania lub wyłączenia harmonogramu operacji na indeksie kliknięcie ikony **Uruchom** pozwala wymusić rozpoczęcie operacji budowania indeksu.
6. Jeśli indeks jest uruchomiony i trzeba go zatrzymać, kliknij ikonę  **Zatrzymaj**.
Potrzeba zatrzymania budowy indeksu przyrostowego zachodzi na przykład wtedy, gdy trzeba wymusić budowanie indeksu głównego ze względu na zmianę typu klasyfikacji w kolekcji.
7. Jeśli podczas budowania indeksu wystąpiły błędy, kliknij ikonę  **Błąd**.
Zostanie wyświetlona strona Zawartość pliku dziennika, na której znajdują się dodatkowe informacje o błędach indeksowania. Na tej stronie można wybrać poszczególne komunikaty o błędach i wyświetlić szczegóły dotyczące danego problemu.

Monitorowanie kolejki indeksu wyszukiwania korporacyjnego

Użytkownik może wyświetlać status wszystkich operacji budowania indeksów w kolejce indeksów, zatrzymać operację budowania indeksu, a także usunąć indeks z kolejki.

Zagadnienia wstępne




Aby administrować kolejką indeksów, wymagana jest rola administratora wyszukiwania korporacyjnego.

Informacje o tym zadaniu

Jednocześnie może być budowanych wiele indeksów, ale w kolejce w danej chwili może znajdować się tylko jeden indeks dla danej kolekcji. Podczas konfigurowania w systemie opcji dotyczących indeksów podaje się liczbę indeksów, które mogą współużytkować kolejkę i współbieżnie indeksować zasoby.

Procedura

Aby monitorować kolejkę indeksów, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Wybierz stronę Indeksowanie.
Wyświetlona zostanie lista kolekcji, których indeksy znajdują się w kolejce indeksów. Dla każdego indeksu można wyświetlić typ budowanego indeksu (przyrostowy lub pełny), godzinę, o której indeks został umieszczony w kolejce, oraz godzinę, o której rozpoczęto budowanie indeksu (jeśli budowanie jest w toku).
3. Aby podjąć administrację wybranego indeksu, kliknij ikonę **Status**.
Na przykład można sprawdzić, ile czasu potrzeba jeszcze na zakończenie operacji budowania indeksu lub ile dokumentów znajduje się w indeksie, bądź wyłączyć harmonogram operacji na indeksie.
4. Aby zatrzymać budowany indeks, kliknij ikonę  **Zatrzymaj**.
Na przykład, jeśli zmieniono reguły kategorii, warto zatrzymać budowanie indeksu przyrostowego, aby wymusić w zamian rozpoczęcie budowy indeksu głównego.
Aby rozpocząć budowanie indeksu po jego zatrzymaniu, odczekaj na wprowadzenie indeksu do kolejki w następnym zaplanowanym terminie albo kliknij ikonę **Status**, aby monitorować indeks, a następnie ikonę  **Uruchom**, aby uruchomić budowanie indeksu.
5. Aby usunąć indeks z kolejki indeksów, kliknij ikonę  **Usuń**.

Monitorowanie serwerów wyszukiwania


Można wyświetlać szczegółowe informacje o statusie działań serwera wyszukiwania dla określonej kolekcji lub szczegółowe informacje o statusie serwerów wyszukiwania w całym systemie wyszukiwania korporacyjnego.

Zagadnienia wstępne


Wszyscy administratorzy wyszukiwania korporacyjnego mogą monitorować serwery wyszukiwania kolekcji, do których mają uprawnienia administracyjne. Aby monitorować wszystkie serwery wyszukiwania w systemie wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego.



Aby zatrzymać lub uruchomić serwer wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora kolekcji dla tej kolekcji.

Procedura

1. Aby monitorować serwery wyszukiwania dla pojedynczej kolekcji, należy wykonać poniższe czynności:
 - a. W widoku Kolekcje odzyskaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**.

- b. Otwórz stronę Wyszukiwanie.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Wyszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

2. Aby monitorować wszystkie serwery wyszukiwania w systemie wyszukiwania korporacyjnego, należy wykonać poniższe czynności:
 - a. Kliknij opcję **System**, aby otworzyć widok System.
 - b. Wybierz stronę Wyszukiwanie.
3. Jeśli serwer wyszukiwania jest zatrzymany i trzeba go uruchomić, kliknij ikonę  **Uruchom**.
4. Jeśli serwer wyszukiwania jest uruchomiony i trzeba go zatrzymać, kliknij ikonę  **Zatrzymaj**.

Jeśli zostanie włączona lub wyłączona pamięć podręczna wyszukiwania, zostanie zmieniona wielkość pamięci podręcznej lub gdy zostaną wprowadzone zmiany do szybkich odsyłaczy, należy zatrzymać i zrestartować serwery wyszukiwania, aby zmiany zostały uwzględnione.
5. Aby wyświetlić podsumowanie czasu przeznaczanego przez serwer wyszukiwania na przetwarzanie żądań wyszukiwania, kliknij opcję **Historia czasu odpowiedzi**.

Raport przedstawia średni czas (w milisekundach) przeznaczony przez serwer wyszukiwania na udzielanie odpowiedzi na żądania wyszukiwania w wybranym dniu. Średni czas odpowiedzi jest wskaźnikiem ogólnej sprawności systemu i zarazem jakości usług. Wydłużenie czasu odpowiedzi może wskazywać, że system jest mocno obciążony. Na przykład duża liczba objętych wyszukiwaniem kolekcji oraz rozmiar kolekcji mogą stanowić dla systemu nadmierne obciążenie.
6. Aby wyświetlić listę najczęściej wprowadzanych zapytań, kliknij opcję **Często zadawane zapytania**.

Raport zawiera słowa kluczowe użyte w 50 najczęściej wprowadzanych zapytaniach z wyszczególnieniem częstotliwości ich wprowadzania.

Przeglądając najczęściej zadawane zapytania, można zorientować się w potrzebach w zakresie tworzenia szybkich odsyłaczy. Szybkie odsyłacze pozwalają podnieść jakość wyszukiwania w przypadku wielu użytkowników. Dzięki temu dokumenty o dużym stopniu adekwatności zawsze będą zwracane w wynikach wyszukiwania.

Można również zamieszczać na stronach portalu korporacyjnego odsyłacze do zasobów odpowiadających tym zapytaniom. Na przykład, jeśli użytkownicy bardzo często wyszukują informacji na temat rozliczania wydatków, warto zamieścić odsyłacz do strony omawiającej procedury rozliczania wydatków na intranetowej stronie głównej.
7. Aby wyświetlić listę ostatnio zadawanych zapytań, kliknij opcję **Ostatnie zapytania**.

Raport przedstawia słowa kluczowe zawarte w 50 ostatnio wprowadzonych zapytaniach.

Przeglądając ostatnie zapytania, można rozpoznawać aktualnie panujące tendencje oraz nagle sytuacje występujące w organizacji. Na przykład można zaobserwować wzrost zainteresowania pewnym tematem. Taki wzrost zainteresowania może świadczyć o tym, że należy utworzyć szybki odsyłacz do stron poświęconych temu tematowi albo że informacje na ten temat należy udostępnić użytkownikom w inny sposób (na przykład zamieszczając bezpośredni odsyłacz w portalu korporacyjnym).

Zmiana sposobu obliczania statystyk zapytań

Można zmienić sposób, w jaki system oblicza liczbę popularnych zapytań oraz ostatnio przetworzonych zapytań.

Informacje o tym zadaniu



W przypadku monitorowania serwerów wyszukiwania można wybrać opcje wyświetlenia listy 50 najpopularniejszych zapytań oraz listy 50 ostatnio przetworzonych zapytań. Według domyślnej konfiguracji serwera wyszukiwania zapytania, które mają równoważny łańcuch zapytania, lecz różne ustawienia zakresu wyników, są zliczane jako niezależne zapytania. Dlatego licznik zapytań zostanie zwiększony na przykład wtedy, gdy użytkownik kliknie opcję wyświetlenia następnej strony wyników.

Sposób, w jaki system oblicza statystykę zapytań, można zmienić przez edycję pliku `runtime-generic.properties` serwera wyszukiwania. Jeśli parametr **distinctRecentQueryCheck** zostanie ustawiony na wartość `true`, jako zapytania niezależne system będzie zliczać tylko te zapytania, które zwracają początkową stronę wyników.

Aby dokonać edycji właściwości serwera wyszukiwania, należy zalogować się jako administrator wyszukiwania korporacyjnego. Aby zatrzymać lub uruchomić serwer wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora kolekcji dla tej kolekcji.

Procedura

Aby zmienić sposób, w jaki system oblicza statystykę zapytań:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition. W przypadku konfiguracji z wieloma serwerami należy zalogować się na serwerze wyszukiwania.
2. Dokonaj edycji następującego pliku, gdzie `id_kolekcji` jest identyfikatorem kolekcji, która ma zostać zmieniona, natomiast `id_węzła` służy do identyfikacji serwera wyszukiwania, na którym wprowadzane są zmiany:
`ES_NODE_ROOT/master_config/id_kolekcji.runtime.id_węzła/runtime-generic.properties`
3. Ustaw parametr **distinctRecentQueryCheck** na wartość `true`, a następnie zapisz i zamknij plik.
4. W przypadku konfiguracji z wieloma serwerami powtórz wcześniejsze kroki w odniesieniu do drugiego serwera wyszukiwania.
5. Jeśli chcesz zmienić zachowanie statystyki zapytań w odniesieniu do innych kolekcji, powtórz dotychczasowe kroki w odniesieniu do każdej kolekcji, którą chcesz zmienić.
6. Aby zmiany zostały uwzględnione, zaloguj się w konsoli administracyjnej wyszukiwania korporacyjnego i zrestartuj serwery wyszukiwania:
 - a. Kliknij opcję **System**, aby otworzyć widok System.
 - b. Wybierz stronę Wyszukiwanie.
 - c. W odniesieniu do każdego serwera wyszukiwania, który został zmieniony, kliknij przycisk  **Zatrzymaj**.
 - d. W odniesieniu do każdego serwera wyszukiwania, który został zmieniony, kliknij przycisk  **Uruchom**.

Monitorowanie nasłuchiwanie danych

Monitorowanie nasłuchiwanie danych służy do wyświetlania statusu i szczegółów działania klienta aplikacji nasłuchiwanie danych.



Zagadnienia wstępne

Aby monitorować nasłuchiwanie danych, wymagana jest rola administratora wyszukiwania korporacyjnego.

Ważne: Nasłuchiwanie danych nie będzie obsługiwane w przyszłych wersjach. Na potrzeby tworzenia aplikacji klienckich dla wyszukiwania korporacyjnego należy używać interfejsów API wyszukiwania i indeksowania (SI-API) zamiast interfejsów API nasłuchiwanie danych. Poniższe informacje są przeznaczone dla użytkowników, którzy utworzyli już wcześniej aplikacje nasłuchiwanie danych.

Procedura

Aby monitorować program nasłuchiwanie danych, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Na stronie Nasłuchiwanie danych zwróć uwagę na ikony statusu, aby sprawdzić, czy nasłuchiwanie danych jest aktywne czy zatrzymane.
3. Jeśli aplikacja nasłuchiwanie danych jest uruchomiona i trzeba wyświetlić szczegółowe informacje o statusie aktywności aplikacji klienckiej, kliknij ikonę  **Szczegóły**.
Ikony statusu na stronie Szczegóły nasłuchiwanie danych wskazują, czy nasłuchiwanie danych jest uruchomione czy zatrzymane. Na informacje statystyczne składają się: liczba żądań oczekujących na przetwarzanie, aktualny stan każdego wątku zajętego przez żądania aplikacji klienckiej oraz liczba aktywnych wątków znajdujących się w określonym stanie.
4. Jeśli zmieniono numer portu nasłuchiwanie danych lub parametr określający maksymalną liczbę dokumentów, jakie mogą być przechowywane w pamięci tymczasowej, kliknij ikonę  **Restartuj**.

Nasłuchiwanie danych jest uruchamiane podczas startu systemu wyszukiwania korporacyjnego. Restartowanie nasłuchiwanie danych jest potrzebne tylko w przypadku, gdy zmieniono jedną lub obie z tych opcji konfiguracyjnych.

Śledzenie dokumentu

Dokumenty mogą być usuwane z systemu na różnych etapach przetwarzania. Istnieje możliwość określenia opcji powiadamiania o usuwaniu dokumentów oraz o problemach, które spowodowały usunięcie.

Gdy analizator składni napotyka błąd uniemożliwiający analizę dokumentu, informacja o usunięciu dokumentu zostaje zarejestrowana w postaci odpowiedniego komunikatu wraz z kodem przyczyny. (Ten typ błędu nie powoduje usunięcia z indeksu starszych wersji danego dokumentu).

Dokumenty mogą być usuwane podczas etapów indeksowania, a informacje o tym także są rejestrowane. Na przykład identyfikatory URI i ich wzorce mogły zostać jawnie usunięte. Dokument mógł zostać przeszukany przez przeszukiwacz, który został później usunięty. Dokument źródłowy może już nie istnieć (z dokumentem związany jest ujemny kod HTTP) lub kod HTTP związany z dokumentem może być nieznany. Usunięcie jest też możliwe, gdy w dokumencie wymagającym analizy globalnej nie ma informacji o jego ocenie.

Jeśli wiadomo, że pewien dokument został przeszukany, ale nie znajduje się w indeksie, za pomocą konsoli administracyjnej wyszukiwania korporacyjnego można prześledzić drogę tego dokumentu przez system. Szczegółowe raporty mogą wskazać czas, miejsce i przyczynę usunięcia dokumentu. Na przykład raport może zawierać informację, że dokument został nieoczekiwanie usunięty podczas analizy globalnej lub że administrator usunął określony identyfikator URI z indeksu.

Zadania pokrewne

“Wyświetlanie szczegółów dotyczących identyfikatora URI” na stronie 302

Konfigurowanie plików dziennika do śledzenia dokumentów

Aby określić, kiedy, gdzie i dlaczego dany dokument został usunięty z systemu, można skonfigurować pliki dziennika w celu śledzenia informacji o usuwanych dokumentach.

Zagadnienia wstępne

Aby skonfigurować opcje śledzenia usuwanych dokumentów, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

Informacje o tym zadaniu

Aby uniknąć zajmowaniu zbyt dużej części dysku przez pliki dzienników, system prowadzi rotację tych plików, zaczynając nowy plik dziennika przy każdej zmianie daty. Jeśli plik dziennika osiągnie maksymalną dopuszczalną wielkość, system utworzy nowy plik dziennika, nawet jeśli data nie uległa jeszcze zmianie. Po osiągnięciu maksymalnej liczby plików dzienników najstarszy z plików zostanie usunięty, aby umożliwić utworzenie nowego.

Procedura

Aby skonfigurować pliki dziennika do śledzenia dokumentów:

1. Przejdź do edycji kolekcji, wybierz stronę Rejestrowanie, a następnie kliknij opcję **Konfiguruj śledzenie dokumentów**.
2. Na stronie Śledzenie dokumentów zaznacz pole wyboru dotyczące śledzenia dokumentów.
3. Podaj liczbę plików dziennika, które mają być używane do rejestrowania informacji na temat dokumentów usuwanych z systemu. Te pliki dziennika są współużytkowane przez wszystkie sesje, podczas których dokumenty mogą być usuwane.

Wyświetlanie raportów o usuniętych dokumentach

Użytkownik może wyświetlić szczegółowe informacje na temat dokumentów, które zostały usunięte z systemu wyszukiwania korporacyjnego. Informacje te są dostępne tylko pod warunkiem, że dla danej kolekcji włączono śledzenie dokumentów.

Zagadnienia wstępne

Przed wprowadzeniem żądania wyświetlenia raportu o usuniętych dokumentach lub wysłania raportu na wskazany adres e-mail należy się upewnić, że sesje, które mają być źródłami informacji, są aktywne. Na przykład, aby zapoznać się z informacjami na temat dokumentów, które zostały usunięte podczas analizowania lub indeksowania, należy się upewnić, że sesje analizy i indeksowania kolekcji zostały uruchomione.

Zanim możliwe będzie odebranie raportu, należy się upewnić, że dla wyszukiwania korporacyjnego został skonfigurowany serwer poczty. Informacje te należy podać podczas konfigurowania opcji poczty elektronicznej na stronie Rejestrowanie w widoku System.


Informacje o tym zadaniu


Gromadzenie informacji na temat usuwanych dokumentów jest procesem czasochłonnym. Można wybrać opcję umożliwiającą wyświetlenie informacji, a następnie poczekać na ich wyświetlenie. Metodą bardziej efektywną jest wysłanie raportu na podany przez użytkownika adres e-mail.


Gdy dokument został usunięty, raport zawiera datę i godzinę usunięcia, poziom istotności błędu, komponent i sesję, w których wystąpił błąd, oraz komunikat o błędzie.

Procedura

Aby wyświetlić szczegółowe informacje o usuniętych dokumentach:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Ogólne jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby wyświetlić widok umożliwiający monitorowanie kolekcji.

2. Kliknij ikonę  **Szczegóły URI**.
3. Na stronie Szczegóły dotyczące identyfikatorów URI wpisz identyfikator URI, którego mają dotyczyć wyświetlane informacje.
4. Zaznacz pola wyboru odpowiadające rodzajowi informacji, które mają być wyświetlone:

Dokumenty usunięte przez analizator składni

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty podczas analizowania, a jeśli tak, to jaka była tego przyczyna.

Dokumenty usunięte z indeksu

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty w czasie indeksowania lub analizowania, a jeśli tak, to jaka była tego przyczyna.

5. Określ, w jaki sposób chcesz wyświetlić raport:
 - Aby zacząć na wyświetlenie raportu, kliknij opcję **Wyświetl raport**.
 - Aby wysłać raport na podany adres e-mail w celu przejrzania go w późniejszym terminie, kliknij opcję **Wyślij raport**.

Na stronie Wysłanie szczegółowego raportu o identyfikatorach URI w polu **Adres poczty elektronicznej do wysłania powiadomienia** wpisz adres e-mail, na który ma zostać wysłany raport, a następnie kliknij przycisk **Wyślij raport**.

Zadania pokrewne

“Wyświetlanie szczegółów dotyczących identyfikatora URI” na stronie 302

Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 120

Wyświetlanie plików dzienników dotyczących usuniętych dokumentów


Użytkownik może wyświetlić zapisane w dziennikach informacje na temat dokumentów, które zostały usunięte z systemu wyszukiwania korporacyjnego. Informacje te są dostępne tylko pod warunkiem, że dla danej kolekcji włączono śledzenie dokumentów.

Informacje o tym zadaniu


Aby wyświetlić raport o usuniętym dokumencie, należy znać identyfikator URI tego dokumentu. Dziennik usuniętego pliku zawiera datę i godzinę usunięcia, poziom istotności błędu, komponent i sesję, których błąd dotyczył, a także szczegółowy komunikat o błędzie.

Procedura

Aby wyświetlić dzienniki usuniętych dokumentów:

1. W widoku Kolekcje odszukaj kolekcję, którą chcesz monitorować, i kliknij ikonę  **Monitoruj**.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Ogólne jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby wyświetlić widok umożliwiający monitorowanie kolekcji.


2. Kliknij ikonę  **Pliki dzienników usuniętych dokumentów**.
3. Na stronie Pliki dzienników usuniętych dokumentów wybierz dziennik, który ma zostać wyświetlony. Nazwa każdego pliku dziennika wskazuje, czy dokument został usunięty przez analizator składni (pd) czy podczas budowania indeksu (in), i zawiera datę utworzenia pliku. Jeśli tworzony jest więcej niż jeden plik dziennika tego samego typu o tej samej dacie, przyrostek liczbowy wskaże kolejność, w której dany plik został w danym dniu utworzony. Na przykład:

```
dropped_doc_in_20060525.log  
dropped_doc_pd_20060524.log (zawiera najnowsze wpisy o tej dacie)  
dropped_doc_pd_20060524.log.1  
dropped_doc_pd_20060524.log.2 (zawiera najstarsze wpisy o tej dacie)
```

4. Kliknij przycisk **Wyświetl dziennik**.

Dla każdego komunikatu wyświetlanego na stronie Zawartość pliku dziennika podawana jest data i godzina wygenerowania komunikatu, jego poziom istotności, nazwa sesji zgłaszającej komunikat, identyfikator komunikatu i treść błędu.

Za pomocą odpowiednich przycisków można przechodzić do pierwszej, ostatniej, poprzedniej i następnej strony pliku dziennika. Można też podać numer strony i przejść bezpośrednio do niej.

5. Aby przejrzeć szczegółowe informacje na temat danego komunikatu, kliknij przycisk  **Szczegóły**.

Na stronie Szczegóły komunikatu dziennika wyświetlana jest nazwa hosta serwera wyszukiwania korporacyjnego, na którym pojawił się komunikat, nazwa pliku, który spowodował błąd, nazwa funkcji i numer wiersza, w którym wystąpił błąd, identyfikator procesu i identyfikator wątku.

Klikając przyciski, można przechodzić do następnego i poprzedniego komunikatu w pliku dziennika.

Alerty i pliki dziennika

Użytkownik może wybrać typy komunikatów, które mają być zapisywane w dzienniku dla kolekcji i dla systemu, określić opcje dotyczące tworzenia i wyświetlania plików dzienników oraz opcje otrzymywania alertów oraz wiadomości e-mail dotyczących komunikatów.

Podczas wykonywania zwykłych operacji komponenty wyszukiwania korporacyjnego zapisują komunikaty dziennika we wspólnym pliku dziennika. Znajduje się on w katalogu `ES_NODE_ROOT/logs` na serwerze indeksowania. Wspólny plik dziennika można wyświetlić, korzystając z konsoli administracyjnej.

Po wystąpieniu problemu, takiego jak błąd komunikacji w sieci, komponent zapisuje odpowiednie komunikaty dziennika w katalogu `logs` na serwerze, na którym jest zainstalowany. Aby wyświetlić te lokalne pliki dzienników, należy użyć na komputerze lokalnym przeglądarki plików, takiej jak program narzędziowy `tail` w systemie UNIX. Do wyświetlenia tego typu plików dzienników nie można użyć konsoli administracyjnej.

Podczas konfigurowania plików dziennika użytkownik może wybrać typy komunikatów, które mają być rejestrowane (np. komunikaty o błędach lub ostrzeżenia), określić częstotliwość usuwania starych plików dzienników w celu zwolnienia miejsca dla nowych, określić maksymalną wielkość plików dzienników, a także wybrać język, w jakim mają być zapisywane komunikaty. Istnieje również możliwość określenia opcji dotyczących otrzymywania wiadomości e-mail po wystąpieniu określonych zdarzeń lub po zapisaniu w dzienniku określonych komunikatów lub komunikatów określonego typu.

Podczas monitorowania plików dzienników użytkownik może wybrać pliki dzienników, które mają zostać otwarte. Treść pliku dziennika można filtrować, aby wyświetlone zostały tylko komunikaty o określonym poziomie istotności (na przykład tylko komunikaty o błędach) lub komunikaty, które zostały wygenerowane przez konkretną sesję wyszukiwania korporacyjnego. Podczas wyświetlania pliku dziennika możliwe jest wyświetlenie szczegółowych informacji o poszczególnych komunikatach. Na przykład użytkownik może wyświetlić nazwę funkcji, która wygenerowała określony komunikat, a także, w razie potrzeby, inne informacje ułatwiające podjęcie odpowiednich działań naprawczych.

Pojęcia pokrewne



Komunikaty dla wyszukiwania korporacyjnego

Alerty

Wyszukiwanie korporacyjne można skonfigurować tak, aby za każdym razem, gdy zostanie wykryte określone zdarzenie, w pliku dziennika zapisywane były odpowiednie komunikaty.

W wyzwalanych przez zdarzenia komunikatach nazywanych alertami dostępne są informacje o wystąpieniu interesujących dla użytkownika warunków, takich jak użycie przez określony zasób zbyt dużego obszaru pamięci. Podczas konfigurowania alertów dla wyszukiwania korporacyjnego należy określić warunki, które mają być monitorowane przez system. Za każdym razem, gdy wystąpi określony warunek, system automatycznie zapisze odpowiedni komunikat w pliku dziennika.

Aby otrzymywać bezpośrednio powiadomienia o wystąpieniu określonych warunków, należy wybrać odpowiednie opcje wysyłania wiadomości e-mail podczas rejestrowania monitorowanych komunikatów.

Istnieje możliwość skonfigurowania alertów związanych ze zdarzeniami na poziomie kolekcji i zdarzeniami występującymi na poziomie systemu. Dla zdarzeń na poziomie kolekcji system może wykonywać następujące operacje:

- monitorowanie liczby dokumentów przeszukiwanych przez każdego z przeszukiwaczy i generowanie komunikatu alertu po osiągnięciu maksymalnej dozwolonej liczby dokumentów
- monitorowanie liczby dokumentów dodawanych do indeksu związanego z używanymi kolekcjami i generowanie komunikatu alertu po osiągnięciu maksymalnej dozwolonej liczby dokumentów
- wysyłanie użytkownikowi informacji o przekroczeniu zdefiniowanego limitu czasu podczas uzyskiwania odpowiedzi na zapytanie wyszukiwania

Dla zdarzeń na poziomie systemu można monitorować miejsce na dyskach na każdym serwerze wyszukiwania korporacyjnego i generować odpowiedni komunikat alertu, gdy ilość wolnego miejsca na dysku jest mała.

Konfiguracja alertów na poziomie kolekcji

Skonfigurowanie alertów gwarantuje, że w przypadku wystąpienia pewnych zdarzeń na poziomie kolekcji komunikaty są zapisywane w pliku dziennika. Możliwe jest również otrzymywanie wiadomości e-mail informujących o tym, że zdarzenia te zostały zarejestrowane.

Zagadnienia wstępne

Aby skonfigurować alerty dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

Procedura

Aby skonfigurować alerty na poziomie kolekcji:

1. Przejdź do edycji kolekcji, wybierz stronę Rejestrowanie, a następnie kliknij opcję **Konfiguruj alerty**.
2. Jeśli system ma monitorować liczbę dokumentów przeszukiwanych przez przeszukiwacza, należy wykonać poniższe czynności:
 - a. Zaznacz pole wyboru **Gdy liczba dokumentów przeszukanych przez dowolny przeszukiwacz osiągnie procent maksymalnej dozwolonej liczby**.
 - b. W polu **Procent** określ, w którym momencie komunikat ma być rejestrowany. Liczbę tę należy określić jako procent maksymalnej liczby dokumentów, jaką przeszukiwacz może przeszukać (parametr **Maksymalna liczba dokumentów do przeszukania** jest definiowany podczas konfigurowania właściwości przeszukiwacza). Wartość domyślna to 90 procent.

Ponieważ dla poszczególnych przeszukiwaczy można skonfigurować różne limity, dla każdego przeszukiwacza rejestrowane są osobne komunikaty. Na przykład, jeśli ustawiony jest domyślny próg alertów, umożliwienie przeszukiwaczowi DB2 przeszukania 2 000 000 dokumentów, a przeszukiwaczowi Notes - przeszukania 1 000 000 dokumentów spowoduje, że w dzienniku znajdzie się jeden komunikat w momencie, gdy przeszukiwacz DB2 przeszuka 1 800 000 dokumentów, natomiast drugi komunikat zostanie dodany po przeszukaniu 900 000 dokumentów przez przeszukiwacza Notes.
3. Jeśli system ma monitorować liczbę dokumentów dodawanych do indeksu, należy wykonać poniższe czynności:
 - a. Zaznacz pole wyboru **Gdy liczba dokumentów w kolekcji osiągnie określony procent szacowanej wielkości**.

- b. W polu **Procent** określ, w którym momencie komunikat ma być rejestrowany. Podana tu liczba oznacza procent szacowanej liczby dokumentów, które według przewidywań będą przechowywane w kolekcji. Wartością domyślną jest 85 procent.
Bieżąca szacowana wielkość kolekcji jest wyświetlana w polu **Limit**. Aby zmienić tę wartość, otwórz stronę Ogólne kolekcji, wybierz opcję konfigurowania opcji ogólnych i w polu **Szacowana liczba dokumentów** podaj nową wartość.
Ważne: Ten limit i szacowana liczba dokumentów skonfigurowana dla kolekcji służą tylko do monitorowania wzrostu kolekcji. Wartości te nie oznaczają bezwzględnego limitu wzrostu wielkości indeksu.
4. Jeśli system ma informować o tym, że czas wymagany na odpowiedź na żądania wyszukiwania przekracza dopuszczalny limit, należy wykonać poniższe czynności:
 - a. Zaznacz pole wyboru **Gdy czas odpowiedzi wyszukiwania przekroczy limit**.
 - b. W polu **Limit** wpisz wyrażony w sekundach czas, który można uznać za maksymalny czas odpowiedzi na żądanie wyszukiwania.
Po przekroczeniu tej wartości system zarejestruje komunikat informujący o tym zdarzeniu. Na przykład przy zachowaniu wartości domyślnej system utworzy komunikat dziennika, gdy średni czas odpowiedzi serwera wyszukiwania osiągnie poziom przynajmniej pięciu sekund.
Typowe czasy odpowiedzi nie przekraczają pół sekundy. Średnie czasy większe niż sekunda mogą wskazywać, że system operacyjny wymaga lepszego dostrojenia pod względem wydajności lub że istnieje problem związany z ustawieniami konfiguracyjnymi serwera wyszukiwania. Na przykład można zwiększyć ilość miejsca przeznaczonego na pamięć podręczną wyszukiwania.
5. Kliknij przycisk **OK**.

Aby otrzymywać wiadomości e-mail informujące o tym, że zdarzenia te zostały zarejestrowane, otwórz stronę **Dziennik** i kliknij opcję **Konfiguracja opcji pocztu elektronicznej dla komunikatów** w celu podania adresu e-mail. Identyfikatory komunikatów dla włączonych alertów zostaną automatycznie dodane do listy identyfikatorów komunikatów, dla których będą wysyłane wiadomości e-mail:

Aby otrzymywanie wiadomości e-mail było możliwe, należy pamiętać o skonfigurowaniu informacji o serwerze poczty. Aby to zrobić, administrator wyszukiwania korporacyjnego musi wybrać opcję **System** na pasku narzędzi, otworzyć stronę **Dziennik**, a następnie kliknąć opcję **Konfiguracja opcji pocztu elektronicznej dla komunikatów**.

Zadania pokrewne

“Odbieranie wiadomości pocztu elektronicznej dotyczących zarejestrowanych komunikatów” na stronie 328

Konfiguracja alertów na poziomie systemu


Skonfigurowanie alertów gwarantuje, że w przypadku wystąpienia pewnych zdarzeń na poziomie systemu komunikaty są zapisywane w pliku dziennika. Możliwe jest również otrzymywanie wiadomości e-mail informujących o tym, że zdarzenia te zostały zarejestrowane.

Zagadnienia wstępne

Alerty na poziomie systemu może konfigurować tylko administrator wyszukiwania korporacyjnego.

Procedura

Aby skonfigurować alerty na poziomie systemu:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Dziennik kliknij opcję **Konfiguruj alerty**.
4. Jeśli system ma monitorować ilość wolnej pamięci dostępnej na serwerach wyszukiwania korporacyjnego, zaznacz pole wyboru **Gdy ilość dostępnego miejsca w systemie plików osiągnie określony procent łącznej przestrzeni w tym systemie plików**.
5. W polu **Procent** określ, w którym momencie system ma powiadamiać o wyczerpywaniu się wolnej pamięci na serwerze. Wartość tę należy określić jako procent całkowitej pamięci systemu plików. Wartość domyślna to 80 procent.
Jeśli system wyszukiwania korporacyjnego jest skonfigurowany na wielu serwerach, dla każdego z nich tworzony jest osobny komunikat w dzienniku. Na przykład jeden komunikat może informować o małej ilości wolnej przestrzeni na serwerze przeszukiwacza, a inne o ograniczeniach dotyczących przestrzeni na serwerach wyszukiwania i indeksowania.
6. Kliknij przycisk **OK**.

Aby otrzymywać wiadomości e-mail informujące o tym, że zdarzenie to zostało zarejestrowane, otwórz stronę Dziennik i kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów** w celu podania adresu e-mail oraz informacji o serwerze poczty.

Zadania pokrewne

“Odbieranie wiadomości poczty elektronicznej dotyczących zarejestrowanych komunikatów” na stronie 328

Konfiguracja plików dziennika

Istnieje możliwość określenia typów komunikatów, które mają być rejestrowane, oraz opcji dotyczących tworzenia plików dzienników.

Zagadnienia wstępne

Aby skonfigurować pliki dzienników na poziomie kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji. Pliki dzienników na poziomie systemu może konfigurować tylko administrator wyszukiwania korporacyjnego.

Informacje o tym zadaniu



Aby uniknąć zajmowaniu zbyt dużej części dysku przez pliki dzienników, system prowadzi rotację tych plików, zaczynając nowy plik dziennika przy każdej zmianie daty. Jeśli plik dziennika osiągnie maksymalną dopuszczalną wielkość, system utworzy nowy plik dziennika, nawet jeśli data nie uległa jeszcze zmianie. Po osiągnięciu maksymalnej liczby plików dzienników najstarszy z plików zostaje usunięty, aby umożliwić utworzenie nowego.

Aby otrzymywać wiadomości e-mail na temat rejestrowanych komunikatów, należy najpierw określić informacje dotyczące sposobu dostarczania tych wiadomości. Następnie należy określić, na temat jakich komunikatów mają być wysyłane wiadomości e-mail.

Procedura

Aby skonfigurować pliki dzienników wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Aby skonfigurować opcje dotyczące tworzenia i rotacji plików dzienników na poziomie systemu:

- a. Kliknij opcję **System**, aby otworzyć widok System.
 - b. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
 - c. Na stronie Rejestrowanie kliknij opcję **Konfiguracja opcji pliku dziennika**. Zostanie wyświetlona strona Opcje pliku dziennika na poziomie systemu.
2. Aby skonfigurować opcje dotyczące tworzenia i rotacji plików dzienników na poziomie kolekcji:
 - a. W widoku Kolekcje odzyskaj kolekcję, dla której chcesz zdefiniować opcje, i kliknij opcję  **Edytuj**.
 - b. Na stronie Rejestrowanie kliknij opcję **Konfiguracja opcji pliku dziennika**. Zostanie wyświetlona strona Opcje pliku dziennika na poziomie kolekcji.
 3. W polu **Typ informacji do rejestrowania** wybierz typy komunikatów, które mają być rejestrowane:

Tylko komunikaty o błędach

Komunikaty o błędach informują o zaistnieniu niepożądanego zachowania, skutkiem czego kontynuacja procesu nie jest możliwa. Należy przedsięwziąć działania mające na celu rozwiązanie problemu.

Komunikaty o błędach i ostrzeżenia

Komunikat ostrzegawczy wskazuje na możliwość wystąpienia konfliktu lub niespójności, jednak nie powodują zatrzymania procesu. Jest to opcja domyślna.

Wszystkie komunikaty

Komunikaty informacyjne udostępniają ogólne informacje na temat systemu lub bieżącego zadania i nie wymagają podejmowania jakichkolwiek działań naprawczych.

Ważne: Wybranie tej opcji może negatywnie wpłynąć na wydajność systemu. Wszystkie komunikaty należy rejestrować tylko w przypadku konieczności rozwiązania problemów lub po uzyskaniu takiego zalecenia z działu wsparcia dla oprogramowania IBM.

4. W polu **Maksymalna wielkość każdego pliku dziennika** wpisz maksymalną wielkość pliku dziennika wyrażoną w megabajtach. Wartością domyślną jest 10 MB.
Po osiągnięciu przez plik dziennika wskazanej wielkości tworzony jest nowy plik dziennika, aż do czasu osiągnięcia maksymalnej dozwolonej liczby plików dzienników. Przechowywanie stosunkowo niewielkich plików dzienników ułatwia ich przeglądanie.
5. W polu **Maksymalna liczba plików dzienników** wpisz maksymalną liczbę plików dzienników, które mają zostać utworzone. Wartością domyślną jest 16.
Jeśli istnieje potrzeba zachowania starszych zarejestrowanych komunikatów do wglądu przez dłuższy czas, należy zwiększyć tę wartość. Jeśli użytkownika bardziej interesują bieżące komunikaty i nie ma potrzeby utrzymywania długiej historii, wartość tę można zmniejszyć.
6. W polu **Domyślne ustawienia narodowe** wybierz język, który ma być stosowany przy rejestrowaniu komunikatów. Wartością domyślną jest English (Angielski).
7. Kliknij przycisk **OK**.
8. Aby zmiany zostały uwzględnione, wprowadź następujące komendy w celu zatrzymania i zrestartowania systemu wyszukiwania korporacyjnego.

```
esdmin system stopall
esdmin system startall
```

Konfigurowanie informacji o serwerze SMTP

Aby odbierać wiadomości e-mail dotyczące działań związanych z wyszukiwaniem korporacyjnym, należy skonfigurować informacje o serwerze SMTP (Simple Mail Transfer Protocol).

Zagadnienia wstępne

Aby skonfigurować informacje o serwerze SMTP, wymagana jest rola administratora wyszukiwania korporacyjnego.


Informacje o tym zadaniu

Kilka funkcji administracyjnych wyszukiwania korporacyjnego umożliwia otrzymywanie wiadomości e-mail. Aby można było odbierać pocztę elektroniczną generowaną przez te funkcje, należy podać informacje o serwerze SMTP:

- Po skonfigurowaniu alertów na poziomie kolekcji lub na poziomie systemu można otrzymywać pocztą elektroniczną powiadomienia o każdym zarejestrowaniu tych komunikatów. Można także otrzymywać powiadomienia pocztą elektroniczną o zarejestrowaniu innych komunikatów, nie tylko tych spowodowanych wystąpieniem monitorowanych zdarzeń.
- Aby zapoznać się ze szczegółowymi informacjami o identyfikatorze URI w indeksie lub o dokumencie, który został usunięty z systemu wyszukiwania korporacyjnego, można odebrać raport pocztą elektroniczną.
- Jeśli podczas monitorowania przeszukiwacza sieci WWW jest włączona opcja tworzenia raportów jego historii, można otrzymywać pocztą elektroniczną powiadomienia o utworzeniu raportów.

Procedura

Aby skonfigurować informacje o serwerze SMTP, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Dziennik kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów**.
4. Na stronie Opcje poczty elektronicznej dla komunikatów systemowych, w polu **Serwer SMTP do dostarczania poczty elektronicznej** wpisz pełną, kwalifikowaną nazwę hosta lub adres IP wykorzystywanego serwera.
System będzie wykorzystywał ten serwer do wysyłania poczty elektronicznej na podane adresy.
5. W polu **Częstotliwość sprawdzania poczty elektronicznej** określ, jak często system ma sprawdzać obecność odpowiednich komunikatów i wysłać dotyczące ich wiadomości e-mail.
System łączy wszystkie wiadomości kierowane na określony adres e-mail w jedną wiadomości i wysyła ją ze zdefiniowaną tutaj częstotliwością.
6. Kliknij przycisk **OK**.

Odbieranie wiadomości poczty elektronicznej dotyczących zarejestrowanych komunikatów

Istnieje możliwość określenia opcji w celu otrzymywania wiadomości e-mail informujących o zarejestrowaniu pewnych komunikatów lub typów komunikatów.

Zagadnienia wstępne

Aby konfigurować opcje poczty elektronicznej dla komunikatów na poziomie systemu, wymagana jest rola administratora wyszukiwania korporacyjnego. Aby konfigurować opcje poczty elektronicznej dla komunikatów na poziomie kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.



Zanim możliwe będzie odbieranie poczty elektronicznej, należy najpierw skonfigurować parametry serwera SMTP, na który poczta będzie dostarczana.

Informacje o tym zadaniu

Konfigurując alerty, można włączyć opcję rejestrowania komunikatów po wystąpieniu określonych zdarzeń. W przypadku włączenia tej opcji można następnie skonfigurować opcje dotyczące automatycznego otrzymywania wiadomości e-mail po zarejestrowaniu odpowiednich komunikatów. Istnieje również możliwość określenia opcji dotyczących otrzymywania wiadomości e-mail w momencie zarejestrowania innych komunikatów, a nie tylko komunikatów wyzwalanych przez zdarzenia.

Procedura

Aby skonfigurować opcje poczty elektronicznej dla komunikatów, należy wykonać poniższe czynności:

1. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o komunikatach systemowych:
 - a. Kliknij opcję **System**, aby otworzyć widok System.
 - b. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
 - c. Na stronie Dziennik kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów**.
 - d. Na stronie Opcje poczty elektronicznej dla komunikatów systemowych zaznacz pole wyboru **Wyślij wiadomość e-mail na temat komunikatów dotyczących systemu**.
 - e. W polu **Adres poczty elektronicznej do odbierania poczty** wpisz jeden lub kilka adresów. Zwykle informacje o komunikatach systemowych powinien otrzymywać administrator wyszukiwania korporacyjnego.
Adresy należy oddzielać przecinkiem. Na przykład:
steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com.
 - f. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o wszystkich zarejestrowanych komunikatach o błędach, zaznacz pole wyboru **Wyślij wiadomość e-mail na temat wszystkich komunikatów o błędach**.
 - g. Jeśli chcesz otrzymywać pocztą elektroniczną tylko powiadomienia o zarejestrowaniu określonych komunikatów systemowych, wpisz identyfikatory tych komunikatów w obszarze **Wyślij wiadomość e-mail na temat wybranych komunikatów**. Każdy identyfikator komunikatu należy wpisać w osobnym wierszu. Na przykład:
FFQC4819E
FFQ00005E
Niektóre identyfikatory komunikatów są wyliczone domyślnie. Opis tych komunikatów można wyświetlić za pomocą przycisku **Pomoc**.
 - h. Kliknij przycisk **OK**.
2. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o komunikatach dla kolekcji:
 - a. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.
 - b. W liście kolekcji odzyskaj kolekcję, którą chcesz skonfigurować, i kliknij opcję  **Edytuj**.

- c. Na stronie Dziennik kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów**.
- d. Na stronie Opcje poczty elektronicznej dla komunikatów kolekcji zaznacz pole wyboru **Wyślij wiadomość e-mail na temat komunikatów dotyczących kolekcji**.
- e. W polu **Adres poczty elektronicznej do odbierania poczty** wpisz jeden lub kilka adresów. Zwykle informacje o komunikatach na poziomie kolekcji powinien otrzymywać administrator kolekcji.
Adresy należy oddzielać przecinkiem. Na przykład:
steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com.
- f. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o wszystkich zarejestrowanych komunikatach o błędach, zaznacz pole wyboru **Wyślij wiadomość e-mail na temat wszystkich komunikatów o błędach**.
- g. Jeśli chcesz otrzymywać pocztą elektroniczną tylko powiadomienia o zarejestrowaniu określonych komunikatów dotyczących kolekcji, wpisz identyfikatory tych komunikatów w obszarze **Wyślij wiadomość e-mail na temat wybranych komunikatów**. Każdy identyfikator komunikatu należy wpisać w osobnym wierszu. Na przykład:
FFQC4819E
FFQ00005E
Niektóre identyfikatory komunikatów są wyliczone domyślnie. Opis tych komunikatów można wyświetlić za pomocą przycisku **Pomoc**.
- h. Kliknij przycisk **OK**.

Pojęcia pokrewne



Komunikaty dla wyszukiwania korporacyjnego

Zadania pokrewne

“Konfiguracja alertów na poziomie kolekcji” na stronie 324

“Konfiguracja alertów na poziomie systemu” na stronie 325

Zmienianie wielkości dziennika zapytań

Wielkość plików dzienników tworzonych na potrzeby przetwarzania zapytań można zwiększać lub zmniejszać, edytując plik konfiguracyjny. Konsola administracyjna nie umożliwia wykonania tego zadania.

Informacje o tym zadaniu

W toku przetwarzania zapytań dane dziennika są zapisywane w pliku *id_kolekcji_OmniFindQueryLog_data.log*, gdzie *id_kolekcji* jest identyfikatorem kolekcji, która ma zostać skonfigurowana, natomiast *data* jest datą utworzenia pliku dziennika. Wielkość pliku dziennika można zwiększyć lub zmniejszyć, zależnie od ilości danych, po zarejestrowaniu której ma zostać utworzony nowy plik dziennika.

Procedura

Aby zmienić wielkość pliku dziennika przetwarzania zapytań:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W systemie wyszukiwania korporacyjnego z wieloma serwerami należy zalogować się na serwerze indeksowania.
2. Otwórz plik *ES_ROOT_NODE/master_config/id_kolekcji.runtime.node1/runtime-generic.properties*.
3. Wyszukaj właściwość **MaxFileSize**. Zwiększ lub zmniejsz jej wartość, aby zwiększyć lub zmniejszyć wielkość plików dzienników, a następnie zapisz zmiany.
4. W systemie wyszukiwania korporacyjnego z jednym serwerem:

- a. Użyj konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania kolekcji, która ma zostać zmieniona, i zatrzymaj serwery wyszukiwania.
 - b. Zrestartuj aplikację ESSearchServer na serwerze WebSphere Application Server.
 - c. W konsoli administracyjnej zrestartuj zatrzymane wcześniej serwery wyszukiwania.
 - d. Otwórz aplikację wyszukującą w nowej przeglądarce.
5. W systemie wyszukiwania korporacyjnego z wieloma serwerami:
- a. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego.
 - b. Wprowadź następujące komendy w celu zrestartowania systemu wyszukiwania korporacyjnego:

```
esadmin system stopall
esadmin system startall
```


Wyświetlanie plików dziennika


Użytkownik może wyświetlać komunikaty zapisywane we wspólnym pliku dziennika przez komponenty systemu i kolekcji. Na stronie tej można także określić filtry w celu wyświetlenia komunikatów o konkretnym poziomie istotności oraz komunikatów z określonych sesji wyszukiwania korporacyjnego.

Zagadnienia wstępne

Wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego mogą przeglądać zawartość plików dzienników dotyczących kolekcji, do administrowania którymi są upoważnieni. Aby móc przeglądać systemowe pliki dziennika, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia dostępu do paska narzędzi **System**.

Procedura


1. Aby wyświetlić pliki dzienników dla pojedynczej kolekcji, należy wykonać poniższe czynności:
 - a. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.
 - b. Na liście kolekcji odszukaj kolekcję, którą chcesz przejrzeć, kliknij przycisk  **Monitor** i otwórz stronę Dziennik.

Wskazówka: W przypadku, gdy wykonywana jest edycja kolekcji i strona Dziennik jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

2. Aby wyświetlić pliki dzienników na poziomie systemu, należy wykonać poniższe czynności:
 - a. Kliknij opcję **System**, aby otworzyć widok System.
 - b. Wybierz stronę Dziennik.
3. W polu **Plik dziennika** wybierz plik dziennika, który ma zostać wyświetlony.
Nazwa każdego pliku dziennika zawiera jego typ (na przykład systemowy lub będący nazwą kolekcji) oraz datę utworzenia tego pliku. Jeżeli tego samego dnia tworzonych jest wiele plików dzienników tego samego typu, liczbowy przyrostek określa kolejność, w jakiej tworzone są te pliki. Na przykład:

```
nazwa_typu_dziennika_20060526.log (zawiera najnowsze wpisy z tego dnia)
nazwa_typu_dziennika_20060526.log.1
nazwa_typu_dziennika_20060526.log.2 (zawiera najstarsze wpisy z tego dnia)
nazwa_typu_dziennika_20060525.log (zawiera najnowsze wpisy z tego dnia)
```

nazwa_typu_dziennika_20060525.log.1
nazwa_typu_dziennika_20060525.log.2
nazwa_typu_dziennika_20060525.log.3 (zawiera najstarsze wpisy z tego dnia)

4. Aby wyświetlać wyłącznie komunikaty o określonych poziomach istotności, zaznacz odpowiednie pola wyboru w polu **Istotność**.
5. Aby wyświetlać wyłącznie komunikaty z określonych sesji, zaznacz odpowiednie pola wyboru w polu **Sesja**.
6. Kliknij przycisk **Wyświetl dziennik**.
Dla każdego komunikatu wyświetlanego na stronie Zawartość pliku dziennika podawana jest data i godzina wygenerowania komunikatu, jego poziom istotności, nazwa sesji zgłaszającej komunikat, identyfikator komunikatu i treść błędu.
Za pomocą odpowiednich przycisków można przechodzić do pierwszej, ostatniej, poprzedniej i następnej strony pliku dziennika. Można też podać numer strony i przejść bezpośrednio do niej.
7. Aby przejrzeć szczegółowe informacje na temat danego komunikatu, kliknij przycisk  **Szczegóły**.
Na stronie Szczegóły komunikatu dziennika wyświetlana jest nazwa hosta serwera wyszukiwania korporacyjnego, na którym pojawił się komunikat, nazwa pliku, który spowodował błąd, nazwa funkcji i numer wiersza, w którym wystąpił błąd, identyfikator procesu i identyfikator wątku.
Klikając przyciski, można przechodzić do następnego i poprzedniego komunikatu w pliku dziennika.

Tworzenie kopii zapasowej i odtwarzanie systemu wyszukiwania korporacyjnego

Skrypty tworzenia i odtwarzania kopii zapasowych umożliwiają utworzenie i odtworzenie kopii zapasowej systemu wyszukiwania korporacyjnego.

Dane, których kopię zapasową tworzą skrypty

Skrypty te służą do tworzenia kopii zapasowych i odtwarzania następujących plików:

- Plików konfiguracyjnych znajdujących się w katalogu ES_NODE_ROOT/master_config.
- Plików bazy danych dla przeszukiwaczy, w tym wszystkich metadanych przeszukiwaczy, na przykład określających czas ostatniego przeszukiwania źródeł danych.
- Wszystkich plików w katalogu ES_NODE_ROOT/data.
- Plików indeksu dla kolekcji skonfigurowanych z innymi niż domyślne katalogami danych.

Tworzenie kopii zapasowej struktury katalogów

Skrypt tworzenia kopii zapasowej tworzy następujące podkatalogi w katalogu określonym podczas jego uruchamiania. Identyfikator administratora wyszukiwania korporacyjnego musi mieć uprawnienia do zapisu w określonym katalogu.

master_config

Zawiera pliki konfiguracyjne znajdujące się w katalogu ES_NODE_ROOT/master_config.

database

Zawiera pliki bazy danych pochodzące z serwera przeszukiwacza.

data

Zawiera pliki indeksów pochodzące z serwera indeksowania.

Wskazówki dotyczące używania

- Można tworzyć kopię zapasową danych z jednego komputera i odtwarzać ją na innym. Jednakże:
 - Nie można odtworzyć plików, których kopia zapasowa została utworzona w produkcie OmniFind Enterprise Edition w systemie, na którym uruchomiona jest inna wersja produktu OmniFind Enterprise Edition.
 - Należy odtworzyć dane w systemie, którego numer wersji jest taki sam, jak numer wersji serwerów wyszukiwania korporacyjnego, lub większy. Na przykład, jeśli tworzona jest kopia zapasowa systemu wyszukiwania korporacyjnego działającego na jednym serwerze, można odtworzyć dane w systemie, który korzysta z dwóch lub czterech serwerów wyszukiwania. Nie można natomiast w systemie, który korzysta z dwóch lub jednego serwera odtworzyć danych, których kopię zapasową utworzono w systemie czteroserwerowym.
 - Plików, których kopie zapasowe zostały utworzone w jednym systemie operacyjnym, nie można odtworzyć w systemie, który działa pod kontrolą innego systemu operacyjnego. Na przykład, jeśli system wyszukiwania korporacyjnego został zainstalowany w systemie AIX, a teraz ma działać w systemie Linux, to na serwerach systemu Linux musi zostać zainstalowany nowy system wyszukiwania korporacyjnego.
- Przed rozpoczęciem tworzenia kopii zapasowej należy zbudować indeks główny, aby została utworzona kopia zapasowa najnowszych indeksowanych danych.

- Wszystkie ustawienia dotyczące katalogu instalacji (ES_INSTALL_ROOT), katalogu danych (ES_NODE_ROOT) oraz identyfikatora i hasła administratora wyszukiwania korporacyjnego muszą być takie same w systemie, którego kopia zapasowa jest tworzona i w tym, w którym będą odtwarzane dane.
- W przypadku konfiguracji z wieloma serwerami należy utworzyć kopię zapasową i odtworzyć system z serwera indeksowania wyszukiwania korporacyjnego. Ponieważ wszystkie dane przeszukiwacza znajdują się w bazach danych na serwerze przeszukiwacza, skrypty uruchamiają komendy zdalne w celu utworzenia kopii zapasowej i odtworzenia danych przeszukiwacza.
- Aby można było utworzyć kopie plików wyszukiwania korporacyjnego w innym katalogu, należy zapewnić wystarczającą ilość wolnego miejsca na dysku. Skrypty tworzenia kopii zapasowych i odtwarzania nie sprawdzają plików.
- Podczas działania skryptów tworzenia i odtwarzania kopii zapasowych wszystkie sesje systemowe są zatrzymywane. W celu uniknięcia prezentacji niepoprawnych lub niespójnych informacji systemowych nie należy korzystać z konsoli administracyjnej wyszukiwania korporacyjnego podczas działania tych skryptów.
- Jeśli z powodu nienaprawialnego błędu wystąpi awaria systemu, konieczne będzie ponownie zainstalowanie produktu OmniFind Enterprise Edition, a następnie uruchomienie skryptu odtwarzania.

Tworzenie kopii zapasowej wyszukiwania korporacyjnego

Kopię zapasową systemu wyszukiwania korporacyjnego tworzy się za pomocą skryptu `esbackup.sh` dla systemów AIX, Linux lub Solaris albo skryptu `esbackup.bat` dla systemów Microsoft Windows.

Ograniczenia

Identyfikator administratora wyszukiwania korporacyjnego musi mieć uprawnienia do zapisu w katalogu określonym podczas uruchamiania skryptu tworzenia kopii zapasowej.

Podczas działania skryptów tworzenia i odtwarzania kopii zapasowych wszystkie sesje systemowe są zatrzymywane. W celu uniknięcia prezentacji niepoprawnych lub niespójnych informacji systemowych nie należy korzystać z konsoli administracyjnej wyszukiwania korporacyjnego podczas działania tych skryptów.

Ważne: Po naciśnięciu klawiszy Ctrl+C w celu przerwania działania skryptu tworzenia kopii zapasowej system wchodzi w niespójny stan. Aby uruchomić wszystkie sesje usługi i wszelkie działające sesje, które zatrzymano podczas tworzenia kopii zapasowej, należy wpisać następującą komendę:

```
esadmin system startall
```

Procedura

Aby utworzyć kopię zapasową wyszukiwania korporacyjnego:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Jeśli wspólna warstwa komunikacyjna (CCL) wyszukiwania korporacyjnego nie działa, uruchom ją:

```
Systemy AIX, Linux lub Solaris
startccl.sh -bg
```

Wiersz komend systemu Windows

```
startccl
```

Narzędzie administracyjne usług systemu Windows

Aby uruchomić warstwę CCL w tle:

- a. Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
 - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
3. W przypadku konfiguracji z wieloma serwerami upewnij się, że warstwa CCL jest uruchomiona na każdym z nich. W razie potrzeby powtórz poprzednie kroki.
 4. Uruchom tworzenie kopii zapasowej, gdzie *katalog_kopii_zapasowej* jest katalogiem, w którym zostanie umieszczona kopia zapasowa:

Systemy AIX, Linux lub Solaris

```
esbackup.sh katalog_kopii_zapasowej
```

Wiersz komend systemu Windows

```
esbackup.bat katalog_kopii_zapasowej
```

Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 369

Odtwarzanie systemu wyszukiwania korporacyjnego

Po ponownym zainstalowaniu programu OmniFind Enterprise Edition, aby odtworzyć system wyszukiwania korporacyjnego w systemach AIX, Linux lub Solaris można użyć skryptu *esrestore.sh*, a w systemach Microsoft Windows - skryptu *esrestore.bat*.

Ograniczenia

Podczas działania skryptów tworzenia i odtwarzania kopii zapasowych wszystkie sesje systemowe są zatrzymywane. W celu uniknięcia prezentacji niepoprawnych lub niespójnych informacji systemowych nie należy korzystać z konsoli administracyjnej wyszukiwania korporacyjnego podczas działania tych skryptów.

Nie można odtworzyć plików, których kopia zapasowa została utworzona w produkcie OmniFind Enterprise Edition w systemie, na którym uruchomiona jest inna wersja produktu OmniFind Enterprise Edition. Ponadto system, na którym odtwarzane są dane, musi mieć tą samą lub większą liczbę serwerów wyszukiwania korporacyjnego niż system, w którym wykonywana była kopia bezpieczeństwa.

Procedura

Aby odtworzyć system wyszukiwania korporacyjnego:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
2. Jeśli wspólna warstwa komunikacyjna (CCL) wyszukiwania korporacyjnego nie działa, uruchom ją:

Systemy AIX, Linux lub Solaris

```
startccl.sh -bg
```

Wiersz komend systemu Windows

```
startccl
```

Narzędzie administracyjne usług systemu Windows

Aby uruchomić warstwę CCL w tle:

- a. Uruchom okno usług systemu Windows: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**.
 - b. Kliknij prawym przyciskiem myszy opcję **IBM OmniFind Enterprise Edition**, a następnie kliknij opcję **Uruchom**.
3. W przypadku konfiguracji z wieloma serwerami upewnij się, że warstwa CCL jest uruchomiona na każdym z nich. W razie potrzeby powtórz poprzednie kroki.
 4. Zatrzymaj usługę controller:
`esadmin stop`
 5. Odtwórz dane wyszukiwania korporacyjnego, przy czym `katalog_kopii_bezpieczeństwa` jest katalogiem, w którym znajduje się kopia bezpieczeństwa plików:

Systemy AIX, Linux lub Solaris

```
esrestore.sh katalog_kopii_bezpieczeństwa
```

Wiersz komend systemu Windows

```
esrestore.bat katalog_kopii_bezpieczeństwa
```

Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 369

Eksportowanie i importowanie konfiguracji kolekcji

Istnieje możliwość eksportowania oraz importowania konfiguracji poszczególnych kolekcji. Eksportowane oraz importowane są tylko pliki konfiguracji kolekcji, a nie dane kolekcji.

Kolekcję można wyeksportować z jednego systemu wyszukiwania korporacyjnego, a następnie zaimportować ją do innego systemu wyszukiwania korporacyjnego. Niemniej jednak kolekcje można eksportować oraz importować tylko między systemami, w których działa ta sama wersja wyszukiwania korporacyjnego. Na przykład można wyeksportować kolekcję z systemu w wersji 8.4, a następnie kolekcję tę zaimportować do systemu w wersji 8.4 lub w wersji 8.4.0.150, ale nie można wyeksportować kolekcji z systemu w wersji 8.3, a zaimportować do systemu w wersji 8.4.

W przypadku eksportowania kolekcji można ją zaimportować tylko do systemu wyszukiwania korporacyjnego działającego w tym samym systemie operacyjnym. Na przykład nie można wyeksportować kolekcji z systemu wyszukiwania korporacyjnego, który jest zainstalowany w systemie Linux, a zaimportować go do systemu wyszukiwania korporacyjnego, który jest zainstalowany w systemie Windows.

Eksportowanie kolekcji

Do eksportowania kolekcji służy komenda `esadmin export`, umożliwiająca eksportowanie kolekcji pojedynczo. Konsola administracyjna wyszukiwania korporacyjnego nie umożliwia eksportowania kolekcji.

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W konfiguracji z wieloma serwerami należy zalogować się na serwerze indeksowania (komenda nie zadziała w razie próby jej uruchomienia z dowolnego innego serwera wyszukiwania korporacyjnego).
2. Wprowadź następujące komendy:

```
esadmin export -help  
esadmin export -cid id_kolekcji [-fname nazwa_pliku_eksportu] [-verbose]
```

Gdzie:

-help

Podaje informacje pomocy dla komendy.

-cid *id_kolekcji*

Określa identyfikator kolekcji do wyeksportowania.

Wskazówka: W celu ustalenia identyfikatora kolekcji, która ma zostać wyeksportowana, można się posłużyć konsolą administracyjną wyszukiwania korporacyjnego lub użyć komendy **esadmin report collections**.

-fname *nazwa_pliku_eksportu*

Określa ścieżkę do pliku eksportu. Jeśli nazwa pliku nie jest podana jako bezwzględna, to zostaje przyjęty katalog ES_NODE_ROOT/dump. Jeśli ta opcja zostanie pominięta, w katalogu ES_NODE_ROOT/dump zostanie utworzony plik z zastosowaniem następującej konwencji nazewnictwa:

_export_rrrrMMdd_GGmmssz.zip

gdzie:

rrrrMMdd

To bieżący rok, miesiąc i dzień uruchomienia komendy eksportu.

GGmmss

To bieżąca godzina, minuta i sekunda uruchomienia komendy eksportu.

z To przesunięcie czasu w bieżącej strefie czasowej względem czasu GMT w chwili uruchomienia komendy eksportu. Na przykład standardowy czas pacyficzny to czas GMT -0800.

-verbose

Powoduje wyświetlenie informacji, które mogą pomóc w rozwiązaniu problemów z operacją eksportu.

Importowanie kolekcji

Do importowania kolekcji służy komenda **esadmin import**, umożliwiająca importowanie kolekcji pojedynczo. Konsola administracyjna wyszukiwania korporacyjnego nie umożliwia importowania kolekcji.

1. Zaloguj się jako administrator wyszukiwania korporacyjnego. W konfiguracji z wieloma serwerami należy zalogować się na serwerze indeksowania (komenda nie zadziała w razie próby jej uruchomienia z dowolnego innego serwera wyszukiwania korporacyjnego).
2. Wprowadź następujące komendy:

```
esadmin import -help
```

```
esadmin import -fname nazwa_pliku_importu
```

```
[-cid nowy_id_kolekcji]
```

```
[-name nowa_nazwa_kolekcji]
```

```
[-colDataDir nowy_katalog_danych_kolekcji]
```

```
[-force]
```

```
[-verbose]
```

Gdzie:

-help

Podaje informacje pomocy dla komendy.

-fname *nazwa_pliku_importu*

Określa ścieżkę do pliku importu. Jeśli nazwa pliku nie jest podana jako bezwzględna, to zostaje przyjęty katalog ES_NODE_ROOT/dump.

-cid *nowy_id_kolekcji*

Określa nowy identyfikator kolekcji, jeśli kolekcja musi być zaimportowana z innym identyfikatorem.

Wskazówka: W celu ustalenia identyfikatora kolekcji, która ma zostać zaimportowana, można się posłużyć konsolą administracyjną wyszukiwania korporacyjnego lub użyć komendy **esadmin report collections**.

-name *nowa_nazwa_kolekcji*

Określa nową nazwę kolekcji, jeśli kolekcja musi być zaimportowana z inną nazwą.

-colDataDir *nowy_katalog_danych_kolekcji*

Określa katalog danych dla kolekcji. W razie pominięcia zostanie przypisany katalog domyślny.

-force

Wymusza import kolekcji już istniejącej w systemie docelowym. Kolekcję do zastąpienia system ustala na podstawie identyfikatora importowanej kolekcji.

Ważne:

- Po zaimportowaniu kolekcji nie można ponownie zaimportować kolekcji z tym samym identyfikatorem, chyba że zostanie użyta ta opcja w celu wymuszenia importu.
- Kolekcja, która jest importowana przez wymuszenie, nie zachowuje mechanizmów analizy tekstu, słowników ani wtyczek przeszukiwacza. Są one zastępowane informacjami zawartymi w importowanej kolekcji.
- Wszelkie przeszukiwacze powiązane z kolekcją zostają usunięte i zastąpione przeszukiwaczami określonymi w pliku importu. Ponieważ zostają usunięte przeszukiwacze, wszystkie jeszcze nie przeanalizowane metadane i dokumenty przeszukiwacza również zostają usunięte. Po zaimportowaniu kolekcji należy ponownie przeszukać wszystkie dokumenty, aby na nowo wygenerować te dane dla zaimportowanych przeszukiwaczy.

-verbose

Powoduje wyświetlenie informacji, które mogą pomóc w rozwiązaniu problemów z operacją importu.

Wskazówki dotyczące używania

Mechanizmy analizy tekstu i słowniki

Wszelkie mechanizmy analizy tekstu i słowniki (takie jak słowniki synonimów, słów pomijanych i kryteriów) powiązane z kolekcją w systemie eksportu nie są eksportowane wraz z danymi konfiguracyjnymi kolekcji. Eksportowane są tylko powiązania nazw z mechanizmami analizy tekstu i słownikami.

Podczas importu, jeśli w systemie docelowym już istnieje mechanizm analizy tekstu lub słownik o tej samej nazwie, to zostaje powiązany z importowaną kolekcją. W przeciwnym razie wyświetlany jest komunikat ostrzegawczy sygnalizujący, że powiązania nie można ustanowić. Dlatego należy przesłać mechanizmy analizy tekstu i słowniki do serwera docelowego i użyć tych samych nazw, które były używane w importowanej kolekcji.

Jeśli w systemie docelowym nie istnieje mechanizm analizy tekstu lub słownik o tej samej nazwie, to powiązanie zostaje zerwane. Jeśli powiązanie zostanie zerwane,

kolekcja będzie poprawnie funkcjonować w systemie docelowym, ale nie będzie korzystać z odpowiednich mechanizmów analizy tekstu lub słowników.

Modele pamięci

Model pamięci skonfigurowany dla systemu eksportu nie zostaje zachowany. Model pamięci skonfigurowany dla systemu docelowego pozostaje taki sam. W razie próby zaimportowania kolekcji do systemu skonfigurowanego dla mniejszego modelu pamięci wyświetlany jest komunikat ostrzegawczy. Zachowanie kolekcji może być niezgodne z oczekiwaniami i może mieć poważne konsekwencje w zakresie użycia zasobów w mniejszym systemie. Aby uniknąć problemów, kolekcję należy zaimportować do systemu, w którym używany jest taki sam lub większy model pamięci niż w systemie eksportu.

Przeszukiwacze

Należy zadbać, aby w systemie docelowym zostały zainstalowane i skonfigurowane elementy określone jako wymagania wstępne przeszukiwacza.

Zaimportowane przeszukiwacze nie działają, gdy nie są dostępne przeszukiwane przez nie źródła danych. Na przykład, jeśli kolekcja zawiera przeszukiwacz systemu plików przeszukujący lokalny system plików, to przeszukiwacz ten nie może wykonywać przeszukiwania, gdy w systemie docelowym nie istnieje taka sama struktura katalogów.

Wtyczki przeszukiwaczy nie są eksportowane. Gdy kolekcja jest importowana, wyświetlany jest komunikat ostrzegawczy, po czym importowanie jest kontynuowane aż do jego ukończenia. Po zaimportowaniu kolekcji należy zaktualizować właściwości przeszukiwacza i powiązać jego wtyczki. Dla zapewnienia poprawnego funkcjonowania przeszukiwacza należy zadbać, aby wtyczki przeszukiwacza były wdrożone w systemie docelowym.

Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 369

Integracja z produktem Lotus Notes wersja 8

Można rozszerzyć możliwości wyszukiwania wdrożonych klientów IBM Lotus Notes, wersja 8, wdrażając wtyczkę wyszukiwania Lotus Notes produktu OmniFind Enterprise Edition. Wtyczka ta umożliwi użytkownikom wyszukiwanie w kolekcjach wyszukiwania korporacyjnego za pomocą paska wyszukiwania klienta Lotus Notes.

Aby zintegrować możliwości wyszukiwania korporacyjnego z produktem Lotus Notes, należy utworzyć serwis aktualizacyjny wtyczki. Wtedy użytkownicy będą mogli instalować wtyczkę w swoich lokalnych systemach klienta Lotus Notes, wersja 8.

Tworzenie serwisu aktualizacyjnego wtyczki wyszukiwania korporacyjnego

Aby zintegrować wyszukiwanie korporacyjne z produktem Lotus Notes, wersja 8, na serwerze WWW w danej organizacji należy wdrożyć serwis aktualizacyjny Eclipse wyszukiwania korporacyjnego. Serwis aktualizacyjny umożliwi użytkownikom wdrażanie wtyczki wyszukiwania korporacyjnego w lokalnych instalacjach klienta Lotus Notes.

Informacje o tym zadaniu

Pliki wymagane do utworzenia serwisu aktualizacyjnego wyszukiwania korporacyjnego są udostępnione w pliku `com.ibm.es.notes.search.plugin_8.4.0.150.zip`.

Procedura

Aby utworzyć serwis aktualizacyjny Eclipse wyszukiwania korporacyjnego:

1. Skopiuj plik `com.ibm.es.notes.search.plugin_8.4.0.150.zip` na serwer WWW w organizacji.
2. Rozpakuj zawartość pliku zip w podkatalogu katalogu współużytkowanego w ścieżce instalacyjnej serwera WWW.

Na przykład, aby wdrożyć serwis aktualizacyjny na serwerze WWW używanym do wyszukiwania korporacyjnego, zawartość pliku zip rozpakuj w katalogu `C:\Program Files\IBM\HTTP Server\htdocs\en_US\omnifind`.

Instalowanie wtyczki wyszukiwania korporacyjnego w kliencie Lotus Notes, wersja 8

Aby przetwarzać zapytania przy użyciu funkcji wyszukiwania korporacyjnego, do paska wyszukiwania Lotus Notes można dodać wtyczkę wyszukiwania korporacyjnego.

Informacje o tym zadaniu

W prawym górnym rogu klienta Lotus Notes znajduje się pasek wyszukiwania. Pasek wyszukiwania ma kilka wtyczek wyszukiwania, które są dostarczane wraz z podstawową instalacją klienta Lotus Notes. Po zainstalowaniu wtyczki wyszukiwania korporacyjnego wyszukiwanie w kolekcjach wyszukiwania korporacyjnego można przeprowadzać z poziomu tego paska wyszukiwania.

Procedura

Aby zainstalować wtyczkę wyszukiwania korporacyjnego w kliencie Lotus Notes:

1. Otwórz plik NOTES_INSTALL_ROOT\notes.ini, gdzie NOTES_INSTALL_ROOT to zwykle ścieżka c:\lotus\notes.
2. Na końcu pliku dodaj właściwość o nazwie OMNIFIND_ENTERPRISE_EDITION_SEARCH_SERVER_URL i jako wartość określ `http://nazwa_hosta:port/ESSearchApplication/search.do?q=`, gdzie *nazwa_hosta* jest nazwą hosta serwera wyszukiwania korporacyjnego, a *port* to port serwera WWW.
3. Zapisz i zamknij plik.
4. Otwórz aplikację kliencką Lotus Notes.
5. Wybierz kolejno opcje: **Plik** → **Aplikacja** → **Zainstaluj**.
6. Wybierz przełącznik **Szukaj nowych opcji do zainstalowania**.
7. Kliknij opcję **Dodaj zdalną lokalizację**.
8. Podaj nazwę dla tego nowego miejsca, na przykład wyszukiwanie korporacyjne. W polu URL wprowadź główny adres URL serwisu aktualizacyjnego wyszukiwania korporacyjnego. Na przykład, jeśli plik com.ibm.es.notes.search.plugin_8.4.0.150.zip został rozpakowany do podkatalogu omnifind w głównym katalogu serwera IBM HTTP Server dla języka angielskiego, to adresem URL, który należy podać, jest `http://nazwa_hosta:port/omnifind/`.
9. Kliknij przycisk **Zakończ**, aby zapisać nowe miejsce zdalne. .
10. Zaznacz pole wyboru obok nazwy miejsca i kliknij przycisk **Dalej**.
11. Zapoznaj się z umową licencyjną, zaznacz pole wyboru **Akceptuję warunki umowy licencyjnej**, a następnie kliknij przycisk **Dalej**.
12. Kliknij przycisk **Zakończ**, aby zainstalować wtyczkę. Jeśli otrzymasz ostrzeżenie, że instalowana jest niepodpisana wtyczka, wybierz przełącznik **Zainstaluj wtyczkę** i kliknij przycisk **OK**.
13. Gdy pojawi się monit, kliknij przycisk **Tak**, aby zrestartować klienta Lotus Notes.

Teraz będzie można kliknąć menu obok paska wyszukiwania produktu Lotus Notes i jako opcję wyszukiwania wybrać **OmniFind Enterprise Edition**. Jeśli zostanie wprowadzony termin zapytania, to po kliknięciu przycisku **Wyszukiwanie** żądanie zostanie skierowane do serwera wyszukiwania korporacyjnego, a wyniki zostaną przedstawione w oknie przeglądarki WWW produktu Lotus Notes.

Integracja z produktem WebSphere Portal

Integrując portlet wyszukiwania korporacyjnego z produktem IBM WebSphere i konfigurując portal WebSphere tak, żeby jako domyślnego mechanizmu wyszukiwania używał portletu wyszukiwania, można rozszerzyć możliwości wyszukiwania tego portalu.

Punkty integracji

Program instalacyjny systemu OmniFind Enterprise Edition zawiera skrypty integrujące mechanizmy wyszukiwania korporacyjnego z portalem WebSphere. Po uruchomieniu tych skryptów system wyszukiwania korporacyjnego może zostać zintegrowany z portalem WebSphere na kilka sposobów:

Portlet wyszukiwania korporacyjnego

W portalu WebSphere Portal istnieje pojedynczy punkt dostępu umożliwiający wzajemne oddziaływanie między aplikacjami, treścią, procesami i użytkownikami. Środowisko portalu WebSphere umożliwi nowym aplikacjom, nazywanym portletami, integrację z portalem bez wpływania na pozostałe aplikacje działające w portalu.

Jeśli do portalu WebSphere dołączony zostanie portlet wyszukiwania korporacyjnego, to do wyszukiwania dokumentów w kolekcjach wyszukiwania korporacyjnego i pracy z wynikami można będzie użyć interfejsu portalu WebSphere. Definiując odpowiednio ustawienia konfiguracyjne portalu WebSphere Portal, można zapewnić, że wygląd i działanie portletu wyszukiwania korporacyjnego będzie takie samo, jak pozostałych portletów w używanym środowisku portalu WebSphere Portal.

Centrum wyszukiwania portalu WebSphere Portal

Centrum wyszukiwania portalu WebSphere Portal stanowi centralny punkt początkowy dla wyszukiwania we wszystkich źródłach, które zostały udostępnione dla wyszukiwania przy użyciu portalu WebSphere Portal. Centrum wyszukiwania i uniwersalny portlet wyszukiwania pozwalają na wyszukiwanie w treści portalu WebSphere i dowolnych innych kolekcjach, które zostały zarejestrowane w centrum wyszukiwania.

Jeśli uruchomione zostaną skrypty integrujące wyszukiwanie korporacyjne z portalem WebSphere w wersji 5.1, to do interfejsu centrum wyszukiwania dodawana jest strona wyszukiwania korporacyjnego. Aby wyszukiwać wyłącznie w kolekcjach wyszukiwania korporacyjnego, należy wybrać tę stronę albo wprowadzić zapytanie wyszukujące w kolekcjach wyszukiwania korporacyjnego oraz innych kolekcjach dostępnych w centrum wyszukiwania.

Jeśli uruchomione zostaną skrypty integrujące wyszukiwanie korporacyjne z portalem WebSphere w wersji 6, to funkcje wyszukiwania korporacyjnego dodawane są jako usługa stowarzyszona, z której można skorzystać do wyszukiwania w kolekcjach wyszukiwania korporacyjnego oraz innych kolekcjach dostępnych w centrum wyszukiwania.

Pasek wyszukiwania portalu WebSphere

Pasek wyszukiwania znajduje się w prawym górnym rogu wszystkich kompozycji interfejsu portalu WebSphere. Domyślne działanie tego paska polega na kierowaniu wszystkich żądań wyszukiwania do domyślnego mechanizmu centrum wyszukiwania. Aby do przetwarzania zapytań użyć bardziej rozbudowanych funkcji

wyszukiwania korporacyjnego, można zmienić to domyślne zachowanie, tak aby wszystkie żądania wyszukiwania kierowane były do portletu wyszukiwania korporacyjnego.

Przeszukiwacze WebSphere Portal i Web Content Management

Aby do indeksu wyszukiwania korporacyjnego dodać serwisy portalu WebSphere oraz serwisy IBM Workplace Web Content Management, należy skorzystać z konsoli administracyjnej wyszukiwania korporacyjnego i skonfigurować w niej przeszukiwacze WebSphere Portal i Web Content Management. Do wyszukiwania w poindeksowanych treściach można użyć portletu wyszukiwania korporacyjnego albo aplikacji wyszukiwania.

Przeszukiwacz WebSphere Portal może przeszukiwać serwisy portalu WebSphere w wersji 5.1 lub w wersji 6. Przeszukiwacz Web Content Management może przeszukiwać serwisy na serwerze WebSphere Portal, wersja 6.

Dokumenty IBM Lotus Quickr

Aby dokumenty IBM Lotus Quickr uwzględnić w indeksie wyszukiwania korporacyjnego, można skonfigurować przeszukiwacz listy adresów początkowych za pomocą konsoli administracyjnej wyszukiwania korporacyjnego. Następnie do wyszukiwania poindeksowanej treści będzie można używać portletu wyszukiwania korporacyjnego w portalu WebSphere Portal lub autonomicznej aplikacji wyszukiwującej.

Przeszukiwacz listy adresów początkowych może przeszukiwać biblioteki treści (dokumentów) Lotus Quickr na serwerze WebSphere Portal, wersja 6.

Korzyści wynikające z integracji

Wyszukiwanie korporacyjne rozszerza środowisko wyszukiwania portalu WebSphere Portal, ponieważ umożliwia wyszukiwanie w szerszym zakresie typów źródeł danych. Za pomocą portletu wyszukiwania korporacyjnego można wyszukiwać w serwisach WWW oraz innych źródłach danych obsługiwanych przez system wyszukiwania korporacyjnego.

Użycie wyszukiwania korporacyjnego wiąże się także z korzyściami dotyczącymi skalowalności. Mechanizm wyszukiwania środowiska Portal jest przydatny w przypadku małych i średnich przedsiębiorstw, w których do obsłużenia obciążenia związanego z wyszukiwaniem i pobieraniem wystarczy pojedynczy serwer. W celu poprawienia jakości wyszukiwania na poziomie przedsiębiorstwa można rozdzielić obciążenie wyszukiwania korporacyjnego na wiele serwerów, spośród których dwa zajmować się będą wyszukiwaniem i pobieraniem.

Skrypty konfiguracyjne do integrowania wyszukiwania korporacyjnego z programem WebSphere Portal

W celu zintegrowania wyszukiwania korporacyjnego z portalem IBM WebSphere można uruchomić skrypty dostarczane wraz z programem instalującym serwer OmniFind Enterprise Edition.

Plik JAR zawierający skrypty konfiguracyjne właściwe dla używanej wersji portalu WebSphere należy skopiować z serwera wyszukiwania korporacyjnego na serwer, na którym zainstalowany jest portal WebSphere. Skrypty konfiguracyjne:

- Instalują pliki EAR pozwalające na korzystanie z wyszukiwania korporacyjnego wewnątrz portalu WebSphere i tworzą przeszukiwacze umożliwiające dodanie do kolekcji wyszukiwania korporacyjnego treści z portalu WebSphere i systemu IBM Workplace Web Content Management.
- Instalują pliki WAR, z których korzysta portlet wyszukiwania korporacyjnego.

- Tworzą strony portalu WebSphere i przypisują do nich pliki portletu wyszukiwania korporacyjnego.
- Kopiują wszystkie potrzebne pliki JAR do katalogów instalacyjnych portalu WebSphere (przed skopiowaniem plików JAR używanych przez system wyszukiwania korporacyjnego wykonywana jest kopia zapasowa wszystkich plików JAR znajdujących się w tych katalogach).
- Zapewniają punkt integracji produktu WebSphere Information Integrator Content Edition umożliwiający wyszukiwanie w dokumentach Portal Document Manager.

Po uruchomieniu tych skryptów trzeba przejść do interfejsu administracyjnego portalu WebSphere, zaktualizować w nim właściwości portletu wyszukiwania i podać informacje na temat serwera wyszukiwania korporacyjnego.

Wskazówki dotyczące używania

- Skrypty konfiguruje wszystkie punkty integracji wyszukiwania korporacyjnego z portalem WebSphere. Na przykład, nie można wybiórczo zainstalować portletu, nie instalując plików EAR obsługujących przeszukiwacze WebSphere Portal iWeb Content Management.
- Jeśli serwer WebSphere Information Integrator Content Edition nie zostanie skonfigurowany, a później podjęta zostanie decyzja o wykorzystaniu portletu wyszukiwanego w dokumentach Portal Document Manager, należy uruchomić skrypt usuwający wyszukiwanie korporacyjne z portalu WebSphere. Później można ponownie uruchomić skrypt konfiguracyjny i podać ścieżkę instalacji serwera WebSphere Information Integrator Content Edition.
- Skrypty zatrzymują i ponownie uruchamiają portal WebSphere. Zaleca się uruchamianie tych skryptów poza normalnymi godzinami pracy, upewniając się tym samym, że wspólnota użytkowników nie zostanie zaskoczona brakiem dostępu do usług portalu.
- Jeśli w czasie pracy skryptów konfiguracyjnych pojawi się jakiś błąd, należy ponownie je uruchomić. Zadania zakończone pomyślnie mogą podczas pierwszej próby zgłaszać różne błędy, ale proces konfiguracji nie jest zatrzymywany i pozostałe zadania powinny zostać wykonane bez przeszkód.
- Po uruchomieniu skryptów konfiguracyjnych, przy pierwszej próbie otwarcia stron portletu wyszukiwania korporacyjnego, strony te mogą otwierać się bardzo powoli, ponieważ system musi skompilować wszystkie pliki JSP (Java Server Pages) związane z tym portletem.

Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 5.1

Aby zintegrować system wyszukiwania korporacyjnego z portalem WebSphere Portal w wersji 5.1.0 lub nowszej, należy użyć skryptu `wp5_install`.

Informacje o tym zadaniu

Pliki wymagane do integracji wyszukiwania korporacyjnego z portalem WebSphere Portal są dostarczane w pliku `es.wp5.install.jar`. Po rozpakowaniu tego pliku wyodrębniane są następujące pliki:

- `ESSearchPortlet.war`
- `ESSearchAdapterPortlet.war`
- `ESSearchAdapter.ear`
- `ESPACServer.ear`
- `esapi.jar`
- `siapi.jar`

- es.security.jar
- Pliki skryptów, programów wsadowych, XML i JACL, które są potrzebne do instalacji

Procedura

Aby zintegrować wyszukiwanie korporacyjne z portalem WebSphere Portal, wersja 5.1:

1. Skopiuj plik **es.wp5.install.jar** z serwera wyszukiwania korporacyjnego do serwera WebSphere Portal, a następnie za pomocą komendy Java **JAR** (lub komendy **TAR**) rozpakuj plik.
2. Opcjonalnie: jeśli ma być obsługiwana integracja z produktem WebSphere Portal Document Manager (PDM), wykonaj jedną z następujących czynności:
 - Uruchom program instalacyjny produktu WebSphere Information Integrator Content Edition, wybierz opcję instalacji samego konektora i zainstaluj konektor PDM na serwerze WebSphere Portal.
 - Utwórz strukturę katalogów produktu WebSphere Information Integrator Content Edition na serwerze WebSphere Portal i skopiuj następujące pliki z istniejącej instalacji produktu WebSphere Information Integrator Content Edition do serwera WebSphere Portal, gdzie **CE_ROOT** określa główny katalog instalacyjny produktu WebSphere Information Integrator Content Edition:

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

3. W wierszu komend uruchom komendę **wp5_install.bat** (w systemie Windows) lub komendę **wp5_install.sh** (w systemie AIX, Linux lub Solaris). Aby zapewnić czytelność, w poniższym przykładzie przedstawiono parametry w oddzielnych wierszach; należy je podać razem z komendą:

```
wp5_install.bat
-WASDir "C:\\Program Files\\WebSphere\\AppServer"
-WASUser wpsbind -WASPassword wpsbind
-WPSDir "C:\\Program Files\\WebSphere\\PortalServer"
-WPSUser wpsadmin -WSPassword wpsadmin
-WPSHost "portalserver.ibm.com:9081"
-IICEDir "C:\\IICE"
```

WASDir

Pełna ścieżka do katalogu instalacyjnego serwera WebSphere Application Server.

WASUser

Nazwa użytkownika będącego administratorem serwera WebSphere Application Server; wymagana tylko wtedy, gdy na serwerze WebSphere Application Server włączono zabezpieczenia globalne.

WASPassword

Hasło użytkownika administracyjnego serwera WebSphere Application Server (jeśli podany został użytkownik).

WPSDir

Pełna ścieżka do katalogu instalacyjnego serwera WebSphere Portal.

WPSUser

Nazwa użytkownika administracyjnego serwera WebSphere Portal.

WSPassword

Hasło podanego użytkownika administracyjnego serwera WebSphere Portal.

WPSHost

Nazwa hosta i numer portu serwera WebSphere Portal.

IICEDir

Pełna ścieżka do katalogu instalacyjnego produktu WebSphere Information Integrator Content Edition; wymagana tylko wtedy, gdy uprzednio skonfigurowano konektor Portal Document Manager na serwerze WebSphere Portal.

4. Po uruchomieniu skryptu (serwer WebSphere Portal zostanie zatrzymany i zrestartowany) zaktualizuj portlet wyszukiwania korporacyjnego, aby identyfikował serwer wyszukiwania:
 - a. Zaloguj się do portalu WebSphere za pomocą identyfikatora i hasła administratora.
 - b. Kliknij opcję **Administrowanie** w prawym górnym rogu.
 - c. Kliknij opcję **Zarządzanie portletami** w obszarze nawigacyjnym z lewej strony, a następnie kliknij opcję **Portlety**.
 - d. Zmień opcję **Szukaj według**, nadając jej wartość **Tytuł zawiera**.
 - e. W polu **Wyszukiwanie** wpisz wyszukiwanie korporacyjne, a następnie kliknij przycisk **Szukaj**.
 - f. Po wyświetleniu nowych ikon z prawej strony kliknij ikonę klucza maszynowego, aby skonfigurować portlet wyszukiwania na potrzeby wyszukiwania korporacyjnego.
 - g. Na liście parametrów portletu zmodyfikuj następujące parametry:

hostname

Podaj pełną nazwę hosta serwera wyszukiwania korporacyjnego.

port Podaj numer portu używany przez serwer aplikacji WebSphere na serwerze wyszukiwania korporacyjnego. Wartością domyślną jest 80 (wartością domyślną dla komunikacji SSL jest 443).

username

Jeśli na serwerze wyszukiwania w serwerze aplikacji WebSphere włączono zabezpieczenia globalne, podaj prawidłową nazwę użytkownika figurującą w rejestrze użytkowników serwera aplikacji WebSphere.

password

Jeśli podano nazwę użytkownika serwera aplikacji WebSphere, podaj powiązane z nią hasło.

protocol

Podaj protokół stosowany do komunikacji między portalem WebSphere a serwerem wyszukiwania. Domyślnym protokołem jest HTTP. Jeśli stosowany ma być protokół SSL, podaj tu protokół HTTPS.

trustStore

Jeśli stosowany jest protokół SSL, podaj pełną ścieżkę (włącznie z nazwą pliku) do miejsca przechowywania certyfikatu SSL.

trustPassword

Jeśli stosowany jest protokół SSL, podaj hasło do pliku wpisanego w opcji trustStore.

ssoCookieName

Podaj nazwę informacji cookie zawierającej ciąg znaków znacznika pojedynczego logowania (SSO). Wartością domyślną jest LtpaToken.

proxyHost

Jeśli do dostępu do serwera wyszukiwania korporacyjnego wymagane jest użycie serwera proxy, podaj pełną nazwę hosta tego serwera.

proxyPort

Jeśli zdefiniowany został serwer proxy, podaj numer portu tego serwera.

proxyUser

Jeśli serwer proxy wymaga prostego uwierzytelniania, podaj tu nazwę użytkownika pozwalającą na zalogowanie się do tego serwera.

proxyPassword

Jeśli podana została nazwa użytkownika serwera proxy, podaj też powiązane z tą nazwą hasło.

h. Kliknij przycisk **OK**, aby zapisać zmiany.

Przystosowywanie paska wyszukiwania portalu WebSphere w wersji 5.1 do wykorzystania wyszukiwania korporacyjnego

Portal WebSphere w wersji 5.1.0 lub nowszej można skonfigurować tak, żeby w momencie, gdy użytkownicy wysyłać będą zapytania za pomocą paska wyszukiwania, korzystał z mechanizmów wyszukiwania korporacyjnego, a nie z domyślnego mechanizmu wyszukiwania portalu WebSphere.

Zagadnienia wstępne

Zanim możliwe będzie przekierowywanie żądań wyszukiwania do wyszukiwania korporacyjnego, konieczne jest uruchomienie skryptu konfiguracyjnego **wp5_install**, który zintegruje wyszukiwanie korporacyjne z portalem WebSphere. Trzeba też zaktualizować parametry portletu wyszukiwania korporacyjnego, wpisując do niego nazwę hosta, numer portu oraz inne informacje na temat serwera wyszukiwania korporacyjnego.

Informacje o tym zadaniu

Pasek wyszukiwania znajduje się w prawym górnym rogu wszystkich kompozycji interfejsu portalu WebSphere. Domyślne zachowanie tego paska polega na kierowaniu wszystkich żądań wyszukiwania do portletu centrum wyszukiwania. Aby do przetwarzania zapytań użyć bardziej rozbudowanych funkcji wyszukiwania korporacyjnego, można zmienić to domyślne zachowanie, tak aby wszystkie żądania wyszukiwania kierowane były do portletu wyszukiwania korporacyjnego.

Po przekierowaniu paska wyszukiwania zmiany będą dotyczyły stron korzystających z tej samej kompozycji portalu WebSphere, z której korzysta portlet wyszukiwania korporacyjnego. Oprócz tego, strony te muszą wywoływać plik **SearchBarInclude.jsp**. Strony korzystające z innych kompozycji lub niewywołujące pliku **SearchBarInclude.jsp** nadal będą stosowały domyślny portlet centrum wyszukiwania.

Po wypełnieniu tego zadania nie można korzystać z centrum wyszukiwania, o ile nie wycofano zmian (na przykład odtworzono oryginalny plik **SearchBarInclude.jsp**).

Procedura

Aby używać portletu wyszukiwania korporacyjnego do obsługi żądań wyszukiwania wpisywanych przez użytkowników do paska wyszukiwania portalu WebSphere:

1. Na serwerze portalu WebSphere przejdź do katalogu **WAS_INSTALL_ROOT/AppServer/installedApps/eswin1/wps.ear/wps.war/themes/html**.
2. Utwórz kopię zapasową pliku **SearchBarInclude.jsp**, kopiując go i nadając mu inną nazwę (na przykład, **SearchBarInclude.jsp.BACKUP**).
3. Dokonaj edycji pliku **SearchBarInclude.jsp** i poniższym tekstem zastąp w nim treść znajdującą się pomiędzy instrukcjami **<wps:if loggedIn="yes">** i **</wps:if>**. W atrybucie **action=** elementu formularza wartość **localhost:9081** zastąp nazwą hosta i numerem portu serwera WebSphere Portal.

WASUser

Nazwa użytkownika będącego administratorem serwera WebSphere Application Server; wymagana tylko wtedy, gdy na serwerze WebSphere Application Server włączono zabezpieczenia globalne.

WASPassword

Hasło użytkownika administracyjnego serwera WebSphere Application Server (jeśli podany został użytkownik).

WPSDir

Pełna ścieżka do katalogu instalacyjnego serwera WebSphere Portal.

WPSUser

Nazwa użytkownika administracyjnego serwera WebSphere Portal.

WPSPassword

Hasło podanego użytkownika administracyjnego serwera WebSphere Portal.

WPSHost

Nazwa hosta i numer portu serwera WebSphere Portal.

Konfigurowanie wyszukiwania korporacyjnego w programie WebSphere Portal, wersja 6

Aby zintegrować system wyszukiwania korporacyjnego z portalem WebSphere w wersji 6, należy użyć skryptu `wp6_install`.

Informacje o tym zadaniu

Pliki wymagane do integracji wyszukiwania korporacyjnego z portalem WebSphere są dostarczane w pliku `es.wp6.install.jar`. Po rozpakowaniu tego pliku wyodrębniane są następujące pliki:

- `ESSearchPortlet.war`
- `ESPACServer.ear`
- `esapi.jar`
- `es.search.provider.jar`
- `es.security.jar`
- Ikony typów plików źródłowych aplikacji wyszukiwania stosowane na stronie wyników dostawcy wyszukiwania.
- Pliki skryptów, programów wsadowych, XML i JACL, które są potrzebne do instalacji

Procedura

Aby zintegrować wyszukiwanie korporacyjne z portalem WebSphere w wersji 6:

1. Skopiuj plik `es.wp6.install.jar` z serwera wyszukiwania korporacyjnego na serwer portalu WebSphere, a następnie użyj polecenia Java **JAR** (lub polecenia **TAR**), aby rozpakować zawartość tego pliku.
2. Opcjonalnie: jeśli ma być obsługiwana integracja z produktem WebSphere Portal Document Manager (PDM), wykonaj jedną z następujących czynności:
 - Uruchom program instalacyjny produktu WebSphere Information Integrator Content Edition, wybierz opcję instalacji samego konektora i zainstaluj konektor PDM na serwerze WebSphere Portal.
 - Utwórz strukturę katalogów produktu WebSphere Information Integrator Content Edition na serwerze WebSphere Portal i skopiuj następujące pliki z istniejącej instalacji

produktu WebSphere Information Integrator Content Edition do serwera WebSphere Portal, gdzie CE_ROOT określa główny katalog instalacyjny produktu WebSphere Information Integrator Content Edition:

```
CE_ROOT/lib/vbr.jar  
CE_ROOT/ejb/vbr_pdm.jar  
CE_ROOT/war/services.war  
CE_ROOT/vbr_services.properties
```

3. W wierszu poleceń uruchom polecenie **wp6_install.bat** (w systemach Windows) lub skrypt **wp6_install.sh** (w systemach AIX, Linux lub Solaris). W poniższym przykładzie dla uzyskania lepszej czytelności poszczególne parametry przedstawione zostały w osobnych wierszach, jednak w praktyce wszystkie te parametry muszą być częścią polecenia:

```
wp6_install.bat  
-WSPProfileDir "C:\\Program Files\\IBM\\WebSphere\\profiles\\wp_profile"  
-WASDir "C:\\Program Files\\IBM\\WebSphere\\AppServer"  
-WASUser wpsbind -WASPassword wpsbind  
-WPSDir "C:\\Program Files\\IBM\\WebSphere\\PortalServer"  
-WPSUser wpsadmin -WSPassword wpsadmin  
-WPSHost "portalserver.ibm.com:9081"  
-IICEDir "C:\\IICE"
```

WSPProfileDir

Pełna ścieżka do katalogu profilu portalu WebSphere. Domyślną ścieżką w systemach AIX jest /usr/IBM/WebSphere/AppServer/profiles/wp_profile, w systemach Linux i Solaris /opt/IBM/WebSphere/AppServer/profiles/wp_profile, a w systemach Windows C:\Program Files\IBM\WebSphere\profiles\wp_profile.

WASDir

Pełna ścieżka do głównego katalogu serwera WebSphere Application Server; wymagana tylko dla systemów AIX, Linux i Solaris. Domyślną ścieżką do głównego katalogu w systemach AIX jest /usr/IBM/WebSphere/AppServer, w systemach Linux i Solaris /opt/IBM/WebSphere/AppServer, a w systemach Windows C:\Program Files\IBM\WebSphere\AppServer.

WASUser

Nazwa użytkownika będącego administratorem serwera WebSphere Application Server; wymagana tylko wtedy, gdy na serwerze WebSphere Application Server włączono zabezpieczenia globalne.

WASPassword

Hasło użytkownika administracyjnego serwera WebSphere Application Server (jeśli podany został użytkownik).

WPSDir

Pełna ścieżka do katalogu instalacyjnego serwera WebSphere Portal.

WPSUser

Nazwa użytkownika administracyjnego serwera WebSphere Portal.

WSPassword

Hasło podanego użytkownika administracyjnego serwera WebSphere Portal.

WPSHost

Nazwa hosta i numer portu serwera WebSphere Portal.

IICEDir

Pełna ścieżka do katalogu instalacyjnego produktu WebSphere Information Integrator Content Edition; wymagana tylko wtedy, gdy uprzednio skonfigurowano konektor Portal Document Manager na serwerze WebSphere Portal.

4. Po uruchomieniu skryptu, zatrzymaniu i zrestartowaniu serwera WebSphere Portal, zaktualizuj portlet wyszukiwania korporacyjnego, aby identyfikował serwer wyszukiwania:
 - a. Zaloguj się do portalu WebSphere za pomocą identyfikatora i hasła administratora.
 - b. W lewym dolnym rogu kliknij opcję **Administracja**.
 - c. Kliknij opcję **Zarządzanie portletami** w obszarze nawigacyjnym z lewej strony, a następnie kliknij opcję **Portlety**.
 - d. Zmień opcję **Szukaj według**, nadając jej wartość **Tytuł zawiera**.
 - e. W polu **Wyszukiwanie** wpisz wyszukiwanie korporacyjne, a następnie kliknij przycisk **Szukaj**.
 - f. Po wyświetleniu nowych ikon z prawej strony kliknij ikonę klucza maszynowego, aby skonfigurować portlet wyszukiwania na potrzeby wyszukiwania korporacyjnego.
 - g. Na liście parametrów portletu zmodyfikuj następujące parametry:

hostname

Podaj pełną nazwę hosta serwera wyszukiwania korporacyjnego.

port

Podaj numer portu używany przez serwer aplikacji WebSphere na serwerze wyszukiwania korporacyjnego. Wartością domyślną jest 80 (wartością domyślną dla komunikacji SSL jest 443).

username

Jeśli na serwerze wyszukiwania w serwerze aplikacji WebSphere włączono zabezpieczenia globalne, podaj prawidłową nazwę użytkownika figurującą w rejestrze użytkowników serwera aplikacji WebSphere.

password

Jeśli podano nazwę użytkownika serwera aplikacji WebSphere, podaj powiązane z nią hasło.

protocol

Podaj protokół stosowany do komunikacji między portalem WebSphere a serwerem wyszukiwania. Domyślnym protokołem jest HTTP. Jeśli stosowany ma być protokół SSL, podaj tu protokół HTTPS.

trustStore

Jeśli stosowany jest protokół SSL, podaj pełną ścieżkę (włącznie z nazwą pliku) do miejsca przechowywania certyfikatu SSL.

trustPassword

Jeśli stosowany jest protokół SSL, podaj hasło do pliku wpisanego w opcji trustStore.

ssoCookieName

Podaj nazwę informacji cookie zawierającej ciąg znaków znacznika pojedynczego logowania (SSO). Wartością domyślną jest `LtpaToken`.

proxyHost

Jeśli do dostępu do serwera wyszukiwania korporacyjnego wymagane jest użycie serwera proxy, podaj pełną nazwę hosta tego serwera.

proxyPort

Jeśli zdefiniowany został serwer proxy, podaj numer portu tego serwera.

proxyUser

Jeśli serwer proxy wymaga prostego uwierzytelniania, podaj tu nazwę użytkownika pozwalającą na zalogowanie się do tego serwera.

proxyPassword

Jeśli podana została nazwa użytkownika serwera proxy, podaj też powiązane z tą nazwą hasło.

h. Kliknij przycisk **OK**, aby zapisać zmiany.

Konfigurowanie centrum wyszukiwania portalu WebSphere w wersji 6 do współpracy z wyszukiwaniem korporacyjnym

Portal WebSphere w wersji 6 można skonfigurować tak, żeby zapytania wpisywane przez użytkowników w centrum wyszukiwania używane były do wyszukiwania w kolekcjach wyszukiwania korporacyjnego.

Ograniczenia

Jeśli przeznaczone do wyszukiwania kolekcje wyszukiwania korporacyjnego są zabezpieczone, to należy uruchomić portlet wyszukiwania korporacyjnego i skonfigurować profil użytkownika. Profil jest szyfrowany i zapisywany w bezpiecznym miejscu przez system wyszukiwania korporacyjnego. Profil musi istnieć, zanim użytkownicy centrum wyszukiwania portalu WebSphere będą mogli przysyłać zapytania wyszukiwania w kolekcjach.

Informacje o tym zadaniu

Centrum wyszukiwania portalu WebSphere w wersji 6 udostępnia wyszukiwanie w wielu kolekcjach. Kolekcje mogą zawierać najróżniejsze rodzaje treści, takie jak biblioteki dokumentów portalu lub treść portalu (strony i portlety). Po uruchomieniu skryptów integrujących wyszukiwanie korporacyjne z portalem WebSphere można tak skonfigurować centrum wyszukiwania, żeby wyszukiwanie odbywało się również w kolekcjach wyszukiwania korporacyjnego.

Procedura

Aby skonfigurować centrum wyszukiwania do wyszukiwania w kolekcjach wyszukiwania korporacyjnego:

1. Zaloguj się do portalu WebSphere za pomocą identyfikatora i hasła administratora.
2. W lewym dolnym rogu kliknij opcję **Administracja**.
3. W obszarze nawigacji kliknij opcję **Zarządzanie wyszukiwaniem**, a następnie kliknij opcję **Zarządzaj wyszukiwaniem**.
4. Kliknij opcję **Usługi wyszukiwania**, a następnie opcję **Nowa usługa wyszukiwania**.
5. W polu **Implementacja usług wyszukiwania** wybierz pozycję **Usługa wyszukiwania korporacyjnego**, a w polu tekstowym **Nazwa usługi** wprowadź nazwę, jaką chcesz nadać tej usłudze.
6. Na liście parametrów zmodyfikuj następujące parametry:

hostname

Podaj pełną nazwę hosta serwera wyszukiwania korporacyjnego.

port

Podaj numer portu używany przez serwer aplikacji WebSphere na serwerze wyszukiwania korporacyjnego. Wartością domyślną jest 80 (wartością domyślną dla komunikacji SSL jest 443).

username

Jeśli na serwerze wyszukiwania w serwerze aplikacji WebSphere włączono zabezpieczenia globalne, podaj prawidłową nazwę użytkownika figurującą w rejestrze użytkowników serwera aplikacji WebSphere.

password

Jeśli podano nazwę użytkownika serwera aplikacji WebSphere, podaj powiązane z nią hasło.

protocol

Podaj protokół stosowany do komunikacji między portalem WebSphere a serwerem wyszukiwania. Domyślnym protokołem jest HTTP. Jeśli stosowany ma być protokół SSL, podaj tu protokół HTTPS.

trustStore

Jeśli stosowany jest protokół SSL, podaj pełną ścieżkę (włącznie z nazwą pliku) do miejsca przechowywania certyfikatu SSL.

trustPassword

Jeśli stosowany jest protokół SSL, podaj hasło do pliku wpisanego w opcji trustStore.

ssoCookieName

Podaj nazwę informacji cookie zawierającej ciąg znaków znacznika pojedynczego logowania (SSO). Wartością domyślną jest `LtpaToken`.

proxyHost

Jeśli do dostępu do serwera wyszukiwania korporacyjnego wymagane jest użycie serwera proxy, podaj pełną nazwę hosta tego serwera.

proxyPort

Jeśli zdefiniowany został serwer proxy, podaj numer portu tego serwera.

proxyUser

Jeśli serwer proxy wymaga prostego uwierzytelniania, podaj tu nazwę użytkownika pozwalającą na zalogowanie się do tego serwera.

proxyPassword

Jeśli podana została nazwa użytkownika serwera proxy, podaj też powiązane z tą nazwą hasło.

7. Kliknij przycisk **OK**, aby zapisać zmiany.

Konfigurowanie paska wyszukiwania portalu WebSphere w wersji 6 do korzystania z wyszukiwania korporacyjnego

Portal WebSphere w wersji 6 lub nowszej można skonfigurować tak, żeby w momencie, gdy użytkownicy wysyłają będą zapytania za pomocą paska wyszukiwania, korzystał z mechanizmów wyszukiwania korporacyjnego, a nie z domyślnego mechanizmu wyszukiwania portalu WebSphere.

Zagadnienia wstępne

Zanim możliwe będzie przekierowywanie żądań wyszukiwania do wyszukiwania korporacyjnego, konieczne jest uruchomienie skryptu konfiguracyjnego **wp6_install**, który zintegruje wyszukiwanie korporacyjne z portalem WebSphere. Trzeba też zaktualizować parametry portletu wyszukiwania korporacyjnego, wpisując do niego nazwę hosta, numer portu oraz inne informacje na temat serwera wyszukiwania korporacyjnego.

Informacje o tym zadaniu

Pasek wyszukiwania znajduje się w prawym górnym rogu wszystkich kompozycji interfejsu portalu WebSphere. Domyślne zachowanie tego paska polega na kierowaniu wszystkich żądań wyszukiwania do portletu centrum wyszukiwania. Aby do przetwarzania zapytań użyć

bardziej rozbudowanych funkcji wyszukiwania korporacyjnego, można zmienić to domyślne zachowanie, tak aby wszystkie żądania wyszukiwania kierowane były do portletu wyszukiwania korporacyjnego.

Po przekierowaniu paska wyszukiwania zmiany będą dotyczyły stron korzystających z tej samej kompozycji portalu WebSphere, z której korzysta portlet wyszukiwania korporacyjnego. Oprócz tego, strony te muszą wywoływać plik `banner_searchControl.jspf`. Strony korzystające z innych kompozycji lub niewywołujące pliku `banner_searchControl.jspf` nadal będą stosowały domyślny portlet centrum wyszukiwania.

Po wypełnieniu tego zadania nie można korzystać z centrum wyszukiwania, o ile nie wycofano zmian (na przykład odtworzono oryginalny plik `banner_searchControl.jspf`).

Procedura

Aby używać portletu wyszukiwania korporacyjnego do obsługi żądań wyszukiwania wpisywanych przez użytkowników do paska wyszukiwania portalu WebSphere:

1. Zatrzymaj instancję serwera aplikacji WebSphere Portal.
2. Na serwerze WebSphere Portal przejdź do katalogu `WPS_PROFILE_ROOT/installedApps/nazwa_węzła/wps.ear/wps.war/themes/html/nazwa_bieżącego_tematu`, gdzie `nazwa_węzła` jest nazwą węzła serwera WebSphere Portal, a `nazwa_bieżącego_tematu` jest nazwą tematu zastosowanego obecnie w serwerze WebSphere Portal. Domyślną nazwą serwera WebSphere Portal jest IBM.
3. Utwórz kopię zapasową pliku `banner_searchControl.jspf`, kopiując go i nadając mu inną nazwę (na przykład `banner_searchControl.jspf.BACKUP`).
4. Poddaj edycji plik `banner_searchControl.jspf` i zastąp jego zawartość poniższym tekstem. W atrybucie `action=` elementu formularza wartość `localhost:10038` zastąp nazwą hosta i numerem portu serwera WebSphere Portal.

```
<%@ taglib uri="/WEB-INF/tld/SearchMenuControl.tld" prefix="searchmenu" %>
<%String ic = (bidiImageRTL == null) ? "icons/scope_search_submit.gif" :
"icons/scope_search_submit"+bidiImageRTL+".gif";%>

<searchmenu:adminlinkinfo name="SEARCH_CENTER">
<div class="searchControl">
<form name="SearchForm" style="margin: 0px;" method="GET"
action="http://localhost:10038/wps/omnifind/portalSearchBar.jsp">
<table border="0" cellpadding="0" cellspacing="0">
<tr>
<td><span class="wpsInstructionText">
<portal-fmt:text key="search.theme.control.label" bundle="nls.engine"/></span></td>
<td valign="middle" style="padding: 0px 4px 0px 4px;">
<input type="text" name="q"></input></td>
<td valign="middle"><input tabIndex="4" valign="middle"
title="<portal-fmt:text key='search.theme.searchresultsicon.alttext' bundle='nls.engine' />"
alt="<portal-fmt:text key='search.theme.searchresultsicon.alttext' bundle='nls.engine' />"
src="<portal-logic:urlFindInTheme file=">"/>" type="image"></input></td>
</tr>
</table>
</form>
</div>
</searchmenu:adminlinkinfo>
```

5. Otwórz plik `banner.jspf` i zapisz go. Ten krok aktualizuje datę modyfikacji pliku, co zapewnia jego ponowną kompilację. Jeśli używana jest kompozycja inna niż domyślna kompozycja portalu WebSphere, krok ten można uznać za opcjonalny.
6. Otwórz plik `Default.jsp` i zapisz go.
7. Zrestartuj instancję serwera aplikacji WebSphere Portal.

Konfigurowanie portletu wyszukiwania korporacyjnego dla produktu Lotus Quickr

Portlet wyszukiwania korporacyjnego w portalu WebSphere Portal, wersja 6, można skonfigurować tak, aby możliwe było wyszukiwanie w źródłach Lotus Quickr.

Zagadnienia wstępne

Uruchom komendę **wp6_install.bat** (w systemie Windows) lub komendę **wp6_install.sh** (w systemie AIX, Linux lub Solaris) i wykonaj procedurę konfigurowania wyszukiwania korporacyjnego w portalu WebSphere Portal, wersja 6.

Procedura

Aby skonfigurować portlet wyszukiwania korporacyjnego w portalu WebSphere Portal, wersja 6, w celu wyszukiwania w źródłach Lotus Quickr:

1. Zaktualizuj parametry portletu w konfiguracji portalu WebSphere Portal:
 - a. Zaloguj się w portalu WebSphere Portal z identyfikatorem i hasłem administratora Lotus Quickr.
 - b. Kliknij opcję **Administrowanie serwisem**, a następnie kliknij opcję **Administrowanie zaawansowane**.
 - c. Kliknij opcję **Zarządzanie portletami** w obszarze nawigacyjnym z lewej strony, a następnie kliknij opcję **Portlety**.
 - d. Zmień opcję **Szukaj według**, nadając jej wartość **Tytuł zawiera**.
 - e. W polu **Wyszukiwanie** wpisz wyszukiwanie korporacyjne, a następnie kliknij przycisk **Szukaj**.
 - f. Po wyświetleniu nowych ikon z prawej strony kliknij ikonę klucza maszynowego, aby skonfigurować portlet wyszukiwania na potrzeby wyszukiwania korporacyjnego.
 - g. Na liście parametrów portletu zmodyfikuj następujące parametry:

hostname

Podaj pełną nazwę hosta serwera wyszukiwania korporacyjnego.

port

Podaj numer portu używany przez serwer aplikacji WebSphere na serwerze wyszukiwania korporacyjnego. Wartością domyślną jest 80 (wartością domyślną dla komunikacji SSL jest 443).

username

Jeśli na serwerze wyszukiwania w serwerze aplikacji WebSphere włączono zabezpieczenia globalne, podaj prawidłową nazwę użytkownika figurującą w rejestrze użytkowników serwera aplikacji WebSphere.

password

Jeśli podano nazwę użytkownika serwera aplikacji WebSphere, podaj powiązane z nią hasło.

protocol

Podaj protokół stosowany do komunikacji między portalem WebSphere a serwerem wyszukiwania. Domyślnym protokołem jest HTTP. Jeśli stosowany ma być protokół SSL, podaj tu protokół HTTPS.

trustStore

Jeśli stosowany jest protokół SSL, podaj pełną ścieżkę (włącznie z nazwą pliku) do miejsca przechowywania certyfikatu SSL.

trustPassword

Jeśli stosowany jest protokół SSL, podaj hasło do pliku wpisanego w opcji trustStore.

ssoCookieName

Podaj nazwę informacji cookie zawierającej ciąg znaków znacznika pojedynczego logowania (SSO). Wartością domyślną jest `LtpaToken`.

proxyHost

Jeśli do dostępu do serwera wyszukiwania korporacyjnego wymagane jest użycie serwera proxy, podaj pełną nazwę hosta tego serwera.

proxyPort

Jeśli zdefiniowany został serwer proxy, podaj numer portu tego serwera.

proxyUser

Jeśli serwer proxy wymaga prostego uwierzytelniania, podaj tu nazwę użytkownika pozwalającą na zalogowanie się do tego serwera.

proxyPassword

Jeśli podana została nazwa użytkownika serwera proxy, podaj też powiązane z tą nazwą hasło.

- h. Kliknij przycisk **OK**, aby zapisać zmiany.
2. Aby uzyskać dostęp do portletu po jego skonfigurowaniu:
 - a. Zaloguj się na serwerze Lotus Quickr.
 - b. W oknie przeglądarki zmień adres URL na następujący: `http://nazwa_hosta:port/lotus/myquickr/ESSearchPortlet`

Usuwanie wyszukiwania korporacyjnego z portalu WebSphere w wersji 6

Aby usunąć system wyszukiwania korporacyjnego z portalu WebSphere w wersji 6, należy skorzystać ze skryptu `wp6_uninstall`.

Informacje o tym zadaniu

Usuując system wyszukiwania korporacyjnego z portalu WebSphere, trzeba pamiętać, że parametry portletu podawane w czasie konfigurowania portletu wyszukiwania korporacyjnego nie są zapisywane.

Po uruchomieniu skryptu zatrzymuje on pracę serwera portalu WebSphere. Po usunięciu systemu wyszukiwania korporacyjnego skrypt ponownie uruchamia serwer portalu WebSphere.

Procedura

Aby usunąć system wyszukiwania korporacyjnego z portalu WebSphere w wersji 6:

W wierszu poleceń uruchom polecenie **wp6_uninstall.bat** (w systemach Windows) lub polecenie **wp6_uninstall.sh** (w systemach AIX, Linux lub Solaris). Aby zapewnić czytelność, w poniższym przykładzie przedstawiono parametry w oddzielnych wierszach; należy je podać razem z komendą:

```
wp6_uninstall.bat
-WSPProfileDir "C:\\Program Files\\IBM\\WebSphere\\AppServer\\profiles\\wp_profile"
-WASDir "C:\\Program Files\\IBM\\WebSphere\\AppServer"
-WASUser wpsbind -WASPassword wpsbind
-WPSDir "C:\\Program Files\\IBM\\WebSphere\\PortalServer"
-WPSUser wpsadmin -WSPassword wpsadmin
-WPSHost "portalserver.ibm.com:9081"
```

WSPProfileDir

Pełna ścieżka do katalogu profilu portalu WebSphere.

WASDir

Pełna ścieżka do głównego katalogu serwera WebSphere Application Server; wymagana tylko dla systemów AIX, Linux i Solaris.

WASUser

Nazwa użytkownika będącego administratorem serwera WebSphere Application Server; wymagana tylko wtedy, gdy na serwerze WebSphere Application Server włączono zabezpieczenia globalne.

WASPassword

Hasło użytkownika administracyjnego serwera WebSphere Application Server (jeśli podany został użytkownik).

WPSDir

Pełna ścieżka do katalogu instalacyjnego serwera WebSphere Portal.

WPSUser

Nazwa użytkownika administracyjnego serwera WebSphere Portal.

WPSPassword

Hasło podanego użytkownika administracyjnego serwera WebSphere Portal.

WPSHost

Nazwa hosta i numer portu serwera WebSphere Portal.

Integracja wyszukiwania korporacyjnego z systemami klastrowymi WebSphere Portal

Portlet wyszukiwania korporacyjnego można skonfigurować tak, aby działał w systemie klastrowym WebSphere Portal, wersja 6.

Wskazówki dotyczące używania

- Przed uruchomieniem skryptów konfiguracyjnych wyszukiwania korporacyjnego należy się upewnić, że działa program Network Deployment Manager serwera WebSphere Application Server oraz że działa każdy z węzłów w klastrze.
- Skrypty konfiguruje wszystkie punkty integracji wyszukiwania korporacyjnego z portalem WebSphere. Na przykład, nie można wybiórczo zainstalować portletu, nie instalując plików EAR obsługujących przeszukiwacze WebSphere Portal i Web Content Management.
- Skrypty zatrzymują i restartują wszystkie instancje serwera WebSphere Portal w klastrze. Wskazane może być uruchomienie tych skryptów poza normalnymi godzinami pracy, aby wspólnota użytkowników nie dotknął brak dostępu do usług portalu.
- Jeśli w czasie pracy skryptów konfiguracyjnych wystąpią błędy, należy je uruchomić ponownie. Zadania zakończone pomyślnie mogą podczas pierwszej próby zgłaszać różne błędy, ale proces konfiguracji nie jest zatrzymywany i pozostałe zadania powinny zostać wykonane bez przeszkód.
- Po uruchomieniu skryptów konfiguracyjnych, przy pierwszej próbie otwarcia stron portletu wyszukiwania korporacyjnego, strony te mogą otwierać się bardzo powoli, ponieważ system musi skompilować wszystkie pliki JSP (Java Server Pages) związane z tym portletem.

Konfigurowanie wyszukiwania korporacyjnego w systemie klastrowym WebSphere Portal

Do integrowania systemu wyszukiwania korporacyjnego z systemem klastrowym WebSphere Portal, wersja 6, służy skrypt `wp6_cluster_install`.

Informacje o tym zadaniu

Pliki wymagane do integracji wyszukiwania korporacyjnego z portalem WebSphere są dostarczane w pliku **es.wp6.install.jar**. Po rozpakowaniu tego pliku wyodrębniane są następujące pliki:

- ESSearchPortlet.war
- ESPACServer.ear
- esapi.jar
- es.search.provider.jar
- es.security.jar
- Pliki skryptów, programów wsadowych, XML i JACL, które są potrzebne do instalacji

Procedura

Aby zintegrować wyszukiwanie korporacyjne z systemem klastrowym WebSphere Portal, wersja 6:

1. Skopiuj plik **es.wp6.install.jar** z serwera wyszukiwania korporacyjnego na każdy węzeł w klastrze, w którym zainstalowany jest portal WebSphere Portal, a następnie rozpakuj plik za pomocą komendy **JAR** Java (lub komendy **TAR**).
2. Opcjonalnie: jeśli ma być obsługiwana integracja z produktem WebSphere Portal Document Manager (PDM), wykonaj jedną z następujących czynności w odniesieniu do każdego węzła w klastrze:
 - Uruchom program instalacyjny produktu WebSphere Information Integrator Content Edition, wybierz opcję instalacji samego konektora i zainstaluj konektor PDM na serwerze WebSphere Portal.
 - Utwórz strukturę katalogów produktu WebSphere Information Integrator Content Edition na serwerze WebSphere Portal i skopiuj następujące pliki z istniejącej instalacji produktu WebSphere Information Integrator Content Edition do serwera WebSphere Portal, gdzie **CE_ROOT** określa główny katalog instalacyjny produktu WebSphere Information Integrator Content Edition:

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

3. W wierszu komend uruchom komendę **wp6_cluster_copyFiles.bat** (w systemie Windows) lub komendę **wp6_cluster_copyFiles.sh** (w systemie AIX, Linux lub Solaris) w odniesieniu do każdego węzła w klastrze, w którym jest zainstalowany portal WebSphere Portal. W poniższych przykładach dla uzyskania lepszej czytelności poszczególne parametry przedstawione zostały w osobnych wierszach, jednak w praktyce wszystkie te parametry muszą być częścią polecenia:

```
wp6_cluster_copyFiles.bat
-WPSDir "C:\\Program Files\\IBM\\WebSphere\\PortalServer"
-WSPProfileDir "C:\\Program Files\\IBM\\WebSphere\\AppServer\\profiles\\wp_profile"

wp6_cluster_copyFiles.sh
-WASDir /opt/IBM/WebSphere/AppServer
-WPSDir /opt/IBM/WebSphere/PortalServer
-WSPProfileDir /opt/IBM/WebSphere/AppServer/profiles/wp_profile
```

4. W wierszu komend uruchom komendę **wp6_cluster_install.bat** (w systemie Windows) lub komendę **wp6_cluster_install.sh** (w systemie AIX, Linux lub Solaris). W poniższym przykładzie dla uzyskania lepszej czytelności poszczególne parametry przedstawione zostały w osobnych wierszach, jednak w praktyce wszystkie te parametry muszą być częścią polecenia:

```
wp6_cluster_install.bat
-WPSClusterName MyCluster
-WSPProfileDir "C:\\Program Files\\IBM\\WebSphere\\profiles\\wp_profile"
-WASDir "C:\\Program Files\\IBM\\WebSphere\\AppServer"
```

```
-WASUser wpsbind
-WASPassword wpsbind
-WPSDir "C:\\Program Files\\IBM\\WebSphere\\PortalServer"
-WPSUser wpsadmin
-WSPassword wpsadmin
-WPSHost "portalserver.ibm.com"
-webServerName webserver1
-webServerNodeName node1
-IICEDir "C:\\Program Files\\IBM\\Content Edition"
```

WPSClusterName

Nazwa klastra, w którym jest zainstalowany portal WebSphere Portal.

WSPProfileDir

Pełna ścieżka do katalogu profilu portalu WebSphere. Domyślną ścieżką w systemach AIX jest /usr/IBM/WebSphere/AppServer/profiles/wp_profile, w systemach Linux i Solaris /opt/IBM/WebSphere/AppServer/profiles/wp_profile, a w systemach Windows C:\\Program Files\\IBM\\WebSphere\\profiles\\wp_profile.

WASDir

Pełna ścieżka do głównego katalogu serwera WebSphere Application Server; wymagana tylko dla systemów AIX, Linux i Solaris. Domyślną ścieżką do głównego katalogu w systemach AIX jest /usr/IBM/WebSphere/AppServer, a w systemach Linux i Solaris /opt/IBM/WebSphere/AppServer.

WASUser

Nazwa użytkownika będącego administratorem serwera WebSphere Application Server; wymagana tylko wtedy, gdy na serwerze WebSphere Application Server włączono zabezpieczenia globalne.

WASPassword

Hasło użytkownika administracyjnego serwera WebSphere Application Server (jeśli podany został użytkownik).

WPSDir

Pełna ścieżka do katalogu instalacyjnego serwera WebSphere Portal.

WPSUser

Nazwa użytkownika administracyjnego serwera WebSphere Portal.

WSPPassword

Hasło podanego użytkownika administracyjnego serwera WebSphere Portal.

WPSHost

Nazwa hosta i numer portu serwera WebSphere Portal.

webServerName

Nazwa definicji serwera WWW, do której należy portal WebSphere Portal.

webServerNodeName

Nazwa węzła serwera WebSphere Application, do którego należy definicja serwera WWW.

IICEDir

Pełna ścieżka do katalogu instalacyjnego produktu WebSphere Information Integrator Content Edition; wymagana tylko wtedy, gdy uprzednio skonfigurowano konektor Portal Document Manager na serwerze WebSphere Portal.

5. Gdy skrypt zakończy działanie, otwórz przeglądarkę WWW i zaloguj się w konsoli administracyjnej WebSphere na serwerze Network Deployment. Zazwyczaj adres to `http://nazwa_hosta:9060/ibm/console`.
6. Rozwiń sekcję **Serwery** i wybierz opcję **Serwery WWW**.

7. Zaznacz pole **Wybierz** obok właściwego serwera WWW, a następnie kliknij przycisk **Generuj wtyczkę**.
8. Zaznacz pole **Wybierz** obok właściwego serwera WWW, a następnie kliknij przycisk **Propaguj wtyczkę**.
9. Wyloguj się z konsoli administracyjnej.
10. Zaktualizuj portlet wyszukiwania korporacyjnego w celu zidentyfikowania serwera wyszukiwania:
 - a. Zaloguj się do portalu WebSphere za pomocą identyfikatora i hasła administratora.
 - b. W lewym dolnym rogu kliknij opcję **Administracja**.
 - c. Kliknij opcję **Zarządzanie portletami** w obszarze nawigacyjnym z lewej strony, a następnie kliknij opcję **Portlety**.
 - d. Zmień opcję **Szukaj według**, nadając jej wartość **Tytuł zawiera**.
 - e. W polu **Wyszukiwanie** wpisz wyszukiwanie korporacyjne, a następnie kliknij przycisk **Szukaj**.
 - f. Po wyświetleniu nowych ikon z prawej strony kliknij ikonę klucza maszynowego, aby skonfigurować portlet wyszukiwania na potrzeby wyszukiwania korporacyjnego.
 - g. Na liście parametrów portletu zmodyfikuj następujące parametry:

hostname

Podaj pełną nazwę hosta serwera wyszukiwania korporacyjnego.

port

Podaj numer portu używany przez serwer aplikacji WebSphere na serwerze wyszukiwania korporacyjnego. Wartością domyślną jest 80 (wartością domyślną dla komunikacji SSL jest 443).

username

Jeśli na serwerze wyszukiwania w serwerze aplikacji WebSphere włączono zabezpieczenia globalne, podaj prawidłową nazwę użytkownika figurującą w rejestrze użytkowników serwera aplikacji WebSphere.

password

Jeśli podano nazwę użytkownika serwera aplikacji WebSphere, podaj powiązane z nią hasło.

protocol

Podaj protokół stosowany do komunikacji między portalem WebSphere a serwerem wyszukiwania. Domyślnym protokołem jest HTTP. Jeśli stosowany ma być protokół SSL, podaj tu protokół HTTPS.

trustStore

Jeśli stosowany jest protokół SSL, podaj pełną ścieżkę (włącznie z nazwą pliku) do miejsca przechowywania certyfikatu SSL.

trustPassword

Jeśli stosowany jest protokół SSL, podaj hasło do pliku wpisanego w opcji trustStore.

ssoCookieName

Podaj nazwę informacji cookie zawierającej ciąg znaków znacznika pojedynczego logowania (SSO). Wartością domyślną jest LtpaToken.

proxyHost

Jeśli do dostępu do serwera wyszukiwania korporacyjnego wymagane jest użycie serwera proxy, podaj pełną nazwę hosta tego serwera.

proxyPort

Jeśli zdefiniowany został serwer proxy, podaj numer portu tego serwera.

proxyUser

Jeśli serwer proxy wymaga prostego uwierzytelniania, podaj tu nazwę użytkownika pozwalającą na zalogowanie się do tego serwera.

proxyPassword

Jeśli podana została nazwa użytkownika serwera proxy, podaj też powiązane z tą nazwą hasło.

- h. Kliknij przycisk **OK**, aby zapisać zmiany.

Usuwanie wyszukiwania korporacyjnego z systemu klastrowego WebSphere Portal

Do usuwania wyszukiwania korporacyjnego z systemu klastrowego WebSphere Portal, wersja 6, służy skrypt `wp6_cluster_uninstall`.

Informacje o tym zadaniu

Usuując system wyszukiwania korporacyjnego z portalu WebSphere, trzeba pamiętać, że parametry portletu podawane w czasie konfigurowania portletu wyszukiwania korporacyjnego nie są zapisywane.

Po uruchomieniu skryptu zatrzymuje on pracę serwera portalu WebSphere. Po usunięciu systemu wyszukiwania korporacyjnego skrypt ponownie uruchamia serwer portalu WebSphere.

Procedura

Aby usunąć wyszukiwanie korporacyjne z systemu klastrowego WebSphere Portal, wersja 6:

1. W wierszu komend uruchom komendę **wp6_cluster_uninstall.bat** (w systemach Windows) lub komendę **wp6_cluster_uninstall.sh** (w systemach AIX, Linux i Solaris) na jednym z węzłów klastra. Aby zapewnić czytelność, w poniższym przykładzie przedstawiono parametry w oddzielnych wierszach; należy je podać razem z komendą:

```
wp6_cluster_uninstall.bat
-WPSClusterName MyCluster
-WSPProfileDir "C:\\Program Files\\IBM\\WebSphere\\profiles\\wp_profile"
-WASDir "C:\\Program Files\\IBM\\WebSphere\\AppServer"
-WASUser wpsbind
-WASPassword wpsbind
-WPSDir "C:\\Program Files\\IBM\\WebSphere\\PortalServer"
-WPSUser wpsadmin
-WSPassword wpsadmin
-WPSHost "portalserver.ibm.com"
-webServerName webserver1
-webServerNodeName node1
```

WPSClusterName

Nazwa klastra, w którym jest zainstalowany portal WebSphere Portal.

WSPProfileDir

Pełna ścieżka do katalogu profilu portalu WebSphere. Domyślną ścieżką w systemach AIX jest `/usr/IBM/WebSphere/AppServer/profiles/wp_profile`, w systemach Linux i Solaris `/opt/IBM/WebSphere/AppServer/profiles/wp_profile`, a w systemach Windows `C:\\Program Files\\IBM\\WebSphere\\profiles\\wp_profile`.

WASDir

Pełna ścieżka do głównego katalogu serwera WebSphere Application Server; wymagana tylko dla systemów AIX, Linux i Solaris. Domyślną ścieżką do głównego katalogu w systemach AIX jest `/usr/IBM/WebSphere/AppServer`, a w systemach Linux i Solaris `/opt/IBM/WebSphere/AppServer`.

WASUser

Nazwa użytkownika będącego administratorem serwera WebSphere Application Server; wymagana tylko wtedy, gdy na serwerze WebSphere Application Server włączono zabezpieczenia globalne.

WASPassword

Hasło użytkownika administracyjnego serwera WebSphere Application Server (jeśli podany został użytkownik).

WPSDir

Pełna ścieżka do katalogu instalacyjnego serwera WebSphere Portal.

WPSUser

Nazwa użytkownika administracyjnego serwera WebSphere Portal.

WSPassword

Hasło podanego użytkownika administracyjnego serwera WebSphere Portal.

WPSHost

Nazwa hosta i numer portu serwera WebSphere Portal.

webServerName

Nazwa definicji serwera WWW, do której należy portal WebSphere Portal.

webServerNodeName

Nazwa węzła serwera WebSphere Application, do którego należy definicja serwera WWW.

-
2. Gdy skrypt zakończy działanie, otwórz przeglądarkę WWW i zaloguj się w konsoli administracyjnej WebSphere na serwerze Network Deployment. Zazwyczaj adres to `http://nazwa_hosta:9060/ibm/console`.
3. Rozwiń sekcję **Serwery** i wybierz opcję **Serwery WWW**.
4. Zaznacz pole **Wybierz** obok właściwego serwera WWW, a następnie kliknij przycisk **Generuj wtyczkę**.
5. Zaznacz pole **Wybierz** obok właściwego serwera WWW, a następnie kliknij przycisk **Propaguj wtyczkę**.
6. Wyloguj się z konsoli administracyjnej.

Migracja z portalu WebSphere Portal do wyszukiwania korporacyjnego

System wyszukiwania korporacyjnego jest wyposażony w kreatora migracji, który ułatwia migrowanie kolekcji i systematyk opartych na regułach z programu IBM WebSphere Portal do systemu wyszukiwania korporacyjnego.

W systemie wyszukiwania korporacyjnego systematykę określa się terminem *drzewa kategorii*. Po migracji systematyki można modyfikować drzewo kategorii oraz reguły klasyfikacji za pośrednictwem konsoli administracyjnej wyszukiwania korporacyjnego. Konsola administracyjna pozwala także po migracji zarządzać kolekcjami.

Aby przeprowadzić migrację systematyk i kolekcji, należy uruchomić kreatora migracji na serwerze indeksowania w wyszukiwaniu korporacyjnym.

Migrowanie kolekcji z portalu WebSphere Portal

Aby przeprowadzić migrację kolekcji i systematyk opartych na regułach z produktu WebSphere Portal do systemu wyszukiwania korporacyjnego, należy przygotować kolekcje w programie WebSphere Portal, a następnie wykonać ich migrację za pomocą kreatora migracji.

Zagadnienia wstępne

Jeśli planowana jest migracja systematyk i kolekcji, należy najpierw wykonać migrację plików systematyk, a później dopiero użyć tej procedury do migrowania kolekcji. Taki sposób postępowania zagwarantuje, że migrowane reguły klasyfikacji będą działały z kolekcjami po migracji.

Procedura

Aby wykonać migrację kolekcji (ewentualnie razem z systematyką) z programu WebSphere Portal do systemu wyszukiwania korporacyjnego:

1. W mechanizmie wyszukiwania środowiska WebSphere Portal zatrzymaj wszystkie procesy przeszukiwaczy w kolekcjach, które mają zostać poddane migracji, i zatwierdź lub odrzuć wszystkie oczekujące dokumenty. (W wyszukiwaniu korporacyjnym nie są obsługiwane oczekujące dokumenty).
2. Dla każdej kolekcji, która ma zostać poddana migracji, użyj portletów mechanizmu wyszukiwania środowiska Portal, aby wyeksportować ustawienia do plików XML.
3. Jeśli serwer indeksowania wyszukiwania korporacyjnego jest zainstalowany na oddzielnym serwerze, skopiuj wyeksportowane pliki XML na ten serwer indeksowania.
4. Na serwerze indeksowania wyszukiwania korporacyjnego zaloguj się jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu OmniFind Enterprise Edition.
5. Przejdź do katalogu instalacyjnego wyszukiwania korporacyjnego:

UNIX: `cd $ES_INSTALL_ROOT/bin`

Windows: `cd %ES_INSTALL_ROOT%\bin`

6. Aby wykonać migrację kolekcji z włączonym mechanizmem bezpieczeństwa, wprowadź poniższą komendę w celu uruchomienia kreatora migracji i kliknij przycisk **Dalej**.

UNIX: ./eswpsmigrate.sh
Windows: eswpsmigrate.bat

7. Aby wyłączyć zabezpieczenia na poziomie kolekcji dla migrowanych kolekcji, wprowadź poniższą komendę, aby uruchomić kreatora migracji i kliknij przycisk **Dalej**.

UNIX: ./eswpsmigrate.sh disable.security
Windows: eswpsmigrate.bat disable.security

8. Wybierz opcję **Migrowanie ustawień wyszukiwania z mechanizmu wyszukiwania Portalu produktu WebSphere Portal**, a następnie kliknij przycisk **Dalej**.
9. Wyszukaj katalog, w którym znajdują się wyeksportowane pliki konfiguracyjne mechanizmu wyszukiwania środowiska Portal, wybierz pliki, które mają zostać poddane migracji, a następnie kliknij przycisk **Dalej**. Wybrane pliki konfiguracyjne zostaną przeanalizowane i sprawdzone pod kątem poprawności.
10. Dla każdej kolekcji wprowadź poniższe informacje, a następnie kliknij przycisk **Dalej**, aby rozpocząć migrowanie kolekcji do wyszukiwania korporacyjnego:
 - Nazwa kolekcji, która ma być używana w wyszukiwaniu korporacyjnym.
 - Kryteria określania ważności dokumentu w kolekcji. Współczynnik oceny statycznej może być wyłączony, oparty na datach dokumentów albo na liczbie odsyłaczy prowadzących do danego dokumentu z innych dokumentów internetowych.
 - Typ klasyfikacji, jaka ma zostać użyta dla tej kolekcji. W razie wybrania opcji braku klasyfikacji migracja danych systematyki do systemu wyszukiwania korporacyjnego nie jest wykonywana. Jeśli wybrano kategorie oparte na regułach, systematyka jest migrowana do systemu wyszukiwania korporacyjnego razem z kolekcją.

Jeśli podczas wykonywania migracji wystąpią błędy, należy przejrzeć plik **MigrationWizard.log** znajdujący się w katalogu, w którym zainstalowany jest Kreator migracji.

Teraz można użyć konsoli administracyjnej wyszukiwania korporacyjnego i skonfigurować dodatkowe ustawienia dla zmigrowanych kolekcji.

Wymaganie: Podczas konfigurowania właściwości przeszukiwacza sieci WWW dla zmigrowanej kolekcji należy określić adres e-mail do otrzymywania komentarzy o przeszukiwaczu oraz nazwę agenta użytkownika (w celu uzyskania pomocy w trakcie konfigurowania właściwości przeszukiwacza sieci WWW należy kliknąć przycisk **Pomoc**).

11. Używając konsoli administracyjnej wyszukiwania korporacyjnego, uruchom procesy przeszukiwania, analizowania i indeksowania dla zmigrowanej kolekcji.
12. Po sprawdzeniu, że zmigrowana kolekcja jest dostępna dla operacji przeszukiwania w wyszukiwaniu korporacyjnym, usuń oryginalną kolekcję w mechanizmie wyszukiwania środowiska Portal.
13. Opcjonalnie: Jako administrator portalu WebSphere Portal wykonaj następujące kroki, o ile chcesz, aby użytkownicy mogli przeszukiwać zmigrowaną kolekcję przy użyciu portalu WebSphere Portal.
 - a. Wykonaj wdrożenie portletu wyszukiwania korporacyjnego w lokalnej instalacji portalu WebSphere Portal.

W klastrze serwerów WebSphere Portal należy dokonać wdrożenia na serwerze, na którym zainstalowany jest program WebSphere Application Server Deployment Manager. Program Deployment Manager rozprowadza portlet wyszukiwania korporacyjnego na pozostałe serwery w klastrze serwerów WebSphere Portal.
 - b. Dodaj portlet wyszukiwania korporacyjnego do odpowiednich stron portalu.

W portalu WebSphere Portal dostęp portletu wyszukiwania jest określany na podstawie uprawnień dostępu do konkretnych stron i portletów. Choć ustawienia konfiguracyjne kolekcji są migrowane, administrator serwera WebSphere Portal musi ręcznie określić pozycję portletu.

Ustawienia zmigrowanej kolekcji

Podczas migrowania kolekcji z portalu IBM WebSphere Portal Kreator migracji tworzy domyślne ustawienia dla kolekcji i przeszukiwaczy.

Jeśli to samo ustawienie istnieje w kolekcjach mechanizmu wyszukiwania środowiska Portal i kolekcjach wyszukiwania korporacyjnego, wówczas podczas migrowania kolekcji do wyszukiwania korporacyjnego w kreatorze zostanie użyte ustawienie mechanizmu wyszukiwania środowiska Portal. W przypadku ustawień, które istnieją tylko w wyszukiwaniu korporacyjnym, w kreatorze użyte zostaną ustawienia określone przez użytkownika podczas migrowania kolekcji lub domyślne ustawienia dla kolekcji w wyszukiwaniu korporacyjnym.

Ustawienia istniejące w mechanizmie wyszukiwania środowiska Portal i wyszukiwaniu korporacyjnym

Dla każdej migrowanej przez użytkownika kolekcji Kreator migracji migruje następujące ustawienia:

- Występujące w kolekcji mechanizmu wyszukiwania środowiska Portal serwisy mechanizmu wyszukiwania.
- Język kolekcji.
- Systematyka (lub drzewo kategorii) i reguły dla kategorii opartych na regułach, o ile w kolekcji wyszukiwania korporacyjnego jest używana klasyfikacja oparta na regułach.

Każdy występujący w kolekcji serwis mechanizmu wyszukiwania środowiska Portal jest konsolidowany z przeszukiwaczem sieci WWW wyszukiwania korporacyjnego. Kreator migracji migruje następujące ustawienia przeszukiwacza:

- początkowe adresy URL
- liczba równoległych procesów przeszukiwacza
- głębokość przeszukiwania
- limit czasu (w sekundach) dla operacji pobierania dokumentu
- domyślny zestaw znaków
- reguły przeszukiwania serwisów WWW (włączenia i wykluczenia)

Ustawienia istniejące wyłącznie w wyszukiwaniu korporacyjnym

Podczas migrowania kolekcji należy określić informacje dotyczące tej kolekcji. Kreator migracji migruje te ustawienia i wykorzystuje domyślne ustawienia kolekcji w wyszukiwaniu korporacyjnym do skonfigurowania każdej migrowanej kolekcji.

Użytkownik może zmodyfikować ustawienia konfiguracyjne kolekcji i przeszukiwacza sieci WWW za pomocą konsoli administracyjnej wyszukiwania korporacyjnego. Wartości ujęte w nawiasy () to domyślne ustawienia dla migrowanych danych.

- nazwa kolekcji
- strategia statycznego oceniania dokumentów
- stosowany typ klasyfikacji (na podstawie reguł lub brak)

- ustawienie określające, czy ma być używana pamięć podręczna wyszukiwania, a także określające liczbę wyników zapytań, które mogą zostać zapisane w pamięci podręcznej wyszukiwania (tak, 5000)
- ustawienie określające, czy ma być monitorowany czas uzyskania odpowiedzi z wynikami wyszukiwania i generowany alert, gdy przekroczony zostanie określony limit (tak, 5 sekund)
- ustawienie określające, czy mają być używane mechanizmy kontroli dostępu (nie)
- harmonogram budowy indeksu przyrostowego
- harmonogram budowy indeksu głównego
- poziom szczegółowości informacji w dzienniku (wszystkie komunikaty)

Kreator migracji tworzy także dla każdego przeszukiwacza sieci WWW następujące ustawienia:

- nazwa przeszukiwacza
- opis przeszukiwacza
- maksymalna długość strony
- ustawienia zabezpieczeń dokumentu
- typy MIME (Multipurpose Internet Mail Extensions) dokumentów, które mają zostać przeszukane, o ile mają one zastosowanie dla używanego typu źródła danych

Przed uruchomieniem nowo zmigrowanego przeszukiwacza sieci WWW należy przejrzeć wszystkie właściwości przeszukiwacza, ustawienia przestrzeni przeszukiwania oraz sprawdzić, czy wszystkie wymagane wartości zostały określone (wymagane pola są oznaczone czerwoną gwiazdką). W szczególności należy sprawdzić, czy został określony adres e-mail do otrzymywania komentarzy o przeszukiwaczu oraz nazwa agenta użytkownika dla przeszukiwacza. W celu uzyskania pomocy podczas konfigurowania właściwości przeszukiwacza sieci WWW należy kliknąć przycisk **Pomoc**.

Plik dziennika Kreatora migracji

Kreator migracji zapisuje wszystkie komunikaty w pliku `WpsMigratorLog.log` znajdującym się w katalogu, w którym zainstalowany jest Kreator migracji.

Dla każdej migrowanej kolekcji w pliku dziennika `WpsMigratorLog.log` zapisywane są wartości wszystkich ustawień odczytanych z mechanizmu wyszukiwania środowiska WebSphere Portal, a także informacje o miejscach w kolekcjach wyszukiwania korporacyjnego, do których ustawienia te zostały zaimportowane.

Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

W instalacji wieloserwerowej można wykonywać komendy z dowolnego serwera w systemie. Zaleca się jednak uruchamianie komend z serwera indeksowania. Serwer indeksowania lub serwer kontrolera może uzyskać dostęp do informacji znajdujących się na wszystkich innych serwerach w systemie.

Większość komend ma następujące formaty:

```
esadmin nazwa_komendy argumenty  
esadmin id_sesji dzialanie -opcja
```

Aby uzyskać więcej informacji na temat wszystkich komend, należy wprowadzić komendę `esadmin help`. Aby uzyskać więcej informacji na temat konkretnej komendy, należy wprowadzić komendę `esadmin dzialanie help`.

Komendy esadmin wyszukiwania korporacyjnego

Poniższe komendy należy wprowadzać w jednym wierszu.

Tabela 9. Komendy **esadmin** wyszukiwania korporacyjnego

Komenda	Opis
<code>esadmin system startall</code>	Uruchamia komponenty wyszukiwania korporacyjnego na wszystkich serwerach wyszukiwania korporacyjnego, takich jak serwer WWW, aplikacja ESSearchServer, Centrum informacyjne na serwerach wyszukiwania, sesje przeszukiwaczy na serwerze przeszukiwaczy oraz sesje indeksowania na serwerze indeksowania. Uruchamia warstwę CCL (Common Communication Layer) na lokalnym serwerze. Aby zrestartować warstwę CCL, należy ręcznie zatrzymać i ponownie uruchomić ją na każdym zdalnym serwerze wyszukiwania korporacyjnego. Przykładowa komenda: <code>esadmin system startall</code>
<code>esadmin system stopall</code>	Zatrzymuje komponenty wyszukiwania korporacyjnego na wszystkich serwerach wyszukiwania korporacyjnego, takich jak serwer WWW, aplikacja ESSearchServer, Centrum informacyjne na serwerach wyszukiwania, sesje przeszukiwaczy na serwerze przeszukiwaczy oraz sesje indeksowania na serwerze indeksowania. Na lokalnym serwerze zatrzymuje warstwę CCL. Aby zrestartować warstwę CCL, należy ręcznie zatrzymać i ponownie uruchomić ją na każdym zdalnym serwerze wyszukiwania korporacyjnego. Przykładowa komenda: <code>esadmin system stopall</code>
<code>esadmin system checkall</code>	Sprawdza status wszystkich komponentów wyszukiwania korporacyjnego na wszystkich serwerach wyszukiwania korporacyjnego. Przykładowa komenda: <code>esadmin system checkall</code>

Tabela 9. Komendy **esadmin** wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
<code>esadmin id_sesji_przeszukiwacza start</code>	<p>Rozpoczyna sesję przeszukiwacza. Komenda ta nie powoduje rozpoczęcia żadnych działań związanych z przeszukiwaniem.</p> <p>Przykładowa komenda: <code>esadmin col1.WEB1.esadmin start</code></p> <p>Przykładowe komunikaty i kody powrotu: FFQC5310I WEBCrawler1 (sid: col1.WEB1.esadmin) nie działa. FFQC5314I Wynik: 0</p>
<code>esadmin id_sesji_przeszukiwacza startCrawl</code>	<p>Uruchamia przeszukiwanie.</p> <p>Przykładowa komenda: <code>esadmin col3.DB21.esadmin startCrawl</code></p> <p>Przykładowe komunikaty i kody powrotu: FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0</p>
<code>esadmin id_sesji_przeszukiwacza pause</code>	<p>Wstrzymuje przeszukiwanie.</p> <p>Przykładowa komenda: <code>esadmin col3.DB21.esadmin pause</code></p> <p>Przykładowe komunikaty i kody powrotu: FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0</p>
<code>esadmin id_sesji_przeszukiwacza resume</code>	<p>Wznawia przeszukiwanie.</p> <p>Przykładowa komenda: <code>esadmin col3.DB21.esadmin resume</code></p> <p>Przykładowe komunikaty i kody powrotu: FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0</p>
<code>esadmin id_sesji_przeszukiwacza stopCrawl</code>	<p>Zatrzymuje przeszukiwanie.</p> <p>Przykładowa komenda: <code>esadmin col3.DB21.esadmin stopCrawl</code></p> <p>Przykładowe komunikaty i kody powrotu: FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0</p>
<code>esadmin id_sesji_przeszukiwacza stop</code>	<p>Kończy sesję przeszukiwacza.</p> <p>Przykładowa komenda: <code>esadmin col3.DB21.esadmin stop</code></p> <p>Przykładowe komunikaty i kody powrotu: FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0</p>

Tabela 9. Komendy **esadmin** wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
<pre>esadmin id_sesji_przeszukiwacza getCrawlerStatus</pre>	<p>Pobiera status przeszukiwacza. Zwracane informacje zależą od tego, czy przeszukiwacz jest przeszukiwaczem sieci WWW czy przeszukiwaczem innych źródeł danych.</p> <p>Przykład dla przeszukiwacza sieci WWW:</p> <pre>esadmin col1.WEB1.esadmin getCrawlerStatus</pre> <p>Możliwe kody powrotu i komunikaty dla przeszukiwacza sieci WWW:</p> <pre>FFQC5303I WebCrawler1 (sid: col1.WEB1.esadmin) już działa. PID: 23650</pre> <p>Przykład dla przeszukiwacza innego niż przeszukiwacz sieci WWW:</p> <pre>esadmin col3.DB21.esadmin getCrawlerStatus</pre> <p>Możliwe kody powrotu i komunikaty dla przeszukiwacza innego niż przeszukiwacz sieci WWW:</p> <pre>FFQC5303I db2crawler (sid: db2col.DB2_96945) już działa. PID: 5936</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 374.</p>
<pre>esadmin id_sesji_przeszukiwacza_źródła_danych getCrawlSpaceStatus esadmin id_sesji_przeszukiwacza_WWW getCrawlStatus -selections wartość</pre>	<p>Pobiera ogólny status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW.</p> <p>Przykładowa komenda:</p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatus</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650</pre> <p>Pobiera ogólny status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW.</p> <p>Przykładowa komenda:</p> <pre>esadmin col1.WEB1.esadmin getCrawlStatus</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 374.</p>

Tabela 9. Komendy **esadmin** wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
<pre>esadmin id_sesji_przeszukiwacza_zrodla_danych getCrawlSpaceStatusDetail -ts id_serwera_docelowego esadmin id_sesji_przeszukiwacza_WWW getCrawlDetailsPerSite -url łańcuch -selections liczba -threshold liczba</pre>	<p>Pobiera szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW. Jeśli użytkownik nie poda opcji określającej serwer docelowy, zwrócone zostaną dane na temat wszystkich serwerów docelowych. Na przykład, gdy przeszukiwacz DB2 przeszukuje bazy danych FOUNTAIN i SAMPLE, a użytkownik nie określi serwera docelowego, zwracany jest status wszystkich tabel w bazach danych FOUNTAIN i SAMPLE.</p> <p>Przykładowa komenda:</p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatusDetail -ts FOUNTAIN</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650</pre> <p>Pobiera szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW.</p> <p>Przykładowa komenda:</p> <pre>esadmin col1.WEB1.esadmin getCrawlDetailsPerSite</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 374.</p>
<pre>esadmin monitor getCollectionParserMonitorStatus -cid id_kolekcji</pre>	<p>Pobiera status analizatora składni.</p> <p>Przykładowa komenda:</p> <pre>esadmin monitor getCollectionParserMonitorStatus -cid coll</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Monitor (node1) (sid: monitor) już działa. PID: 12543</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 374.</p>
<pre>esadmin startMain -cid id_kolekcji</pre>	<p>Uruchamia budowanie głównego indeksu.</p> <p>Przykładowa komenda:</p> <pre>esadmin startMain -cid coll</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 25917 FFQC5314I Wynik: 1117671147056</pre>
<pre>esadmin startDelta -cid id_kolekcji</pre>	<p>Uruchamia budowanie indeksu przyrostowego.</p> <p>Przykładowa komenda:</p> <pre>esadmin startDelta -cid coll</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 4548 FFQC5314I Wynik: 1117670603408</pre>

Tabela 9. Komendy **esadmin** wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
<pre>esadmin controller startIndexBuild -cid <i>id_kolekcji</i> -buildType <i>typ</i> -detectChanges</pre>	<p>Uruchom budowanie indeksu głównego lub przyrostowego i określ, że budowanie powinno być realizowane tylko wtedy, gdy nie zostaną wykryte żadne zmiany, które muszą zostać zastosowane w indeksie.</p> <p>Przykładowa komenda:</p> <pre>esadmin controller startIndexBuild -cid col_1 -buildType main -detectChanges</pre>
<pre>esadmin monitor getCollectionIndexMonitorStatus -cid <i>id_kolekcji</i> -buildType [main delta] -numrecords <i>N_ostatnich_rekordów</i></pre>	<p>Pobiera status budowania indeksu głównego lub indeksu przyrostowego. Opcja <code>numrecords</code> prezentuje <i>N</i> ostatnich rekordów statusu tworzenia indeksu. Pominięcie opcji <code>numrecords</code> powoduje, że zwrócony zostaje status 20 ostatnich operacji tworzenia indeksu.</p> <p>Przykładowa komenda:</p> <pre>esadmin monitor getCollectionIndexMonitorStatus -cid col1 -buildType main -numrecords 4</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Monitor (node1) (sid: monitor) już działa. PID: 12649</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 374.</p>
<pre>esadmin startSearch -cid <i>id_kolekcji</i></pre>	<p>Uruchamia procesy serwera wyszukiwania.</p> <p>Przykładowa komenda:</p> <pre>esadmin startSearch -cid col1</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 25917 FFQC5314I Wynik: 0</pre>
<pre>esadmin stopSearch -cid <i>id_kolekcji</i></pre>	<p>Zatrzymuje procesy serwera wyszukiwania.</p> <p>Przykładowa komenda:</p> <pre>esadmin stopSearch -cid col1</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 15292 FFQC5314I Wynik: 0</pre>

Tabela 9. Komendy **esadmin** wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
<pre>esadmin monitor getCollectionSearchMonitorStatus -cid id_kolekcji esadmin id_sesji_menedzera_wyszukiwania getStatus -cid id_kolekcji</pre>	<p>Pobiera status serwera wyszukiwania.</p> <p>Przykładowa komenda:</p> <pre>esadmin monitor getCollectionSearchMonitorStatus -cid coll</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Monitor (node1) (sid: monitor) już działa. PID: 12649</pre> <p>Zwraca szczegółowe informacje o statusie indeksu wyszukiwania dla kolekcji na danym serwerze wyszukiwania. Dla każdego serwera wyszukiwania istnieje jedna sesja menedżera wyszukiwania. Każda sesja menedżera wyszukiwania jest odpowiedzialna za monitorowanie indeksów wyszukiwania i działanie na nich na konkretnym serwerze wyszukiwania.</p> <p>Przykładowa komenda:</p> <pre>esadmin searchmanager.node1 getStatus -cid coll</pre> <p>Przykładowe komunikaty i kody powrotu:</p> <pre>FFQC5303I Search Manager (node1) (sid: searchmanager.node1) już działa. PID: 15711 FFQC5314I Wynik: PID=18390 CacheHits=3 QueryRate=1 Port=44008 SessionId=coll.runtime.node1 CacheHitRate=0.333 ResponseTime=70 Status=1 SessionName=coll.runtime.node1.1</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu”.</p>

Szczegółowe informacje o komendach dotyczących statusu

Niektóre komendy mogą zwracać bardzo obszerne informacje. W niniejszej sekcji opisane są informacje na temat statusu przeszukiwacza lub przestrzeni przeszukiwania, które mogą zostać zwrócone. Tabela w sekcji “Komendy esadmin wyszukiwania korporacyjnego” na stronie 369 zawiera możliwe informacje zwracane przez komendę esadmin. W tej sekcji omówione są informacje zwracane przez następujące komendy:

- status przeszukiwacza sieci WWW
- status przeszukiwacza innego niż przeszukiwacz sieci WWW
- status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW
- status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacz sieci WWW
- szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW
- szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacz sieci WWW
- status analizatora składni

- status operacji tworzenia indeksu
- status serwera wyszukiwania
- szczegółowy status serwera wyszukiwania

Status przeszukiwacza sieci WWW: Komenda pobierająca status przeszukiwacza sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca status przeszukiwacza sieci WWW może zwrócić następujące informacje:

```
FFQC5314I Result: <?xml version='1.0' encoding='UTF-8'?>
<CrawlerStatus>
<CrawlerRunLevel Value="Running"/>
<CrawlerThreadStateDist Count="4" Total="200">
<CrawlerThreadState State="FETCHING" Count="100"/>
. . .
</CrawlerThreadState State="FETCHING" Count=100>
<ActiveBucketList Count="500">
<ActiveBucket URL="http://w3.ibm.com/"
NumActURLs="355"
NumProcURLs="350"
TimeRem="5" Duration="1195"/>
. . .
</ActiveBucketList>
<CrawlRate Value="75"/>
<RecentlyCrawledURLList Count="40">
<RecentlyCrawledURL URL="http://w3.ibm.com/foo.html"/>
<RecentlyCrawledURL URL="http://w3.ibm.com/foo.html"/>
<NumURLsThisSession Value="160000"/>
</CrawlerStatus>
```

W poniższej tabeli opisano każdy element XML (wraz z jego możliwymi atrybutami), który jest zwracany przez komendę pobierającą status przeszukiwacza sieci WWW:

Tabela 10. Informacje o statusie przeszukiwacza sieci WWW

Element	Atrybuty	Opis
CrawlerStatus	<ul style="list-style-type: none"> • CrawlerThreadStateDist • ActiveBucketList • CrawlRate • RecentlyCrawledURLList • NumURLsThisSession 	Status przeszukiwacza.
CrawlerRunLevel Value	<ul style="list-style-type: none"> • Łańcuch (w języku angielskim) “Not started”: Istnieje sesja przeszukiwacza, ale nie odebrała jeszcze komunikatu powodującego rozpoczęcie przetwarzania dokumentów. • “Started”: Przeszukiwacz rozpoczyna działanie. • “Running”: Przeszukiwacz został zainicjowany i uruchomiony i realizuje proces przeszukiwania. • “Paused”: Przeszukiwacz odebrał polecenie wstrzymania przeszukiwania, ale nie zakończenia działania. • “Stopping”: Przeszukiwacz odebrał sygnał nakazujący zatrzymanie i jest w trakcie zatrzymywania działania. • “Error”: Przeszukiwacz jest w stanie błędu niemożliwego do naprawienia i aby wznowił przeszukiwanie, trzeba go zatrzymać i zrestartować. 	Informacje o działaniu wykonywanym przez przeszukiwacz.

Tabela 10. Informacje o statusie przeszukiwacza sieci WWW (kontynuacja)

Element	Atrybuty	Opis
CrawlerThreadState State	Łańcuch (w języku angielskim)	Działanie wątku przeszukiwacza. Pole to zawiera informacje o bieżącym działaniu wątku.
ActiveBucket	<ul style="list-style-type: none"> URL: Łańcuch (specyfikacja adresu URL) Protokół, host i port przeszukiwanych w danej chwili adresów URL. NumActURLs: Liczba całkowita (dodatnia) Liczba adresów URL w porcji w momencie jej udostępnienia (aktywowania) do przeszukiwania. NumProcURLs: Liczba całkowita (nieujemna) Liczba adresów URL w porcji, które do chwili obecnej zostały przetworzone (przeszukane lub odrzucone). TimeRem: Liczba całkowita Liczba sekund, po których zostanie przekroczony limit czasu dla porcji. Duration: Liczba całkowita (nieujemna) Liczba sekund od chwili aktywowania porcji. 	Bieżąca aktywność określonego serwisu WWW.
CrawlRate	<p>Wartość: Liczba całkowita (nieujemna)</p> <p>Liczba stron przeszukiwanych w ciągu jednej sekundy (we wszystkich porcjach łącznie).</p>	Przepustowość przeszukiwacza.
RecentlyCrawledURL	<p>URL: Łańcuch (specyfikacja adresu URL)</p> <p>Łańcuch określający protokół, hosta, port i plik, który został przeszukany.</p>	Strona, która została ostatnio przeszukana.
NumURLsThisSession	Wartość: Liczba całkowita (nieujemna)	Liczba adresów URL, które zostały przeszukane od chwili rozpoczęcia przeszukiwania przez tę instancję (proces) przeszukiwacza.

Status przeszukiwacza innego niż przeszukiwacz sieci WWW: Komenda pobierająca status przeszukiwacza innego niż przeszukiwacz sieci WWW zwraca informacje w postaci dokumentu XML. Komenda **getCrawlerStatus** dla przeszukiwaczy innych niż przeszukiwacze sieci WWW może zwrócić następujące informacje:

```
FFQC5314I Result: <?xml version='1.0' encoding='UTF-8'?>
<GeneralStatus>
<Status>0</Status>
<StatusMessage>Idle</StatusMessage>
<NumberOfServers>1</NumberOfServers>
<NumberOfCompletedServers>1</NumberOfCompletedServers>
<NumberOfTargets>3</NumberOfTargets>
<NumberOfCompletedTargets>3</NumberOfCompletedTargets>
<NumberOfCrawledRecords>115</NumberOfCrawledRecords>
<RunningThreads>0</RunningThreads>
</GeneralStatus>
```

Poniższa tabel zawiera opis elementów XML i ich atrybutów dla każdego przeszukiwacza wyszukiwania korporacyjnego z wyjątkiem przeszukiwacza sieci WWW. Informacje te zwracane są przez komendę pobierającą status przeszukiwacza.

Tabela 11. Informacje o statusie przeszukiwacza dla przeszukiwaczy NNTP, DB2, JDBC i Notes

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacze DB2 i JDBC	Przeszukiwacz Notes
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd
NumberOfServers	Liczba serwerów NNTP w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.
NumberOfCompletedServers	Liczba przeszukanych serwerów NNTP.	Liczba przeszukanych baz danych.	Liczba przeszukanych baz danych.
NumberOfTargets	Liczba grup dyskusyjnych w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.	Liczba widoków i folderów w przestrzeni przeszukiwania.
NumberOfCompletedTargets	Liczba przeszukanych grup dyskusyjnych.	Liczba przeszukanych tabel.	Liczba przeszukanych widoków i folderów.
NumberOfCompletedRecords	Liczba przeszukanych artykułów.	Liczba przeszukanych rekordów.	Liczba przeszukanych dokumentów.
RunningThreads	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.

Tabela 12. Informacje o statusie przeszukiwacza dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd
NumberOfServers	Liczba serwerów Exchange Server w przestrzeni przeszukiwania.	Liczba serwerów Content Manager w przestrzeni przeszukiwania.	Liczba repozytoriów w przestrzeni przeszukiwania.
NumberOfCompletedServers	Liczba przeszukanych serwerów Exchange Server.	Liczba przeszukanych serwerów Content Manager.	Liczba przeszukanych repozytoriów.
NumberOfTargets	Liczba podfolderów w przestrzeni przeszukiwania.	Liczba typów elementów w przestrzeni przeszukiwania.	Liczba klas w przestrzeni przeszukiwania.
NumberOfCompletedTargets	Liczba przeszukanych podfolderów.	Liczba przeszukanych typów elementów.	Liczba przeszukanych klas elementów.
NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.
RunningThreads	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.

Tabela 13. Informacje o statusie przeszukiwacza dla przeszukiwaczy QuickPlace, Domino Document Manager, przeszukiwaczy systemu plików UNIX i przeszukiwaczy systemu plików Windows

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)

Tabela 13. Informacje o statusie przeszukiwacza dla przeszukiwaczy QuickPlace, Domino Document Manager, przeszukiwaczy systemu plików UNIX i przeszukiwaczy systemu plików Windows (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
StatusMessage	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd
NumberOfServers	Liczba lokalizacji w przestrzeni przeszukiwania.	Liczba bibliotek w przestrzeni przeszukiwania.	Ustalona wartość równa 1.
NumberOfCompletedServers	Liczba przeszukanych lokalizacji.	Liczba przeszukanych bibliotek.	0 lub 1, gdy przeszukane są wszystkie podkatalogi.
NumberOfTargets	Liczba baz danych lokalizacji i baz danych pomieszczeń w przestrzeni przeszukiwania.	Liczba szafek w przestrzeni przeszukiwania.	Liczba podkatalogów w przestrzeni przeszukiwania.
NumberOfCompletedTargets	Liczba przeszukanych baz danych lokalizacji i baz danych pomieszczeń.	Liczba przeszukanych szafek.	Liczba przeszukanych podkatalogów.
NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych plików.
RunningThreads	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.

Tabela 14. Informacje o statusie przeszukiwacza dla przeszukiwaczy WebSphere Portal i Web Content Management

Nazwa elementu i atrybutu	Przeszukiwacz WebSphere Portal	Przeszukiwacz Web Content Management
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczynność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd
NumberOfServers	Liczba serwerów w przestrzeni przeszukiwania.	Liczba serwisów w przestrzeni przeszukiwania.
NumberOfCompletedServers	Liczba przeszukanych serwerów.	Liczba przeszukanych serwisów.
NumberOfTargets	Liczba serwerów w przestrzeni przeszukiwania.	Liczba serwisów w przestrzeni przeszukiwania.
NumberOfCompletedTargets	Liczba przeszukanych serwerów.	Liczba przeszukanych serwisów.
NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.
RunningThreads	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.

Status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW: Komenda pobierająca status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca status przestrzeni przeszukiwania sieci WWW może zwrócić następujące informacje:

Tabela 15. Wartości maski wyboru zwracanej przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Bit maski	Wybiera
1	Liczba stron w składnicy surowych danych.
2	Liczba wykrytych serwisów.
4	Liczba serwisów z systemem DNS.

Tabela 15. Wartości maski wyboru zwracanej przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW (kontynuacja)

Bit maski	Wybiera
8	Liczba serwisów bez systemu DNS.
16	Liczba wykrytych adresów URL.
32	Liczba zapisanych unikalnych stron.
64	Liczba przeszukanych adresów URL.
128	Liczba adresów URL, których nie przeszukano.
256	Liczba zaległych adresów URL.
512	Dystrybucja kodów statusu HTTP.

Wszystkie wartości reprezentują sumy łączne dla wszystkich sesji używających bieżącej wewnętrznej bazy danych:

```
<CrawlStatus>
  <NumPagesInRDS Value="5422386"/>
  <NumSitesDiscovered Value="15332"/>
  <NumSitesWithDNS Value="14832"/>
  <NumSitesWithoutDNS Value="500"/>
  <NumURLsDiscovered Value="15222999"/>
  <NumUniquePagesSaved Value="6234789"/>
  <NumURLsCrawled Value="7800422"/>
  <NumURLsUncrawled Value="7422577"/>
  <NumURLsOverdue Value="14000"/>
  <HTTPCodeDist Count="4" Total="1031000"/>
    <HTTPCode Code="200" Count="1000000"/>
    <HTTPCode Code="301" Count="1000"/>
    <HTTPCode Code="404" Count="10000"/>
    <HTTPCode Code="780" Count="20000"/>
  </HTTPCode Code="780" Count="20000"/>
</CrawlStatus>
```

Zwracane dane zawierają wszystkie lub niektóre z następujących elementów (mogą nie zawierać żadnych):

Tabela 16. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Element	Atrybut	Opis
CrawlerStatus	<ul style="list-style-type: none"> • NumPagesInRDS • NumSitesDiscovered • NumSitesWithDNS • NumSitesWithoutDNS • NumURLsDiscovered • NumUniquePagesSaved • NumURLsCrawled • NumURLsUncrawled • NumURLsOverdue • HTTPCodeDist 	Łatwe do szybkiego uzyskania informacji o łącznym stanie przeszukiwania (wszystkie sesje).
NumPagesInRDS	<p>Wartość: Nieujemna liczba całkowita</p> <p>Liczba stron znajdujących się w danej chwili w obszarze pomostowym składnicy surowych danych (pochodzących tylko od danego przeszukiwacza).</p>	Stopień zapełnienia składnicy surowych danych (tylko w wyniku działania danego przeszukiwacza).

Tabela 16. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwacza sieci WWW (kontynuacja)

Element	Atrybut	Opis
NumSitesDiscovered	Wartość: Nieujemna liczba całkowita Liczba hostów wykrytych dzięki przeszukiwaniu (lub pochodzących z listy początkowej).	Miara pokrycia przez przeszukiwacze przeszukiwanej domeny (liczba hostów).
NumSitesWithDNS	Wartość: Nieujemna liczba całkowita Liczba hostów powiązanych z adresami IP (przetłumaczonymi przez przeszukiwacz w tle).	Miara efektywności pobierania przez przeszukiwacz adresów IP hostów, które zostały wykryte na podstawie ich nazw DNS znajdujących się w adresach URL.
NumSitesWithoutDNS	Wartość: Nieujemna liczba całkowita Liczba hostów niepowiązanych z adresami IP (przetłumaczonymi przez przeszukiwacz w tle).	Miara efektywności pobierania przez przeszukiwacz adresów IP hostów, które zostały wykryte na podstawie ich nazw DNS znajdujących się w adresach URL.
NumURLsDiscovered	Wartość: Nieujemna liczba całkowita Liczba unikalnych adresów URL, które odwiedził przeszukiwacz.	Miara pokrycia przez przeszukiwacze przeszukiwanej domeny (liczba adresów URL).
NumUniquePagesSaved	Wartość: Nieujemna liczba całkowita Liczba unikalnych stron zapisanych do składnicy surowych danych w celu dalszego przetworzenia przez inne komponenty wyszukiwania korporacyjnego.	Wkład danego przeszukiwacza do wielkości indeksu.
NumURLsCrawled	Wartość: Nieujemna liczba całkowita Liczba unikalnych adresów URL, które przeszukał przeszukiwacz.	Miara zdolności przeszukiwacza do przetwarzania danych na całej trasie. Liczba ta jest różna od liczby stron zapisanych do składnicy surowych danych, ponieważ nie wszystkie przeszukane strony zostają do niej zapisane.
NumURLsOverdue	Wartość: Nieujemna liczba całkowita Liczba unikalnych adresów URL kwalifikujących się do przeszukania.	Miara zdolności przeszukiwacza do przemieszczania się w przestrzeni sieci WWW.

Status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacze sieci WWW: Komenda pobierająca status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW zwraca informacje w postaci dokumentu XML. Komenda **getCrawlSpaceStatus** dla przeszukiwaczy innych niż przeszukiwacze sieci WWW może zwrócić następujące informacje:

```
FFQC5314I Result: <?xml version='1.0' encoding='UTF-8'?>
<ServerStatus>
  <Server Name ="FOUNTAIN">
    <Status>5</Status>
    <StatusMessage>Scheduled</StatusMessage>
    <NumberOfTargets>1</NumberOfTargets>
    <NumberOfCompletedTargets>1</NumberOfCompletedTargets>
    <NumberOfErrors>0</NumberOfErrors>
    <StartTime>1118354510512</StartTime>
    <EndTime>1118354514386</EndTime>
    <ScheduleConfigured>2</ScheduleConfigured>
    <ScheduleTime>1118393377000</ScheduleTime>
    <TotalTime>3874</TotalTime>
  </Server>
</ServerStatus>
```

Poniższa tabel zawiera opis elementów XML i ich atrybutów dla każdego przeszukiwacza wyszukiwania korporacyjnego z wyjątkiem przeszukiwacza sieci WWW. Informacje te

zwracane są przez komendę pobierającą status przestrzeni przeszukiwania. W przypadku przeszukiwaczy Notes, gdy poziom agregacji jest równy 0, parametr Server@Name ma wartość nazwa serwera + nazwa bazy danych. Gdy poziom agregacji jest równy 1, parametr Server@Name ma wartość nazwa serwera + nazwa katalogu.

Tabela 17. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2, JDBC i Notes

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacze DB2 i	
		JDBC	Przeszukiwacz Notes
Server@Name	Nazwa serwera grup dyskusyjnych	Nazwa bazy danych	Nazwa bazy danych lub nazwa katalogu
Server/Status	Status: (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd 	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd 	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd
Server/StatusMessage	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd
Server/NumberOfTargets	Liczba grup dyskusyjnych w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.	Liczba widoków i folderów lub katalogów w przestrzeni przeszukiwania.
Server/NumberOfCompletedTargets	Liczba przeszukanych grup dyskusyjnych.	Liczba przeszukanych tabel.	Liczba przeszukanych widoków i folderów lub katalogów.
Server/NumberOfErrors	Liczba błędów.	Liczba błędów.	Liczba błędów.
Server/StartTime	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.
Server/EndTime	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.

Tabela 17. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2, JDBC i Notes (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacze DB2 i JDBC	Przeszukiwacz Notes
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona. 	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona. 	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.
Server/ScheduleTime	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.
Server/TotalTime	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.
Server/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0, 1: <ul style="list-style-type: none"> 0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0). 1: Przeszukiwacz Notes przeszukuje dokumenty w trybie katalogu.

Tabela 18. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Server@Name	Nazwa serwera Exchange Server.	Serwery DB2 Content Manager.	Nazwa repozytorium.
Server/Status	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd

Tabela 18. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Server/StatusMessage	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • 5: Zaplanowane • -1: Błąd
Server/NumberOfTargets	Liczba podfolderów w przestrzeni przeszukiwania.	Liczba typów elementów w przestrzeni przeszukiwania.	Liczba klas elementów w przestrzeni przeszukiwania.
Server/NumberOfCompletedTargets	Liczba przeszukanych podfolderów.	Liczba przeszukanych typów elementów.	Liczba przeszukanych klas elementów.
Server/NumberOfErrors	Liczba błędów.	Liczba błędów.	Liczba błędów.
Server/StartTime	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.
Server/EndTime	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> • 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. • 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. • 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona. 	0, 1, 2 <ul style="list-style-type: none"> • 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. • 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. • 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona. 	0, 1, 2 <ul style="list-style-type: none"> • 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. • 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. • 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.
Server/ScheduleTime	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.
Server/TotalTime	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.
Server/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.

Tabela 19. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, przeszukiwaczy systemu plików UNIX i przeszukiwaczy systemu plików Windows

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Server@Name	Katalog lokalizacji	Baza danych biblioteki	Wartość ustalona: localhost.

Tabela 19. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, przeszukiwaczy systemu plików UNIX i przeszukiwaczy systemu plików Windows (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Server/Status	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd
Server/StatusMessage	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd
Server/NumberOfTargets	Liczba baz danych lokalizacji i baz danych pomieszczeń w przestrzeni przeszukiwania.	Liczba szafek w przestrzeni przeszukiwania.	Liczba podkatalogów w przestrzeni przeszukiwania.
Server/NumberOfCompletedTargets	Liczba przeszukanych baz danych lokalizacji i baz danych pomieszczeń.	Liczba przeszukanych szafek.	Liczba podkatalogów w przestrzeni przeszukiwania.
Server/NumberOfErrors	Liczba błędów.	Liczba błędów.	Liczba błędów.
Server/StartTime	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.
Server/EndTime	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona. 	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona. 	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.
Server/ScheduleTime	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.
Server/TotalTime	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.

Tabela 19. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, przeszukiwaczy systemu plików UNIX i przeszukiwaczy systemu plików Windows (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Server/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.

Tabela 20. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy WebSphere Portal i Web Content Management

Nazwa elementu i atrybutu	Przeszukiwacz WebSphere Portal	Przeszukiwacz Web Content Management
Server@Name	Serwer WebSphere Portal	Adres URL zarodka wyszukiwania Web Content Management
Server/Status	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd
Server/StatusMessage	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd 	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane 5: Zaplanowane -1: Błąd
Server/NumberOfTargets	Liczba serwerów w przestrzeni przeszukiwania.	Liczba serwisów w przestrzeni przeszukiwania.
Server/NumberOf CompletedTargets	Liczba przeszukanych serwerów.	Liczba przeszukanych serwisów.
Server/NumberOfErrors	Liczba błędów.	Liczba błędów.
Server/StartTime	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.
Server/EndTime	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona. 	0, 1, 2 <ul style="list-style-type: none"> 0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza. 1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona. 2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.
Server/ScheduleTime	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.

Tabela 20. Informacja o statusie przestrzeni przeszukiwania dla przeszukiwaczy WebSphere Portal i Web Content Management (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz WebSphere Portal	Przeszukiwacz Web Content Management
Server/TotalTime	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.
Server/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.

Szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW:

Komenda pobierająca szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca szczegółowy status przestrzeni przeszukiwania może zwrócić następujące informacje:

Tabela 21. Wartości maski wyboru zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Bit maski	Wybiera
1	Liczba stron w składnicy surowych danych.
2	Liczba wykrytych serwisów.
4	Liczba serwisów z systemem DNS.
8	Liczba serwisów bez systemu DNS.
16	Liczba wykrytych adresów URL.
32	Liczba zapisanych unikalnych stron.
64	Liczba przeszukanych adresów URL.
128	Liczba adresów URL, których nie przeszukano.
256	Liczba zaległych adresów URL.
512	Dystrybucja kodów statusu HTTP.

Przykładowe zwrócone informacje:

```
<CrawlDetailsPerSite>
  <Site URL=http://w3.ibm.com/">
  <NumURLsDiscovered Value="5422386"/>
  <NumURLsOverdue Value="15332"/>
  <NumURLsCrawled Value="15332"/>
  <NumURLsUncrawled Value="15332"/>
  <NumURLsOverdueBy Threshold="604800" Value="14832"/>
  <NumURLsActivated Value="2200"/>
  <LastActivationTime Value="1076227340"/>
  <LastActivationDuration Value="4300"/>
  <IPAddressList Count="1"/>
  <IPAddress Value="9.205.41.33"/>
</IPAddressList>
<RobotsContent>
  robots content. . .
</RobotsContent>
<HTTPCodeDist Count="4" Total="1031000"/>
  <HTTPCode Code="200" Count ="1000000"/>
  <HTTPCode Code="301" Count ="1000"/>
  <HTTPCode Code="404" Count ="10000"/>
  <HTTPCode Code="780" Count="20000"/>
</HTTPCodeDist>
</CrawlDetailsPerSite>
```

Poniższa tabela zawiera opis każdego pola, które jest zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW:

Tabela 22. Szczegółowe informacje o statusie przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Element	Atrybuty	Opis
CrawlDetailsPerSite	<ul style="list-style-type: none"> • LastActivationTime: • LastActivationDuration: • IPAddressList: • RobotsContent: • HTTPCodeDist: 	Łatwe do szybkiego uzyskania informacji o szczegółowym stanie jednego serwisu.
Site	URL	Adres URL strony głównej serwisu.
NumURLsDiscovered	Wartość	Liczba adresów URL wykrytych w danym serwisie.
NumURLsOverdue	Wartość	Liczba adresów URL kwalifikujących się do ponownego przeszukania w danym serwisie.
NumURLsCrawled	Wartość	Liczba adresów URL przeszukanych w danym serwisie.
NumURLsUncrawled	Wartość	Liczba adresów URL, które nie zostały jeszcze przeszukane w danym serwisie.
NumURLsOverdueBy	<p>Próg, wartość: Liczba całkowita (dodatnia lub ujemna)</p> <p>Wartość wskazująca liczbę adresów URL kwalifikujących się do ponownego przeszukania. Próg określa czas oczekiwania adresów URL na ponowne przeszukanie. Próg wyrażany jest przez różnicę liczby sekund w stosunku do obecnego czasu. Ujemne wartości progu oznaczają, że ponowne przeszukiwanie adresów URL powinno być już się odbyć. Dodatnie wartości progu oznaczają, że ponowne przeszukanie adresów URL dopiero się odbędzie.</p>	Liczba adresów URL zakwalifikowanych do ponownego przeszukania co najmniej pewną liczbę sekund temu albo zakwalifikowanych do ponownego przeszukania za określoną liczbę sekund.
NumURLsActivated	Wartość	Liczba adresów URL wczytanych do pamięci podczas ostatniego skanowania danego serwisu i udostępnionych wątkom przeszukiwacza.
LastActivationTime	Wartość	Liczba sekund, które upłynęły od chwili, w której adresy URL danego serwisu zostały po raz ostatni wczytane do pamięci.
LastActivationDuration	Wartość	Liczba sekund, przez które adresy URL danego serwisu były ostatnio wczytane do pamięci i dostępne dla wątków przeszukiwacza.
IPAddressList	IPAddress	Wszystkie znane adresy IP serwera danego serwisu.
IPAddress	Wartość	Adres IPv4 serwera danego serwisu w notacji z kropkami.
RobotsContent	Tekst	Tekst znajdujący się w pliku robots, jeśli istnieje.
HTTPCodeDist	HTTPCode	Rozkład kodów HTTP z operacji pobierania, które usiłowano wykonać z tego serwisu.

Tabela 22. Szczegółowe informacje o statusie przestrzeni przeszukiwania dla przeszukiwacza sieci WWW (kontynuacja)

Element	Atrybuty	Opis
HTTPCode	Kod: Liczba całkowita Kod statusu HTTP lub inny kod wewnętrzny.	Ile razy dany kod statusu HTTP pojawił się w czasie przeszukiwania tego serwisu.

Szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacze sieci WWW: Komenda pobierająca szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW zwraca informacje w postaci dokumentu XML. Komenda **getCrawlSpaceStatusDetail** dla przeszukiwaczy innych niż przeszukiwacze sieci WWW może zwrócić następujące informacje:

```
FFQC5314I Result: <?xml version='1.0' encoding='UTF-8'?>
<TargetStatus>
  <Target Name ="escmgr.crawlerinstances">
    <Status>2</Status>
    <StatusMessage>Completed</StatusMessage>
    <NumberOfRecords></NumberOfRecords>
    <NumberOfCrawledRecords>117</NumberOfCrawledRecords>
    <NumberOfInsertedRecords>21</NumberOfInsertedRecords>
    <NumberOfUpdatedRecords>45</NumberOfUpdatedRecords>
    <StartTime>1118354510727</StartTime>
    <EndTime>1118354514386</EndTime>
    <AggregationLevel>0<AggregationLevel>
  </Target>
</TargetStatus>
```

Tabela 23. Szczegółowe informacje o statusie przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2, JDBC i Notes

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacze DB2 i JDBC	Przeszukiwacz Notes
Target@Name	Nazwa grupy dyskusyjnej	Nazwa tabeli	Nazwa widoku lub folderu
Target@CrawlType	Nie dotyczy.	0,1 (DB2); 0 (baza danych JDBC) • 0: Przeszukiwanie aktywne (normalne) • 1: Przeszukiwanie pasywne (publikowanie zdarzeń DB2)	0
Target/Status	Status: (0, 1, 2, 3, 4, -1) • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd	Status: (0, 1, 2, 3, 4, -1) • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd	Status: (0, 1, 2, 3, 4, -1) • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd

Tabela 23. Szczegółowe informacje o statusie przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2, JDBC i Notes (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacze DB2 i JDBC	Przeszukiwacz Notes
Target/StatusMessage	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd
Target/NumberOf Records	Numer ostatniego artykułu na serwerze.	Liczba przeszukanych rekordów.	Liczba przeszukanych dokumentów.
Target/NumberOf CompletedRecords	Liczba przeszukanych artykułów.	Liczba przeszukanych rekordów.	Liczba przeszukanych dokumentów.
Target/NumberOf InsertedRecords	Liczba nowo opublikowanych artykułów.	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.
Target/NumberOf UpdatedRecords	Nie dotyczy.	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.
Target/NumberOf DeletedRecords	Nie dotyczy.	Liczba usuniętych rekordów.	Liczba usuniętych rekordów.
Target/StartTime	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.
Target/EndTime	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.
Target/TotalTime	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.
Target/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0, 1: <ul style="list-style-type: none"> • 0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym. • Przeszukiwacz przeszukuje dokumenty w trybie katalogu.
Target/LastUpdatedTime	Nie dotyczy.	Czas ostatniej aktualizacji: <ul style="list-style-type: none"> • 0: Przeszukiwanie aktywne (normalne) • 1: Przeszukiwanie pasywne (publikowanie zdarzeń DB2) 	Nie dotyczy.
Target/LastResetTime	Nie dotyczy.	Czas ostatniego resetowania statystyki: <ul style="list-style-type: none"> • 0: Przeszukiwanie aktywne (normalne) • 1: Przeszukiwanie pasywne (publikowanie zdarzeń DB2) 	Nie dotyczy.

Tabela 24. Szczegółowe informacje o statusie przestrzeni przeszukiwania dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Target@Name	Nazwa podfolderu	Nazwa typu elementu	Nazwa klasy elementu
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd
Target/StatusMessage	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd
Target/NumberOfRecords	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.
Target/NumberOfInsertedRecords	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.
Target/NumberOfUpdatedRecords	Nie dotyczy.	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.
Target/NumberOfDeletedRecords	Nie dotyczy.	Liczba usuniętych rekordów.	Liczba usuniętych rekordów.
Target/StartTime	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.
Target/EndTime	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.
Target/TotalTime	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.
Target/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.
Target/LastUpdatedTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/LastResetTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.

Tabela 25. Szczegółowe informacje o statusie przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, przeszukiwaczy systemu plików UNIX i przeszukiwaczy systemu plików Windows

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Target@Name	Nazwa bazy danych lokalizacji lub bazy danych pomieszczeń	Nazwa bazy danych szafek	Nazwa podkatalogu
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane -1: Błąd 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane -1: Błąd 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane -1: Błąd
Target/StatusMessage	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane -1: Błąd 	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane -1: Błąd 	<ul style="list-style-type: none"> 0: Nieprzeszukane 1: Przeszukiwanie 2: Zakończone (niezaplanowane) 3: Oczekiwanie 4: Wstrzymane -1: Błąd
Target/NumberOf Records	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/NumberOf CompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych plików.
Target/NumberOf InsertedRecords	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.
Target/NumberOf UpdatedRecords	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.
Target/NumberOf DeletedRecords	Liczba usuniętych rekordów.	Liczba usuniętych rekordów.	Liczba usuniętych rekordów.
Target/StartTime	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.
Target/EndTime	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.
Target/TotalTime	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.
Target/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.
Target/LastUpdatedTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/LastResetTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.

Tabela 26. Szczegółowe informacje o statusie przestrzeni przeszukiwania dla przeszukiwaczy WebSphere Portal i Web Content Management

Nazwa elementu i atrybutu	Przeszukiwacz WebSphere Portal	Przeszukiwacz Web Content Management
Target@Name	Nazwa serwera WebSphere Portal	Adres URL zarodka wyszukiwania reprezentującego serwis
Target@CrawlType	0	0
Target/Status	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd
Target/StatusMessage	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd 	<ul style="list-style-type: none"> • 0: Nieprzeszukane • 1: Przeszukiwanie • 2: Zakończone (niezaplanowane) • 3: Oczekiwanie • 4: Wstrzymane • -1: Błąd
Target/NumberOf Records	Nie dotyczy.	Nie dotyczy.
Target/NumberOf CompletedRecords	Łączna liczba przeszukanych rekordów.	Łączna liczba przeszukanych rekordów.
Target/NumberOf InsertedRecords	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.
Target/NumberOf UpdatedRecords	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.
Target/NumberOf DeletedRecords	Liczba usuniętych rekordów.	Liczba usuniętych rekordów.
Target/StartTime	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.
Target/EndTime	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.
Target/TotalTime	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.
Target/AggregationLevel	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.	0: Przeszukiwacz przeszukuje dokumenty w trybie normalnym.
Target/LastUpdatedTime	Nie dotyczy.	Nie dotyczy.
Target/LastResetTime	Nie dotyczy.	Nie dotyczy.

Status analizatora składni: Komenda pobierająca status analizatora składni zwraca informacje w postaci dokumentu XML. Komenda pobierająca status analizatora składni może zwrócić następujące informacje:

```
FFQC5314I Result:
<Monitor Type="Parser">
<ParserStatus>
  <Status>1</Status>
  <State>Parsing</State>
  <NumberOfDocsToBeIndexed>231974</NumberOfDocsToBeIndexed>
  <ParseRate>0</ParseRate>
  <ParseRateMBPerHour>0</ParseRateMBPerHour>
  <NumberOfCpmThreads>3</NumberOfCpmThreads>
  <ParserServiceSession>parserservice.1</ParserServiceSession>
</ParserStatus>
```

```

<CrawlerStatus>
  <Name>WEBCrawler1</Name>
  <Crawlerid>coll.WEB1.esadmin</Crawlerid>
  <Type>WEB</Type>
  <ParserStatus>1</ParserStatus>
  <NumberOfDocsAlreadyParsed>29</NumberOfDocsAlreadyParsed>
</CrawlerStatus>
</Monitor>

```

Poniższa tabela zawiera opis elementów XML z informacjami zwracanymi przez komendę pobierającą status analizatora składni:

Tabela 27. Elementy zwracane przez komendę pobierającą status analizatora składni

Element	Opis
Status	<ul style="list-style-type: none"> 0: Sesja analizatora składni dla tej kolekcji została zatrzymana; 1: Sesja analizatora składni dla tej kolekcji działa.
State	<p>Możliwe są następujące stany: Inicjowanie, Bezczynność, Restart, Analizowanie, Zatrzymane, Wstrzymane, Wznawianie i Brak dostępu do usług analizatora (NoParserServiceIsAvailable).</p> <p>Status Initializing oznacza, że analizator składni zaczyna działanie i inicjuje swój stan.</p> <p>Status Bezczynność oznacza, że analizator składni znajduje się w uśpieniu od <i>N</i> minut i oczekuje na dokumenty od przeszukiwaczy zajmujących się tą kolekcją. Wartością domyślną czasu uśpienia jest 300 sekund.</p> <p>Status Restart wskazuje, że analizator oczekuje na zrestartowanie wirtualnej maszyny języka Java, w której realizuje się analizę składni lub podział na leksemy. Maszyna JVM, w której realizuje się analizę składni lub podział na leksemy, działa w oddzielnej sesji i w niej odbywa się ostateczne przetwarzanie dokumentów.</p> <p>Status Parsing wskazuje, że analizator składni przetwarza dokumenty.</p> <p>Status Wstrzymane oznacza, że praca analizatora składni została wstrzymana przez sesję tworzenia indeksu dla tej kolekcji.</p> <p>Status Wznawianie oznacza, że sesja budowania indeksu dla tej kolekcji zmieniła status analizatora składni ze statusu Wstrzymanie na status Analizowanie.</p> <p>Status NoParserServiceIsAvailable wskazuje, że nie można przetwarzać dokumentów danej kolekcji, ponieważ nie jest dostępna żadna maszyna JVM, w której realizuje się analizę składni lub podział na leksemy. Ten status oznacza, że wszystkie maszyny JVM realizujące analizę składni lub dzielenie na leksemy wykorzystywane są przez inne kolekcje.</p>
NumberOfDocsToBeIndexed	Liczba dokumentów danej kolekcji, które znajdują się w składnicy. Wartość ta obejmuje również dokumenty oznaczone do usunięcia w czasie następnego budowania indeksu.
ParseRate	Szybkość analizy wyrażona jako liczba dokumentów na sekundę.
ParseRateMBPerHour	Szybkość analizy wyrażona w megabajtach na godzinę.
NumberOfCpmThreads	Liczba wątków CPM używanych do przetwarzania dokumentów tej kolekcji przez maszynę JVM, w której realizuje się analizę składni lub podział na leksemy.
ParserServiceSession	Nazwa maszyny JVM służącej do analizy składni lub podziału na leksemy, która przetwarza dokumenty dla danej kolekcji. Pole to jest dostępne tylko wtedy, gdy analizator składni znajduje się w stanie Parsing.
Name	Nazwa przeszukiwacza.
Crawlerid	Identyfikator utworzony przez system dla tego przeszukiwacza.
Type	Typ przeszukiwacza (przeszukiwacz sieci WWW, NNTP, DB2 itp.).

Tabela 27. Elementy zwracane przez komendę pobierającą status analizatora składni (kontynuacja)

Element	Opis
ParserStatus	<ul style="list-style-type: none"> • 0: Dokumenty z tego przeszukiwacza nie są analizowane (sesja analizatora składni została zatrzymana); • 1: Dokumenty z tego przeszukiwacza są analizowane (sesja analizatora składni jest aktywna).
NumberOfDocsAlreadyParsed	Liczba dokumentów z tego przeszukiwacza, która została już przeanalizowana.

Status operacji tworzenia indeksu: Komenda pobierająca status operacji tworzenia indeksu zwraca informacje w postaci dokumentu XML. Komenda pobierająca status operacji tworzenia indeksu może zwrócić następujące informacje:

```
<?xml version="1.0"?>
<Monitor Type="MainIndexHistory" Count="1">
  <IndexStatus Id="1">
    <StartTime>1131987633901</StartTime>
    <Progress>0</Progress>
    <CurrentPhase>0</CurrentPhase>
    <TotalPhase>3</TotalPhase>
    <IndexCopyTime>49822</IndexCopyTime>
    <CurrentServer>0</CurrentServer>
    <TotalServer>0</TotalServer>
    <IndexBuildTime>46158</IndexBuildTime>
    <Status>0</Status>
    <JobID>1131987633899</JobID>
    <MessagesAvailable>>false</MessagesAvailable>
    <StopTime>1131987734199</StopTime>
    <TotalTime>100298</TotalTime>
    <NumberOfDocuments>43</NumberOfDocuments>
  </IndexStatus>
  <CurrentIndexWildcardSupport/>
  <NextIndexWildcardSupport Type="None" Limit="0"/>
  <ScheduleStatus>
    <Status>1</Status>
    <ScheduledTime Enabled="false"></ScheduledTime>
  </ScheduleStatus>
</Monitor>
```

Poniższa tabela zawiera opis każdego elementu XML z informacjami zwracanymi przez komendę pobierającą status operacji tworzenia indeksu:

Tabela 28. Elementy zwracane przez komendę pobierającą status operacji tworzenia indeksu

Element	Opis
IndexStatusId	Identyfikator statusu indeksu.
StartTime	Czas w sekundach od roku 1970 opisujący moment rozpoczęcia tego budowania indeksu. Aby obliczyć bieżący czas reprezentowany przez tę wartość, należy użyć wzoru 1 stycznia 1970 %2B StartTime. Więcej informacji na temat czasu epoki znaleźć można pod adresem http://en.wikipedia.org/wiki/Unix_epoch .
Progress	Procent zaawansowania danego budowania indeksu.
CurrentPhase	<ul style="list-style-type: none"> • 1: faza przepisywania składnicy danych • 2: faza analizy globalnej • 3: faza tworzenia indeksu
TotalPhase	Liczba faz budowania indeksu. Wartość ta jest aktualnie równa 3.
IndexCopyProgress	Procent wykonania kopii indeksu. Proces kopiowania indeksu kopiuje utworzony indeks z serwera tworzenia indeksu na serwery wyszukiwania.
CurrentServer	Serwer wyszukiwania, na który kopiowany jest indeks.

Tabela 28. Elementy zwracane przez komendę pobierającą status operacji tworzenia indeksu (kontynuacja)

Element	Opis
TotalServer	Liczba serwerów wyszukiwania, na które ma zostać skopiowany indeks.
IndexCopyTime	Całkowity czas kopiowania indeksu na wszystkie serwery wyszukiwania.
IndexBuildTime	Całkowity czas wszystkich faz budowania indeksu.
Status	<ul style="list-style-type: none"> • 0: tworzenie i kopiowanie indeksu • -1: niepowodzenie żądania budowania indeksu • 1: trwa tworzenie i/lub kopiowanie indeksu
JobID	Unikalny identyfikator przypisywany każdemu żądaniu budowania indeksu.
MessagesAvailable	Wartość boolowska wskazująca na ewentualną dostępność komunikatów o błędach (w przypadku wystąpienia błędu).
StopTime	Czas zakończenia operacji tworzenia indeksu (wszystkie fazy) oraz jego kopiowania.
TotalTime	Przedział czasu pomiędzy godziną rozpoczęcia i zakończenia.
NumberOfDocuments	Liczba dokumentów w indeksie.
CurrentIndexWildcardSupport	Ustawienie znaków zastępczych, które ma zostać użyte podczas następnej operacji tworzenia indeksu. Możliwe wartości to: None, QueryExpansion lub IndexExpansion.
ScheduleStatus	<ul style="list-style-type: none"> • 0 jeśli dla danej kolekcji i typu indeksu nie aktywowano harmonogramu. • 1 jeśli dla danej kolekcji i typu indeksu aktywowano harmonogram.
ScheduledTimeEnabled	Czas w sekundach od roku 1970 oznaczający moment następnego uruchomienia budowania indeksu dla tej kolekcji i typu indeksu. Aby obliczyć czas reprezentowany przez tę wartość, należy użyć wzoru 1 stycznia 1970 %2B ScheduledTimeEnabled. Więcej informacji na temat czasu epoki znaleźć można pod adresem http://en.wikipedia.org/wiki/Unix_epoch .

Status serwera wyszukiwania: Komenda pobierająca status serwera wyszukiwania zwraca informacje w postaci dokumentu XML. Komenda pobierająca status serwera wyszukiwania może zwrócić następujące informacje:

```
FFQC5314I Result: <?xml version="1.0"?>
<Monitor Type="Search" Count="1">
<SearchStatus Name="Search Manager (node1)" SearchID=
"searchmanager.node1" HostName="myComputer.svl.ibm.com">
<Status>1</Status>
</SearchStatus>
</Monitor>
```

Poniższa tabela zawiera opis elementów XML z informacjami zwracanymi przez komendę pobierającą status serwera wyszukiwania:

Tabela 29. Elementy zwracane przez komendę pobierającą status serwera wyszukiwania

Element	Opis
SearchStatusName	Nazwa i identyfikator sesji menedżera wyszukiwania, która monitoruje i konserwuje indeks wyszukiwania dla danej kolekcji.
HostName	Nazwa hosta serwera, na którym działa indeks wyszukiwania.
Status	<ul style="list-style-type: none"> • 0 jeśli indeks wyszukiwania dla danej kolekcji nie działa. • 1 jeśli indeks wyszukiwania dla danej kolekcji działa.

Szczegółowy status serwera wyszukiwania: Komenda pobierająca status serwera wyszukiwania może zwrócić następujące informacje:

```
FFQC5303I Search Manager (node1) (sid: searchmanager.node1)
już działa. PID: 15711
FFQC5314I Wynik: PID=18390
CacheHits=3
QueryRate=1
Port=44008
SessionId=coll.runtime.node1
CacheHitRate=0.333
ResponseTime=70
Status=1
SessionName=coll.runtime.node1.1
```

Poniższa tabela zawiera opis elementów z informacjami zwracanymi przez komendę pobierającą szczegółowy status serwera wyszukiwania:

Tabela 30. Elementy zwracane przez komendę pobierającą szczegółowy status serwera wyszukiwania

Element	Opis
CacheHits	Liczba wyników pobranych z pamięci podręcznej wyszukiwania.
QueryRate	Liczba zapytań odebranych w ostatnim przedziale czasu. Domyślnie przedział czasu ma długość pięciu minut.
Port	Numer portu używanego przez indeks wyszukiwania do nasłuchiwania i odbierania zapytań.
SessionId	Identyfikator sesji indeksu wyszukiwania danej kolekcji.
CacheHitRate	Liczba wyników pobranych z pamięci podręcznej wyszukiwania wyrażona jako procent wszystkich wyników wyszukiwania.
ResponseTime	Średni czas odpowiedzi (w milisekundach) dla określonego przedziału czasu. (Wartością domyślną jest pięć minut).
Status	<ul style="list-style-type: none"> • 0 jeśli indeks wyszukiwania dla danej kolekcji nie działa. • 1 jeśli indeks wyszukiwania dla danej kolekcji działa.
SessionName	Nazwa sesji indeksu wyszukiwania danej kolekcji.

Kody powrotu komend esadmin

Komendy **esadmin** mogą zwracać następujące kody:

Tabela 31. Kody powrotu komend esadmin

Kod	Name	Opis
0	CODE_ERROR_NONE	Wykonanie komendy zakończyło się pomyślnie.
102	CODE_ERROR_INSTANTIATION_EXCEPTION	Wystąpił błąd podczas tworzenia instancji programu obsługi komendy.
103	CODE_ERROR_ACCESS_EXCEPTION	Podczas tworzenia instancji programu obsługi komendy wystąpił błąd niedozwolonego dostępu.
104	CODE_ERROR_EXECUTE_EXCEPTION	
105	CODE_ERROR_THROWABLE	
106	CODE_ERROR_NO_SUCH_METHOD	
107	CODE_ERROR_INVALID_SESSION	
108	CODE_ERROR_INVALID_PARAMETER	
109	CODE_ERROR_SESSION_NOT_RUNNING	

Uzyskiwanie identyfikatorów sesji

Do wyświetlania listy komponentów wyszukiwania korporacyjnego oraz odpowiadających im identyfikatorów sesji służy komenda `esadmin check`. Poniższa tabela zawiera listę najczęstszych sesji, ich identyfikatorów, serwera, na którym działają, oraz ich stanów.

Tabela 32. Przykłady nazw sesji, serwerów, na których je uruchomiono, identyfikatorów sesji i stanów sesji

Sesja	Serwer, na którym działa sesja	Identyfikator sesji	Stan sesji
configmanager	serwer indeksowania	10433	Started
controller	serwer indeksowania	10464	Started
customcommunication	serwer indeksowania	Nie dotyczy	Nie dotyczy
wykrywanie	serwer indeksowania	10649	Started
monitor	serwer indeksowania	10682	Started
parserservice	serwer indeksowania	10718	Started
resource.node1	serwer indeksowania	10759	Started
samplecpp	serwer indeksowania	10827	Started
sampletest	serwer indeksowania	10857	Started
scheduler	serwer indeksowania	10889	Started
searchmanager.node1	serwer indeksowania	10927	Started
utilities.node1	serwer indeksowania	10384	Started

Pojęcia pokrewne

“Monitorowanie działań związanych z wyszukiwaniem korporacyjnym” na stronie 301

“Tworzenie kopii zapasowej i odtwarzanie systemu wyszukiwania korporacyjnego” na stronie 333



Komunikaty dla wyszukiwania korporacyjnego



Komunikaty dla wyszukiwania korporacyjnego

Zadania pokrewne

“Monitorowanie przeszukiwaczy” na stronie 304

“Uruchamianie systemu wyszukiwania korporacyjnego” na stronie 293

“Zatrzymywanie systemu wyszukiwania korporacyjnego” na stronie 295

“Administrowanie serwerami wyszukiwania pracującymi w trybie autonomicznym” na stronie 298

Rozróżnianie wielkości liter w wyszukiwaniu korporacyjnym

Komponenty wyszukiwania korporacyjnego, takie jak składnia zapytań, szybkie odsyłacze, nazwy pól itd. różnie traktują wielkości liter.

Składnia zapytania

W czasie wyszukiwania różnice w wielkości liter nie są uwzględniane z wyjątkiem podanych niżej przypadków:

Nazwy elementów XML i nazwy ich atrybutów

Wielkości liter są rozróżniane. W wartościach terminów i atrybutów nie uwzględnia się różnic w wielkości liter, nawet w zapytaniach XML. Na przykład w tym dokumencie:

```
<książka>
  <Autor>
    <Nazwisko>Ferdynand</Nazwisko>
    <Kontakt Typ="eMail">ferdi@nand.org</Kontakt Typ>
    <Kontakt Typ="Telefon">+1 408 876 4242</Kontakt Typ>
  </Autor>
</książka>
```

Poniższe zapytania nie zwrócą przedstawionego dokumentu:

- @xmlxp::'autor[Nazwisko ftcontains ("Ferdynand")]'
- @xmlxp::'//kontakt[@typ="eMail"]'
- @xmlf2::'<autor><Nazwisko>Ferdynand</nazwisko></autor>'
- @xmlf2::'<KONTAKT TYP="email">ferdy</kontakt>'

Natomiast poniższe zapytania zwrócą przykładowy dokument:

- @xmlxp::'Autor[Nazwisko ftcontains ("ferdynand")]'
- @xmlxp::'//Kontakt[@Typ="email"]'
- @xmlf2::'<Autor><Nazwisko>ferdynand</Autor><Nazwisko>'
- @xmlf2::'<Kontakt Typ="email">ferdy</Kontakt>'

Listy kontroli dostępu (Access control list - ACL)

Wielkości liter są rozróżniane.

Adresy URL w warunkach docid: i samegroupas:

Wielkości liter są rozróżniane. Jednak w adresach URL w warunkach site: lub url: wielkości liter nie są rozróżniane. Na przykład w dokumencie o adresie <http://www.here.com/HR/>:

- Zapytanie docid:<http://www.here.com/hr> nie zwróci tego dokumentu.
- Z kolei zapytania url:hr i url:HERE zwrócą rozpatrywany dokument.
- Podobnie zapytanie site:HERE.com również zwróci ten dokument.

Identyfikatory kategorii i identyfikatory systematyzujące

Wielkości liter nie są rozróżniane. Na przykład w zapytaniu identyfikator_systematyzujący::identyfikator_kategorii wielkości liter w obu identyfikatorach nie mają żadnego znaczenia. Zapytanie RuleBased::c42 wyszuka kategorię c42 w systematyce opartej na regułach, a jednocześnie wyszuka też kategorię rulebased::C42. Nie można co prawda przeprowadzać wyszukiwania po nazwie kategorii, ale można wyszukiwać po identyfikatorze kategorii.

Zasięgi

Wielkości liter nie są rozróżniane. Na przykład zapytania Scope:BADANIA i scope::badania zwróca dokumenty z zasięgu o nazwie Badania.

Warunki ze znakami wieloznacznymi

Wielkości liter nie są rozróżniane. Na przykład warunek Fer*n*d jest równoznaczny z warunkiem fer*n*d.

Nazwy pól

Wielkości liter nie są rozróżniane. Na przykład zapytania Title:Wydatki, TITLE:wydatki i title:wydatki są sobie równoważne. Wielkości liter nie są rozróżniane we wszystkich nazwach pól, nawet jeżeli wywodzą się one z pliku odwzorowań XML. Trzeba jednak pamiętać, że zewnętrzne źródła mogą uwzględniać różnice wielkości liter w nazwach pól zgodnie ze swoją semantyką zapytań.

Szybkie odsyłacze

Wielkości liter nie są rozróżniane.

Klasyfikatory oparte na regułach

Reguły dotyczące adresów URL uwzględniają różnice wielkości liter, ale reguły dotyczące treści dokumentów - nie.

Zwinięte identyfikatory URI i definicje kryteriów oparte na wzorcach URI

W identyfikatorach URI uwzględniane są różnice wielkości liter, ale już w zwiniętych identyfikatorach URI dotyczących nazw grup nie uwzględnia się wielkości liter. Nie można zdefiniować dwóch zwiniętych identyfikatorów URI dotyczących nazw grup, które różnić się będą wyłącznie wielkością liter.

Nazwy pól

Wielkości liter nie są rozróżniane. Ta reguła dotyczy nazw pól stosowanych w zapytaniach, odwzorowaniach XML, definicjach klas kryteriów oraz wszystkich innych interfejsach administracyjnych, w których używane są nazwy pól. Jeśli w konfiguracji interfejsu nazwa pola zapisana zostanie wielkimi literami lub w sposób mieszany, to system może przekształcić ją do zapisu małymi literami i tak będzie wyświetlana podczas następnego przeglądania konfiguracji. Oprócz tego nazwa pola może być interpretowana przez zewnętrzne źródła w sposób uwzględniający różnice wielkości liter.

Słowniki

Wielkości liter nie są rozróżniane. Ta reguła dotyczy słowników synonimów, słowników słów zatrzymujących, słowników propozycji pisowni i słowników kryteriów.

Nazwy i identyfikatory kolekcji

Wielkości liter są rozróżniane. Jeżeli w administracyjnej funkcji API lub funkcji API wyszukiwania podana zostanie nazwa lub identyfikator kolekcji, to wyszukana zostanie kolekcja o nazwie zawierającej litery o dokładnie takiej samej wielkości. Z drugiej strony, mimo że w identyfikatorach kolekcji rozróżniane są wielkości liter, nie można zdefiniować dwóch kolekcji o identyfikatorach różniących się wyłącznie wielkością liter. To samo ograniczenie dotyczy też nazw przeszukiwaczy oraz identyfikatorów i nazw źródeł danych.

Aplikacje wyszukujące

Wielkości liter są rozróżniane. Nie można zdefiniować dwóch nazw lub identyfikatorów aplikacji, które różnić się będą wyłącznie wielkością liter.

Dokumentacja wyszukiwania korporacyjnego

Dokumentacja produktu OmniFind Enterprise Edition jest dostępna w formacie PDF oraz HTML.

Program instalacyjny produktu OmniFind Enterprise Edition automatycznie instaluje Centrum informacyjne, które zawiera wersje HTML dokumentacji wyszukiwania korporacyjnego. W środowisku z wieloma serwerami Centrum informacyjne instalowane jest na obu serwerach wyszukiwania. Jeśli Centrum informacyjne nie zostanie zainstalowane, kliknięcie przycisku pomocy spowoduje otwarcie Centrum informacyjnego w serwisie WWW firmy IBM.

Aby wyświetlić zainstalowaną dokumentację PDF w wybranym języku, należy przejść do katalogu `główny_katalog_instalacyjny_ES/docs/wersja_językowa/pdf`. Na przykład, aby przejrzeć dokumentację w języku angielskim, należy przejść do katalogu `główny_katalog_instalacyjny_ES/docs/en_US/pdf`.

Aby uzyskać dostęp do dokumentacji PDF we wszystkich dostępnych wersjach językowych, należy przejść do serwisu dokumentacji produktu OmniFind Enterprise Edition, wersja 8.5.

Aby uzyskać dostęp do plików produktu do pobrania, pakietów poprawek, not technicznych oraz Centrum informacyjnego, należy przejść do serwisu OmniFind Enterprise Edition Support.

Poniższa tabela zawiera wykaz dostępnej dokumentacji, nazwy plików oraz ich położenie.

Tabela 33. Dokumentacja dotycząca wyszukiwania korporacyjnego

Tytuł	Nazwa pliku	Położenie
Centrum informacyjne		http://publib.boulder.ibm.com/infocenter/discover/v8r5/
<i>Podręcznik instalowania wyszukiwania korporacyjnego</i>	iiysi.pdf	<code>główny_katalog_instalacyjny_ES/docs/wersja_językowa/pdf/</code>
<i>Quick Start Guide</i> (ten dokument jest dostępny również w postaci drukowanej w języku angielskim, francuskim i japońskim)	OmniFindEE850_qsg_dwuliterowe oznaczenie języka.pdf	<code>główny_katalog_instalacyjny_ES/docs/wersja_językowa/pdf/</code>
<i>Administrowanie wyszukiwaniem korporacyjnym</i>	iiysa.pdf	<code>główny_katalog_instalacyjny_ES/docs/wersja_językowa/pdf/</code>
<i>Programming Guide and API Reference for Enterprise Search</i>	iiysp.pdf	<code>główny_katalog_instalacyjny_ES/docs/en_US/pdf/</code>
<i>Troubleshooting Guide and Messages Reference</i>	iiysm.pdf	<code>główny_katalog_instalacyjny_ES/docs/wersja_językowa/pdf/</code>
<i>Text Analysis Integration</i>	iiyst.pdf	<code>główny_katalog_instalacyjny_ES/docs/wersja_językowa/pdf/</code>
<i>Plug-in for Google Desktop Search</i>	iiysg.pdf	<code>główny_katalog_instalacyjny_ES/docs/wersja_językowa/pdf/</code>

Ułatwienia dostępu

Ułatwienia dostępu pomagają użytkownikom niepełnosprawnym fizycznie, na przykład z upośledzeniem ruchowym lub wzrokowym, efektywnie korzystać z oprogramowania.

Firma IBM dąży do tego, aby dostarczać produkty, które mogą być efektywnie używane przez wszystkich, niezależnie od wieku i sprawności.

Ułatwienia dostępu

Poniższa lista zawiera opis najważniejszych opcji ułatwień dostępu w programie OmniFind Enterprise Edition:

- Operacje wykonywane tylko przy użyciu klawiatury
- Interfejsy używane przez lektory ekranowe

Centrum informacyjne OmniFind Enterprise Edition i pokrewne publikacje są dostosowane do wykorzystywania opcji ułatwień dostępu. Opis opcji ułatwień dostępu możliwych do wykorzystania w centrum informacyjnym zawiera temat http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/topic/com.ibm.classify.nav.doc/dochome/accessibility_info.htm.

Nawigacja przy użyciu klawiatury

Ten produkt korzysta ze standardowych klawiszy nawigacyjnych systemu Microsoft Windows.

Do nawigowania i przechodzenia dalej w programie instalacyjnym produktu OmniFind Enterprise Edition można używać również skrótów klawiszowych.

Tabela 34. Skróty klawiszowe w programie instalacyjnym

Działanie	Skrót
Podświetlenie przełącznika	Klawisz strzałki
Wybór przełącznika	Klawisz Tab
Podświetlenie przycisku	Klawisz Tab
Wybór przycisku	Klawisz Enter
Przejsięcie do kolejnego lub poprzedniego okna albo anulowanie	Podświetlenie odpowiedniego przycisku klawiszem Tab i naciśnięcie klawisza Enter
Deaktywacja aktywnego okna	Ctrl + Alt + Esc

Interfejsy

Interfejsy użytkownika dla konsoli administracyjnej, przykładowej aplikacji wyszukiwającej i programu dostosowującego aplikację wyszukiwającą są interfejsami opartymi na przeglądarce, które można wyświetlać w przeglądarce Microsoft Internet Explorer lub Mozilla FireFox. Listę skrótów klawiszowych i innych funkcji ułatwień dostępu przeglądarki Internet Explorer lub FireFox można znaleźć w pomocy elektronicznej tych przeglądarek.

Dodatkowe informacje na temat ułatwień dostępu

Publikacje produktu OmniFind Enterprise Edition w formacie PDF (Adobe Portable Document Format) można przeglądać przy użyciu programu Adobe Acrobat Reader. Pliki

PDF dostępne są na dysku CD dostarczanym z produktem oraz w sieci WWW pod adresem <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27010938>.

IBM i ułatwienia dostępu

Serwis IBM Human Ability and Accessibility Center zawiera więcej informacji na temat zaangażowania firmy IBM w dostarczanie rozwiązań z zakresu ułatwień dostępu.

Glosariusz terminów związanych z wyszukiwaniem korporacyjnym

W niniejszym glosariuszu zostały zdefiniowane terminy wykorzystywane w interfejsie i dokumentacji wyszukiwania korporacyjnego.

lista kontroli dostępu (access control list, ACL)

W bezpieczeństwie komputerowym: lista powiązana z obiektem identyfikującym wszystkie podmioty posiadające dostęp do obiektu i ich prawa dostępu.

rola administracyjna

Klasyfikacja użytkownika przypisująca mu dostęp.

mechanizm analizy

Patrz mechanizm analizy tekstu.

wyniki analizy

Informacje generowane przez adnotatory. Wyniki analizy są zapisywane w strukturze danych nazywanej wspólną strukturą analizy. Wyniki analizy uzyskane przy użyciu niestandardowych mechanizmów analizy tekstu (adnotatorów) również można udostępnić do wyszukiwania, włączając je do indeksu wyszukiwania korporacyjnego.

adnotacja

Informacja na temat zakresu tekstu. Na przykład adnotacja może wskazywać, że dany zakres tekstu reprezentuje nazwę firmy. W architekturze UIMA (Unstructured Information Management Architecture) adnotacja to szczególny rodzaj struktury cech.

adnotator

Komponent programowy, który wykonuje określone zadania analizy lingwistycznej i generuje oraz zapisuje adnotacje. Adnotator jest komponentem logiki analizy w mechanizmie analizy.

wyszukiwanie boolowskie

Wyszukiwanie, w ramach którego jeden lub więcej terminów wyszukiwania jest połączonych przy użyciu takich operatorów, jak AND, NOT i OR.

klasa kryteriów

Obiekt zawierający specyfikacje, które mogą mieć wpływ na względną ocenę dokumentu w wynikach wyszukiwania.

słowo-kryterium

Słowo, które może mieć wpływ na względną ocenę dokumentu w wynikach wyszukiwania. Podczas przetwarzania zapytań ważność dokumentu zawierającego słowo-kryterium może zostać zwiększona lub zmniejszona, w zależności od predefiniowanego wyniku dla danego słowa.

drzewo kategorii

Hierarchia kategorii.

certyfikat

Cyfrowy dokument wiążący klucz publiczny z tożsamością posiadacza tego certyfikatu, a tym samym umożliwiającą uwierzytelnienie posiadacza certyfikatu. Certyfikaty są wydawane przez ośrodki certyfikacji i są przez nie cyfrowo podpisane.

ośrodek certyfikacji

Zaufana niezależna organizacja lub firma, która wydaje certyfikaty cyfrowe służące do tworzenia podpisów cyfrowych i par kluczy publicznych-prywatnych. Ośrodek certyfikacji gwarantuje, że wydawane certyfikaty są unikalne i że posługujące się nimi podmioty są tymi, za które się podają.

normalizacja znaku

Proces, w ramach którego różne formy znaku (takie jak wielka litera lub znaki diakrytyczne) są sprowadzane do wspólnej postaci.

klityka Słowo, które składniowo funkcjonuje oddzielnie, ale fonetycznie jest połączone z innym słowem. Klityka może być zapisana łącznie lub osobno w stosunku do słowa, z którym jest powiązana. Typowe przykłady klityki to ostatnie człony kontrakcji w języku angielskim (*wouldn't* lub *you're*).

kolekcja

Zbiór źródeł danych oraz opcji do przeszukiwania, analizowania, indeksowania i wyszukiwania w tych źródłach danych.

wspólna struktura analizy (common analysis structure, CAS)

Struktura służąca do przechowywania treści i metadanych dokumentu oraz wszystkich wyników analizy dostarczonych przez mechanizm analizy tekstu. Całość wymiany danych podczas analizy tekstu jest obsługiwana przy użyciu wspólnej struktury analizy.

konsument wspólnej struktury analizy (konsument CAS)

Konsument, który poddaje przetwarzaniu końcowemu wyniki analizy przechowywane we wspólnej strukturze analizy. Konsument indeksuje na przykład zawartość wspólnej struktury analizy w mechanizmie wyszukiwania lub zapełnia relacyjną bazę danych konkretnymi wynikami analizy.

wspólna warstwa komunikacyjna (common communication layer, CCL)

Infrastruktura komunikacyjna jednocząca różne komponenty (kontroler, analizator składni, przeszukiwacz, serwer indeksowania) produktu OmniFind Enterprise Edition.

wyodrębnianie pojęcia

Funkcja analizy tekstu, która identyfikuje istotne elementy słownikowe (na przykład ludzi, miejsca lub produkty) w dokumentach tekstowych, i tworzy listę tych elementów. Patrz także wyodrębnianie kompozycji.

przeźreń przeszukiwania

Zestaw źródeł, które są zgodne z określonymi wzorcami (na przykład adresami URL, nazwami baz danych, ścieżkami w systemie plików, nazwami domen i adresami IP), odczytywanych przez przeszukiwacz w celu pobrania elementów do poindeksowania.

przeszukiwacz

Program, który pobiera dane ze źródeł danych i zbiera informacje, które można wykorzystać do tworzenia indeksów wyszukiwania.

referencja

Szczegółowe informacje uzyskiwane podczas uwierzytelniania i opisujące danego użytkownika, jego wszystkie powiązania z grupami i inne atrybuty tożsamości związane z zabezpieczeniami. Referencje mogą służyć do wykonywania całego szeregu usług, takich jak autoryzacja, kontrola i delegowanie. Na przykład, dane logowania użytkownika (identyfikator i hasło) są referencjami, umożliwiającymi uzyskiwanie dostępu do konta.

niestandardowy mechanizm analizy tekstu

Mechanizm analizy tekstu utworzony przy użyciu pakietu do tworzenia

oprogramowania (SDK) z użyciem architektury UIMA (Unstructured Information Management Architecture), który może zostać dodany do zestawu standardowych mechanizmów analizy tekstu wyszukiwania korporacyjnego (określanych również mianem podstawowych adnotatorów wyszukiwania korporacyjnego). Patrz także mechanizm analizy tekstu.

źródło danych

Dowolne repozytorium danych, z którego można pobrać dokumenty, takie jak sieć WWW, relacyjne i nierelacyjne bazy danych oraz systemy zarządzania treścią.

typ źródła danych

Zgrupowanie źródeł danych według protokołu wykorzystywanego do uzyskiwania dostępu do danych.

składnica danych

Struktura danych, w której dokumenty są przechowywane w postaci przeanalizowanej.

budowanie indeksu przyrostowego

W systemie wyszukiwania korporacyjnego: proces dodawania nowych informacji do istniejącego indeksu. Patrz termin przeciwstawny budowanie indeksu głównego.

usuwanie z kolejki

Służy do usuwania elementów z kolejki.

znak diakrytyczny

Znacznik wskazujący zmianę wartości fonetycznej znaku lub kombinacji znaków.

detektor

Funkcja przeszukiwacza, która określa, które źródła danych są dostępne dla przeszukiwacza w celu pobrania informacji.

nazwa wyróżniająca

Nazwa jednoznacznie identyfikująca pozycję w katalogu. Nazwa wyróżniająca składa się z rozdzielonych przecinkami par atrybut:wartość. Jest to również zestaw par nazwa-wartość (takich jak CN=nazwisko osoby oraz C=kraj lub region), które jednoznacznie identyfikują podmiot w certyfikacie cyfrowym.

model DOM (Document Object Model)

System, w którym dokument o określonej strukturze (taki jak plik XML) jest wyświetlany jako drzewo obiektów, do którego można uzyskać dostęp i aktualizować je programowo.

szafka Domino Document Manager

Baza danych produktu Domino Document Manager służąca do organizowania dokumentów. W szafkach przechowywane są bazy danych Domino.

biblioteka Domino Document Manager

Baza danych produktu Domino Document Manager, która jest punktem wejścia do produktu Domino Document Manager.

Domino Internet Inter-ORB Protocol (DIIOP)

Zadanie serwera działające na serwerze i współpracujące z obiektowym brokerem zapytań Domino w celu umożliwienia komunikacji między apletami Java tworzonymi przy użyciu klas Java produktu Notes i serwera Domino. Użytkownicy przeglądarek i serwery Domino używają protokołu DIIOP do komunikacji i wymiany danych obiektowych.

ocena dynamiczna

Typ oceny, w którym warunki w zapytaniu są analizowane pod kątem wyszukiwanych dokumentów w celu określenia oceny wyników. Patrz także punktacja oparta na tekście. Patrz termin przeciwstawny ocena statyczna.

podsumowywanie dynamiczne

Typ podsumowywania, w którym warunki wyszukiwania są podświetlone, a wyniki wyszukiwania zawierają frazy, które najlepiej odpowiadają pojęciom dokumentu wyszukiwanego przez użytkownika. Patrz termin przeciwstawny podsumowywanie statyczne.

wpisywanie do kolejki

Umieszczenie komunikatu lub elementu w kolejce.

administrator wyszukiwania korporacyjnego

Rola administracyjna, która umożliwia użytkownikowi administrowanie całością systemu wyszukiwania korporacyjnego.

podstawowe adnotatory wyszukiwania korporacyjnego

Zestaw standardowych mechanizmów analizy tekstu służący w wyszukiwaniu korporacyjnym do domyślnego przetwarzania w zakresie analizy dokumentów.

znak zmiany znaczenia

Znak, który wyłącza lub wybiera specjalne znaczenie jednego lub więcej występujących po nim znaków.

zewnętrzne źródło danych

Źródło danych do stowarzyszenia, które nie jest przeszukiwane, analizowane ani indeksowane przez produkt OmniFind Enterprise Edition. Operacje wyszukiwania w zewnętrznych źródłach danych są delegowane do odpowiednich interfejsów API zapytań tych źródeł danych.

ścieżka do cechy

Ścieżka, przy użyciu której sięga się do wartości cechy w strukturze cechy architektury UIMA (Unstructured Information Management Architecture).

struktura cechy

Bazowa struktura danych reprezentująca wynik analizy tekstu. Struktura cechy to struktura typu atrybut-wartość. Każda struktura cechy ma określony typ, a każdy typ ma określony zestaw poprawnych cech lub atrybutów, bardzo podobnie jak w przypadku klasy Java.

wyszukiwanie stowarzyszone

Funkcja wyszukiwania umożliwiająca wyszukiwanie w wielu usługach wyszukiwania i zwracająca skonsolidowaną listę wyników wyszukiwania.

stowarzyszenie

Proces łączenia systemów nazewnictwa, tak aby zagregowany system mógł przetwarzać nazwy złożone obejmujące te systemy nazewnictwa.

pole

Obszar, do którego wprowadzane są poszczególne kategorie danych lub informacje sterujące.

wyszukiwanie dla pól

Zapytanie, które jest ograniczone do konkretnego pola.

tekst swobodny

Tekst nieustrukturyzowany składający się ze słów lub zdań.

wyszukiwanie pełnotekstowe

Wyszukiwanie, w którym wyszukiwany jest termin w postaci dowolnego tekstu.

indeks pełnotekstowy

Struktura danych, która zawiera odwołania do elementów danych w celu umożliwienia wyszukiwania dokumentów zawierających terminy z zapytania.

wyszukiwanie rozmyte

Wyszukiwanie, które zwraca słowa wymawiane w sposób podobny do wymowy słowa będącego warunkiem wyszukiwania.

wyszukiwanie hybrydowe

Połączenie wyszukiwania boolowskiego z wyszukiwaniem pełnotekstowym.

zarządzanie tożsamościami

Zestaw funkcji API wyszukiwania korporacyjnego sterujących dostępem do zabezpieczonych danych i umożliwiających użytkownikom przeszukiwanie kolekcji bez konieczności podawania identyfikatora użytkownika i hasła dla każdego repozytorium w kolekcji.

indeks Patrz indeks pełnotekstowy.

kolejka indeksów

Lista żądań budowania indeksów głównych i przyrostowych.

wyodrębnianie informacji

Typ wyodrębniania pojęć, które automatycznie rozpoznaje w dokumentach tekstowych istotne elementy słownikowe, takie jak nazwy, terminy i wyrażenia.

adres IP

Unikalny adres urządzenia lub jednostki logicznej w sieci korzystającej ze standardu protokołu IP.

Java Database Connectivity (JDBC)

Niezależny od bazy danych przemysłowy standard łączności między platformą Java a szerokim zakresem baz danych. Interfejs JDBC udostępnia - oparty na języku SQL - interfejs API poziomu wywołania, umożliwiający dostęp do baz danych.

JavaScript

Skryptowy język WWW używany w przeglądarkach i na serwerach WWW.

strony JSP (JavaServer Pages)

Technologia skryptowa serwerów umożliwiająca dynamiczne osadzanie kodu języka Java w ramach stron WWW (plików HTML) i jego wykonywanie po udostępnieniu strony w celu zwrócenia klientowi dynamicznej treści.

wirtualna maszyna języka Java (JVM)

Programowa implementacja procesora wykonującego skompilowany kod języka Java (aplety i aplikacje).

Katakana

Zestaw znaków składający się z symboli używanych w jednym z dwóch najczęściej używanych fonetycznych alfabetów japońskich, służący przede wszystkim do fonetycznego zapisu słów w językach obcych.

plik bazy danych kluczy

Patrz plik kluczy.

plik kluczy

W bezpieczeństwie komputerowym: plik zawierający klucze publiczne, klucze prywatne, zaufane ośrodki certyfikacji oraz certyfikaty. Patrz także magazyn kluczy.

magazyn kluczy

Plik kluczy zawierający zarówno klucze publiczne przechowywane w postaci certyfikatów osoby podpisującej, jak i klucze prywatne przechowywane w certyfikatach osobistych.

identyfikacja języka

W wyszukiwaniu korporacyjnym: funkcja wyszukiwania określająca język dokumentu.

lemat Podstawowa postać słowa. Lematy są bardzo istotne w przypadku języków o rozbudowanej fleksji, takich jak język polski.

lematyzacja

Proces określania formy podstawowej i różnych form fleksyjnych danego słowa. Na przykład, wyszukiwanie słowa *mysz* zwróci także dokumenty zawierające słowa takie jak *myszki*, *myszka*, a wyszukiwanie słowa *śpię* zwróci także dokumenty zawierające słowa *śpi*, *śpisz*, *spał* czy *spać*.

powinowactwo leksykalne

Związek między bliskimi znaczeniowo słowami wyszukiwanymi w dokumencie. Powinowactwo leksykalne jest używane do obliczania trafności wyniku.

biblioteka

Obiekt systemowy pełniący rolę katalogu innych obiektów. Patrz również biblioteka Domino Document Manager.

ligatura

Co najmniej dwa znaki połączone tak, że wizualnie stanowią jeden znak. Na przykład, *ff* oraz *ffi* są znakami, które można przedstawić jako ligatury.

Lightweight Directory Access Protocol (LDAP)

Otwarty protokół używający protokołu TCP/IP w celu zapewnienia dostępu do katalogów obsługujących model X.500 i nienarzucający związanych z zasobami wymagań bardziej złożonego protokołu X.500 Directory Access Protocol (DAP). Na przykład, protokołu LDAP można używać do szukania osób, organizacji i innych zasobów w katalogach internetowych lub intranetowych.

wyszukiwanie lingwistyczne

Typ wyszukiwania, w którym dokumenty są przeglądane, pobierane i indeksowane z terminami zredukowanymi do form podstawowych (na przykład termin *myszki* jest indeksowany jako *mysz*) lub rozwiniętymi do form podstawowych (jak w przypadku słów złożonych).

analiza odsyłaczy

Metoda oparta na analizie odsyłaczy hipertekstowych między dokumentami i używana do określania, które strony w kolekcji są istotne dla użytkowników.

lokalny program stowarzyszający

W aplikacji wyszukiwania korporacyjnego: obiekt kliencki utworzony przez interfejsy API wyszukiwania i indeksowania, umożliwiające użytkownikom wyszukiwanie zestawów heterogenicznych kolekcji i uzyskiwanie jednego zestawu wyników wyszukiwania.

lokalizacja Lotus QuickPlace

Miejsce w sieci WWW udostępniane przez produkt Lotus QuickPlace i umożliwiające rozproszonym geograficznie uczestnikom współpracę nad projektami oraz komunikowanie się w trybie z połączeniem w zabezpieczonym obszarze roboczym o określonej strukturze.

pomieszczenie Lotus QuickPlace

Wydzielony obszar lokalizacji Lotus QuickPlace przeznaczony tylko dla autoryzowanych członków o wspólnych zainteresowaniach i potrzebie wspólnej pracy.

budowanie indeksu głównego

W wyszukiwaniu korporacyjnym: proces budowania całego indeksu. Patrz termin przeciwstawny budowanie indeksu przyrostowego.

znak maskujący

Znak, który jest wykorzystywany do reprezentowania opcjonalnych znaków na

początku, w środku i na końcu wyszukiwanego terminu. Znaki maskujące są zwykle używane do znajdowania w indeksie odmian terminu. Patrz również znak zastępczy.

typ MIME

Internetowy standard identyfikowania typu obiektu przesyłanego przez sieć Internet.

monitor

Użytkownik funkcji wyszukiwania korporacyjnego z uprawnieniami do obserwowania procesów na poziomie kolekcji.

znak nowego wiersza

Znak sterujący powodujący przejście o jeden wiersz w dół na wydruku lub podczas wyświetlania.

podział na ciągi n-znakowe

Metoda analizy traktująca jako pojedyncze słowo nakładające się sekwencje podanej liczby znaków, w przeciwieństwie do opartego na kodowaniu Unicode podziału wg znaków odstępu, w którym podział na słowa następuje według znaków spacji.

dyrektywa blokująca wykorzystywanie odsyłaczy

Dyrektywa na stronie WWW z instrukcją dla robotów (takich jak przeszukiwacz sieci WWW), aby nie kontynuowały działania korzystając z odsyłaczy znajdujących się na tej stronie.

dyrektywa blokująca indeksowanie

Dyrektywa na stronie WWW z instrukcją dla robotów (takich jak przeszukiwacz sieci WWW), aby nie uwzględniały treści tej strony w indeksie.

Notes remote procedure call (NRPC)

Mechanizm komunikacyjny produktu Lotus Notes używany do całej komunikacji między produktami Notes.

operator

Użytkownik funkcji wyszukiwania korporacyjnego z uprawnieniami do obserwowania, uruchamiania i zatrzymywania procesów na poziomie kolekcji.

wyszukiwanie parametryczne

Typ wyszukiwania, w którym wyszukiwane są obiekty zawierające liczbowe wartości lub atrybuty, takie jak data, liczba całkowita lub inne typy danych liczbowych z określonego zakresu.

analizator składni

Program interpretujący dokumenty, które zostały dodane do składnicy danych wyszukiwania korporacyjnego. Analizator składni wyodrębnia informacje z dokumentów i przygotowuje je do indeksowania, wyszukiwania i pobierania.

sterownik analizatora składni

W wyszukiwaniu korporacyjnym: usługa, w ramach której dokumenty są dostarczane do usługi analizatora składni. Na każdą kolekcję przypada jeden sterownik analizatora składni. Usługa sterownika analizatora składni danej kolekcji odpowiada analizatorowi składni tej kolekcji w konsoli administracyjnej wyszukiwania korporacyjnego.

usługa analizatora składni

Usługa wyszukiwania korporacyjnego obsługująca całość analizy dokumentów oraz przetwarzania w zakresie analizy tekstu w kolekcjach dokumentów. W każdej chwili działa przynajmniej jedna usługa analizatora składni.

lokalizacja

Wirtualne miejsce widoczne w portalu, w którym osoby i grupy osób spotykają się, aby ze sobą współpracować. W portalu każdy użytkownik ma osobistą lokalizację do pracy prywatnej, a poszczególni użytkownicy i grupy mają dostęp do różnych

lokalizacji współużytkowanych, które mogą być lokalizacjami publicznymi lub zastrzeżonymi. Patrz również lokalizacja Lotus QuickPlace.

ocena popularności

Typ oceny, która zwiększa istniejącą ocenę dokumentu w oparciu o jego popularność.

Portal Document Manager (PDM)

Zapewnia użytkownikom jedno centralne repozytorium dokumentów dla celów współpracy zespołowej. Administratorom daje możliwość efektywnego zarządzania swoimi dokumentami, a także umożliwia im określanie sposobu interakcji użytkowników z informacjami.

archiwum mechanizmu przetwarzania

Plik archiwum zip o rozszerzeniu .pear zawierający mechanizm analizy architektury UIMA (Unstructured Information Management Architecture) oraz wszystkie zasoby wymagane do korzystania z tego mechanizmu na potrzeby niestandardowej analizy w ramach wyszukiwania korporacyjnego.

bliskie wyszukiwanie

Wyszukiwanie tekstowe, które zwraca wynik, gdy dwa lub więcej pasujących terminów występuje w określonej odległości od siebie, na przykład w tym samym zdaniu lub akapicie.

serwer proxy

Serwer działający jako pośrednik dla żądań WWW protokołu HTTP udostępniany przez aplikację lub serwer WWW. Serwer proxy działa jak odpowiednik serwerów treści w przedsiębiorstwie.

szybki odsyłacz

Powiązanie między identyfikatorem URI (Uniform Resource Identifier) a słowami kluczowymi lub frazami.

ocenie

Przypisywanie wartości całkowitej do każdego dokumentu w wynikach wyszukiwania zapytania. Kolejność dokumentów w wynikach wyszukiwania jest oparta na trafności dla danego zapytania. Wyższa ocena oznacza lepszą zgodność. Patrz także ocena dynamiczna i ocena statyczna.

składnica danych surowych

Struktura danych, w której przeszukane dokumenty są przechowywane, zanim zostaną wysłane do analizatora składni. Przeszukiwacze zapisują dane w składnicy danych surowych, natomiast analizator składni odczytuje z niej dane. Dokumenty, które zostały już przeanalizowane, są usuwane ze składnicy danych surowych. Pojęcia tego nie należy mylić ze składnicą danych.

adnotator wyrażeń regularnych

Oprogramowanie wykrywające jednostki informacji w tekście dokumentu, na przykład numery produktów, w oparciu o wyrażenia regularne, opisujące dokładne wzorce wyszukiwane w tekście dokumentu. Jeśli jedno z wyrażeń regularnych jest zgodne z pewnymi częściami tekstu dokumentu, adnotator wyrażeń regularnych tworzy odpowiednie adnotacje odnoszące się do stwierdzonej zgodności lub jej części. Te opatrzone adnotacjami wyrażenia są następnie zapisywane albo w indeksie wyszukiwania korporacyjnego przy użyciu pliku odwzorowań indeksu, albo w bazie danych obsługującej interfejs JDBC przy użyciu pliku odwzorowań bazy danych.

zdalny program stowarzyszający

Program stowarzyszający serwera, stowarzyszający zestaw obiektów dostępnych do wyszukiwania.

Robots Exclusion Protocol

Protokół umożliwiający administratorom serwisów WWW wskazanie robotom odwiedzającym ich serwis, które części tego serwisu nie powinny być odwiedzane.

pomieszczenie

Program umożliwiający użytkownikom tworzenie dokumentów przeznaczonych do czytania przez inne osoby, odpowiadanie na komentarze innych osób oraz przeglądanie statusu i terminów projektu. Użytkownicy mogą również brać udział w rozmowie sieciowej z innymi osobami z tego samego pomieszczenia. Patrz również pomieszczenie Lotus QuickPlace.

kategoria oparta na regule

Kategorie tworzone w oparciu o reguły określające, które dokumenty są powiązane z danymi kategoriami. Można na przykład zdefiniować reguły, według których dokumenty zawierające pewne słowa lub ich niezawierające albo dokumenty zgodne ze wzorcem identyfikatora URI (Uniform Resource Identifier) będą wiązane z konkretnymi kategoriami.

aplikacja wyszukiwująca

W wyszukiwaniu korporacyjnym: program, który przetwarza zapytania, przeszukuje indeks, zwraca wyniki wyszukiwania oraz pobiera dokumenty źródłowe.

pamięć podręczna wyszukiwania

Bufor, który przechowuje dane i wyniki poprzednich żądań wyszukiwania.

mechanizm wyszukiwania

Program, który akceptuje żądanie wyszukiwania i zwraca użytkownikowi listę dokumentów.

pliki indeksu wyszukiwania

Zestaw plików, w których przechowywany jest indeks w mechanizmie wyszukiwania.

wyniki wyszukiwania

Lista dokumentów, które są zgodne z żądaniem wyszukiwania.

Secure Sockets Layer (SSL)

Protokół zabezpieczeń zapewniający poufność komunikacji. Protokół SSL umożliwia aplikacjom działającym w architekturze klient/serwer komunikację w sposób zapobiegający podsłuchiwaniu i fałszowaniu wiadomości oraz manipulowaniu nimi.

token zabezpieczeń

Informacje o tożsamości i zabezpieczeniach, które są używane do autoryzacji dostępu do dokumentów w kolekcji. Różne typy źródeł danych obsługują różne typy tokenów zabezpieczeń. Do przykładów należą role użytkowników, identyfikatory użytkowników, identyfikatory grup oraz inne informacje, które można wykorzystać do kontroli dostępu do treści.

strona listy adresów początkowych

W produkcie WebSphere Portal jest to strona XML zawierająca odsyłacze do stron, które są dostępne w danym portalu. Na podstawie listy adresów początkowych przeszukiwacze identyfikują dokumenty do przeszukania. Strona listy adresów początkowych zawiera również metadane, które razem z przeszukanymi dokumentami są przechowywane w indeksie wyszukiwania korporacyjnego.

początkowy adres URL (Uniform Resource Locator)

Punkt początkowy przeszukiwania.

segmentacja

Podział tekstu na odrębne jednostki leksykalne. Przetwarzanie nie oparte na

słownikach obejmuje podział na ciągi n-znakowe oraz według białych znaków, natomiast obsługa oparta na słownikach obejmuje segmentację według słów, zdań i akapitów oraz lematyzację.

wyszukiwanie semantyczne

Typ wyszukiwania słów kluczowych wykorzystujący analizę lingwistyczną i kontekstową. Patrz także analiza tekstu.

serwlet Program Java działający na serwerze WWW i rozszerzający funkcjonalność serwera przez generowanie dynamicznej treści w odpowiedzi na żądania klienta WWW. Serwlety są powszechnie stosowane do podłączania baz danych do sieci WWW.

szyld Łańcuch następujących po sobie leksemów (słów) wydzielonych ze zdania. Na przykład w zdaniu "To jest pewne bardzo krótkie zdanie" 3-wyrazowymi szyldami (czyli trigramami) są:

To jest pewne
jest pewne bardzo
pewne bardzo krótkie
bardzo krótkie zdanie

Szyldów używa się w lingwistyce statystycznej. Na przykład jeśli dwa różne teksty mają dużo wspólnych szyldów, teksty te prawdopodobnie w pewien sposób są sobie pokrewne.

strona błędu programowego

Typ strony WWW zawierającej informacje wyjaśniające dlaczego nie można wyświetlić żądanej strony WWW. Na przykład, zamiast zwracania samego kodu błędu, serwer HTTP może wyświetlić stronę zawierającą szczegółowe wyjaśnienia dotyczące tego kodu błędu.

ocena statyczna

Typ oceniania, w którym na ocenę wpływają czynniki dotyczące ocenianych dokumentów, takie jak data, liczba odsyłaczy wskazujących dokument itd. Patrz termin przeciwstawny ocena dynamiczna.

podsumowywanie statyczne

Typ podsumowania, w którym wyniki wyszukiwania zawierają określone zapisane podsumowanie z dokumentu. Patrz termin przeciwstawny podsumowywanie dynamiczne.

wyszukiwanie rdzenia

Patrz wyszukiwanie rdzenia słowa.

słowo zatrzymania

Powszechnie używane słowo, takie jak *w*, *na* lub *oraz*, które jest ignorowane przez aplikację wyszukującą.

usuwanie słów zatrzymania

Proces usuwania z zapytania słów ignorowanych w celu pominięcia popularnych słów i zwrócenia bardziej precyzyjnych wyników.

podsumowywanie

Proces włączania nienadmiarowych zdań do wyników wyszukiwania w celu skróconego opisu treści dokumentu. Patrz także podsumowywanie dynamiczne i podsumowywanie statyczne.

słownik synonimów

Słownik umożliwiający użytkownikom wyszukiwanie synonimów dla terminów użytych w zapytaniu podczas wyszukiwania w kolekcji.

systematyka

Klasyfikacja obiektów według grup opartych na podobieństwie. W wyszukiwaniu korporacyjnym systematyka organizuje dane według kategorii i podkategorii. Patrz także drzewo kategorii.

analiza tekstu

Proces wyodrębniania z tekstu znaczenia i innych informacji mający na celu poprawę wydajności pobierania danych w kolekcji. Patrz także wyszukiwanie semantyczne.

mechanizm analizy tekstu

Komponent programowy odpowiedzialny za znajdowanie i reprezentowanie kontekstu i treści semantycznej w tekście.

wartościowanie danych tekstowych

Proces przypisywania dokumentowi wartości całkowitej, która określa trafność dokumentu względem terminów podanych w zapytaniu. Wyższa wartość całkowita oznacza większą zgodność z zapytaniem. Patrz także ocena dynamiczna.

segmentacja tekstu

Patrz segmentacja.

wyodrębnianie kompozycji

Typ wyodrębniania pojęcia, które automatycznie rozpoznaje w dokumentach tekstowych istotne elementy słownikowe w celu wyodrębnienia kompozycji lub tematu dokumentu. Patrz także wyodrębnianie pojęcia.

token Podstawowe jednostki tekstu indeksowane przez wyszukiwanie korporacyjne. Leksemami mogą być słowa języka lub inne jednostki tekstu odpowiednie do indeksowania.

dzielenie na leksemy

Proces dzielenia materiału wejściowego na leksemy.

tokenizer

Program segmentacji tekstu skanujący tekst i określający, czy i kiedy seria znaków może zostać uznana za leksem.

znak końcowy

Znak zajmujący ostatnie miejsce w słowie.

system typów

System typów definiuje typy obiektów (struktury cech), które mogą być wykrywane w dokumencie przez mechanizm analizy tekstu. System typów definiuje wszystkie możliwe struktury cech pod względem typów i cech. W systemie typów można zdefiniować dowolną liczbę różnych typów. System typów jest specyficzny dla domeny i aplikacji.

segmentacja przy użyciu znaków spacji, tabulacji lub nowego wiersza oparta na kodzie Unicode

Metoda dzielenia na leksemy, w której do odróżniania leksemów od znaków separujących stosowane są właściwości znaków w kodowaniu Unicode.

identyfikator URI

Minimalny łańcuch znaków identyfikujący zasób abstrakcyjny lub fizyczny.

adres URL

Unikalny adres zasobu informacyjnego dostępnego w sieci, takiej jak Internet. Adres URL składa się ze skróconej nazwy protokołu, przy użyciu którego odbywa się dostęp do zasobu informacyjnego, oraz z informacji, na podstawie których zasób ten jest odnajdowany za pomocą danego protokołu.

Architektura zarządzania informacjami bez określonej struktury (UIMA)

Architektura firmy IBM definiująca środowisko do implementowania systemów na potrzeby analizy danych bez określonej struktury.

agent użytkownika

Aplikacja, która przegląda sieć WWW i pozostawia o sobie informacje w odwiedzanych serwisach WWW. W ramach wyszukiwania korporacyjnego przeszukiwacz sieci WWW jest agentem użytkownika.

przeszukiwacz sieci WWW

Typ przeszukiwacza eksplorującego sieć WWW przez pobieranie dokumentów WWW i podążanie według odsyłaczy w obrębie kolejnych dokumentów.

wyszukiwanie terminów ważonych

Zapytanie, w którym pewne terminy otrzymują większe znaczenie.

znak zastępczy

Znak, który jest wykorzystywany do reprezentowania opcjonalnych znaków na początku, w środku lub na końcu wyszukiwanego terminu.

wyszukiwanie rdzenia słowa

Proces lingwistycznej normalizacji, w ramach której warianty słowa są redukowane do wspólnej postaci. Na przykład słowa takie jak *połączenia*, *połączeniowy* i *połączony* są redukowane do słowa *połącz*.

XML Path Language (XPath)

Język przeznaczony do jednoznacznej identyfikacji lub obsługi części danych źródłowych XML w celu ich wykorzystania w technologiach związanych z językiem XML, takich jak analizatory składni XSLT, XQuery i XML. Język XPath jest standardem World Wide Web Consortium.

Uwagi i znaki towarowe

Uwagi

Niniejsza publikacja została przygotowana z myślą o produktach i usługach oferowanych w Stanach Zjednoczonych.

IBM może nie oferować w innych krajach produktów, usług lub opcji, omawianych w tej publikacji. Informacje o produktach i usługach dostępnych w danym kraju można uzyskać od lokalnego przedstawiciela firmy IBM. Jakakolwiek wzmianka na temat produktu, programu lub usługi firmy IBM nie oznacza, że może być zastosowany jedynie ten produkt, ten program lub ta usługa firmy IBM. Zamiast nich można zastosować ich odpowiednik funkcjonalny, pod warunkiem, że nie narusza to praw własności intelektualnej firmy IBM. Jednakże cała odpowiedzialność za ocenę przydatności i sprawdzenie działania produktu, programu lub usługi pochodzących od producenta innego niż IBM spoczywa na użytkowniku.

IBM może posiadać patenty lub złożone wnioski patentowe na towary i usługi, o których mowa w niniejszej publikacji. Używanie tego dokumentu nie daje żadnych praw do tych patentów. Wnioski o przyznanie licencji można zgłaszać na piśmie pod adresem:

IBM Director of Licensing
IBM Corporation
North Castle Drive Armonk, NY
10504-1785
U.S.A.

Zapytania dotyczące zestawów znaków dwubajtowych (DBCS) należy kierować do lokalnych działów własności intelektualnej firmy IBM (IBM Intellectual Property Department) lub wysłać je na piśmie na adres:

IBM World Trade Asia Corporation Licensing
2-31 Roppongi 3-chome,
Minato-ku
Tokyo 106-0032, Japan

Poniższy akapit nie obowiązuje w Wielkiej Brytanii, a także w innych krajach, w których jego treść pozostaje w sprzeczności z przepisami prawa miejscowego: FIRMA INTERNATIONAL BUSINESS MACHINES CORPORATION DOSTARCZA TĘ PUBLIKACJĘ W TAKIM STANIE, W JAKIM SIĘ ZNAJDUJE "AS IS" BEZ UDZIELANIA JAKICHKOLWIEK GWARANCJI (W TYM TAKŻE RĘKOJMI), WYRAŻNYCH LUB DOMNIEMANYCH, A W SZCZEGÓLNOŚCI DOMNIEMANYCH GWARANCJI PRZYDATNOŚCI HANDLOWEJ ORAZ PRZYDATNOŚCI DO OKREŚLONEGO CELU LUB GWARANCJI, ŻE PUBLIKACJA NIE NARUSZA PRAW OSÓB TRZECICH. Ustawodawstwa niektórych krajów nie dopuszczają zastrzeżeń dotyczących gwarancji wyraźnych lub domniemanych w odniesieniu do pewnych transakcji, w takiej sytuacji powyższe zdanie nie ma zastosowania.

Informacje zawarte w niniejszej publikacji mogą zawierać nieścisłości techniczne lub błędy drukarskie. Informacje te są okresowo aktualizowane, a zmiany te zostaną ujęte w kolejnych wydaniach tej publikacji. IBM zastrzega sobie prawo do wprowadzania ulepszeń i/lub zmian w produktach i/lub programach opisanych w tej publikacji w dowolnym czasie, bez wcześniejszego powiadomienia.

Jakiegokolwiek wzmianki na temat stron internetowych nie należących do firmy IBM zostały podane jedynie dla wygody użytkownika i nie oznaczają, że firma IBM w jakikolwiek sposób firmuje te strony. Materiały dostępne na tych stronach nie są częścią materiałów opracowanych dla tego produktu IBM, a użytkownik korzysta z nich na własną odpowiedzialność.

IBM ma prawo do korzystania i rozpowszechniania informacji przysłanych przez użytkownika w dowolny sposób, jaki uzna za właściwy, bez żadnych zobowiązań wobec ich autora.

Licencjodawcy tego programu, którzy chcieliby uzyskać informacje na temat programu w celu: (i) wdrożenia wymiany informacji między niezależnie utworzonymi programami i innymi programami (łącznie z tym opisywanym) oraz (ii) wykorzystywania wymienianych informacji, powinni skontaktować się z:

IBM Corporation
J46A/G4
555 Bailey Avenue
San Jose, CA 95141-1003
U.S.A.

Informacje takie mogą być udostępnione, o ile spełnione zostaną odpowiednie warunki, w tym, w niektórych przypadkach, uiszczenie odpowiedniej opłaty.

Licencjonowany program opisany w tym dokumencie oraz wszystkie inne licencjonowane materiały dostępne dla tego programu są dostarczane przez IBM na warunkach określonych w Umowie IBM z Klientem, Międzynarodowej Umowie Licencyjnej IBM na Program lub w innych podobnych umowach zawartych między IBM i użytkownikami.

Wszelkie dane dotyczące wydajności zostały zebrane w kontrolowanym środowisku. W związku z tym rezultaty uzyskane w innych środowiskach operacyjnych mogą się znacząco różnić. Niektóre pomiary mogły być dokonywane na systemach będących w fazie rozwoju i nie ma gwarancji, że pomiary te wykonane na ogólnie dostępnych systemach dadzą takie same wyniki. Niektóre z pomiarów mogły być estymowane przez ekstrapolację. Rzeczywiste wyniki mogą być inne. Użytkownicy powinni we własnym zakresie sprawdzić odpowiednie dane dla ich środowiska.

Informacje dotyczące produktów innych firm zostały uzyskane od dostawców tych produktów z opublikowanych przez nich zapowiedzi lub innych powszechnie dostępnych źródeł. Firma IBM nie testowała tych produktów i nie może potwierdzić dokładności pomiarów wydajności, kompatybilności ani żadnych innych danych związanych z tymi produktami. Pytania dotyczące możliwości produktów innych firm należy kierować do dostawców tych produktów.

Jakiegokolwiek wzmianki na temat kierunków rozwoju firmy IBM mogą ulec zmianie lub anulowaniu bez uprzedzenia i dotyczą jedynie ogólnych celów i założeń.

Wszelkie ceny podawane przez IBM są propozycjami cen detalicznych; ceny te są aktualne i podlegają zmianom bez wcześniejszego powiadomienia. Ceny podawane przez dealerów mogą być inne.

Informacje te podano tylko w celu planowania. Wszelkie podane tu informacje mogą zostać zmienione zanim opisywane produkty staną się dostępne.

Publikacja ta zawiera przykładowe dane i raporty używane w codziennej działalności biznesowej. W celu kompleksowego zilustrowania tej działalności podane przykłady

zawierają nazwy osób, firm i ich produktów. Wszystkie te nazwiska/nazwy są fikcyjne i jakakolwiek ich zbieżność z prawdziwymi nazwiskami/nazwami jest całkowicie przypadkowa.

LICENCJA NA PRAWA AUTORSKIE:

Niniejsza publikacja zawiera przykładowe aplikacje w kodzie źródłowym, ilustrujące techniki programowania w różnych systemach operacyjnych. Użytkownik może kopiować, modyfikować i rozpowszechniać te programy przykładowe w dowolnej formie bez uiszczania opłat, w celu rozbudowy, użytkowania, handlowym lub w celu rozpowszechniania aplikacji zgodnych z aplikacyjnym interfejsem programowym dla tego systemu operacyjnego, dla którego napisane były programy przykładowe. Programy przykładowe nie zostały gruntownie przetestowane. Firma IBM nie może zatem gwarantować lub sugerować niezawodności, użyteczności i funkcjonalności tych programów.

Każda kopia lub dowolna część programów przykładowych, albo też dowolna praca pochodna, musi zawierać poniższą informację o prawach autorskich:

© (nazwa_firmy) (rok). Części niniejszego kodu pochodzą z programów przykładowych firmy IBM Corp. © Copyright IBM Corp. _rok_lub_lata_. Wszelkie prawa zastrzeżone.

Częściami niniejszego produktu są:

- Oracle® Outside In Content Access, Copyright © 1992, 2008, Oracle. Wszelkie prawa zastrzeżone.
- Materiały licencjonowane IBM XSLT Processor Licensed Materials - własność firmy IBM © Copyright IBM Corp., 1999-2008. Wszelkie prawa zastrzeżone.

Znaki towarowe

Informacje na temat znaków towarowych firmy IBM można znaleźć pod adresem <http://www.ibm.com/legal/copytrade.shtml>.

Poniższe nazwy są znakami towarowymi lub zastrzeżonymi znakami towarowymi innych firm:

Adobe, Acrobat, Portable Document Format (PDF), PostScript i wszystkie pokrewne znaki towarowe Adobe są znakami towarowymi lub zastrzeżonymi znakami towarowymi Adobe Systems Incorporated w Stanach Zjednoczonych i/lub innych krajach.

Intel, logo Intel, Intel Inside, logo Intel Inside, Intel Centrino, logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium i Pentium są znakami towarowymi lub zastrzeżonymi znakami towarowymi firmy Intel Corporation lub firm zależnych w Stanach Zjednoczonych i innych krajach.

Java i wszystkie pokrewne znaki towarowe Java są znakami towarowymi Sun Microsystems, Inc. w Stanach Zjednoczonych i/lub innych krajach.

Linux jest zastrzeżonym znakiem towarowym Linusa Torvaldsa w Stanach Zjednoczonych i/lub innych krajach.

Microsoft, Windows, Windows NT i logo Windows są znakami towarowymi firmy Microsoft Corporation w Stanach Zjednoczonych i/lub innych krajach.

UNIX jest zastrzeżonym znakiem towarowym The Open Group w Stanach Zjednoczonych i innych krajach.

Nazwy innych firm, produktów i usług mogą być znakami towarowymi lub znakami usług firm trzecich.

Indeks

A

- adapter pętli zwrotnej, konfiguracja 27
- administrator kolekcji
 - konfiguracja roli 259
 - opis 258
- administrator wyszukiwania korporacyjnego
 - konfiguracja roli 259
 - opis 258
 - zmiana hasła na pojedynczym serwerze 20
 - zmiana hasła na wielu serwerach 21
- AdminLinkBarInclude.jsp, plik 348
- adnotatory 143
- adresy IP
 - obsługa w protokole IPv6 28
- Adresy IP
 - adapter pętli zwrotnej 27
 - obsługa podwójnych 27
- adresy IP, zmienianie 25
- agenty użytkownika 89
- aktywne serwisy WWW, monitorowanie 305, 307
- alerty
 - czasy odpowiedzi wyszukiwania 324
 - indeksowane dokumenty 324
 - konfiguracja serwera SMTP 328
 - limity indeksu 301
 - na poziomie kolekcji 301, 324
 - na poziomie systemu 325
 - opcje poczty elektronicznej 324, 325
 - opis 323
 - otrzymywanie wiadomości e-mail 328
 - przeszukane dokumenty 324
 - wolna pamięć na serwerach 325
- alerty o wolnej pamięci 325
- analiza globalna
 - analiza tekstu zakotwiczenia 255, 271
 - opis 7
 - wykrywanie duplikatów dokumentów 184, 255, 260
- analiza tekstu
 - mechanizmy analizy tekstu 144, 145
 - odwzorowywanie elementów XML 146
 - wspólne struktury analizy 147, 149
- analiza tekstu zakotwiczenia
 - analiza globalna 271
 - indeksowanie dokumentów 272
 - opis 255
 - zabezpieczenia kolekcji 271
- analizator składni ASCII 158
- analizatory składni
 - analiza terminów złożonych 150
 - analizator składni ASCII 158
 - monitorowanie 313
 - nieznane typy dokumentów 158
 - obsługiwane języki 167
 - obsługiwane typy dokumentów Stellent 163
 - opis 5, 131
 - pliki bez rozszerzeń 158
 - podział na ciągi n-znakowe 170

- analizatory składni (*kontynuacja*)
 - przetwarzanie lingwistyczne 167
 - reguły wymiany HTML 158, 159
 - rodzime wyszukiwanie XML 151
 - rozpoznawanie języków 168
 - rozpoznawanie strony kodowej 169
 - status systemu 313
 - typy analizowanych dokumentów 155
 - typy dokumentów dla analizatorów składni Stellent 160
 - typy dokumentów dla usług analizatora składni 155, 157
 - uruchamianie 313
 - usuwanie znaków nowego wiersza 170
 - usuwanie znaków spacji, tabulacji i nowego wiersza 170
 - wątki 150
 - wybór typu analizatora 153
 - wykrywanie formatu dokumentów 153
 - zadania analizy danych 5
 - zatrzymywanie 313
- aplikacja ESSearchServer
 - restartowanie 238, 239
 - wyłączenie zabezpieczenia 288
- aplikacje korporacyjne
 - aplikacja ESSearchServer 288
 - ESAdmin, aplikacja 288
 - ESSearchApplication, aplikacja 288
- aplikacje wyszukujące
 - dostosowywanie 239
 - identyfikatory aplikacji 260
 - niestandardowe 219
 - opis 12
 - powiązanie z kolekcjami 220
 - powiązanie ze źródłami zewnętrznymi 254
 - przykład 219, 220
 - uruchamianie 246
 - uzyskiwanie dostępu 246
 - zabezpieczenia na poziomie kolekcji 260
- automatyczne rozpoznawanie języki 168
- strony kodowe 169
- autoryzacja, opis 257

B

- banner.jspf 354
- banner_searchControl.jspf 354
- bazy danych DB2
 - dostęp jako źródło zewnętrzne 251
 - dostęp za pomocą przeszukiwaczy DB2 47
 - dostęp za pomocą przeszukiwaczy JDBC 64
- budowanie indeksów 173
 - opis 173
 - planowanie 175
 - równoległe 176
 - startIndexBuild, komenda 177
 - status systemu 315

- budowanie indeksów (*kontynuacja*)
 - uruchamianie 315
 - współbieżne 176
 - wykrywanie zmian 177
 - zatrzymywanie 179, 315

C

- ccl.properties, plik 25
- Centrum wyszukiwania portalu WebSphere
 - opis 343
 - wersja 6, integracja z systemem wyszukiwania korporacyjnego 353
- chiński
 - podział na ciągi n-znakowe 170
 - usuwanie znaków nowego wiersza 170
- ciąg znaków kontekstu zabezpieczeń użytkownika 265
- config.properties, plik 276, 277
- dostosowywanie 239
- edycja 238
- klonowanie 241
- opisy właściwości 222
- cookie, przeszukiwanie sieci WWW
 - format 104
 - konfigurowanie 105
 - opis 104
- cookies.ini, plik
 - format 104
 - konfigurowanie 105
 - opis 104
- crawl.rules, plik 105
- crawler_rdb_plugin.xml, plik 68
- czasy odpowiedzi wyszukiwania
 - alerty o 324
 - monitorowanie 316
- często zadawane zapytania
 - obliczanie 317
- często zadawane zapytania, monitorowanie 316

D

- daty przeszukanych dokumentów
 - konfigurowanie przeszukiwaczy 108
- Default, aplikacja wyszukująca 238
- Default.jsp, plik 348, 354
- distinctRecentQueryCheck, parametr 317
- dokumentacja
 - HTML 401
 - PDF 401
 - znajdowanie 401
- dokumentacja wyszukiwania korporacyjnego w formacie HTML 401
- dokumentacja wyszukiwania korporacyjnego w formacie PDF 401
- dokumenty HTML
 - analizowanie 158, 159
 - przeszukiwanie 141
 - reguły wymiany 158, 159

- dokumenty usunięte
 - konfiguracja pliku dziennika 320
 - opis 319
 - pliki dzienników 321
 - raporty na temat 320
- dokumenty XML
 - przeszukiwanie 138
 - rodzime wyszukiwanie XML 151
- domeny Lotus Domino 282
- domeny Windows 285
- drzewa kategorii
 - migrowanie z portalu WebSphere Portal 365
 - opis 134
- Dublin Core, elementy 141
- dyrektywy blokujące indeksowanie
 - konfigurowanie 108
 - opis 107
- dyrektywy blokujące wykorzystywanie
 - odsyłaczy
 - konfigurowanie 108
 - opis 107

E

- edycja
 - aplikacje nasłuchiwania danych 116
 - kolekcje 34
 - przestrzeń przeszukiwania 40
 - właściwości aplikacji wyszukującej 222, 238
 - właściwości przeszukiwacza 40
- elementy XML
 - odwzorowywanie na pola wyszukiwania 138
 - odwzorowywanie na wspólną strukturę analizy 146
 - przeszukiwanie 137, 146
- es.cfg, plik 20, 21, 25, 27, 247, 276, 277
- es.search.provider.jar, plik 350, 358
- es.security.jar, plik 345, 350, 358
- es.wp5.install.jar, plik 345
- es.wp6.install.jar, plik 350, 358
- ES_INSTALL_ROOT, opis 20, 21
- ES_NODE_ROOT, opis 20, 21
- es_special_field.default_field, pole zastrzeżone 216
- es_special_field.default_metadata_field, pole zastrzeżone 216
- es_special_field.regular_text, pole zastrzeżone 216
- esadmin startSearch, komenda 298
- esadmin stopIndex, komenda 179
- esadmin stopSearch, komenda 298
- esadmin system startall, komenda 369
- esadmin system stopall, komenda 369
- ESAdmin, aplikacja
 - logowanie 19
 - wyłączanie zabezpieczenia 288
- esapi.jar, plik 345, 350, 358
- esbackup.bat, skrypt 334
- esbackup.sh, skrypt 334
- eschangeproxy, komenda 248
- eschangepw, skrypt 20, 21
- eschangetrustpw, komenda 247
- eschangewaspw, komenda
 - konfiguracja jednoserverowa 276

- eschangewaspw, komenda (*kontynuacja*)
 - konfiguracja wieloserwerowa 277
- escrm.sh, skrypt 57
- escrm.vbs, skrypt 59
- escrnote.sh, skrypt 76
- escrvbr.sh, skrypt 45
- escrvbr.vbs, skrypt 46
- ESPACServer.ear, plik 345, 350, 358
- esrestore.bat, skrypt 335
- esrestore.sh, skrypt 335
- ESSearchAdapter.ear, plik 345
- ESSearchApplication, aplikacja
 - config.properties, plik 222, 238
 - uruchamianie 246
 - wyłączanie zabezpieczenia 288
- ESSearchPortlet.war, plik 345, 350, 358
- ESSearchRegistrationPortlet.war, plik 345

F

- firewalle, przeszukiwanie dokumentów
 - serwera Exchange 281
- followindex.rules, plik
 - konfigurowanie 108
 - opis 107
- fragmenty XML, rodzime wyszukiwanie XML 151

G

- global.rules, plik 105
- globalna przestrzeń przeszukiwania sieci WWW 105
- głębokość ścieżki URL 93
- główne indeksy
 - opis 7

H

- hasło administratora
 - zmiana na pojedynczym serwerze 20
 - zmiana na wielu serwerach 21
- hasło, administrator wyszukiwania korporacyjnego 20, 21
- historia czasu odpowiedzi, monitorowanie 316
- HTTPS, konfiguracja serwera wyszukiwania 247

I

- identyfikator kolekcji, reguły składni 33
- identyfikatory aplikacji 260
- identyfikatory kolekcji 36
- identyfikatory sesji, wyszukiwanie korporacyjne 369
- identyfikatory URI
 - formaty w wyszukiwaniu korporacyjnym 120
 - reguły kategorii 132, 136
 - szybkie odsyłacze 203
 - usuwanie z indeksu 192
 - wpływanie na wyniki statyczne 211, 212
 - wświetlanie szczegółów o 302
 - zasięgi 188, 189

- identyfikatory URI (*kontynuacja*)
 - zwinęte w wynikach wyszukiwania 190, 191

indeksy

- alerty o 324
- działanie znaków wieloznacznych 187
- formaty URI 120
- monitorowanie 315
- opis 7, 173
- planowanie 174
- równoległe operacje budowania 176
- startIndexBuild, komenda 177
- tekst zakotwiczenia 272
- usuwanie identyfikatorów URI 179, 192
- usuwanie z kolejki 315
- włączanie harmonogramu 175, 315
- współbieżne operacje budowania 176
- wykrywanie zmian 177
- wyłączanie harmonogramu 175, 315
- zasięgi 179, 188
- zmiana harmonogramu 175
- znaki wieloznaczne 179, 185, 187
- zwinęte identyfikatory URI 179, 190, 191

indeksy główne

- opis 173
- planowanie 174
- współbieżne operacje budowania 176
- wykrywanie zmian 177

indeksy przyrostowe

- opis 173
- planowanie 174
- współbieżne operacje budowania 176
- wykrywanie zmian 177

integracja z produktem WebSphere Portal

- es.wp5.install.jar, plik 345
- es.wp6.install.jar, plik 350, 358
- Lotus Quickr 343
- opis 343
- skrypty konfiguracyjne 344
- system klastrowy 358
- Web Content Management 343
- wp5_install, skrypt 345
- wp6_cluster_install, skrypt 358
- wp6_install, skrypt 350

interfejsy API

- opis 11
- wyszukiwanie i indeksowanie 11, 219
- IOCP, konfiguracja przeszukiwacza 81

J

japoński

- podział na ciągi n-znakowe 170
- usuwanie znaków nowego wiersza 170

JavaScript, obsługa przez przeszukiwacze sieci WWW 92

JDBC, źródła zewnętrzne

- edycja 251
- konfiguracja 251
- sterowniki JDBC 251
- usuwanie 251

języki

- automatyczne rozpoznawanie 168
- kody dwuznakowe 167
- obsługiwane 167, 168
- przeszukiwanie 167

języki ideograficzne 167
języki z tekstem złożonym 167
języki ze zwykłym tekstem 167

K

kategorie
drzewa kategorii 134
formaty URI 120
migrowanie z portalu WebSphere Portal 365
oparte na regułach 132, 135
opis 132
przeszukiwanie 132
tworzenie 136
typ klasyfikacji 135
zagnieżdżone podkategorie 134
kategorie oparte na regułach
opis 132
tworzenie 136
wybór typu klasyfikacji 135
klastry
Portal WebSphere Portal 358
klasy kryteriów
konfiguracja 214, 215
odzworowywanie pól na 214
opis 213
wartości domyślne 216
wykrywanie duplikatów dokumentów 213
zapytania o niskiej znajdawalności 213, 216
zapytania o wysokiej znajdawalności 213, 216
klonowanie
aplikacje wyszukujące 241
przeszukiwacze 37
kodowanie jednobajtowe 169
kodowanie Unicode 169
kodowanie wielobajtowe 169
kody powrotu, wyszukiwanie korporacyjne 369
kody statusu HTTP
odebrane przez przeszukiwacze sieci WWW 308
raport o przeszukiwaczu sieci WWW 308
kolejka indeksów 315
kolekcje
analizowanie 131
domyślne ustawienia migracji 367
edycja 34
MigrationWizard.log, plik 365
migrowanie z portalu WebSphere Portal 365
monitorowanie 302
określanie identyfikatora 36
opis 3
pomijanie kontroli dostępu na poziomie dokumentu 290
powiązanie z aplikacjami wyszukującymi 220
przeszukiwanie 180
robocze 31
serwery wyszukiwania 193
sposoby tworzenia 31
status systemu 302
stowarzyszenie 31

kolekcje (*kontynuacja*)
szacowanie wielkości 33
szacowanie zasobów 301
tworzenie przy użyciu Kreator kolekcji 31
tworzenie przy użyciu widoku Kolekcje 33
usuwanie 35
wykrywanie duplikatów dokumentów 184
zabezpieczenia 260
zabezpieczenia duplikatów dokumentów 260
zabezpieczenia przy użyciu identyfikatora aplikacji 260
zabezpieczenia tekstu zakotwiczenia 271
kolekcje stowarzyszone 31
Kolekcje, widok
opis 15
tworzenie kolekcji 33
komenda esadmin 369
komendy, wyszukiwanie korporacyjne 369
komunikaty o błędach
konfiguracja serwera SMTP 328
otrzymywanie wiadomości e-mail 326, 328
wyświetlanie dzienników usuniętych dokumentów 321
wyświetlanie plików dziennika 331
konektor Java dla programu DB2 Content Manager 57, 59
konfiguracja dziennika zapytań 330
konfiguracja opcji Directory Assistance 285
konfiguracja serwera SMTP 328
konfiguracja serwera zaufanego 282
konfiguracja użytkownika programu Domino, przeszukiwacze QuickPlace 284
konfigurowanie aplikacji wyszukujących 238, 239
konsola administracyjna
interfejs 15
logowanie 19
opis 9
podsumowanie zadań 15
kopia zapasowa systemu 333, 334
koreański
analiza terminów złożonych 150
podział na ciągi n-znakowe 170
Kreator kolekcji 31
Kreator migracji
domyślne ustawienia kolekcji 367
domyślne ustawienia przeszukiwaczy 367
kolekcje 365
opis 365
plik dziennika 368
systematyki oparte na regułach 365
uruchamianie 365

L

LDAP, źródła zewnętrzne
edycja 251
konfigurowanie 251
usuwanie 251
logowanie się do konsoli administracyjnej 19

Lotus Notes
instalacja wtyczki 341
integracja z wyszukiwaniem korporacyjnym 341
pasek wyszukiwania 341
serwis aktualizacyjny wtyczki 341
Lotus Quickr
integracja z produktem WebSphere Portal 343, 356
konfiguracja przeszukiwacza QuickPlace 82
lista adresów początkowych, konfiguracja przeszukiwacza 86

M

maksymalny odstęp czasu przed ponownym przeszukiwaniem 98
mechanizmy analizy tekstu
dodawanie do systemu 144
odzworowywanie elementów XML 146
odzworowywanie wspólnej struktury analizy na relacyjne bazy danych 149
odzworowywanie wyników analizy na indeks 147
odzworowywanie wyników analizy na relacyjne bazy danych 149
opis 143
powiązanie z kolekcjami 145
migracja
kolekcje 365
systematyka oparta na regułach 365
minimalny odstęp czasu przed ponownym przeszukiwaniem 98
moduł bos.iocp.rte 81
moduł portu zakończenia we/wy, konfiguracja przeszukiwacza 81
monitor
konfiguracja roli 259
opis 258
monitorowanie
Aktywne serwisy przeszukiwacza sieci WWW 307
analizatory składni 313
często zadawane zapytania 316, 317
dokumenty usunięte 320, 321
historia czasu odpowiedzi 316
kolekcje 302
nasłuchiwanie danych 318
ostatnie zapytania 316, 317
pliki dzienników 331
przeszukiwacze 304
przeszukiwacze sieci WWW 305
serwery wyszukiwania 316, 317
Szczegółowe informacje o wątkach przeszukiwacza sieci WWW 306
szczegóły dotyczące URI 302
Szybkość przeszukiwania przeszukiwacza sieci WWW 307
wyszukiwanie korporacyjne 301

N

najczęstsze wyniki
wykresy słupkowe 243

- nasłuchiwanie danych
 - konfigurowanie 116
 - monitorowanie 318
 - restartowanie 116, 318
- newHtmlTagReplacement, parametr 158
- niestandardowa analiza tekstu
 - mechanizmy analizy tekstu 144, 145
 - odzworowywanie elementów XML 146
 - odzworowywanie wspólnej struktury analizy na indeks 147
 - odzworowywanie wspólnej struktury analizy na relacyjną bazę danych 149
 - odzworowywanie wyników analizy na indeks 147
 - odzworowywanie wyników analizy na relacyjną bazę danych 149
 - opis 143
- nieznane typy dokumentów 158
- nodes.ini, plik 25, 247
- numer portu, wyszukiwanie korporacyjne 24

O

- obsługa lingwistyczna
 - kody języków 167
 - niestandardowa analiza tekstu 143
 - podział na ciągi n-znakowe 170
 - rodzime wyszukiwanie XML 151
 - rozpoznawanie języków 168
 - rozpoznawanie strony kodowej 169
 - słowniki kryteriów 209
 - słowniki słów ignorowanych 197
 - słowniki synonimów 194
 - ustawienia narodowe 167
 - usuwanie znaków spacji, tabulacji i nowego wiersza 170
 - wyszukiwanie semantyczne 143, 151
- ocena dynamiczna 205
- ocena statyczna
 - opis 206
 - w migrowanych kolekcjach 365
 - włączanie dla kolekcji 33
- ocena wyników wyszukiwania dynamiczne 205
 - klasy kryteriów 213, 214, 215, 216
 - odtworzenie wartości domyślnych 207
 - opis 205
 - runtime.properties, plik 207
 - słowniki kryteriów 211
 - statyczna 206, 207
 - wartościowanie danych tekstowych 205
 - wzorce URI 211, 212
- odstępy czasu przed ponownym przeszukiwaniem przy użyciu przeszukiwaczy sieci WWW 98
- odtworzenie systemu 333, 335
- odtworzenie wyszukiwania korporacyjnego 333
- odzworowanie
 - elementów XML na wspólną strukturę analizy 146
 - poła wyszukiwania HTML 141
 - poła wyszukiwania XML 138
 - pół na klasy kryteriów 214
 - wspólne struktury analizy na indeks 147
 - wspólnej struktury analizy na relacyjne bazy danych 149

- odzworowanie (*kontynuacja*)
 - wyników analizy na relacyjne bazy danych 149
- ograniczanie przestrzeni przeszukiwania sieci WWW 93
- OmniFind Enterprise Edition
 - analizatory składni 5
 - aplikacje wyszukujące 12
 - diagram przepływu danych 12
 - identyfikatory sesji 369
 - integracja z produktem Lotus Notes 341
 - integracja z produktem WebSphere Portal 343
 - interfejsy API 11
 - kody powrotu 369
 - komendy 369
 - komponenty 3
 - konfiguracja numeru portu 24
 - konsola administracyjna 9
 - obsługa podwójnych adresów IP 27
 - obsługa protokołu IPv6 28
 - przegląd informacji 1
 - serwery indeksowania 7
 - serwery przeszukiwaczy 4
 - serwery wyszukiwania 8
 - zmiana hasła na pojedynczym serwerze 20
 - zmiana hasła na wielu serwerach 21
 - zmienianie adresów IP 25
- operator
 - dla wyników wyszukiwania 180
 - dopasowanie pełne 180
 - poła pozwalające na sortowanie 180
 - treść dokumentu 180
 - wyszukiwanie dla pół 180
 - wyszukiwanie parametryczne 180
 - wyszukiwanie pełnotekstowe 180
- operator
 - konfiguracja roli 259
 - opis 258
- Oracle, bazy danych
 - dostęp jako źródło zewnętrzne 251
 - dostęp za pomocą przeszukiwaczy DB2 47
 - dostęp za pomocą przeszukiwaczy JDBC 64
- ostatnie zapytania
 - obliczanie 317
- ostatnie zapytania, monitorowanie 316
- ostatnio przeszukane adresy URL, monitorowanie 305

P

- pamięć podręczna wyszukiwania
 - konfigurowanie 194
 - opis 194
- parserdriver.collection.properties, plik 158
- parserTypes.cfg, plik 153, 155, 158
- pasek wyszukiwania, portal WebSphere
 - wersja 5.1, przekierowanie do wyszukiwania korporacyjnego 348
 - wersja 6, przekierowanie do wyszukiwania korporacyjnego 354
- planowanie
 - budowanie indeksów 174, 175
 - przeszukiwacze 37, 41
- plik CCLServer_date.log 24
- pliki archiwów
 - formaty URI 120
 - obsługiwane formaty 119
 - przeszukiwanie 119
- pliki dzienników
 - dla śledzenia dokumentów 320
 - dzienniki zapytań 330
 - filtrowanie 331
 - konfiguracja serwera SMTP 328
 - konfiguracja wielkości 330
 - Kreator migracji 368
 - maksymalna wielkość 326
 - monitorowanie 321, 331
 - opcje poczty elektronicznej 328
 - opis 323
 - położenie domyślne 323
 - poziomy istotności 326
 - rotacja 326
 - wyświetlanie 331
 - wyświetlanie usuniętych dokumentów 321
- pliki EAR
 - aplikacja ESSearchServer 288
 - ESAdmin, aplikacja 288
 - ESSearchApplication, aplikacja 288
- pliki kluczy 247
- początkowe adresy URL dla przeszukiwaczy sieci WWW 93, 99
- podstawowe uwierzytelnianie HTTP 101
- podsumowania
 - dostosowywanie 201
 - dynamiczne 200
- podsumowania dokumentów
 - dostosowywanie 201
 - edycja właściwości dla 201
- podsumowanie zadań, konsola administracyjna 15
- podsumowywanie dynamiczne 200
- podział na ciągi n-znakowe 170
- pojedyncze logowanie - zabezpieczenie konfiguracja 270
 - zarządzanie tożsamościami 269
- poła metadanych, wykres słupkowy najczęstszych wyników 243
- poła parametryczne
 - opis 180
 - sortowanie liczbowe 180
- poła pozwalające na sortowanie
 - sortowanie liczbowe 180
 - sortowanie łańcuchowe 180
- poła wyników wyszukiwania, opis 180
- poła wyszukiwania HTML
 - Dublin Core, elementy 141
 - odzworowywanie elementów 141
 - opis 141
 - tworzenie 141
- poła wyszukiwania XML
 - odzworowywanie elementów 137, 138, 146
 - opis 137, 146
 - tworzenie 138
- poła wyszukiwania z pełną zgodnością, opis 180
- poła, odzworowywanie na klasy kryteriów 214

- ponowne przeszukiwanie adresów URL w najbliższym możliwym terminie 99
- Portal WebSphere Portal
 - Centrum wyszukiwania, opis 343
 - domyślne ustawienia migracji 367
 - integracja z wyszukiwaniem korporacyjnym 343
 - migracja drzewa kategorii 365
 - migracja kolekcji 365
 - migracja systematyki 365
 - pasek wyszukiwania, opis 343
 - Skrypty konfiguracyjne wyszukiwania korporacyjnego 344
 - system klastrowy 358
 - wersja 5.1, konfiguracja paska wyszukiwania 348
 - wersja 5.1, skrypty integracyjne 345
 - wersja 5.1, usuwanie wyszukiwania korporacyjnego 349
 - wersja 6, konfiguracja centrum wyszukiwania 353
 - wersja 6, konfiguracja paska wyszukiwania 354
 - wersja 6, skrypty integracji 350
 - wersja 6, usuwanie systemu wyszukiwania korporacyjnego 357
- portlet Wyszukiwanie, wdrażanie
 - es.wp5.install.jar, plik 345
 - es.wp6.install.jar, plik 350, 358
 - wp5_install, skrypt 345
 - wp5_uninstall, skrypt 349
 - wp6_cluster_install, skrypt 358
 - wp6_cluster_uninstall 362
 - wp6_install, skrypt 350
 - wp6_uninstall, skrypt 357
- portlety
 - integracja z klastrami portalu WebSphere Portal 358
 - integracja z portalem WebSphere Portal 5.1 345
 - integracja z portalem WebSphere wersja 6 350, 356
 - konfigurowanie dla źródeł Lotus Quickr 356
 - opis 343
 - usuwanie z klastrów WebSphere Portal 362
 - usuwanie z portalu WebSphere w wersji 5.1 349
 - usuwanie z portalu WebSphere w wersji 6 357
 - wyszukiwanie korporacyjne 343
- powiadomienia pocztą elektroniczną
 - dotycząca komunikatów 328
 - dotyczące alertów 328
 - konfiguracja serwera SMTP 328
- prawa dostępu
 - opis 257
 - pojedyncze logowanie - zabezpieczenie 269
 - sprawdzanie bieżącego użytkownika 264
 - wyłączanie dla kolekcji 290
 - wymagania dla produktu Lotus Domino 282
 - wymagania dla systemów plików Windows 285
- prawa dostępu (*kontynuacja*)
 - zabezpieczenia na poziomie dokumentu 263
 - zarządzanie tożsamościami 265, 268
- profile użytkowników
 - konfiguracja 270
 - opis 268
- program WebSphere II Event Publisher Edition, konfiguracja przeszukiwacza DB2 51
- program WebSphere MQ, konfiguracja przeszukiwacza DB2 54
- program WebSphere MQ, konfiguracja serwera przeszukiwacza 50
- protokół DIIOP, konfiguracja przeszukiwacza 80
- protokół IPv6 28
- protokół NRPC, konfiguracja przeszukiwacza 76, 78
- przepływ danych, system wyszukiwania korporacyjnego 12
- przestrzeń przeszukiwania
 - alerty o 324
 - edycja 40
 - konfiguracja przeszukiwacza sieci WWW 93
 - opis 4
- przeszukiwacze
 - aplikacje nasłuchiwanie danych 116
 - Content Edition 42, 44
 - DB2 47
 - DB2 Content Manager 56
 - Domino Document Manager 61
 - domyślne ustawienia migracji 367
 - edycja przestrzeni przeszukiwania 40
 - edycja właściwości przeszukiwacza 40
 - Exchange Server 63, 281
 - formaty URI 120
 - JDBC, baza danych 64, 66, 68
 - lista adresów początkowych 86
 - łączenie typów przeszukiwaczy 37
 - monitorowanie 304
 - NNTP 72
 - Notes 73, 75
 - obsługa zewnętrznych 116
 - opis 4
 - planowanie 37, 41
 - pliki archiwów 119
 - Portal WebSphere Portal 112, 113
 - przegląd konfiguracji 37
 - QuickPlace 82
 - sieć WWW 88
 - status systemu 304
 - system plików UNIX 87
 - system plików Windows 114
 - tworzenie 39
 - usuwanie 41
 - wartości początkowe 39
 - wartości podstawowe 37
 - Web Content Management 110, 113
 - włączanie zabezpieczeń na poziomie dokumentu 37
 - wtyczki 117
 - zabezpieczenia na poziomie dokumentu 261
- przeszukiwacze baz danych JDBC
 - formaty URI 120
- przeszukiwacze baz danych JDBC (*kontynuacja*)
 - konfiguracja 64
 - obsługiwane sterowniki 64
 - przeszukiwanie wielu tabel 68
 - wtyczka do przeszukiwania wielu tabel 68
- Przeszukiwacze baz danych JDBC
 - przeszukiwanie wielu tabel 66
 - wtyczka do przeszukiwania wielu tabel 66
- przeszukiwacze Content Edition
 - formaty URI 120
 - konfiguracja 42
 - konfigurowanie w systemie operacyjnym AIX 45
 - konfigurowanie w systemie operacyjnym Linux 45
 - konfigurowanie w systemie Windows 46
 - konfigurowanie w Środowisku Operacyjnym Solaris 45
 - tryb bezpośredni 44
 - tryb serwera 44
- przeszukiwacze DB2
 - formaty URI 120
 - instalacja programu WebSphere MQ w systemie AIX 50
 - instalacja programu WebSphere MQ w systemie Linux 50
 - instalacja programu WebSphere MQ w systemie Solaris 50
 - instalacja programu WebSphere MQ w systemie Windows 50
 - konfiguracja 47
 - konfiguracja programu WebSphere II Event Publisher Edition 51
 - konfiguracja programu WebSphere MQ 54
 - konfiguracja publikowania zdarzeń 50
 - konfigurowanie w systemie AIX 50
 - konfigurowanie w systemie Linux 50
 - konfigurowanie w systemie Solaris 50
 - konfigurowanie w systemie Windows 50
 - publikowanie zdarzeń 47
 - WebSphere II Classic Federation 55
- przeszukiwacze DB2 Content Manager
 - formaty URI 120
 - konfiguracja 56
 - konfigurowanie w systemie operacyjnym AIX 57
 - konfigurowanie w systemie operacyjnym Linux 57
 - konfigurowanie w systemie Windows 59
 - konfigurowanie w Środowisku Operacyjnym Solaris 57
- przeszukiwacze Domino Document Manager
 - formaty URI 120
 - konfiguracja 61
 - konfiguracja IOCP 81
 - konfiguracja protokołu DIIOP 80
 - konfigurowanie w systemie operacyjnym AIX 76
 - konfigurowanie w systemie operacyjnym Linux 76
 - konfigurowanie w systemie Windows 78
 - konfigurowanie w Środowisku Operacyjnym Solaris 76

- przeszukiwacze Domino Document Manager (kontynuacja)
 - NRPC, protokół 76, 78
 - przeszukiwacze Exchange Server
 - formaty URI 120
 - konfiguracja 63
 - zabezpieczone dokumenty 281
 - przeszukiwacze listy adresów początkowych
 - formaty URI 120
 - integracja z produktem WebSphere Portal 343, 356
 - konfiguracja 86
 - przeszukiwacze NNTP, konfigurowanie 72
 - przeszukiwacze Notes
 - formaty URI 120
 - konfiguracja 73
 - konfiguracja IOCP 81
 - konfiguracja protokołu DIIOP 80
 - konfiguracja zabezpieczeń na poziomie dokumentu 282
 - konfigurowanie w systemie operacyjnym AIX 76
 - konfigurowanie w systemie operacyjnym Linux 76
 - konfigurowanie w systemie operacyjnym Windows 78
 - konfigurowanie w Środowisku Operacyjnym Solaris 76
 - NRPC, protokół 76, 78
 - reguły odwzorowywania pól 75
 - sprawdzanie poprawności bieżących referencji 282
 - wskazówki na temat używania 75
 - zaufany serwer Lotus Domino 282
 - przeszukiwacze QuickPlace
 - formaty URI 120
 - konfiguracja IOCP 81
 - konfiguracja opcji Directory Assistance 285
 - konfiguracja protokołu DIIOP 80
 - konfiguracja użytkownika programu Domino 284
 - konfigurowanie w systemie operacyjnym AIX 76
 - konfigurowanie w systemie operacyjnym Linux 76
 - konfigurowanie w systemie operacyjnym Windows 78
 - konfigurowanie w Środowisku Operacyjnym Solaris 76
 - NRPC, protokół 76, 78
 - zabezpieczenia użytkownika lokalnego 284
 - Przeszukiwacze QuickPlace
 - konfiguracja 82
 - przeszukiwacze sieci WWW
 - agenty użytkownika 89
 - aktywne serwisy 305, 307
 - cookie, format 104
 - cookie, konfiguracja informacji 105
 - daty przeszukanych dokumentów 108
 - dyrektywy blokujące indeksowanie 107, 108
 - dyrektywy blokujące wykorzystywanie odsyłaczy 107, 108
 - followindex.rules, plik 107, 108
 - globalna przestrzeń przeszukiwania 105
 - historia przeszukiwacza 305
 - informacje cookie 104
 - przeszukiwacze sieci WWW (kontynuacja)
 - JavaScript, obsługa 92
 - konfiguracja 88
 - monitorowanie 305
 - odstęp czasu przed ponownym przeszukiwaniem 98
 - ograniczanie przestrzeni przeszukiwania 93
 - ostatnio przeszukane adresy URL 305
 - początkowe adresy URL 93, 99
 - przeszukiwanie adresów URL w najbliższym możliwym terminie 99
 - reguły przeszukiwania 93
 - robots.txt, pliki 89, 90
 - serwery proxy 103
 - serwisy WWW zabezpieczone hasłem 101, 102
 - status adresu URL 305
 - status systemu 305
 - strony błędów programowych 99
 - szczegóły serwisu 305
 - szczegóły wątków 305, 306
 - szybkość przeszukiwania 307
 - tworzenie raportów o 308
 - przeszukiwacze systemu plików UNIX
 - formaty URI 120
 - konfiguracja 87
 - przeszukiwacze systemu plików Windows
 - formaty URI 120
 - konfiguracja 114
 - konfiguracja zabezpieczeń na poziomie dokumentu 285
 - przeszukiwacze Web Content Management
 - formaty URI 120
 - konfiguracja 110
 - kopiowanie adresów URL serwisów 113
 - przeszukiwacze WebSphere Portal
 - formaty URI 120
 - konfiguracja 112
 - przeszukiwanie
 - dokumenty HTML 141
 - dokumenty XML 137, 138, 146
 - kategorie 132
 - kolekcje 180
 - szybkie odsyłacze 203
 - przeszukiwanie adresów URL w najbliższym możliwym terminie 99
 - przetwarzanie tekstu
 - adnotatory 143
 - mechanizmy analizy tekstu 143
 - wspólne struktury analizy 143
 - przykładowa aplikacja wyszukująca
 - config.properties, plik 222, 238
 - funkcje wyszukiwania 219, 220
 - HTTPS, wymuszanie 247
 - instalacja domyślna 238
 - klonowanie 241
 - opis 12, 220
 - wyłączanie zabezpieczenia 288
 - wymuszanie SSL 247
 - przyrostowe indeksy
 - opis 7
 - publikowanie zdarzeń
 - konfiguracja przeszukiwacza DB2 51, 54
 - konfigurowanie w systemie operacyjnym AIX 50
 - publikowanie zdarzeń (kontynuacja)
 - konfigurowanie w systemie operacyjnym Linux 50
 - konfigurowanie w systemie Windows 50
 - konfigurowanie w Środowisku Operacyjnym Solaris 50
 - opis 47
- ## R
- ranking dokumentów
 - odtworzenie wartości domyślnych 207
 - raporty o szczegółach serwisu
 - opis 305
 - tworzenie 308
 - raporty z historii przeszukiwacza
 - opis 305
 - raport o serwisie 308
 - raporty o kodach statusu HTTP 308
 - tworzenie 308
 - reguły domenowe dla przeszukiwaczy sieci WWW 93
 - reguły dotyczące adresów dla przeszukiwaczy sieci WWW 93
 - reguły dotyczące adresów IP dla przeszukiwaczy sieci WWW 93
 - reguły dotyczące przedrostków dla przeszukiwaczy sieci WWW 93
 - reguły kategorii
 - konfigurowanie 136
 - treść dokumentu 132, 136
 - wzorce URI 132, 136
 - reguły wymiany HTML 158, 159
 - rejestr użytkownika LDAP 274
 - removeCjNewlineChars, opcja 170
 - removeCjNewlineCharsMode, opcja 170
 - Robots Exclusion, protokół
 - agent użytkownika, identyfikacja 89
 - zgodność przeszukiwacza sieci WWW 90
 - robots.txt, pliki
 - agent użytkownika, identyfikacja 89
 - zgodność przeszukiwacza sieci WWW 90
 - rodzime wyszukiwanie XML 151
 - role administracyjne
 - administrator kolekcji 258, 259
 - administrator wyszukiwania korporacyjnego 258, 259
 - konfigurowanie 259
 - monitor 258, 259
 - operator 258, 259
 - opis 258
 - rozszerzenia plików
 - obsługiwane analizatory składni Stellent 163
 - obsługiwane przez analizatory składni kolekcji 155, 157
 - obsługiwane przez analizatory składni Stellent 160
 - wykluczanie z przestrzeni przeszukiwania w sieci WWW 93
 - rozwińnięcie indeksu
 - opis 185
 - wpływ na czas budowania indeksu 187
 - wpływ na rozmiar indeksu 187

rozwińnięcie zapytania
 opis 185
 wpływ na czas budowania indeksu 187
 wpływ na rozmiar indeksu 187
równoległe operacje budowania indeksu 176
runtime-generic.properties, plik 201, 317

S

Search Application Customizer
 config.properties, plik 239
 uruchamianie 239
SearchBarInclude.jsp, plik 348
serwery analizatorów składni
 konfiguracja wątku 150
serwery indeksowania
 uruchamianie 293
 zatrzymywanie 295
serwery proxy 103
 konfiguracja serwera wyszukiwania 248
serwery proxy HTTP 103
serwery przeszukiwaczy
 uruchamianie 293, 304
 zatrzymywanie 295, 304
serwery wyszukiwania
 często zadawane zapytania 316, 317
 historia czasu odpowiedzi 316
 konfiguracja HTTPS 247
 konfiguracja serwera proxy 248
 konfiguracja SSL 247
 monitorowanie 316, 317
 obliczanie liczby zapytań 317
 opis 8, 193
 ostatnie zapytania 316, 317
 pamięć podręczna wyszukiwania 194
 ponowne wdrażanie słowników 199
 powiązanie słowników kryteriów 210,
 211
 powiązanie słowników słów
 ignorowanych 198, 199
 powiązanie słowników synonimów 196
 słowniki kryteriów 209
 słowniki słów ignorowanych 197
 słowniki synonimów 194, 196
 status systemu 316
 uruchamianie 293, 298, 316
 zatrzymywanie 295, 298, 316
serwery wyszukiwania korporacyjnego
 obsługa podwójnych adresów IP 27
 obsługa protokołu IPv6 28
 zmienianie adresów IP 25
serwisy WWW zabezpieczone hasłem 101
 podstawowe uwierzytelnianie HTTP 101
 uwierzytelnianie oparte na
 formularzu 102
SI-API (interfejs API wyszukiwania i
 indeksowania) 11, 219
siapi.jar, plik 345
składnia zapytania XML, rodzima 151
skrypt escrdb2.sh 50
skrypt escrdb2.vbs 50
skrypt escrnote.vbs 78
skrypty
 esbackup.bat 334
 esbackup.sh 334
 escrcm.sh 57
 escrcm.vbs 59

skrypty (*kontynuacja*)
 escrdb2.sh 50
 escrdb2.vbs 50
 escrnote.sh 76
 escrnote.vbs 78
 escrvbr.sh 45
 escrvbr.vbs 46
 esrestore.bat 335
 esrestore.sh 335
 startccl 335
skrypty konfiguracyjne
 Portal WebSphere Portal 344
skrypty odtwarzania
 opis 333
 uruchamianie 335
skrypty tworzenia kopii zapasowych
 opis 333
 uruchamianie 334
słowa kluczowe w szybkich odsyłaczach 203
słowniki kryteriów
 dodawanie do systemu 210
 opis 209
 ponowne wdrażanie 199
 powiązanie z kolekcją 211
słowniki słów ignorowanych
 dodawanie do systemu 198
 opis 197
 ponowne wdrażanie 199
 powiązanie z kolekcją 199
słowniki synonimów
 dodawanie do systemu 196
 opis 194
 ponowne wdrażanie 199
 powiązanie z kolekcją 196
Solaris, system operacyjny
 konfiguracja przeszukiwacza DB2 50
sprawdzanie poprawności bieżących
 referencji 264, 282, 285
sprawdzanie poprawności zapytań 264
SQL Server, bazy danych
 dostęp za pomocą przeszukiwaczy
 DB2 47
 dostęp za pomocą przeszukiwaczy
 JDBC 64
SSL, konfiguracja serwera
 wyszukiwania 247
startccl, skrypt 335
startIndexBuild, komenda 177
status systemu
 analizatory składni 313
 budowanie indeksów 315
 kolekcje 302
 przeszukiwacze 304
 przeszukiwacze sieci WWW 305
 serwery wyszukiwania 316
Stellent, analizator składni
 domyślne typy dokumentów 163
 opis 153
 powiązanie typów dokumentów 160
 typy analizowanych dokumentów 155
stellent.properties, plik 160
stellenttypes.cfg, plik 160
stellentTypes.cfg, plik 153
sterowniki JDBC
 dla przeszukiwaczy bazy danych
 JDBC 64
 dla źródeł zewnętrznych JDBC 251

strona Analiza składni, opis 15
strona Dziennik, opis 15
strona Indeks, opis 15
strona Przeszukiwanie, opis 15
strona Wyszukiwanie, opis 15
strony błędów programowych, przeszukiwacze
 sieci WWW 99
strony kodowe
 automatyczne rozpoznawanie 169
 obsługiwane 169
system operacyjny AIX
 konfiguracja przeszukiwacza Content
 Edition 45
 konfiguracja przeszukiwacza DB2 50
 konfiguracja przeszukiwacza DB2 Content
 Manager 57
 konfiguracja przeszukiwacza Domino
 Document Manager 76
 konfiguracja przeszukiwacza Notes 76
 konfiguracja przeszukiwacza
 QuickPlace 76
 konfiguracja publikowania zdarzeń 50
system operacyjny Linux
 konfiguracja przeszukiwacza Content
 Edition 45
 konfiguracja przeszukiwacza DB2 50
 konfiguracja przeszukiwacza DB2 Content
 Manager 57
 konfiguracja przeszukiwacza Domino
 Document Manager 76
 konfiguracja przeszukiwacza Notes 76
 konfiguracja przeszukiwacza
 QuickPlace 76
 konfiguracja publikowania zdarzeń 50
 Środowisko Operacyjne Solaris
 konfiguracja publikowania
 zdarzeń 50
systematyka oparta na regułach, migrowanie z
 portalu WebSphere Portal 365
systematyki, migrowanie z portalu WebSphere
 Portal 365
szacowanie zasobów systemowych 301
szczegóły dotyczące URI
 dokumenty usunięte 320
 monitorowanie 302
szczegóły wątków, monitorowanie 305
szybkie odsyłacze
 formaty URI 120
 opis 203
 przeszukiwanie 203
 tworzenie 203
szybkość przeszukiwania,
 monitorowanie 307

Ś

śledzenie dokumentu
 konfiguracja pliku dziennika 320
 opis 319
 pliki dzienników 321
 raporty 320
 włączanie 320
 wyłączanie 320
Środowisko Operacyjne Solaris
 konfiguracja przeszukiwacza Content
 Edition 45

Środowisko Operacyjne Solaris (*kontynuacja*)
konfiguracja przeszukiwacza DB2 Content
Manager 57
konfiguracja przeszukiwacza Domino
Document Manager 76
konfiguracja przeszukiwacza Notes 76
konfiguracja przeszukiwacza
QuickPlace 76

T

tar, pliki
formaty URI 120
przeszukiwanie 119
terminy złożone, analizowanie 150
treść dokumentu, opis 180
tryb bezpośredni, repozytoria Content
Edition 44
tryb serwera, repozytoria Content Edition 44
tworzenie
kategorie oparte na regułach 136
kolekcje 31, 33
pola wyszukiwania HTML 141
pola wyszukiwania XML 138
przeszukiwacze 39
raporty o przeszukiwaczach sieci
WWW 308
szybkie odsyłacze 203
zasiegi 189
tworzenie kopii zapasowej wyszukiwania
korporacyjnego 333, 334
typ klasyfikacji
oparte na regułach 132
wybór 33, 135
typy dokumentów
analizowanie 155
dla analizatorów składni Stellent 160
obsługiwane przez analizatory składni
Stellent 163
usługi analizatora składni 155, 157
wykrywanie 153
typy MIME, włączanie do przestrzeni
przeszukiwania w sieci WWW 93
typy przeszukiwaczy
łączenie w kolekcji 37
wartości podstawowe 37
typy źródeł danych
bazy danych DB2 47, 64, 251
bazy danych Domino Document
Manager 61
bazy danych Notes 73, 75
bazy danych QuickPlace 82
CA-Datacom, bazy danych 47
DB2 for iSeries, bazy danych 47
DB2 for z/OS 47
foldery publiczne serwera Exchange
Server 63
grupy dyskusyjne NNTP 72
IMS, bazy danych 47
Informix, bazy danych 47
JDBC, bazy danych 64, 66, 68, 251
Lotus Quickr for Domino 82
Lotus Quickr for WebSphere Portal 86
obsługa zewnętrznych 2, 11
obsługiwane w wyszukiwaniu
korporacyjnym 2
Oracle, bazy danych 47, 64, 251

typy źródeł danych (*kontynuacja*)
relacyjne bazy danych 47
repozytoria Content Edition 42, 44
serwisy WebSphere Portal 112
serwisy WWW 88
Software AG Adabas, bazy danych 47
SQL Server, bazy danych 47, 64
Sybase, bazy danych 47
systemy plików UNIX 87
systemy plików Windows 114
typy elementów DB2 Content
Manager 56
VSAM, bazy danych 47
Web Content Management, serwisy 110

U

UIMA
dodawanie mechanizmów analizy tekstu do
systemu 144
odwzorowywanie elementów XML 146
odwzorowywanie wspólnej struktury
analizy na indeks 147
odwzorowywanie wspólnej struktury
analizy na relacyjne bazy danych 149
odwzorowywanie wyników analizy na
indeks 147
odwzorowywanie wyników analizy na
relacyjne bazy danych 149
opis 143
powiązanie z kolekcjami 145
wspólne struktury analizy 147, 149
ułatwienia dostępu w produkcie 403
uruchamianie
analizatory składni 313
aplikacje wyszukujące 246
budowanie indeksów 315
Kreator migracji 365
nasłuchiwanie danych 318
Search Application Customizer 239
serwery przeszukiwaczy 304
serwery wyszukiwania 298, 316
serwery wyszukiwania
korporacyjnego 293
USC, ciąg znaków 265
ustawienia narodowe
analizowanie 167
przeszukiwanie 167
usuwanie
indeksy z kolejki 315
kolekcje 35
przeszukiwacze 41
usuwanie identyfikatorów URI z indeksu 192
usuwanie znaków nowego wiersza 170
usuwanie znaków spacji, tabulacji i nowego
wiersza 170
uwierzytelnianie
opis 257
wyłączanie dla aplikacji
korporacyjnych 288
uwierzytelnianie oparte na formularzu 101,
102
użytkownik serwera WebSphere Application
Server
hasło dla konfiguracji
jednoserverowej 276

użytkownik serwera WebSphere Application
Server (*kontynuacja*)
hasło dla konfiguracji
wieloserwerowej 277

V

vbr_access_services.jar, plik 45, 46

W

wartościowanie danych tekstowych 205
ważność dokumentu
klasy kryteriów 213, 216
odtworzenie wartości domyślnych 207
słowniki kryteriów 211
statyczna 206
w migrowanych kolekcjach 365
włączanie dla kolekcji 33
wzorce URI 211, 212
wątki
analizator składni 150
przeszukiwacz sieci WWW 306
Web Content Management
integracja z produktem WebSphere
Portal 343
WebSphere II Classic Federation 55
WebSphere Portal, klastry
skrypty integracyjne 358
usuwanie wyszukiwania
korporacyjnego 362
wytyczne na temat integracji 358
WebSphere, zabezpieczenia globalne
Search Application Customizer 239
właściwości aplikacji wyszukującej 238
wyłączanie 288
widok Monitor, opis 15
widok System, opis 15
widok Zabezpieczenia, opis 15
wiele tabel o określonej strukturze,
wtyczka 66, 68
Windows
obsługa protokołu IPv6 28
Windows, system operacyjny
konfiguracja przeszukiwacza 78
konfiguracja przeszukiwacza Content
Edition 46
konfiguracja przeszukiwacza DB2 50
konfiguracja przeszukiwacza DB2 Content
Manager 59
konfiguracja publikowania zdarzeń 50
właściwości przeszukiwacza
edycja 40
opis 4
włączanie harmonogramów operacji na
indeksach 175
wp5_install, skrypt 345
wp5_uninstall, skrypt 349
wp6_cluster_install, skrypt 358
wp6_cluster_uninstall, skrypt 362
wp6_install, skrypt 350
wp6_uninstall, skrypt 357
WpsMigratorLog.log, plik 368
wspólne struktury analizy
odwzorowywanie elementów XML 146
odwzorowywanie na indeks 147

- wspólne struktury analizy (*kontynuacja*)
 - odzworowywanie na relacyjne bazy danych 149
 - opis 143
- współbieżne operacje budowania indeksu 176
- współczynniki kryteriów
 - dla klas kryteriów 214, 216
 - dla słowników kryteriów 209
 - dla wzorców URI 211, 212
 - konfiguracja klasy kryteriów 213, 215
- wtyczka dla produktu Lotus Notes
 - instalacja 341
 - serwis aktualizacyjny 341
- wtyczki
 - przeszukiwacze baz danych JDBC 68
 - Przeszukiwacze baz danych JDBC 66
 - przeszukiwanie wielu tabel o określonej strukturze 66, 68
- wtyczki dla przeszukiwaczy 117
- wtyczki przeszukiwaczy 117
- wykresy słupkowe
 - klasy Java dla najczęstszych wyników 243
 - najczęstsze wyniki 243
- wykrywanie 4
- wykrywanie duplikatów dokumentów
 - analiza globalna 184, 260
 - konfiguracja klasy kryteriów 213
 - opis 184, 255
 - włączanie zabezpieczeń 260
- wyłączanie harmonogramów operacji na indeksach 175
- wyniki wyszukiwania
 - dostosowywanie podsumowań 201
 - grupowanie 190, 191
 - konfiguracja klasy kryteriów 213, 214, 215, 216
 - konfiguracja wzorca URI 211
 - końcowe filtrowanie 262
 - ocena 212
 - ocena dynamiczna 205
 - ocena statyczna 206
 - opis 205
 - podsumowania 201
 - podsumowywanie dynamiczne 200
 - rozwijanie znaków wieloznacznych 187
 - wartościowanie danych tekstowych 205
 - wstępne filtrowanie 262
 - znaki wieloznaczne 185
 - zwijanie 190, 191, 291
- wyszukiwanie dla pól
 - opis 180
 - sortowanie łańcuchowe 180
- wyszukiwanie i indeksowanie, interfejs API 11, 219
- wyszukiwanie korporacyjne
 - analizatory składni 5, 131
 - aplikacje wyszukujące 12
 - diagram przepływu danych 12
 - formaty URI 120
 - identyfikatory sesji 369
 - integracja z produktem Lotus Notes 341
 - integracja z produktem WebSphere Portal 343
 - interfejsy API 11
 - kody powrotu 369

- wyszukiwanie korporacyjne (*kontynuacja*)
 - komendy 369
 - komponenty 3
 - konfiguracja numeru portu 24
 - konsola administracyjna 9
 - monitorowanie 301
 - odtworzenie z kopii zapasowej 335
 - pliki dzienników 323
 - przegląd informacji 1
 - role administracyjne 258
 - serwery indeksowania 7, 173
 - serwery przeszukiwaczy 4, 37
 - serwery wyszukiwania 8, 193
 - skrypty odtwarzania 333
 - skrypty tworzenia kopii zapasowych 333
 - tworzenie kopii zapasowej 334
 - uruchamianie serwerów 293
 - uruchamianie serwerów wyszukiwania 298
 - zabezpieczenia 255
 - zabezpieczenia na poziomie dokumentu 261
 - zabezpieczenia na poziomie kolekcji 260
 - zatrzymywanie serwerów 293, 295
 - zatrzymywanie serwerów wyszukiwania 298
- wyszukiwanie semantyczne 143, 146, 151
- wyszukiwanie tekstowe, opis 180
- wyświetlanie
 - dzienniki usuniętych dokumentów 321
 - pliki dzienników 331
 - szczegóły dotyczące URI 302

X

- XPath, rodzime wyszukiwanie XML 151

Z

- zabezpieczenia
 - analiza tekstu zakotwiczenia 271
 - dokumenty Lotus Domino 282
 - domeny Windows 285
 - globalne serwera aplikacji WebSphere 273, 274
 - identyfikatory aplikacji wyszukującej 260
 - instalacja jednoserwerowa 276
 - konfiguracja wieloserwerowa 277
 - konfigurowanie używania protokołu HTTPS do wyszukiwania 247
 - konfigurowanie używania protokołu SSL do wyszukiwania 247
 - na poziomie dokumentu 261, 262, 263, 264, 270, 290
 - na poziomie kolekcji 260, 290
 - opis 255
 - pojedyncze logowanie - obsługa 269
 - pomijanie kontroli dostępu na poziomie dokumentu 290
 - prawa dostępu 257
 - profile użytkowników 268
 - rejestr użytkownika LDAP 274
 - role administracyjne 259
 - uwierzytelnianie 257, 288
 - WebSphere, zabezpieczenia globalne 288
- zabezpieczenia (*kontynuacja*)
 - włączanie dla kolekcji 33, 255
 - włączanie dla wyszukiwania korporacyjnego 273
 - wtyczki przeszukiwaczy 117
 - wykrywanie duplikatów dokumentów 260
 - wyłączanie dla aplikacji korporacyjnej 288
 - zarządzanie tożsamościami 265, 270
 - zwinęte wyniki wyszukiwania 291
- zabezpieczenia na poziomie dokumentu
 - dla dokumentów Lotus Domino 282
 - dla systemów plików Windows 285
 - dokumenty Lotus Domino 282
 - konfiguracja przeszukiwacza 37
 - kontrola dostępu w indeksie 263
 - końcowe filtrowanie wyników 262
 - opis 255, 261
 - pojedyncze logowanie - obsługa 269
 - profile użytkowników 268
 - sprawdzanie poprawności bieżących referencji 264
 - sprawdzanie poprawności w czasie rzeczywistym 264
 - wstępne filtrowanie wyników 262
 - wtyczki przeszukiwaczy 117
 - zarządzanie tożsamościami 265, 270
 - znaczniki bezpieczeństwa 263
- zabezpieczenia na poziomie kolekcji
 - analiza tekstu zakotwiczenia 271
 - identyfikatory aplikacji 260
 - opis 255, 260
 - włączanie 33
 - wykrywanie duplikatów dokumentów 260
- zabezpieczenia użytkownika lokalnego, przeszukiwacze QuickPlace 284
- zapytania o niskiej znajdawalności
 - domyślne współczynniki kryteriów 216
 - opis 213
- zapytania o wysokiej znajdawalności
 - domyślne współczynniki kryteriów 216
 - opis 213
- zarządzanie tożsamościami
 - ekstrakcja grup 265
 - konfiguracja 270
 - kontekst zabezpieczeń użytkownika 265
 - opis 265
 - pojedyncze logowanie - obsługa 269
 - profile użytkowników 268
 - wyłączanie 265
 - XML, ciąg znaków zapytania 265
- zasięgi
 - formaty URI 120
 - opis 188
 - przeszukiwanie 188
 - tworzenie 189
 - wzorce URI 188, 189
- zasoby systemowe
 - szacowanie 301
- zatrzymywanie
 - analizatory składni 313
 - budowanie indeksów 315
 - serwery przeszukiwaczy 304
 - serwery wyszukiwania 298, 316

- zatrzymywanie (*kontynuacja*)
 - serwery wyszukiwania
 - korporacyjnego 293, 295
- zaufane serwery Lotus Domino 282
- zewnętrzne przeszukiwacze
 - aplikacje nasłuchiwania danych 116
 - konfigurowanie 116
- zip, pliki
 - formaty URI 120
 - przeszukiwanie 119
- znaczniki bezpieczeństwa
 - konfiguracja przeszukiwacza 263
 - wyłączanie dla kolekcji 290
 - zabezpieczenia na poziomie dokumentu 263
- znaki wieloznaczne
 - rozwińcie indeksu 185, 187
 - rozwińcie zapytania 185, 187
 - w zapytaniach 185
- zwinięte identyfikatory URI
 - konfigurowanie 191
 - ograniczenia związane z bezpieczeństwem 291
 - opis 190
- zwinięte wyniki wyszukiwania
 - konfigurowanie 191
 - ograniczenia związane z bezpieczeństwem 291
 - opis 190

Ż

- źródła zewnętrzne
 - konfiguracja 251
 - opis 251
 - powiązanie z aplikacjami wyszukiwującymi 254
 - zabezpieczenia przy użyciu identyfikatora aplikacji 260

IBM



Java[™]
COMPATIBLE

SC85-0162-04

