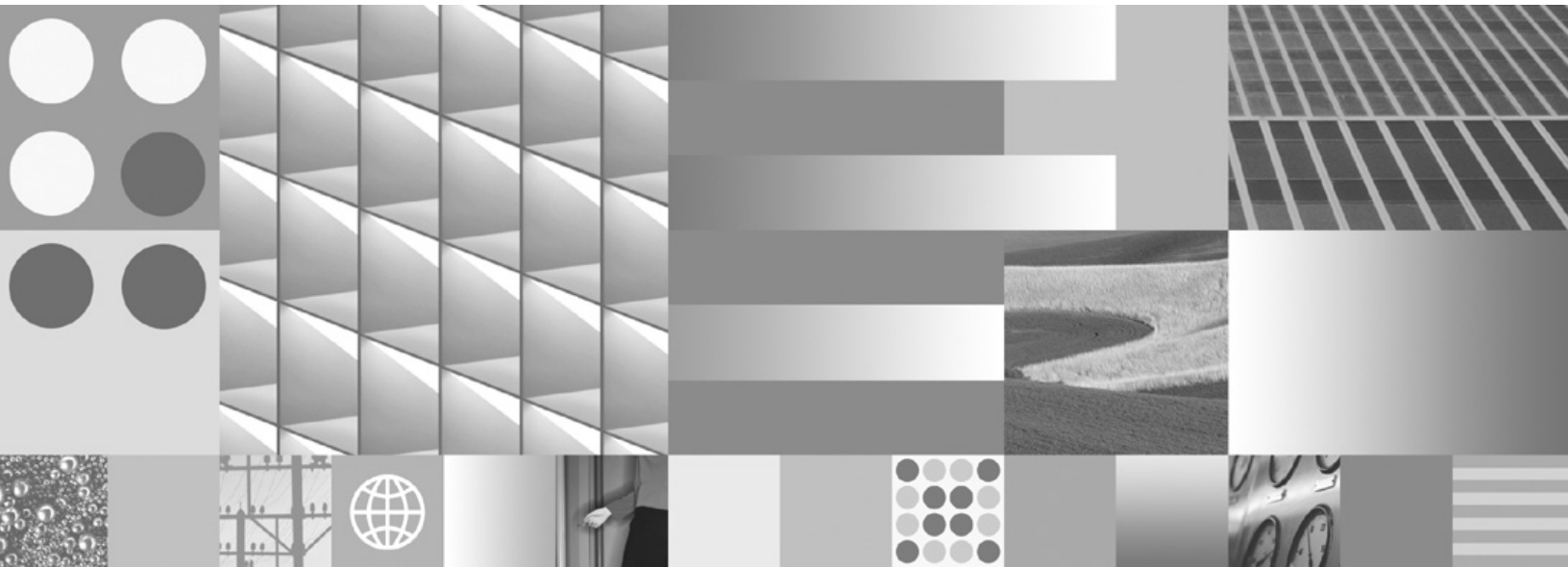


Izvajanje skrbništva nad iskanjem v podjetju



Izvajanje skrbništva nad iskanjem v podjetju

Opomba

Pred uporabo teh informacij in izdelka, ki ga opisujejo, preberite "Opombe in blagovne znamke" na strani 381.

Opomba k izdaji

Ta izdaja velja za različico 8, izdajo 5, raven popravkov 0 izdelka IBM OmniFind Enterprise Edition (številka izdelka 5724-C74) ter za vse nadaljnje izdaje in ravni popravkov, dokler ne bo v novih izdajah navedeno drugače.

S tem ko IBM-u pošljete informacije, IBM-u dodelite neizključne pravice, da vaše informacije distribuira na katerikoli način, za katerega meni, da je primeren, ne da bi si s tem naprtil kakršnekoli obveznosti do vas.

© Copyright International Business Machines Corporation 2004, 2008. Vse pravice pridržane.

Kazalo

ibm.com in povezani viri vii

Kako poslati komentarje vii

Stik z IBM-om viii

Kaj je iskanje v podjetju? 1

Tipi izvorov podatkov, ki jih podpira iskanje v podjetju 2

Pregled komponente za iskanje v podjetju 3

 Pajki iskanja v podjetju 4

 Razčlenjevalniki iskanja v podjetju 4

 Indeksi iskanja v podjetju 6

 Iskalni strežniki iskanja v podjetju 7

 Konzola za upravljanje iskanja v podjetju 8

 Nadzorovanje sistema iskanja v podjetju 8

 Datoteke dnevnika iskanja v podjetju 9

 Prilagajanje iskanja v podjetju 9

 Vzorčna iskalna aplikacija za iskanje v podjetju 10

Tok podatkov iskanja v podjetju 11

Skrbnišтво nad sistemom iskanja v podjetju 13

Prijava na skrbniško konzolo 16

Spreminjanje gesla skrbnika sistema iskanja v podjetju v konfiguraciji z enim strežnikom 17

Spreminjanje gesla skrbnika sistema iskanja v podjetju v konfiguraciji z več strežniki 18

Številke vrat TCP, ki se uporabljajo za iskanje v podjetju 20

 Spreminjanje številke vrat za sistem iskanja v podjetju 21

Spreminjanje gostiteljskih imen ali naslovov IP strežnika za iskanje v podjetju 22

Konfiguriranje podpore za dvojne naslove IP 24

Omogočitev podpore za protokol IPv6 25

Zbirke iskanja v podjetju 29

Izdelava zbirke s pomočjo čarovnika za zbirke 29

Izdelava zbirke s pogledom Zbirke 30

Urejanje zbirke 32

Brisanje zbirke 33

Ugotavljanje ID-ja zbirke 33

Skrbnišťvo nad pajkom 35

Izdelava pajka 36

Urejanje lastnosti pajka 37

Urejanje prostora za pajkanje 38

Brisanje pajka 38

Razporedi pajka 39

Pajki Content Edition 39

 Način neposrednega dostopa do repozitorijev Content Edition 41

 Način strežniškega dostopa do repozitorijev WebSphere II Content Edition 42

Pajki DB2 44

 Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2 46

 Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2 47

 Konfiguriranje programa WebSphere Information Integrator Event Publisher Edition za pajke DB2 48

 Konfiguriranje programa WebSphere MQ za pajke DB2 50

 Pajkanje baz podatkov DB2 na klasičnem strežniku vira podatkov 51

Pajki DB2 Content Manager 52

 Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2 Content Manager 53

 Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2 Content Manager 54

Pajki Domino Document Manager 56

Pajki strežnika Exchange Server 58

Pajki baze podatkov JDBC 58

 Preslikave razmerij za baze podatkov JDBC 60

 Pajkanje več strukturiranih tabel baze podatkov JDBC 62

Pajki NNTP 65

Pajki Notes 66

 Nasveti za pajkanje baz podatkov Lotus Domino 68

 Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino 69

 Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino 71

 Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP 72

 Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino 73

Pajki QuickPlace 75

Pajki seznama semen 78

Pajki datotečnega sistema UNIX 79

Spletni pajki 80

 Konfiguriranje uporabniškega posrednika 81

 Kako spletni pajek uporablja protokol za izključevanje robotov 82

 Podpora za JavaScript 83

 Pravila za omejevanje prostora za spletno pajkanje 84

 Preskušanje povezav z URL-ji s spletnim pajkom 88

 Nastavitve intervala vnovičnega pajkanja v spletnem pajku 88

 Možnosti za obiskovanje URL-jev s spletnim pajkom 89

 Kako spletni pajek obravnava strani s programskimi napakami 89

 Podpora za pajkanje zaščitenih spletnih mest 90

 Spletna mesta, ki jih strežejo strežniki proxy 92

 Skrbnišťvo nad piškotki 93

 Konfiguriranje globalnega prostora za spletno pajkanje 94

 Napotki za prepoved sledenja in indeksiranja 96

 Preglasitev ukazov no-follow in no-index na spletnih straneh 96

 Konfiguriranje datuma, ki ga bo spletni pajek uporabljal za že pajkane dokumente 97

Pajki Web Content Management 98

Pajki WebSphere Portal 100

 Kopiranje URL-ja za pajkanje iz programa WebSphere Portal 101

Pajki datotečnega sistema Windows	102
Konfiguriranje podpore za aplikacije poslušalnika podatkov	104
Prilagojeni vtičniki za pajke	105
Podpora za pajkanje arhivskih datotek	107
Formati URI v indeksu iskanja v podjetju	107

Skrbnišтво nad razčlenjevalnikom . . . 117

Delo s kategorijami	118
Kategorije, ki temeljijo na pravilih	118
Drevesa kategorij	120
Izbira vrste kategorizacije	121
Konfiguriranje kategorij	121
Delo z iskalnimi polji XML	122
Iskalna polja XML	123
Preslikovanje elementov XML v iskalna polja	123
Delo z iskalnimi polji HTML	126
Iskalna polja HTML	126
Preslikava elementov metapodatkov HTML v iskalna polja	126
Prilagojena obdelava besedila	128
Dodajanje motorjev za analizo besedila v sistem	129
Povezovanje motorja za analizo besedila z zbirko	130
Preslikovanje elementov XML v strukturo splošne analize	131
Preslikava strukture splošne analize v indeks	132
Preslikava strukture splošne analize v relacijsko bazo podatkov	133
Konfiguriranje niti za storitev razčlenjevalnika	134
Omogočanje zahtevnejše analize zloženek	135
Omogočanje podpore za iskanje po izvornem XML	135

Odkrivanje formata dokumentov . . . 137

Privzeti podprti tipi dokumentov	138
Tipi dokumentov, povezani z razčlenjevalniki zbirke in razčlenjevalniki Stellent	138
Povezovanje tipov dokumentov z razčlenjevalnikom zbirke	139
Privzeta pravila za storitev razčlenjevalnika zbirke	141
Razčlenjevanje neznanih tipov dokumentov	141
Spreminjanje pravil zamenjave za nekatere oznake HTML	142
Privzeta pravila zamenjave za HTML	143
Povezovanje tipov dokumentov z razčlenjevalnikom Stellent	144
Privzeta pravila za razčlenjevanje za razčlenjevalnike Stellent	146

Jezikovna podpora in podpora za kodno stran 151

Samodejno zaznavanje jezika	152
Samodejno zaznavanje kodne strani	153
Jezikovna analiza za kitajske, japonske in korejske dokumente	154
Segmentacija n-gram	154
Odstranjevanje presledkov iz besedila	154

Skrbnišтво nad indeksi 157

Razpored gradenj indeksov	158
Spreminjanje razporeda indeksiranja	159

Omogočanje in onemogočanje razporedov indeksiranja	159
Konfiguriranje sočasnih gradenj indeksa	160
Izgradnja indeksov samo v primeru odkritih sprememb	161
Zaustavljanje gradnje indeksov	162
Možnosti, ki vplivajo na iskalni pogled indeksa	163
Možnosti indeksiranja za iskanje dokumentov	163
Odkrivanje podvojenih dokumentov	167
Univerzalni znaki v poizvedbah	168
Območja	171
Konfiguriranje območij	171
Skrčeni URI-ji	172
Skrčenje URI-jev v rezultatih iskanja	173
Odstranitev URI-jev iz indeksa	174

Skrbnišтво nad iskalnim strežnikom 177

Iskalni predpomnilniki	177
Konfiguriranje iskalnega predpomnilnika	178
Prilagojeni slovarji sopomenk	178
Dodajanje slovarjev sopomenk v sistem	180
Povezovanje slovarja sopomenk z zbirko	180
Prilagojeni slovarji zaustavitvenih besed	181
Dodajanje slovarjev zaustavitvenih besed sistemu	182
Povezovanje slovarja zaustavitvenih besed z zbirko	182
Vnovična razmestitev slovarjev po meri	183
Dinamično povzemanje	184
Prilaganje povzetkov dokumentov na skrbniški konzoli	184
Prilaganje povzetkov dokumentov z urejanjem lastnosti	185
Delo s hitrimi povezavami	186
Hitre povezave	186
Konfiguriranje hitrih povezav	186

Rangiranje dokumentov 189

Ocenjevanje na osnovi besedila	189
Statično rangiranje	190
Obnavljanje privzetih vrednosti za statično rangiranje dokumentov	191
Prilagojeni slovarji poviševalnih besed	192
Dodajanje slovarjev poviševalnih besed v sistem	194
Povezovanje slovarja poviševalnih besed z zbirko	194
Rangiranje dokumentov, ki temelji na vzorcih URI	195
Vplivanje na ocene dokumentov, ki ustrezajo vzorcem URI-jev	195
Rangiranje dokumentov, ki temelji na poviševalnih razredih	196
Preslikava polj v poviševalne razrede	198
Konfiguriranje poviševalnih faktorjev za poviševalne razrede	198
Privzete vrednosti poviševalnega razreda	199

Iskalne aplikacije za iskanje v podjetju 201

Povezovanje iskalnih aplikacij z zbirkami	202
Funkcije vzorčne iskalne aplikacije	202
Lastnosti iskalnih aplikacij	203
Urejanje lastnosti vzorčne iskalne aplikacije	218
Prilaganje iskalnih aplikacij	220
Kloniranje vzorčne iskalne aplikacije	221
Analiziranje najboljših rezultatov	223

Dostopanje do iskalnih aplikacij	226
Konfiguriranje iskalnih strežnikov, da sprejmejo samo zaščitene (SSL) zahteve	226
Konfiguriranje iskalnih strežnikov za sprejemanje zahtev prek strežnika proxy	228

Podpora za zunanje vire 229

Dodajanje zunanjih virov v sistem	229
Povezovanje iskalnih aplikacij z zunanjimi viri	231

Zaščita iskanja v podjetju 233

Namestitvena zaščita	234
Overjanje v primerjavi z nadzorom dostopa	235
Skrbniške vloge	235
Konfiguriranje skrbniških uporabnikov	236
Zaščita na ravni zbirke	237
Analiza podvojenih dokumentov in zaščita zbirke	237
Identifikatorji iskalne aplikacije	237
Zaščita na ravni dokumenta	238
Pred- in po-filtriranje rezultatov iskanja	239
Potrjevanje s shranjenimi žetoni zaščite	240
Potrjevanje trenutnih poverilnic med obdelavo poizvedbe	241
Analiza sidrnega besedila	247
Omogočanje zaščite za iskanje v podjetju	248
Konfiguriranje globalne zaščite in uporabniškega registra LDAP na aplikacijskem strežniku WebSphere Application Server	249
Omogočanje zaščite za sistem iskanja v podjetju za en sam strežnik	251
Omogočanje zaščite za sistem iskanja v podjetju za več strežnikov	252
Zahteve za nastavitve pajka za podporo zaščite	253
Preverjanje dostopa do zaščitene dokumentov strežnika Exchange Server	256
Zagotavljanje zaščite na ravni dokumenta za dokumente Lotus Domino	256
Zagotavljanje zaščite na ravni dokumenta za dokumente datotečnega sistema Windows	259
Onemogočanje zaščite za iskanje v podjetju	262
Onemogočanje zaščite za aplikacijo podjetja na aplikacijskem strežniku WebSphere Application Server	262
Onemogočanje zaščite na ravni dokumenta	264
Onemogočanje zaščite za skrčene rezultate iskanja	265

Zagon in zaustavitev sistema iskanja v podjetju 267

Zagon sistema iskanja v podjetju	267
Zaustavitev sistema iskanja v podjetju	269
Krmiljenje, katere komponente bodo zagnane ali zaustavljene	270
Izvajanje skrbništva nad iskalnimi strežniki v samostojnem načinu	271

Nadziranje dejavnosti iskanja v podjetju 273

Ocenjevanje števila dokumentov v zbirki	273
Nadziranje zbirke	274
Prikaz podrobnosti o URI-jih	274

Nadziranje pajkov	275
Prikaz podrobnosti o dejavnostih spletnega pajka	276
Podrobnosti o nitih spletnega pajka	277
Aktivna mesta spletnega pajka	278
Hitrost pajkanja spletnega pajka	278
Izdelava poročil za spletnega pajka	279
Statusne kode HTTP, vrnjene spletnemu pajku	280
Nadziranje razčlenjevalnika	284
Nadziranje dejavnosti indeksa za zbirko	285
Nadziranje čakalne vrste indeksov za iskanje v podjetju	285
Nadziranje iskalnih strežnikov	286
Spreminjanje načina izračuna statističnih podatkov poizvedb	287
Nadziranje poslušalnika podatkov	288
Sledenje dokumentov	288
Konfiguriranje datotek dnevnika za sledenje dokumentov	289
Prikazovanje poročil o odstranjenih dokumentih	289
Prikazovanje datotek dnevnika o odstranjenih dokumentih	290

Datoteke dnevnika in opozorila 293

Opozorila	293
Konfiguriranje opozoril na ravni zbirke	294
Konfiguriranje opozoril na ravni sistema	295
Konfiguriranje datotek dnevnika	296
Konfiguriranje informacij strežnika SMTP	297
Sprejemanje elektronske pošte o zabeleženih sporočilih	298
Spreminjanje velikosti dnevnika poizvedb	299
Prikaz datotek dnevnika	300

Varnostno kopiranje in obnavljanje sistema iskanja v podjetju 303

Izdelava varnostne kopije sistema iskanja v podjetju	304
Obnovitev sistema iskanja v podjetju	305
Izvažanje in uvažanje konfiguracij zbirke	306

Integracija v Lotus Notes različice 8 309

Izdelava spletnega mesta s posodobitvami za vtičnik za iskanje v podjetju	309
Nameščanje vtičnika za iskanje v podjetju v odjemalca Lotus Notes različice 8	309

Integracija s portalom WebSphere Portal 311

Nastavitveni skripti za integracijo sistema iskanja v podjetju v program WebSphere Portal	312
Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 5.1	313
Konfiguriranje iskalne vrstice izdelka WebSphere Portal različice 5.1 za uporabo sistema iskanja v podjetju	315
Odstranjevanje sistema iskanja v podjetju iz izdelka WebSphere Portal različice 5.1	317
Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 6	318
Konfiguriranje Centra za iskanje izdelka WebSphere Portal različice 6 za sistem iskanja v podjetju	320

Konfiguriranje iskalne vrstice izdelka WebSphere Portal različice 6 za uporabo sistema iskanja v podjetju	322
Nastavitev portalskega programčka za iskanje v podjetju za Lotus Quickr	323
Odstranjevanje sistema iskanja v podjetju iz izdelka WebSphere Portal različice 6	324
Integracija iskanja v podjetju z gručnimi sistemi	
WebSphere Portal	325
Nastavitev iskanja v podjetju v gručnem sistemu	
WebSphere Portal	326
Odstranjevanje iskanja v podjetju iz gručnega sistema	
WebSphere Portal	329

Selitev iz izdelka WebSphere Portal v iskanje v podjetju. 331

Selitev zbirke s portala WebSphere Portal	331
Nastavitve preseljene zbirke	332
Datoteka dnevnika čarovnika za selitev	334

Ukazi za iskanje v podjetju, povratne kode in ID-ji sej 335

Občutljivost na velike in male črke pri iskanju v podjetju 363

Dokumentacija iskanja v podjetju . . . 365

Pripomočki za osebe s posebnimi potrebami 367

Pojmovnik izrazov iskanja v podjetju 369

Opombe in blagovne znamke. 381

Opombe	381
Blagovne znamke	383

Stvarno kazalo 385

ibm.com in povezani viri

Podpora in dokumentacija za izdelek so na voljo na spletnem mestu [ibm.com](http://www.ibm.com).

Podpora in pomoč

Podpora za izdelek je na voljo na spletu.

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>

Informacijski center

Dokumentacijo izdelka si lahko ogledate s spletnim brskalnikom v informacijskem centru, ki temelji na tehnologiji Eclipse. Oglejte si informacijski center na naslovu <http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/>.

Publikacije v formatu PDF

Datoteke PDF si lahko ogledate v zaslonski obliki s programom Adobe Acrobat Reader za svoj operacijski sistem. Če tega programa nimate nameščenega, ga lahko prenesete z Adobeovega spletnega mesta na naslovu <http://www.adobe.com>.

Oglejte si naslednja spletna mesta s publikacijami v obliki PDF:

Izdelek	Naslov spletnega mesta
OmniFind Enterprise Edition, Version 8.5	http://www.ibm.com/support/docview.wss?rs=63&uid=swg27010938
OmniFind Discovery Edition, Version 8.4	http://www.ibm.com/support/docview.wss?rs=3035&uid=swg27008552
OmniFind Yahoo! Edition, Version 8.4	http://www.ibm.com/support/docview.wss?rs=3193&uid=swg27008932

Kako poslati komentarje

Vaše povratne informacije so zelo pomembne, saj nam pomagajo pri podajanju natančnih in visoko kakovostnih informacij.

Svoje komentarje pošljite z zaslonskim obrazcem za komentarje, ki ga najdete na naslovu https://www14.software.ibm.com/webapp/iwm/web/signup.do?lang=en_US&source=swg-rcf.

Stik z IBM-om

Če želite stopiti v stik z IBM-ovo servisno službo za stranke v Združenih državah Amerike ali v Kanadi, pokličite na številko 1-800-IBM-SERV (1-800-426-7378).

Če vas zanima, katere možnosti servisiranja so na voljo, pokličite eno od naslednjih števil:

- V Združenih državah Amerike: 1-888-426-4343
- V Kanadi: 1-800-465-9600

Dodatne informacije o stiku z IBM-om boste našli na spletni strani stika z IBM-om na naslovu <http://www.ibm.com/contact/us/>.

Kaj je iskanje v podjetju?

Sistem iskanja v podjetju nudi obsežne zmožnosti za preiskovanje kakršnegakoli števila strukturiranih in nestrukturiranih izvorov podatkov z eno samo poizvedbo. Hitri odzivni časi poizvedbe in združen, urejen nastali niz, ki temelji na obširni analizi besedila, ne omogočata le iskanja dokumentov, pač pa tudi povzetje pomena iz vsebine dokumentov.

Komponente za iskanje v podjetju, ki jih namestite skupaj z izdelkom IBM OmniFind Enterprise Edition, zbirajo informacije v celotnem podjetju. Z vnosom poizvedbe v spletni pregledovalnik lahko sočasno preiskujete lokalne in oddaljene baze podatkov, sisteme za sodelovanje, sisteme za upravljanje vsebine, datotečne sisteme ter notranja in zunanja spletna mesta.

Sistem iskanja v podjetju, ki je oblikovan tako, da se popolnoma integrira z obstoječimi sistemi, obravnava logistiko, potrebno za zbiranje podatkov iz različnih izvorov in indeksiranje podatkov za hitro pridobivanje. Z uveljavitvijo lingvistične analize in drugih tipov analize v podatkih lahko iskanje v podjetju poda rezultate iskanja visoke ravni ustreznosti. Za preiskovanje različnih tipov repozitorijev vam ni treba poznati različnih vmesnikov.

Dodate lahko podporo za preiskovanje izvorov podatkov, ki jih ne želite vključiti v indeks iskanja v podjetju. Z zmožnostjo zveznega iskanja, ki je del iskanja v podjetju, lahko preiskujete zunanje izvore sočasno s preiskovanjem indeksiranih izvorov podatkov.

Kakovost iskanja

OmniFind Enterprise Edition podpira arhitekturo UIMA (IBM Unstructured Information Management Architecture), s pomočjo katere zagotovi, da uporabniki najdejo iskane informacije. UIMA je odprt sestav, ki definira splošen, standardni vmesnik za analiziranje besedila. Z obsežno semantično analizo lahko iskanje v podjetju identificira koncepte, latentne pomene, odnose, dejstva in druge ustrezne podatke, ki so pogosto skriti v nestrukturiranem besedilu. Informacije, ki jih povzamete med analizo, lahko uporabite za izboljšanje kakovosti rezultatov iskanja ali za izboljšanje kakovosti drugih aplikacij, kot so na primer aplikacije poslovne inteligence in rudarjenja po podatkih.

Zaščita

Zaščita je integralni element iskanja v podjetju. Uporabljajo jo lahko samo tisti uporabniki, ki imajo pooblastilo za upravljanje sistema. S pomočjo varnostnih mehanizmov, ki jih vključuje IBM WebSphere Application Server, lahko konfigurirate skrbniške vloge in določite, kateri uporabniki bodo imeli dostop do raznih skrbniških funkcij.

Podate lahko tudi možnosti za povezovanje žetonov zaščite s podatki med zbiranjem le-teh. Če vaše iskalne aplikacije omogočajo zaščito, lahko uporabite žetone, ki so z dokumenti shranjeni v indeksu, za uveljavitev nadzora dostopa in zagotovitev, da lahko samo uporabniki z ustreznimi priporočili poizvedujejo po podatkih in prikažejo rezultate iskanja.

Za določene tipe izvorov podatkov lahko konfigurirate možnosti za preverjanje prijavnih priporočil uporabnika s trenutnim nadzorom dostopa med obdelavo poizvedbe. Ta dodatna plast zaščite zagotavlja, da so pooblastila uporabnika preverjena v realnem času z izvirnim izvorom podatkov. Ta možnost omogoča tudi zaščito pred primerki, v katerih se priporočila uporabnika spremenijo po indeksiranju dokumenta in njegovih žetonov zaščite.

Vadnica za izdelek

Na naslovu <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html> je na voljo spletna vadnica. Vadnica opisuje posamezne korake, potrebne za konfiguracijo in namestitve izdelka, prikazuje načine iskanja po posameznih tipih podatkovnih izvorov ter opisuje, kako s pomočjo aplikacijskega programerskega vmesnika izdelka razširite iskanje v podjetju. Čeprav se vadnica nanaša na starejšo različico izdelka OmniFind Enterprise Edition, je večina zasnov in postopkov še vedno aktualna.

S tem povezani pojmi

“Pregled komponente za iskanje v podjetju” na strani 3

“Tok podatkov iskanja v podjetju” na strani 11

Zaščita iskanja v podjetju



Integracija analize besedila po meri



Osnovni koncepti, uporabljeni pri obdelavi analize besedila

S tem povezane povezave

“Tipi izvorov podatkov, ki jih podpira iskanje v podjetju”

Tipi izvorov podatkov, ki jih podpira iskanje v podjetju

Za preiskovanje raznolikih tipov izvorov podatkov je na voljo vnaprej definirana podpora.

Potem ko namestite IBM OmniFind Enterprise Edition, lahko začnete z zbiranjem podatkov iz naslednjih tipov podatkovnih izvorov:

- Tipi postavk IBM DB2 Content Manager (dokumenti, viri in postavke)
- Baze podatkov IBM DB2
- Baze podatkov IBM Domino Document Manager (prej znan kot Domino.Doc)
- Baze podatkov IBM Lotus Notes
- Baze podatkov IBM Lotus QuickPlace
- Knjižnice vsebine IBM Lotus Quickr
- Repozitoriji IBM WebSphere Information Integrator Content Edition vključno z izdelki Documentum, FileNet Panagon Content Services, FileNet P8 Content Manager, Hummingbird Document Management (DM), Microsoft SharePoint, OpenText Livelink Enterprise Server in WebSphere Portal Document Manager (PDM)
- Tabele vzdevkov IBM WebSphere Information Integrator za številne tipe sistemov baz podatkov, vključno s sistemi IBM DB2 for z/OS, IBM Informix, Microsoft SQL Server, Oracle in Sybase
- Spletna mesta IBM WebSphere Portal
- Spletna mesta IBM Workplace Web Content Management
- Javne mape Microsoft Exchange Server
- Baze podatkov Microsoft SQL Server
- Datotečni sistemi Microsoft Windows
- Novičarske skupine NNTP (Network news transfer protocol)
- Baze podatkov Oracle
- Datotečni sistemi UNIX
- Spletna mesta na internetu ali intranetu

Dodate lahko tudi podporo za preiskovanje naslednjih tipov zunanjih izvorov brez dodajanja dokumentov iz teh izvorov v indeks iskanja v podjetju:

- Baze podatkov, ki podpirajo protokol JDBC (povezljivost baze podatkov Java (samo sistemi baze podatkov DB2 in Oracle). Za vsako tabelo, ki jo omogočite za iskanje, je izdelan ločen zunanji izvor.
- Strežniki LDAP (Lightweight Directory Access Protocol). Za vsak strežnik LDAP je izdelan po en zunanji izvor.

Najnovejše informacije o podprtih tipih izvorov podatkov in podprte različice izdelka najdete na strani sistemskih zahtev na spletnem mestu OmniFind Enterprise Edition Support.

S tem povezani pojmi

Podpora za zunanje izvore

“Tok podatkov iskanja v podjetju” na strani 11

Pregled komponente za iskanje v podjetju

Komponente iskanja v podjetju zbirajo podatke iz celotnega podjetja, analizirajo, razčlenijo in kategorizirajo informacije ter izdelajo indeks, po katerem lahko iščejo uporabniki.

Zbirka za iskanje v podjetju predstavlja niz izvorov, ki jih lahko uporabniki preiščejo z eno samo poizvedbo. Pri izdelavi zbirke podate, katere izvore želite vključiti, in konfigurirate možnosti, ki določajo, kako lahko uporabniki preiskujejo indeksirane podatke.

Izdelate lahko več zbirk, vsaka med njimi pa lahko vsebuje podatke iz različnih izvorov podatkov. Tako lahko na primer izdelate zbirko, ki vključuje dokumente iz baz podatkov IBM DB2 Universal Database, IBM Lotus Notes in IBM DB2 Content Manager. Če uporabniki preiščejo to zbirko, lahko rezultati iskanja vključujejo dokumente iz vsakega izvora podatkov.

Podpora za zvezno iskanje uporabnikom omogoča, da z eno samo poizvedbo preiščejo več kot eno zbirko. Rezultati iskanja lahko vsebujejo dokumente iz vseh zbirk in zunanjih izvorov v sistemu za iskanje v podjetje.

Izdelava in upravljanje zbirke vključuje naslednje dejavnosti:

Zbiranje podatkov

Komponente *pajka* zbirajo dokumente iz izvorov podatkov, in sicer neprekinjeno ali glede na urnik, ki ga podate. S pogostim pajkanjem zagotovite, da imajo uporabniki vedno dostop do najnovejših informacij.

Analiziranje podatkov

Komponente *razčlenjevalnika* povzamejo besedilo iz dokumentov in v vsakem dokumentu, v katerem pajka pajek, opravijo lingvistično analizo in druge vrste analiz. Podrobna analiza vsebine izboljša kakovost rezultatov iskanja.

Indeksiranje podatkov

Komponente *indeksiranja* se izvajajo po razporedu in v indeks dodajajo informacije o novih in spremenjenih dokumentih. Komponente indeksiranja opravljajo tudi globalno analizo dokumentov v zbirki, s čimer izboljšajo kakovost rezultatov iskanja.

Iskanje podatkov

Komponente *iskanja* preiščejo indeks in delajo z iskalnimi aplikacijami, da obdelajo iskalne zahteve in vrnejo rezultate iskanja.

Druge komponente OmniFind Enterprise Edition omogočajo, da podate nastavitve zaščite, nadzorujete dejavnost sistema in odpravite težave, ki se pojavijo. Izdelek nudi tudi delovno vzorčno iskalno aplikacijo, ki jo lahko uporabite kot predlogo za izdelavo lastnih iskalnih aplikacij.

S tem povezani pojmi

Skrbništvo nad pajkom
Skrbništvo nad razčlenjevalnikom
Skrbništvo nad indeksi
Skrbništvo nad iskalnim strežnikom
“Tok podatkov iskanja v podjetju” na strani 11

Pajki iskanja v podjetju

Pajki iskanja v podjetju zbirajo dokumente iz izvorov podatkov, da je dokumente mogoče analizirati, indeksirati in preiskati.

Komponenta pajka, ki je na voljo z izdelkom OmniFind Enterprise Edition, ima naslednje funkcije:

- Pri konfiguriranju pajka procesi *odkrivanja* poiščejo informacije o izvorihi, ki so na voljo za pajkanje, kot so na primer imena vseh pogledov in map v bazi podatkov Lotus Notes ali imena vseh datotečnih sistemov na strežniku UNIX.
- Ko izberete izvore, v katerih želite pajkati in zaženete pajka, komponente pajka zberejo podatke iz izvorov, da je podatke mogoče analizirati in indeksirati.

Ena zbirka ima lahko več pajkov, vsak pajek pa je oblikovan tako, da zbira podatke iz določenega tipa izvora podatkov. Tako lahko na primer izdelate tri pajke, ki v eni zbirki združijo podatke iz datotečnih sistemov, baz podatkov Notes in relacijskih baz podatkov. Nastavite pa lahko tudi več pajkov istega tipa in zanje nastavite različne urnike pajkanja glede na to, kako pogosto se spreminjajo podatki, v katerih pajka posamezen pajek.

Ko zaženete spletnega pajka, se le-ta neprekinjeno izvaja. Podati mu morate URL-je (uniform resource locators), v katerih želite pajkati, pajek pa se občasno vrne, da preveri nove ali spremenjene podatke. Druge tipe pajkov lahko zaženete in zaustavite ročno ali pa nastavite urnike pajkanja. Ko določite urnik pajka, podate, kdaj se bo prvič zagnal in kako pogosto se mora vrniti v izvore podatkov, kjer izvede pajkanje v novih in spremenjenih dokumentih.

Lastnosti pajka so niz pravil, ki določajo vedenje določenega pajka pri pajkanju. Tako na primer podate pravila, ki nadzorujejo, kako pajek uporablja sistemske vire. Niz izvorov, ki so primerni za pajkanje, sestavlja *prostor za pajkanje* pajka. Po izdelavi pajka lahko kadarkoli uredite lastnosti pajka in spremenite način, na katerega pajek zbira podatke. Uredite lahko tudi prostor za pajkanje in spremenite urnik pajka, dodate nove izvore ali odstranite izvore, ki jih ne želite več preiskovati.

S tem povezani pojmi

Skrbništvo nad pajkom

S tem povezana opravila

Nadzorovanje pajkov

Razčlenjevalniki iskanja v podjetju

Razčlenjevalnik iskanja v podjetju analizira dokumente, ki jih je zbral pajek, in jih pripravi za indeksiranje.

Komponenta razčlenjevalnika, ki je na voljo z izdelkom OmniFind Enterprise Edition, analizira vsebino in metapodatke dokumentov. Rezultate analize shrani v podatkovno shrambo, do katere lahko dostopa komponenta za indeksiranje. Razčlenjevalnik izvaja naslednje naloge:

- Povzame besedilo iz kateregakoli formata, v katerem je dokument. Tako na primer povzame besedilo iz oznak v dokumentih XML in HTML. Razčlenjevalnik s pomočjo izdelka Oracle (prej znan pod imenom Stellent) Outside In Content Access for IBM

OmniFind Enterprise Edition ekstrahira besedilo iz dvojiških formatov, kot so na primer dokumenti Microsoft Word in Adobe Acrobat PDF (prenosljivi format dokumentov).

- Zazna kodiranja nabora znakov vsakega dokumenta. Preden razčlenjevalnik izvede lingvistično analizo, uporabi te informacije za pretvorbo vsega besedila v Unicode.
- Zazna izvorni jezik vsakega dokumenta.
- Povzame besedilo in doda žetone za izboljšanje branja podatkov. V tej fazi izvede razčlenjevalnik naslednje naloge:
 - Normalizacija podatkov, kot je na primer normalizacija velikih črk in diakritičnih znakov, kot je na primer nemški preglas.
 - Analiza strukture odstavkov, stavkov, besed in presledkov. Prek lingvistične analize razčlenjevalnik razstavi sestavljene besede in dodeli žetone, ki omogočajo iskanje slovarjev in sopomenk.
- Uveljavi pravila razčlenjevanja, ki jih podate za zbirko. Pri konfiguriranju razčlenjevalnika lahko konfigurirate naslednje dejavnosti razčlenjevanja:

Pravila preslikave polj za dokumente XML in HTML

Ta možnost uporabnikom omogoča, da iščejo strukturirano in nestrukturirano vsebino v dokumentih XML in HTML. Če preslikate elemente XML ali elemente metapodatkov HTML v iskalna polja v indeksu iskanja v podjetju, lahko uporabniki podajo v poizvedbah imena polj in iščejo določene dele dokumentov XML in HTML. Poizvedbe, ki iščejo določena polja, lahko podajo natančnejše rezultate iskanja kot proste besedilne poizvedbe, ki preiščejo vso vsebino dokumentov.

Kategorije

Ta možnost uporabnikom omogoča za iščejo v dokumentih s kategorijami, v katere spadajo dokumenti. Uporabniki lahko tudi izberejo kategorije v rezultatih iskanja in pregledujejo samo tiste dokumente, ki spadajo v isto kategorijo.

Pri izdelavi zbirke lahko podate, ali želite uporabiti kategorizacijo. S kategorijami, *ki temeljijo na pravilih*, so dokumenti povezani s kategorijami glede na pravila, ki jih definirate. Kategorije, ki temeljijo na pravilih, lahko konfigurirate z zbirkami iskanja v podjetju, ki jih izdelate, in z zbirkami, ki jih preselite iz portala IBM WebSphere.

Analiza besedila po meri

Razvijalci aplikacij lahko izdelajo programe za analizo po meri, ki izvajajo kompleksno lingvistično analizo podatkov, ki jih morate preiskati. Te programe lahko vključite v sistem iskanja v podjetju in jih uporabite za označevanje vsebine zbirk. Z indeksiranjem opomb omogočite zbirke za semantično iskanje.

Uporabniki lahko tako na primer iščejo izraze poizvedbe, ki se pojavijo eden blizu drugega ali v istem stavku, ali pa iščejo odnose med izrazi poizvedbe. Uporabnik lahko na primer potrebuje dokumente, v katerih je omenjen IBM-ov prodajalec s priimkom Šeruga in ne IBM-ov inženir s priimkom Šeruga.

Podpora za n-gram segmentacijo

Za izboljšanje branja dokumentov, napisanih v kitajščini, japonsščini ali korejščini, lahko omogočite metodo n-gram segmentacije leksične analize. Ta oblika analize za ločevanje besed ne uporablja presledkov. Po izdelavi zbirke načina segmentacije ni več mogoče spremeniti.

Podpora za preiskovanje dokumentov XML z izvornim iskanjem XML

Izvirno iskanje XML lahko z iskanjem označevanja XML nudi natančnejše rezultate iskanja. Tako lahko na primer poizvedba podaja, da se mora beseda pojaviti v določenem elementu XML.

Razredi za povišanje točk relativne pomembnosti polj

Če preslikate polja v razrede povišanja, lahko vplivate na to, kako so dokumenti razvrščeni v rezultatih iskanja. Rezultat polj z naslovom lahko na primer povišate, če želite zagotoviti, da bodo v primeru, če se izraz poizvedbe pojavi v naslovu, dokumenti, ki vsebujejo ta izraz v svojem naslovu, v rezultatih iskanja uvrščeni višje.

S tem povezani pojmi

Delo s kategorijami

Iskalna polja XML

Iskalna polja HTML



Integracija analize besedila po meri



Analiza besedila, vključenega v iskanje v podjetju

S tem povezana opravila

Nadzorovanje razčlenjevalnika

Indeksi iskanja v podjetju

Komponente za indeksiranje iskanja v podjetju se izvajajo po razporedu in v indeks dodajajo informacije o novih in spremenjenih dokumentih.

Za zagotovitev, da imajo uporabniki vedno dostop do najnovejših informacij v izvorihi, ki jih preiskujejo, vključuje izdelava indeksa dve stopnji:

Izdelava glavnega indeksa

Med izdelavo glavnega indeksa je znova izdelan celoten indeks, tako da ima struktura optimalno ureditev. Procesi indeksiranja preberejo vse podatke, ki so jih zbrali pajki in jih je razčlenil razčlenjevalnik.

Izdelava delta indeksov

Med izdelavo delta indeksa so v indeks dodane informacije, v katerih je bilo izvedeno pajkanje od zadnje izdelave glavnega indeksa.

Pri konfiguriranju indeksnih možnosti za zbirko podate urnike za izdelavo glavnih in delta indeksov. Pogostost, s katero izdelate indekse, je odvisna od sistemskih virov in od tega, ali viri, ki jih indeksirate, vsebujejo statično ali dinamično vsebino.

Za zagotovitev razpoložljivosti novih informacij načrtujte pogosto izdelavo delta indeksov. Občasno načrtujte tudi izdelavo glavnega indeksa, v katerem bodo združene vse nove informacije, bo analizirana nova vsebina in optimizirana zmogljivost indeksa.

Procese indeksiranja lahko zaženete tudi, ne da bi jih načrtovali. Če na primer spremenite določena pravila razčlenjevanja in želite, da so te spremembe na voljo za iskalne aplikacije, lahko po vnovičnem pajkanju in razčlembi podatkov zaženete glavni indeks, namesto da bi čakali, da se izdelava indeksa začne ob zanj načrtovanem času.

Za nadzorovanje uporabe virov lahko nadzorujete, koliko zbirk lahko souporablja indeksne procese in sočasno predložite zahteve za izdelavo indeksa. S sočasno izdelavo indeksov zagotovite, da z izdelavo zelo velikega glavnega indeksa ne boste blokirali izdelav delta indeksov za druge zbirke. Izdelava indeksov lahko zahteva veliko virov, zato morate za velike sisteme nadzorovati obremenitev sistema in prilagoditi urnike za izdelavo glavnega in delta indeksa.

Pri izdelavi indeksa opravijo indeksni procesi globalno analizo dokumentov. V tej fazi so uveljavljeni algoritmi za določitev podvojenih dokumentov, analiziranje povezovalne

strukture dokumentov in za izvajanje posebne obdelave sidrnega besedila (besedila, ki opisuje ciljno stran v hiperbesedilni povezavi) v spletnih dokumentih.

Podate lahko možnosti za naslednje dejavnosti indeksiranja:

- Če želite uporabnikom omogočiti uporabo univerzalnih znakov, lahko v indeks vdelate podporo za razširitev izrazov poizvedbe ali podate, da želite med obdelavo poizvedbe razširiti izraze poizvedbe. Odločitev, ki jo sprejmete, je kompromis med porabo virov in odzivnim časom poizvedbe.
- Konfigurirate lahko tudi območja. *Območje* omogoča, da omejite informacije, so uporabnikom prikazane v zbirki. Tako lahko na primer izdelate eno območje, ki vključuje URI-je dokumentov v oddelku za tehnično podporo, drugo območje pa za URI-je dokumentov v kadrovskem oddelku. Če nudi iskalna aplikacija podporo za območja, lahko uporabniki preiščejo in pridobijo dokumente samo iz teh podnizov zbirke.
- Podate lahko možnosti za skrčenje dokumentov v rezultatih iskanja, ki imajo enako predpono URI. Podate lahko tudi ime skupine, tako da so lahko dokumenti z različnimi priponami URI v rezultatih iskanja skrčeni skupaj.
- Po izdelavi indeksa lahko odstranite URI-je, za katere ne želite, da uporabniki v njih izvajajo iskanje.

S tem povezani pojmi

Skrbništvo nad indeksi

Univerzalni znaki v poizvedbah

Območja

Skrčeni URI-ji

Rangiranje dokumentov, ki temelji na vzorcih URI

S tem povezana opravila

Razporejanje izdelave indeksov

Konfiguriranje sočasne izdelave indeksov

Odstranitev URI-jev iz indeksa

Nadzorovanje dejavnosti indeksov za zbirko

Nadzorovanje čakalne vrste indeksov za iskanje v podjetju

Iskalni strežniki iskanja v podjetju

Iskalni strežniki iskanja v podjetju delajo z iskalnimi aplikacijami, da obdelajo poizvedbe, preiščejo indeks in vrnejo rezultate iskanja.

Iskalni strežniki iskanja v podjetju se namestijo pri namestitvi izdelka OmniFind Enterprise Edition. Pri konfiguriranju iskalnih strežnikov za zbirko lahko podate možnosti za način preiskovanja zbirke:

- Konfigurirate lahko iskalni predpomnilnik, ki bo vseboval pogosto zahtevane rezultate iskanja. Iskalni predpomnilnik lahko izboljša zmogljivost iskanja in pridobivanja.
- Za iskanje dokumentov v zbirki lahko podate privzet jezik.
- Če izdelajo razvijalci aplikacij slovarje po meri, jih lahko povežete z zbirkami:
 - Če uporabniki izvajajo poizvedbo v zbirki, ki uporablja *slovar sopomenk*, bodo dokumenti, ki vsebujejo sopomenke izrazov poizvedbe, vključeni v rezultate iskanja.
 - Če uporabniki izvajajo poizvedbo v zbirki, ki uporablja *slovar zaustavitvenih besed*, bodo zaustavitvene besede odstranjene iz poizvedbe pred njeno obdelavo.
 - Če uporabniki izvajajo poizvedbo v zbirki, ki uporablja *slovar pospeševalnih besed*, je pomembnost dokumentov, ki vsebujejo pospeševalne besede, povečana ali zmanjšana glede na faktor pospeševanja, ki je povezan z besedo v slovarju.

- Če vnaprej določite, da so določeni dokumenti ustrezni za določene poizvedbe, lahko konfigurirate hitre povezave. *Hitra povezava* povezuje določen URI z določenimi ključnimi besedami in frazami. Če vsebuje poizvedba katerekoli ključne besede ali fraze, ki jih podate v definiciji hitre povezave, je povezan URI samodejno vrnjen v rezultatih iskanja.

V konfiguraciji z več strežniki je zaščita pred napakami na voljo na ravni zbirke in ne samo na ravni strežnika. Če zbirka na enem od iskalnih strežnikov zaradi kakršnegakoli vzroka postane nerazpoložljiva, so poizvedbe za to zbirko samodejno usmerjene na drug iskalni strežnik.

S tem povezani pojmi

Iskalne aplikacije za iskanje v podjetju

Iskalni predpomnilniki

Slovarji sopomenk po meri

Slovarji zaustavitvenih besed po meri

Slovarji pospeševalnih besed po meri

Hitre povezave

S tem povezana opravila

Nadzorovanje iskalnih strežnikov

Konzola za upravljanje iskanja v podjetju

Konzola za upravljanje iskanja v podjetju se izvaja v pregledovalniku, kar pomeni, da lahko upravni uporabniki do nje dostopijo kadarkoli in s kateregakoli mesta. Mehanizmi zaščite zagotavljajo dostop samo tistim uporabnikom, ki imajo dostop do upravnih funkcij.

Konzola za upravljanje iskanja v podjetju je nameščena na iskalne strežnike pri namestitvi OmniFind Enterprise Edition.

Konzola za upravljanje vključuje čarovnike, ki vam pomagajo pri izvedbi številnih osnovnih upravnih nalog. Čarovnik za zbirke vam na primer pomaga izdelati zbirko in omogoča, da shranite delo v načinu osnutka. Čarovniki pajka so specifični za tip izvora podatkov, in vam pomagajo izbrati izvore, v katerih želite uporabnikom omogočiti iskanje.

Za druge upravne naloge lahko izberete posamezne postavke, ki jih želite upravljati. Če na primer urejate zbirko, lahko izberete stran Indeks, na kateri spremenite urnik indeksiranja ali pa izberete stran Razčlenjevanje, na kateri spremenite pravilo za razčlenjevanje dokumentov XML.

S tem povezani pojmi

Upravljanje sistema za iskanje v podjetju

Upravne vloge

S tem povezana opravila

Prijava na konzolo za upravljanje

Nadzorovanje sistema iskanja v podjetju

Konzolo za upravljanje iskanja v podjetju lahko uporabite za nadzorovanje sistemskih dejavnosti in za prilagoditev operacij, ko je to potrebno.

Ko namestite OmniFind Enterprise Edition in izdelate vsaj eno zbirko, si lahko ogledate podrobne statistične podatke za vsako glavno dejavnost (pajkanje, razčlenjevanje, indeksiranje in iskanje). Podatki vključujejo informacije o povprečnih odzivnih časih in poteku, kot je na primer, v koliko dokumentih je bilo med določeno sejo pajkanja ali izdelave indeksa izvedeno pajkanje ali indeksiranje.

Večino dejavnosti lahko zaustavite in zaženete. Dejavnost lahko na primer začasno ustavite, spremenite njeno konfiguracijo ali odpravite težavo, nato pa obdelavo znova zaženete, ko ste pripravljeni omogočiti nadaljevanje dejavnosti.

Konfigurirate lahko tudi opozorila, ki omogočajo sprejemanje elektronske pošte o določenih dejavnostih pri vsakem pojavu nadzorovanega dogodka. Tako lahko na primer prejmete opozorilo, če odzivni čas iskanja preseže podani prag.

Če je bil dokument zbrisan iz sistema iskanja v podjetju, lahko sledite dokumentu in ugotovite, kdaj, kje in zakaj je bil zbrisan. Zgodi se lahko na primer, da razčlenjevalnik ne more razčleniti dokumenta ali da skrbnik dokument odstrani iz indeksa.

S tem povezani pojmi

Nadziranje dejavnosti iskanja v podjetju

Zagon in zaustavitev sistema za iskanje v podjetju

Datoteke dnevnika iskanja v podjetju

Datoteke dnevnika so izdelane za posamezne zbirke in za seje na ravni sistema.

Pri konfiguriranju možnosti beleženja za zbirko iskanja v podjetju ali za sistem podate tipe sporočil, ki jih želite beležiti, kot so na primer sporočila o napakah in opozorilna sporočila. Podate lahko tudi, kako pogosto naj sistem zavrti starejše datoteke dnevnika, da naredi prostor za novejša sporočila. Izberete lahko možnosti za sprejemanje elektronske pošte o specifičnih sporočilih (vključno z opozorili) ali o vseh sporočilih o napaki.

Ko si ogledujete datoteke dnevnika, izberete datoteko dnevnika, ki si jo želite ogledati. Ime datoteke vključuje informacije o tem, kdaj je bila datoteka izdelana in katera komponenta je izdala sporočila. Podate lahko tudi filtre prikaza. Izberete lahko na primer samo prikaz sporočil o napaki ali samo sporočil iz določene seje iskanja v podjetju.

S tem povezani pojmi

Datoteke dnevnikov in opozorila

Opozorila



Sporočila za iskanje v podjetju

S tem povezana opravila

Konfiguriranje datotek dnevnika

Konfiguriranje informacij strežnika SMTP

Sprejemanje elektronske pošte o zabeleženih sporočilih

Prikaz datotek dnevnika

Prilagajanje iskanja v podjetju

Aplikacijski programerski vmesniki iskanja v podjetju omogočajo izdelavo iskalnih aplikacij po meri, aplikacij po meri za posodobitev vsebine zbirk, programov po meri za analizo besedila in slovarjev po meri za sopomenke, zaustavitvene besede in pospeševalne besede.

Po namestitvi izdelka OmniFind Enterprise Edition so za razširitev zbirk iskanja v podjetju na voljo naslednje družine API-jev:

API iskanja in indeksiranja (SI-API)

Ta API uporabite za izdelavo iskalnih aplikacij po meri in upravnega vmesnika po meri.

Vtičniki pajka

API-je vtičnikov uporabite za dodajanje metapodatkov v dokumente pri pajkanju v dokumentih ali za povezavo žetonov zaščite, ki uveljavljajo poslovna in zaščitna pravila vašega podjetja.

Branje informacij lahko izboljšate z integracijo programov po meri za lingvistično analizo z zbirkami iskanja v podjetju. Ko v sistem dodate motorje za analizo besedila po meri, lahko povežete motorje z zbirkami. Ko uporabniki izvajajo poizvedbo v zbirki, lahko izkoriščajo prednosti, ki jih nudijo povezave besed, ki jih vaši programi po meri vdelaajo v indeks. Tako lahko na primer uporabniki iščejo koncepte in povezave med izrazi in ne samo izrazov samih.

Branje informacij lahko izboljšate tudi z integracijo slovarjev po meri, ki na primer odražajo akronime, okrajšave in izraze besednjaka, ki so specifični za vašo industrijo. Ko slovarje dodate v sistem, jih lahko povežete z zbirkami. Ko uporabniki izvajajo poizvedbo v zbirki, lahko izkoriščajo naslednje prednosti:

- Če vključuje poizvedba besede, ki so definirane kot sopomenke, bodo dokumenti, ki vključujejo sopomenke izrazov poizvedbe, vključeni v rezultate iskanja.
- Če vključuje poizvedba zaustavitvene besede, bodo le-te odstranjene iz poizvedbe, tako da v rezultatih iskanja ne bodo vrnjeni neustrezni dokumenti.
- Če vključuje poizvedba pospeševalne besede, bodo dokumenti, ki vsebujejo pospeševalne besede, v rezultatih iskanja uvrščeni višje ali nižje glede na pospeševalno vrednost, ki je povezana z besedo v slovarju.

S tem povezani pojmi

Iskalne aplikacije za iskanje v podjetju

Slovarji sopomenk po meri

Slovarji zaustavitvenih besed po meri

Slovarji pospeševalnih besed po meri



Pregled API-ja za iskanje in indeksiranje



Vtičniki pajka

Vzorčna iskalna aplikacija za iskanje v podjetju

Vzorčno iskalno aplikacijo za iskanje v podjetju je mogoče uporabiti kot predlogo za razvoj iskalnih aplikacij po meri.

Vzorčna iskalna aplikacija se namesti pri namestitvi izdelka OmniFind Enterprise Edition, in prikazuje večino funkcij za iskanje in pridobivanje, ki so na voljo za iskanje v podjetju. Aplikacija je tudi delovni primer, ki omogoča, da preiščete vse aktivne zbirke in zunanje izvore v sistemu iskanja v podjetju. Vzorčno aplikacijo lahko uporabite za preskus novih zbirk in zunanjih izvorov, preden le-te omogočite za uporabnike.

Vzorčna iskalna aplikacija prikazuje podporo za zvezno iskanje, tako da omogoča preiskovanje ene ali več zbirk ali zunanjih izvorov.

Za določene tipe pajkov lahko s pomočjo komponente za upravljanje identitet iskanja v podjetju preverite trenutna priporočila, ko uporabniki dostopijo do iskalne aplikacije. Če je domena, ki jo želite preiskati, zaščiten z enotno prijavo (SSO), lahko uporabite za preverjanje uporabnika v celotni iskalni seji mehanizme SSO. V nasprotnem primeru lahko komponenta za upravljanje identitet šifrira in shrani uporabniška priporočila v profilu in priporočila uporabi za izključitev prepovedanih dokumentov iz rezultatov iskanja.


Če želite prilagoditi vzorčno iskalno aplikacijo, lahko uporabite prilagojevalnik iskalnih aplikacij - grafični uporabniški vmesnik - ki omogoča takojšen ogled učinka sprememb. Iskalno aplikacijo lahko prilagodite tudi tako, da uredite konfiguracijsko datoteko za aplikacijo.

Za izdelavo iskalne aplikacije po meri uporabite API za iskanje in indeksiranje iskanja v podjetju.

S tem povezani pojmi

Iskalne aplikacije za iskanje v podjetju

Funkcije vzorčne iskalne aplikacije

 Pregled API-ja za iskanje in indeksiranje

S tem povezana opravila

Dostopanje do iskalnih aplikacij

Urejanje lastnosti vzorčne iskalne aplikacije

Prilagajanje iskalnih aplikacij

Tok podatkov iskanja v podjetju

Komponente iskanja v podjetju, ki jih namestite z izdelkom OmniFind Enterprise Edition, za zagotovitev pretoka podatkov v sistemu tesno medsebojno sodelujejo.

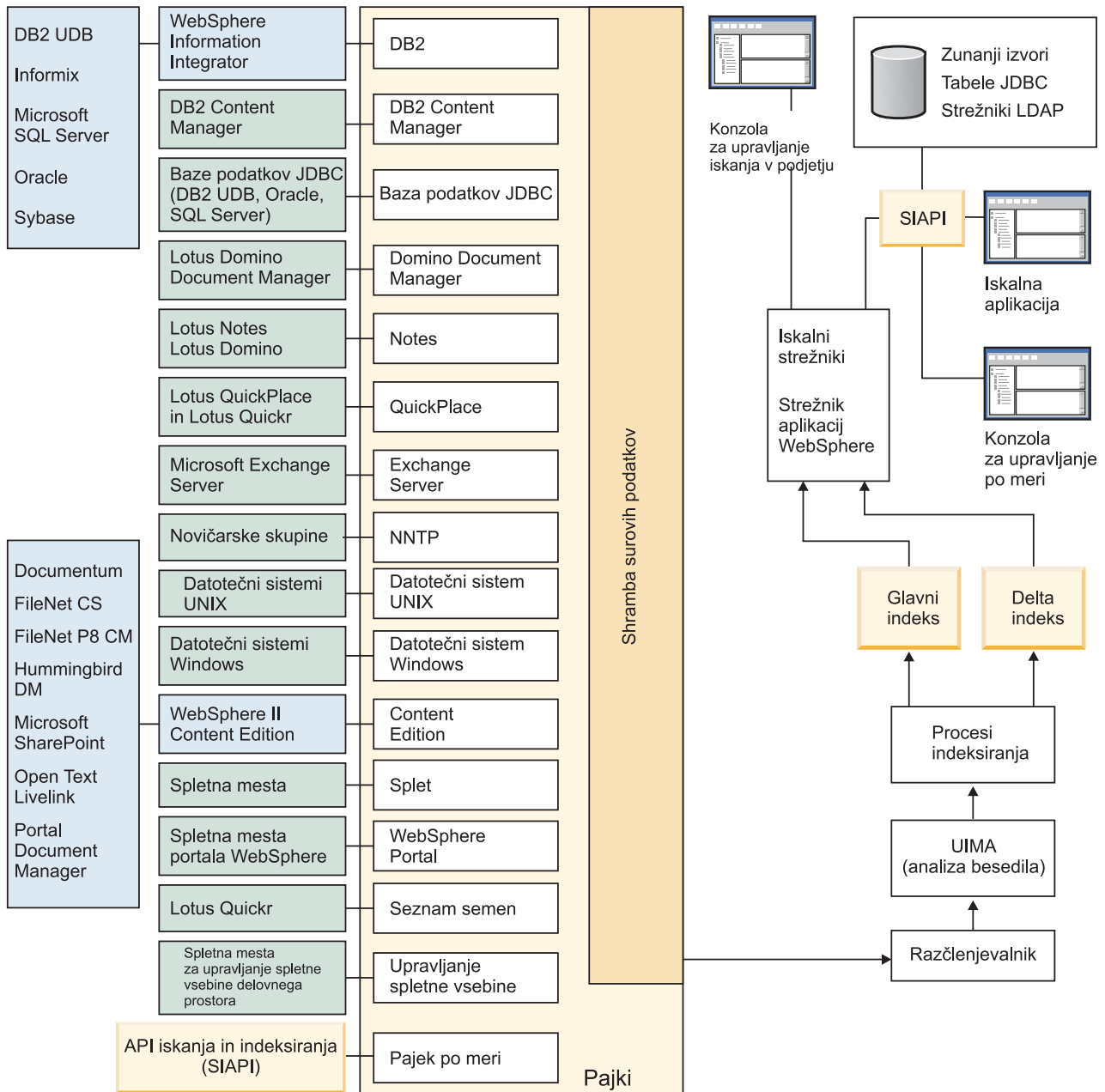
Pajki zbirajo dokumente iz izvorov podatkov v celotnem podjetju. Razčlenjevalnik povzame koristne informacije iz pajkanih dokumentov in ustvari žetone, ki lahko na primer povezujejo dokumente s kategorijami in pomagajo določiti ustreznost dokumentov glede na določbe zahteve za iskanje. Indeks hrani podatke za učinkovito pridobivanje.

Z uporabo spletnega pregledovalnika in iskalne aplikacije uporabniki preiskujejo indeksirane zbirke in zunanje izvore. Iskalna aplikacija lahko prikaže seznam rezultatov, ki ga lahko uporabniki kliknejo v pregledovalniku, aplikacija pa je lahko tudi bolj sofisticirana in vrne dinamično ustvarjeno vsebino, ki temelji na informacijah iz različnih izvorov.

Tako lahko na primer iskalna aplikacija kataloga prilagodi prikaz izdelkov, ki ustrezajo zahtevi za iskanje. Ena sama poizvedba lahko išče v dokumentih iz različnih tipov izvorov podatkov, kot je na primer kombinacija dokumentov iz repozitorijev IBM DB2 Content Manager in Lotus Notes.

Skrbniki določijo, kateri podatki bodo zbrani in kako bodo pajkani, razčlenjeni, indeksirani in preiskani. Z nadzorovanjem sistemske dejavnosti skrbniki tudi izvajajo prilagoditve, da optimizirajo prepustnost podatkov.

Naslednji diagram prikazuje tok informacij v sistemu iskanja v podjetju.



Slika 1. Tok podatkov v sistemu iskanja v podjetju

S tem povezani pojmi

“Kaj je iskanje v podjetju?” na strani 1

“Pregled komponente za iskanje v podjetju” na strani 3

S tem povezane povezave

“Tipi izvorov podatkov, ki jih podpira iskanje v podjetju” na strani 2

Skrbnništvo nad sistemom iskanja v podjetju

Skrbniško konzolo iskanja v podjetju lahko uporabite za izdelavo in izvajanje skrbništva nad zbirkami in zunanjimi viri, zagon in zaustavitev komponent, nadziranje sistemske dejavnosti in datotek dnevnika, konfiguriranje skrbniških uporabnikov, povezovanje iskalnih aplikacij z zbirkami in zunanjimi viri in podajanje informacij za uveljavitev zaščite.

Nasvet: Na naslovu <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html> je na voljo spletna vadnica. Vadnica opisuje posamezne korake, potrebne za konfiguriranje in namestitev izdelka, prikazuje načine iskanja po posameznih tipih virov podatkov ter opisuje, kako s pomočjo aplikacijskega programerskega vmesnika izdelka razširite iskanje v podjetju. Čeprav se vadnica nanaša na starejšo različico izdelka OmniFind Enterprise Edition, je večina zasnov in postopkov še vedno aktualna.


Podrobnejša navodila o načinih konfiguriranja pajkov in omogočitvi zaščite v majhnih, srednjih in velikih organizacijah poiščite v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Pogled Zbirke

Za izdelavo prve zbirke in izvajanje skrbništva nad sistemom se ravnajte po naslednjih smernicah.

Prijava

Prijavite se na skrbniško konzolo iskanja v podjetju. Prikaz Zbirke je vstopna točka za izdelovanje zbirk in izvajanje skrbništva nad njimi.

Nasvet: Če potrebujete pomoč za skrbniško konzolo, lahko kadarkoli kliknete **Pomoč** na orodjarni ali **Pomoč za to stran**. Če so za sporočila na voljo podrobne razlage in ukrepi, lahko za prikaz podrobnosti kliknete ikono  **Dodatne informacije** na koncu sporočila.


Izdelava zbirke

Za izdelavo nove, prazne zbirke izberite enega od naslednjih pristopov:

- Če želite izdelati zbirko s pomočjo čarovnika za zbirke, kliknite **Čarovnik za zbirke** in sledite pozivom čarovnika.
- Če želite izdelati zbirko s pomočjo prikaza Zbirke, kliknite **Izdelaj zbirko**, izpolnite polja na strani Izdelava zbirke, nato pa kliknite **Potrdi**.

Konfiguriranje zbirke

Novo, prazno zbirko morate urediti, da ji lahko dodate vsebino in podate možnosti, ki določajo, kako želite pajkati v podatkih in omogočiti preiskovanje podatkov.

Kliknite  **Uredi** za vašo novo zbirko, nato pa izberite stran za podajanje možnosti za zbirko.

Opozorilo: Če se želite pomakniti na predhodno prikazano stran ali osvežiti informacije na skrbniški konzoli, kliknite **Prejšnji** in **Osveži** na skrbniški konzoli iskanja v podjetju. Če kliknete **Nazaj** ali **Osveži** v spletnem brskalniku, to lahko povzroči neskladnost rezultatov in možno izgubo podatkov.

- Na strani Splošno lahko podate možnosti, ki veljajo za celotno zbirko:
 - Splošne možnosti lahko uredite, če želite spremeniti ime ali opis zbirke ali spremeniti pričakovano velikost zbirke.

- Prikažete lahko informacije o zbirki, ki jih ne morete spreminjati, kot sta na primer ID zbirke in način statičnega rangiranja za dokumente rangiranja v rezultatih iskanja.
- Če ste pri izdelavi zbirke zanjo omogočili zaščito, lahko omogočite ali onemogočite krmilne elemente za zaščito na ravni dokumentov.
- Na strani Pajkanje konfigurirajte vsaj enega pajka.

Ena zbirka lahko vsebuje podatke iz različnih virov podatkov. Za vsak vir podatkov, ki ga želite vključiti, morate konfigurirati vsaj enega pajka. Pri izdelavi pajka vam pri konfiguriranju pomaga čarovnik, specifičen za tip podatkov, v katerih pajkate.
- Na strani Razčlenjevanje lahko konfigurirate možnosti, ki določajo, kako bodo razčlenjeni pajkani podatki, da omogočite učinkovito iskanje v njih:
 - Podate lahko, ali bodo dokumenti XML razčlenjeni tako, da bo iskanje v njih mogoče z izvirnim iskanjem XML.
 - Dokumente lahko povežete s kategorijami, ki uporabnikom omogočajo preiskovanje podniza zbirke ali pregledovanje dokumentov rezultatov iskanja s kategorijami, v katere spadajo.
 - Elemente XML in elemente metapodatkov HTML lahko preslikate v iskalna polja v indeksu, kar uporabnikom omogoča, da v poizvedbah podajo imena polj in preiščejo specifične dele dokumentov.
 - Če ste v sistem iskanja v podjetju dodali motorje za analizo besedila po meri, lahko izberete enega med njimi za uporabo z zbirko, nato pa podate možnosti za obdelavo besedila, ki izboljšajo zmožnost pridobivanja informacij in podpirajo semantično iskanje.
 - Polja lahko povežete s poviševalnimi razredi, ki vplivajo na to, kako so polja, ki se ujemajo z izrazi poizvedbe, uvrščena v rezultatih iskanja.
- Na strani Indeksiranje konfigurirajte razporede za gradnjo indeksa. Načrtujte pogoste gradnje indeksov, da bodo imeli vaši uporabniki vedno dostop do najnovejših informacij. Izvedete lahko tudi naslednje neobvezne dejavnosti:
 - Uporabnikom omogočite, da v izrazih poizvedbe podajo univerzalne znake.
 - Konfigurirate območja, da bodo uporabniki namesto vseh dokumentov v indeksu preiskali omejen del zbirke.
 - Skrčite rezultate iskanja, tako da bodo dokumenti iz istega vira v rezultatih iskanja skrčeni.
 - Dodelite poviševalne faktorje, ki vplivajo na to, kako bodo dokumenti, ki se ujemajo z vzorcem URI, uvrščeni v rezultatih iskanja.
 - Odstranite URI-je iz indeksa. Morda boste morali uporabnikom po izdelavi zbirke preprečiti prikaz določenih dokumentov.
- Na strani Iskanje lahko podate možnosti za preiskovanje dokumentov v zbirki:
 - Prihranite lahko prostor predpomnilnika za rezultate iskanja in spremenite privzeti jezik zbirke.
 - Če ste v sistem iskanja v podjetju dodali slovarje po meri za sopomenke, zaustavitvene besede in poviševalne besede, lahko uporabite slovarje, ki bodo uporabljeni, ko bodo uporabniki preiskovali zbirko.
 - Podate lahko dolžino za prikaz povzetkov dokumentov v rezultatih iskanja.
 - Če želite, da se določeni URI-ji samodejno prikažejo v rezultatih iskanja, če vključuje poizvedba določene ključne besede ali fraze, lahko konfigurirate hitre povezave.
- Na strani Dnevnik lahko opravite naslednje dejavnosti:

- Podate možnosti za tipe sporočil, ki jih želite beležiti in podate, kako pogosto želite reciklirati datoteke dnevnika.
- Podate možnosti za sprejemanje opozoril o dejavnosti zbirke. Opozorilo vas lahko na primer obvesti, da je povprečen odzivni čas iskanja presegel podano omejitev.
- Podate možnosti za sprejem elektronske pošte vsakič, ko je zabeleženo določeno sporočilo ali določen tip sporočila.
- Podate možnosti za beleženje informacij, ki omogočajo, da ugotovite, kdaj, kje in zakaj je bil dokument izbrisan iz sistema iskanja v podjetju.

Zagon komponent

Ko podate vire podatkov za pajkanje in možnosti za zbiranje in preiskovanje podatkov, lahko zaženete procese za gradnjo zbirke. Pri tem je zelo pomemben vrstni red, v katerem zaženete komponente. Pajki morajo izvesti pajkanje dokumentov, preden je podatke mogoče razčleniti in indeksirati, glavni indeks pa mora biti zgrajen, preden lahko iskalni strežniki obdelajo zahteve za iskanje.

Prikaz zunanjih virov

Če želite preiskati vire podatkov, ne da bi jih pajkali ali indeksirali, lahko na orodjarni kliknete **Zunanji viri** in podate možnosti, ki omogočajo preiskovanje virov podatkov. Podati morate informacije, ki omogočajo dostopnost baz podatkov JDBC (Java Database Connectivity) in strežnikov LDAP (Lightweight Directory Access Protocol) za iskanje v podjetju. Ko povežete zunanje vire z iskalnimi aplikacijami, lahko uporabniki preiskujejo te vire sočasno s preiskovanjem zbirk s podatki, ki so bili pajkani, razčlenjeni in indeksirani.

Prikaz sistema

Če ste član s skrbniško vlogo v iskanju v podjetju, na orodjarni kliknite ikono **Sistem**, ki omogoča izvedbo naslednjih dejavnosti. Skrbniki, operaterji in nadzorniki zbirke lahko dostopijo do tega prikaza samo, če jim skrbnik zbirke dodeli dovoljenje.

- V sistem dodajte motorje za analizo besedila po meri.
- V sistem dodajte slovarje po meri za sopomenke, zaustavitvene besede in poviševalne besede.
- Podajte, koliko zbirk lahko vzporedno gradi indekse in ali se lahko gradnje glavnega indeksa za eno zbirko izvajajo sočasno z gradnjami delta indeksa.
- Konfigurirajte opozorila za dogodke na ravni sistema.
- Podajte možnosti za beleženje sporočil, ki jih ustvarijo seje na ravni sistema.
- Podajte informacije o poštnem strežniku, da boste lahko sprejemali elektronsko pošto o dejavnostih iskanja v podjetju.

Prikaz zaščite

Če ste član s skrbniško vlogo v iskanju v podjetju, lahko za podajanje možnosti zaščite kliknete **Zaščita**. Skrbniki, operaterji in nadzorniki zbirke ne morejo dostopiti do tega prikaza.


Če za strežnik WebSphere Application Server omogočite zaščito, lahko v pogledu Zaščita konfigurirate skrbniške vloge. S konfiguriranjem skrbniških vlog lahko več uporabnikom omogočite izvajanje skrbništva sistemom in hkrati omejite dostop vsakega uporabnika do specifičnih funkcij in zbirk.

Prikaz Zaščita lahko uporabite tudi za konfiguriranje možnosti za upravljanje identitet. Tako lahko na primer podate možnosti za shranjevanje poverilnic uporabnikov v profilih, ki jih je

mogoče uporabiti za preverjanje trenutnih poverilnic uporabnika med obdelavo poizvedbe. Če je vir, ki ga želite preiskati, zaščiten z enotno prijavo, lahko podate tudi možnosti za uporabo načinov overjanja z enotno prijavo za preverjanje trenutnih poverilnic uporabnika med obdelavo poizvedbe.

Dokler ne izdelate lastnih iskalnih aplikacij, lahko preiskujete vse zbirke in zunanje vire s pomočjo vzorčne iskalne aplikacije. Ko izdelate iskalno aplikacijo po meri, uporabite prikaz Zaščita za povezavo aplikacije z zbirkami in zunanjimi viri, ki jih lahko preišče.

Prikaz nadzorovanja

Ikono  **Nadzornik** lahko kadarkoli kliknete za nadziranje sistema ali komponent zbirke. Če vam skrbniška vloga dopušča zagon in zaustavitev procesov komponent med njihovim nadziranjem, lahko naredite tudi to.

S tem povezani pojmi

“Skrbnišvo nad pajkom” na strani 35

“Nadziranje dejavnosti iskanja v podjetju” na strani 273

S tem povezana opravila

“Zagon sistema iskanja v podjetju” na strani 267

“Izvajanje skrbništva nad iskalnimi strežniki v samostojnem načinu” na strani 271

“Zaustavitev sistema iskanja v podjetju” na strani 269

“Izdelava zbirke s pomočjo čarovnika za zbirke” na strani 29

“Izdelava zbirke s pogledom Zbirke” na strani 30

Prijava na skrbniško konzolo

Če želite izvajati skrbništvo nad sistemom iskanja v podjetju, morate v spletnem brskalniku podati URL, nato pa se prijaviti na skrbniško konzolo.

Preden začnete

Prijaviti se morate z ID-jem uporabnika, ki ima pooblastila za dostop do skrbniške konzole za iskanje v podjetju:

- Če na aplikacijskem strežniku WebSphere ne omogočite globalne zaščite, lahko do skrbniške konzole pride samo skrbnik iskanja v podjetju, ki ste ga podali ob namestitvi izdelka OmniFind Enterprise Edition.
- Če za izdelek WebSphere Application Server omogočite globalno zaščito, lahko s skrbniško konzolo za iskanje v podjetju konfigurirate skrbniške vloge. Konfigurirani ID-ji uporabnikov morajo obstajati v uporabniškem registru WebSphere Application Server. S konfiguriranjem skrbniških vlog večjemu številu uporabnikov omogočite prijavo na skrbniško konzolo, lahko pa krmilite, do katerih funkcij in zbirk bodo imeli posamezni skrbniški uporabniki dostop.

Postopek

Če se želite prijaviti na skrbniško konzolo za iskanje v podjetju, naredite naslednje:

1. V spletni brskalnik vnesite URL skrbniške konzole. Primer:

`http://SearchServer.com/ESAdmin/`

`SearchServer.com` je gostiteljsko ime strežnika za iskanje v podjetju.

Odvisno od konfiguracije vašega spletnega strežnika boste morda morali podati tudi številko vrat. Primer:

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

2. Na uvodni strani vnesite ID uporabnika in geslo ter kliknite **Prijava**.

Sistem prikaže pogled Zbirke, ki predstavlja vaše izhodišče za izvajanje skrbništva nad sistemom in zbirkami. Če uporabljate skrbniške vloge, so dejanja, ki jih lahko izvedete, in zbirke, ki jih vidite, odvisne od skrbniške vloge, ki jo imate.

Če seja nekaj časa ni dejavna, vas sistem samodejno odjavi. Če želite še naprej izvajati skrbništvo nad sistemom, se znova prijavite.

Ko končate izvajati skrbništvo nad zbirkami, se lahko s klikom na **Odjava** odjavite s konzole. Nato se lahko prijavite z drugima ID-jem in geslom ali pa zaprete spletni brskalnik in zapustite skrbniško konzolo.

S tem povezani pojmi

“Skrbniške vloge” na strani 235

Spreminjanje gesla skrbnika sistema iskanja v podjetju v konfiguraciji z enim strežnikom

Geslo skrbnika sistema iskanja v podjetju je shranjeno v šifrirani obliki. Če ga želite spremeniti, zaženite skript **eschangepw**.

Omejitve

Gesla lahko vključujejo naslednje posebne znake:

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

Če v sistemih AIX, Linux ali Solaris podate geslo, ki vključuje posebne znake, morate celotno geslo ograditi v opuščajeh.

Primer: 'mojegeslo@%\$'

Če v sistemu Windows podate geslo, ki vključuje posebne znake, morate celotno geslo ograditi z navednicami.

Primer: "moj?+!gesl"

O tej nalogi

Geslo začetnega ID-ja skrbnika sistema iskanja v podjetju podate pri nameščanju izdelka OmniFind Enterprise Edition.

Če želite spremeniti geslo, morate zagnati skript **eschangepw** in s tem spremembo razširiti po celotnem sistemu iskanja v podjetju. Namestitveni program izdelka naredi dve spremenljivki okolja, ki jih lahko uporabite s skriptom **eschangepw**:

ES_INSTALL_ROOT

Namestitveni imenik sistema iskanja v podjetju.

ES_NODE_ROOT

Podatkovni imenik sistema iskanja v podjetju. Geslo za ID skrbnika sistema iskanja v podjetju je shranjeno v datoteki **es.cfg** v tem imeniku.

Postopek

Če želite geslo skrbnika sistema iskanja v podjetju spremeniti v konfiguraciji z enim strežnikom, storite naslednje:

1. Prijavite se kot skrbnik sistema iskanja v podjetju in le-tega zaustavite:
esadmin system stopall

Pomembno: Ko se sistem zaustavi, uporabniki ne bodo mogli predložiti nobene zahteve za iskanje.

2. Geslo za ID skrbnika sistema iskanja v podjetju spremenite s pomočjo ukazov operacijskega sistema (v AIX, Linux ali Solaris) ali s pripomočkom za spreminjanje gesel (v Windows).
3. Zaženite naslednji skript, pri čemer je *NovaVrednost* geslo, ki ste ga podali v koraku 2:

AIX, Linux ali Solaris

```
eschangepw.sh NovaVrednost
```

Windows

```
eschangepw NovaVrednost
```

4. Ponovno zaženite sistem iskanja v podjetju:
esadmin system startall

S tem povezane povezave

“Ukazi za iskanje v podjetju, povratne kode in ID-ji sej” na strani 335

Spreminjanje gesla skrbnika sistema iskanja v podjetju v konfiguraciji z več strežniki

Geslo skrbnika sistema iskanja v podjetju je shranjeno v šifrirani obliki. Če ga želite spremeniti, uporabite skript **eschangepw** in ga s tem spremenite v vseh računalnikih v sistemu iskanja v podjetju.

Omejitve

Gesla lahko vključujejo naslednje posebne znake:

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

Če v sistemih AIX, Linux ali Solaris podate geslo, ki vključuje posebne znake, morate celotno geslo ograditi v opuščajje.

Primer: 'mojegeslo@\$%'

Če v sistemu Windows podate geslo, ki vključuje posebne znake, morate celotno geslo ograditi z navednicami.

Primer: "moj?+!gesl"

O tej nalogi

Geslo za ID skrbnika sistema iskanja v podjetju, ki ga podate na začetku, pri nameščanju izdelka OmniFind Enterprise Edition, mora biti enako v vseh strežnikih sistema iskanja v podjetju.

Če želite spremeniti geslo in ga razširiti po celotnem sistemu iskanja v podjetju, skript **eschangepw** zaženite v vseh računalnikih, ki uporabljajo iskanje v podjetju. Spodnji postopek priporoča vrstni red za spreminjanje gesla v vseh strežnikih. Temu vrstnemu redi ni treba slediti, vendar pa morate zahtevane korake dokončati v vseh tipih strežnikov.

Namestitveni program izdelava dve spremenljivki okolja, ki jih lahko uporabite s skriptom **eschangepw**:

ES_INSTALL_ROOT

Namestitveni imenik sistema iskanja v podjetju.

ES_NODE_ROOT

Podatkovni imenik sistema iskanja v podjetju. Geslo za ID skrbnika sistema iskanja v podjetju je shranjeno v datoteki `es.cfg` v tem imeniku.

Postopek

Če želite geslo skrbnika sistema iskanja v podjetju spremeniti v konfiguraciji z več strežniki, storite naslednje:

1. V indeksni strežnik sistema iskanja v podjetju se prijavite kot skrbnik sistema iskanja v podjetju in le-tega zaustavite:

```
esadmin system stopall
```

Pomembno: Ko se sistem zaustavi, uporabniki ne bodo mogli predložiti nobene zahteve za iskanje.

- a. Geslo za ID skrbnika sistema iskanja v podjetju spremenite s pomočjo ukazov operacijskega sistema (v AIX, Linux ali Solaris) ali s pripomočkom za spreminjanje gesel (v Windows).
- b. Zaženite naslednji skript, pri čemer je *NovaVrednost* geslo, ki ste ga podali v koraku 1a:

AIX, Linux ali Solaris

```
eschangepw.sh NovaVrednost
```

Windows

```
eschangepw NovaVrednost
```

2. V ostalih računalnikih sistema iskanja v podjetju opravite naslednje korake:

- a. Prijavite se kot skrbnik iskanja v podjetju.
- b. Zaustavite splošno komunikacijsko plast (CCL) sistema iskanja v podjetju:

AIX, Linux ali Solaris

```
stopccl.sh
```

Ukazni poziv Windows

```
stopccl
```

Skrbniška orodja storitev operacijskega sistema Windows

- 1) Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - 2) Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaustavi**.
- c. Geslo za ID skrbnika sistema iskanja v podjetju spremenite s pomočjo ukazov operacijskega sistema (v AIX, Linux ali Solaris) ali s pripomočkom za spreminjanje gesel (v Windows). To geslo se mora ujemati z geslom, ki ste ga podali v koraku 1a.
 - d. Zaženite naslednji skript, pri čemer je *NovaVrednost* geslo, ki ste ga podali v koraku 1a:

AIX, Linux ali Solaris

```
eschangepw.sh NovaVrednost
```

Windows

```
eschangepw NovaVrednost
```

- e. Znova zaženite CCL:

AIX, Linux ali Solaris
startccl.sh -bg

Ukazni poziv Windows
startccl

Skrbniška orodja storitev operacijskega sistema Windows

Če želite CCL zagnati v ozadju, storite naslednje:

- 1) Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - 2) Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Lastnosti**.
 - 3) Kliknite zavihek **Prijava**.
 - 4) Geslo spremenite tako, da podate novo vrednost gesla in nato kliknete **Potrdi**.
 - 5) Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
3. V indeksni strežnik sistema iskanja v podjetju se prijavite kot skrbnik sistema iskanja v podjetju in le-tega znova zaženite:

```
esadmin system startall
```

S tem povezane povezave



Setting the encrypted administrator password to be the same on all servers

“Ukazi za iskanje v podjetju, povratne kode in ID-ji sej” na strani 335

Številke vrat TCP, ki se uporabljajo za iskanje v podjetju

Preglejte privzete številke vrat, ki se uporabljajo v sistemu iskanja v podjetju, da se boste izognili navkrižjem pri vratih, ko konfigurirate vire ali dodeljujete številke vrat drugim aplikacijam.

Če konfigurirate požarni zid, morate izrecno omogočiti dostop do posameznih številke vrat. Zagotoviti morate tudi, da so vsi strežniki iskanja v podjetju znotraj požarnega zidu.

Tabela 1. Konfiguriranje številke vrat za iskanje v podjetju

Ime vrat	Številke vrat	Mesto konfiguracije
Skupna komunikacijska plast (CCL)	6002	ES_NODE_ROOT/nodeinfo/es.cfg in ES_NODE_ROOT/master_config/nodes.ini na vseh strežnikih za iskanje v podjetju
HTTP na iskalnih strežnikih	80	HTTP_SERVER_ROOT/conf/http.conf na iskalnih strežnikih
Skrbniška konzola WebSphere Application Server različice 6	9060	Na iskalnih strežnikih
Skrbniška konzola WebSphere Application Server različice 5.1	9090	Na iskalnih strežnikih
pajek DB2	6000, 6001, 6002, 60003, 50000	Na strežniku pajkanja
Informacijski center	8888	Na iskalnih strežnikih
Anonimna ali dinamična vrata za CCL, prenose datotek (ESFTP) in kopiranje indeksa	49152 do 65535	Na vseh iskalnih strežnikih podjetja

Tabela 1. Konfiguriranje številke vrat za iskanje v podjetju (nadaljevanje)

Ime vrat	Številke vrat	Mesto konfiguracije
Omrežni strežnik Apache Derby	1527	Na strežniku pajkanja
Komunikacije po meri	8890	Na strežniku pajkanja
Povezave oddaljenega odjemalca s strežnikom DB2 (uporabljen samo z izdajami, starejšimi od OmniFind Enterprise Edition različice 8.4)	50000	Na strežniku pajkanja
Baza podatkov WebSphere Information Integrator Content Edition	6001 (6002 kot nadomestna)	V skrbniški konzoli WebSphere Information Integrator Content Edition
Spojnik proxy za klic oddaljene metode (RMI - remote method invocation) za WebSphere Information Integrator Content Edition	1251 (vrata RMI)	V skrbniški konzoli WebSphere Information Integrator Content Edition

Spreminjanje številke vrat za sistem iskanja v podjetju

Če je številka vrat, ki jo sistem iskanja v podjetju uporablja za komuniciranje, v navzkrižju s številko vrat, ki jo uporablja drug izdelek, morate spremeniti številko vrat sistema iskanja v podjetju.

O tej nalogi

Pri nameščanju izdelka OmniFind Enterprise Edition morate podati številko vrat za sistem iskanja v podjetju. V konfiguraciji z več strežniki morate za vse strežnike podati isto številko vrat.

Če številka vrat ni na voljo (številka vrat je lahko na primer dodeljena drugemu izdelku v istem strežniku), pride do navzkrižja, sistem pa v datoteko `CCLServer_datum.log` zabeleži sporočilo o napaki, pri čemer `datum` podaja datum, ko je sistem izdelal datoteko dnevnika:

```
FFQ00273W Prišlo je do notranjega opozorila - sporočilo o izjemnem stanju: {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)
```

Postopek

Če želite spremeniti številko vrat, ki jo uporablja sistem iskanja v podjetju, storite naslednje:

1. V računalniku, kjer želite spremeniti številko vrat, se prijavite kot skrbnik sistema iskanja v podjetju in le-tega zaustavite:

```
esadmin system stopall
```

Pomembno: Ko se sistem zaustavi, uporabniki ne bodo mogli predložiti nobene zahteve za iskanje.

2. Uredite datoteko `ES_NODE_ROOT/nodeinfo/es.cfg`, poiščite naslednjo lastnost, podajte novo vrednost za številko vrat, datoteko shranite in jo zaprite:

```
CCLPort=nova_številka_vrat
```

3. Ponovno zaženite plast za skupno komunikacijo (CCL) za iskanje v podjetju:

AIX, Linux ali Solaris

startccl.sh

Ukazni poziv Windows

startccl

Skrbniška orodja storitev operacijskega sistema Windows

Če želite zagnati CCL v ozadju:

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
4. V indeksni strežnik se prijavite kot skrbnik sistema iskanja v podjetju in zaustavite CCL:

AIX, Linux ali Solaris

stopccl.sh

Ukazni poziv Windows

stopccl

Skrbniška orodja storitev operacijskega sistema Windows

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaustavi**.
5. Uredite datoteko ES_NODE_ROOT/nodeinfo/es.cfg :
 - a. Poiščite naslednjo lastnost, pri čemer je *ime_racunalnika* ime računalnika, v katerem ste v koraku 2 na strani 21 spremenili številko vrat. *N* v lastnosti *nodeN* je številka, ki označuje strežnik sistema iskanja v podjetju.
`nodeN.destination=ime_racunalnika`
 - b. Poiščite naslednjo podlastnost, podajte isto številko, kot ste jo podali za strežnik v koraku 2 na strani 21, datoteko shranite in jo zaprite:
`nodeN.port=nova_stevilka_vrat`
 6. Ponovno zaženite sistem iskanja v podjetju:

esadmin system startall

S tem povezane povezave

“Ukazi za iskanje v podjetju, povratne kode in ID-ji sej” na strani 335

“Številke vrat TCP, ki se uporabljajo za iskanje v podjetju” na strani 20

Spreminjanje gostiteljskih imen ali naslovov IP strežnika za iskanje v podjetju

Gostiteljska imena in naslove IP, ki jih v skladu s konfiguracijo uporabljajo strežniki za iskanje v podjetju, lahko spremenite.

Morda boste želeli na primer spremeniti naslov IP, če imate več omrežnih vmesniških kartic (NIC) na vsakem od strežnikov za iskanje v podjetju ter ugotovite, da je indeksni strežnik konfiguriran za uporabo počasnega omrežja. Nastavitve v konfiguracijskih datotekah lahko spremenite in indeksnemu strežniku dovolite uporabo hitrejšega omrežja.

Nasvet: Če ne želite urejati konfiguracijskih datotek, lahko znova namestite OmniFind Enterprise Edition in podate nova gostiteljska imena ali naslove IP ob zagonu namestitvenega programa.

Postopek:

Če želite spremeniti gostiteljska imena ali naslove IP, naredite naslednje:

1. Prijavite se kot skrbnik iskanja v podjetju. Če uporabljate konfiguracijo z več strežniki, se prijavite v katerikoli strežnik za iskanje v podjetju.
2. Zaustavite seje iskanja v podjetju:
esadmin system stopall
3. Zaustavite skupno komunikacijsko plast (CCL). Če uporabljate konfiguracijo z več strežniki, z eno od naslednjih metod zaustavite CCL na vsakem od strežnikov za iskanje v podjetju:

AIX, Linux ali Solaris

```
stopccl.sh
```

Ukazni poziv Windows

```
stopccl
```

Skrbniška orodja storitev operacijskega sistema Windows

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaustavi**.
4. Na vsakem od strežnikov za iskanje v podjetju opravite naslednje korake:
 - a. Uredite datoteko ES_INSTALL_ROOT/configurations/ccl.properties in s parametrom **es_server_hostName** podajte gostiteljsko ime ali naslov IP, ki ga želite uporabiti za ta strežnik.
 - b. Uredite datoteko ES_NODE_ROOT/nodeinfo/es.cfg in s parametrom **LocalHostName** podajte gostiteljsko ime ali naslov IP, ki ga želite uporabiti za ta strežnik.
 - c. Uredite datoteko ES_NODE_ROOT/master_config/nodes.ini in zamenjajte vse pojavitve parametra **nodeN.destination** z gostiteljskim imenom ali naslovom IP, ki ga želite uporabiti za ta strežnik.
 - d. Uredite datoteko ES_NODE_ROOT/config/nodes.ini in zamenjajte vse pojavitve parametra **nodeN.destination** z gostiteljskim imenom ali naslovom IP, ki ga želite uporabiti za ta strežnik.
 5. Znova zaženite CCL. Če uporabljate konfiguracijo z več strežniki, z eno od naslednjih metod znova zaženite CCL na vsakem od strežnikov za iskanje v podjetju:

AIX, Linux ali Solaris

```
startccl.sh -bg
```

Ukazni poziv Windows

```
startccl
```

Skrbniška orodja storitev operacijskega sistema Windows

Če želite CCL zagnati v ozadju, storite naslednje:

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
6. Znova zaženite seje iskanja v podjetju:
esadmin system startall

Konfiguriranje podpore za dvojne naslove IP

Če so strežniki, na katere ste namestili iskanje v podjetju, konfigurirani za podporo dvojnih naslovov IP, morate ročno konfigurirati strežnike iskanja v podjetju za delovanje v tem okolju.

Skrbnik namizja lahko na primer namesti Microsoft Loopback Adapter in izdela navidezno omrežje, ki podpira omrežne zahteve, ki jih postavljajo nekateri izdelki, na primer Microsoft SQL Server.

Postopek:

Če želite konfigurirati sistem iskanja v podjetju tako, da ga bo mogoče izvajati na strežnikih, ki podpirajo dvojne naslove IP, naredite naslednje:

1. Prijavite se kot skrbnik iskanja v podjetju. Če uporabljate konfiguracijo z več strežniki, se prijavite v poljuben strežnik za iskanje v podjetju.
2. Zaustavite seje iskanja v podjetju:
esadmin system stopall
3. Zaustavite skupno komunikacijsko plast (CCL). Če uporabljate konfiguracijo z več strežniki, z eno od naslednjih metod zaustavite CCL na vsakem od strežnikov za iskanje v podjetju:

AIX, Linux ali Solaris

```
stopccl.sh
```

Ukazni poziv Windows

```
stopccl
```

Skrbniška orodja storitev operacijskega sistema Windows

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaustavi**.
4. Na vsakem strežniku za iskanje v podjetju, ki podpira naslove IP, uredite datoteko `ES_NODE_ROOT/nodeinfo/es.cfg` in dodajte parameter z imenom **LocalIPAddress**. Za vrednost dodelite naslov IP, ki ga lahko razreši DNS.
 5. Znova zaženite CCL. Če uporabljate konfiguracijo z več strežniki, z eno od naslednjih metod znova zaženite CCL na vsakem od strežnikov za iskanje v podjetju:

AIX, Linux ali Solaris

```
startccl.sh -bg
```

Ukazni poziv Windows

```
startccl
```

Skrbniška orodja storitev operacijskega sistema Windows

Če želite CCL zagnati v ozadju, storite naslednje:

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
6. Znova zaženite seje iskanja v podjetju:
esadmin system startall

Omogočitev podpore za protokol IPv6

Če želite omogočiti podporo za naslove in URL-je, ki upoštevajo protokol IP različice 6 (IPv6), morate definirati spremenljivko okolja, ki sistem iskanja v podjetju napoti, naj uporabi samo naslove vtičnic IPv6.

Praden začnete

1. Preverite, ali so vrednosti za lastnost **LocalHostName** in lastnost **LocalIPAddress** (če je podana) v %ES_CFG% bodisi ime gostitelja ali veljaven naslov IPv6 lokalnega računalnika. Če je treba, lahko te vrednosti ročno spremenite in shranite %ES_CFG%.
2. Preverite, ali je vrednost lastnosti **DerbyServerHostName** v %ES_CFG% ime gostitelja. Ta vrednost ne sme biti naslov IPv4 ali naslov IPv6. Če spremenite to lastnost, shranite datoteko %ES_CFG%.
3. Preverite, ali so vrednosti za lastnosti **destination** in **serverhost** v datoteki ES_NODE_ROOT/master_config/nodes.ini na indeksnem strežniku bodisi veljaven naslov IPv6 ali veljavno ime gostitelja. Vrednosti ne smejo biti naslov IPv4. Če spremenite to datoteko, shranite spremembe.

Omejitve

Podpora za protokol IPv6 je na voljo samo v sistemih za iskanje v podjetju, ki jih namestite na strežnike Windows 2003 Servers. Potem ko omogočite podporo za naslove IPv6, iskanje v podjetju ne bo več uporabljalo naslovov IPv4 za komunikacije z vtičnicami.

Podprti brskalniki

Brskalniki Internet Explorer in Mozilla Firefox obravnavajo naslove IPv6 na različnih način.

Mozilla Firefox

Če želite zagnati konzolo za skrbništvo nad iskanjem v podjetju ali iskalno aplikacijo, lahko v URL vnesete naslov IPv6 ali ime gostitelja. Primer:

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

Internet Explorer

Če želite zagnati konzolo za skrbništvo nad iskanjem v podjetju ali iskalno aplikacijo, ne vnesite naslova IPv6 v URL. Uporabiti morate naslednjo obliko in se prepričati, da se ime gostitelja preslika v naslov IPv6 v DNS ali datoteko c:\windows\system32\etc\hosts:

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

Če vnesete naslov IPv6 v URL, se prikaže sporočilo **Napaka neveljavne skladnje**. Dodatne informacije o tej omejitvi poiščite v temi <http://support.microsoft.com/kb/325414>.

Podpora za podatkovne izvore IPv4

Podatkovne izvore, ki se izvajajo na strežniku IPv4, podpirajo pajki za iskanje v podjetju. Če konfigurirate spletne pajke in podate začetne URL-je v obliki naslova IPv6, ne pozabite URL-jev zapisati v oklepajih. Primer:

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

Postopek

Če želite omogočiti podporo za protokol IPv6, naredite naslednje:

1. V indeksni strežnik sistema za iskanje v podjetju se prijavite kot skrbnik sistema za iskanje v podjetju in le-tega zaustavite:
esadmin system stopall

Pomembno: Ko se sistem zaustavi, uporabniki ne bodo mogli predložiti nobene zahteve za iskanje.

2. Zaustavite strežnik s splošno komunikacijsko plastjo (common communication layer - CCL) na vseh strežnikih iskanja v podjetju:
 - a. Zaženite storitve Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaustavi**.
3. Definirajte **ES_IPV6=1** v spremenljivkah systemskega okolja. Pri namestitvi iskanja v podjetju na več strežnikov naredite ta korak na iskalnih strežnikih.
4. Posodobite lastnosti po meri navideznega računalnika Java Virtual machine (JVM) na aplikacijskem strežniku WebSphere Application Server. Pri namestitvi iskanja v podjetju na več strežnikov naredite te korake na iskalnih strežnikih.
 - a. Če se aplikacijski strežnik ESSearchServer izvaja na aplikacijskem strežniku WebSphere Application Server, ga zaustavite.
 - b. Odprite skrbniško konzolo WebSphere Application Server in se pomaknite do okna z lastnostmi po meri za navidezni računalnik Java Virtual Machine. Izberite **Strežniki** → **Aplikacijski strežniki** → **ESSearchServer** → **Java in upravljanje procesov** → **Definicija procesa** → **Navidezni računalnik Java** → **Lastnosti po meri**.
 - c. Konfigurirajte naslednje lastnosti:

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

Če lastnost po meri še ni navedena, izdelajte novo lastnost in vnesite njeno ime v polje **Ime** in veljavno vrednost v polje **Vrednost**.
 - d. Če izvajate WebSphere Application Server različice 6.1, dodajte lastnost **LocalIPAddress** datoteki ES_NODE_ROOT/nodeinfo/es.cfg, da bo vsebovala naslov IPv6 za ta iskalni strežnik.
 - e. Znova zaženite aplikacijo ESSearchServer.
5. Če izvajate WebSphere Application Server različice 6.1, opravite naslednje korake. Pri namestitvi iskanja v podjetju na več strežnikov naredite te korake na iskalnih strežnikih.
 - a. Uredite datoteko httpd.conf za IBM HTTP Server.
 - b. Odstranite znak # z začetka naslednje vrstice, da odstranite oznako za komentar izpred navodil, s čimer omogočite strežniku IBM HTTP Server spremljanje povezav IPv6 na vratih 80:

```
# Listen [::]:80
```
 - c. Znova zaženite strežnik IBM HTTP Server.
6. Znova zaženite CCL na vseh strežnikih iskanja v podjetju:
 - a. Zaženite storitve Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
 - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
7. Ponovno zaženite sistem iskanja v podjetju:
esadmin system startall
8. Preverite datoteko ES_NODE_ROOT/node/logs/ccl0.log in preverite, ali se je strežnik CCL zagnal brez napak. Če je IPv6 pravilno omogočen, se zabeležijo sporočila, podobna naslednjemu:

```
INFO: CCL server ready for business. Now, waiting for external requests.  
CCL host name is fe80::250:56ff:feb4:27d1  
CCL host dot.ip address is fe80:0:0:0:250:56ff:feb4:27d1  
CCL server port is 6002  
Total JVM Runtime memory is 33022Kb  
Current free memory is 21346Kb  
java.net.preferIPv6Addresses=true  
java.net.preferIPv4Stack=false  
INFO: Session "TraceDaemonSession" was attached PID=4992
```

9. Z naslednjimi koraki preverite, ali aplikacije za iskanje in skrbništvo brezhibno delujejo:
 - a. Zaženite iskalne strežnike za poljubno zbirko, nato pa odprite vzorčno iskalno aplikacijo in predložite poizvedbo. S tem korakom preverite, ali obdelava predloženih poizvedb deluje, tudi če v indeksu ni dokumentov za iskanje.
 - b. Odprite skrbniško konzolo in preverite, ali se je mogoče brez težav prijaviti.

Onemogočitev podpore za protokol IPv6

1. Preberite razdelek **Preden začnete** in v vsakem primerku naslova IPv6 zamenjajte vrednost z imenom gostitelja ali naslovom IPv4.
2. Opravite postopek, s katerim omogočite podporo za protokol IPv6, vendar z naslednjimi izjemami:
 - V koraku 3 na strani 26 odstranite spremenljivko okolja **ES_IPV6**.
 - V koraku 4 na strani 26 odstranite lastnosti po meri za JVM, ki so bile nastavljene za aplikacijo ESSearchServer na strežniku WebSphere Application Server.
3. Odstranite lastnost **LocalIPAddress** iz datoteke ES_NODE_ROOT/nodeinfo/es.cfg, če vsebuje naslov IPv6.
4. Opravite postopek, opisan v koraku 9, da preverite skrbniške in iskalne aplikacije.

Zbirke iskanja v podjetju

Zbirka iskanja v podjetju vsebuje celoten niz virov, ki jih lahko uporabniki preiščejo z eno poizvedbo. S pomočjo združevanja (federation) pa lahko uporabniki z eno poizvedbo preiščejo več zbirk.

Pri izdelavi zbirke podate možnosti, ki veljajo za celotno zbirko. Zbirka je prazna, dokler vanjo ne dodate vsebine.

Zbirke lahko dodate v sistem iskanja v podjetju na dva načina:

- Če ne poznate skrbniške konzole iskanja v podjetju ali če šele spoznavate način skupnega delovanja komponent zbirke, lahko za izdelavo zbirke uporabite čarovnika. Čarovnik za zbirke vam pomaga pri izvedbi nalog in omogoča, da svoje delo med postopkom izdelovanja shranite kot osnutek.
- Če poznate skrbniško konzolo, lahko izdelate zbirke tako, da izberete določene strani, nad katerimi želite izvajati skrbništvo, v pogledu Zbirke.

Po izdelavi zbirke s pomočjo krmilnih elementov v pogledu Zbirke lahko urejate in nadzirate zbirko, sistem iskanja v podjetju in možnosti zaščite.

Združitev zbirke

Če je podpora za združevanje vgrajena v iskalno aplikacijo, lahko uporabniki sočasno preiskujejo več zbirk. Združevanje omogoča tudi, da presežete omejitev velikosti za zbirko, ki je 20 000 000 dokumentov. Uporabniki lahko na primer preiščejo dve zbirki, od katerih vsaka vsebuje 20 000 000 dokumentov.

Kakovost iskanja je odvisna od ujemanj za posamezne zbirke, ki so nato združena v končen nastali niz. Rezultat je enak kot pri predložitvi dveh ločenih iskanj in nato združitvi in rangiranju rezultatov.

S tem povezana opravila

“Nadziranje zbirke” na strani 274

Izdelava zbirke s pomočjo čarovnika za zbirke

Če je za vas iskanje v podjetju novo, vam lahko pri izdelavi zbirke pomaga čarovnik. Čarovnik nudi podrobnosti o vsakem koraku v postopku in omogoča, da sproti shranite nastavitve.

Preden začnete

Zbirke lahko izdelujejo samo člani skrbniške vloge za iskanje v podjetju.

Za dodajanje vsebine zbirki ali podajanje možnosti razčlenjevanja, indeksiranja ali iskanja v zbirki morate biti skrbnik iskanja v podjetju ali skrbnik zbirke za zbirko.

Omejitve

Čarovnika za zbirke lahko uporabite za izdelavo naslednjih vrst pajkov:

- Content Edition
- DB2


- DB2 Content Manager
- Exchange Server
- Notes
- datotečnega sistema UNIX
- spleta
- datotečnega sistema Windows

O tej nalogi

Ko izdelujete zbirko, jo lahko shranite v stanju osnutka. Ko je zbirka v stanju osnutka, jo lahko spremeni vsak skrbnik, ki ima pooblastilo za urejanje zbirke. Tako morda želite, da bi skrbnik zbirke, ki ima izkušnje z viri Lotus Notes konfiguriral pajka Notes. Pozneje lahko skrbnik zbirke, ki ima izkušnje s sistemi UNIX uredi osnutek zbirke tako, da konfigurira pajka datotečnega sistema UNIX.

Postopek

Izdelava zbirke s pomočjo čarovnika za zbirke:

1. Kliknite **Zbirke**, da odprete pogled Zbirke.
2. Kliknite **Čarovnik za zbirke**.
3. Sledite navodilom v čarovniku, da izdelate prazno zbirko in vanjo dodate vsebino. Konfigurirati morate splošne informacije o zbirki in izdelati vsaj enega pajka. Za preostale konfiguracijske možnosti lahko sprejmete privzete vrednosti ali pa za vašo novo zbirko podate možnosti.
4. Če želite zbirko shraniti, preden ste jo dokončali, kliknite **Shrani kot osnutek**. Vaša zbirka je z drugimi osnutki zbirke navedena v pogledu Zbirke. Če ste za zbirko omogočili zaščito, je poleg imena zbirke prikazana ikona  **Zaščita zbirke je omogočena**.
5. Če se želite vrniti na zbirko, ki jo še izdelujete, v pogledu Zbirke kliknite **Vrni se na čarovnika**.
6. Kliknite **Dokončaj**, da izdelate zbirko. Vaša zbirka je z drugimi zbirkami navedena v pogledu Zbirke.

Ko izdelate zbirko, morate zagnati postopke za pajkanje, razčlenjevanje, indeksiranje in iskanje v zbirki. Dokler niste pripravljeni, da bi zbirko povezali z iskalno aplikacijo, ki bo iskala po njej, lahko za iskanje v vaši novi zbirki uporabljate vzorčno iskalno aplikacijo (z imenom Default (Privzeta)).

Izdelava zbirke s pogledom Zbirke

S pogledom Zbirke lahko izdelate prazno zbirko. Nato lahko zbirko uredite, da podate možnosti za dodajanje podatkov zbirki in omogočite iskanje po zbirki.

Predn začnete

Zbirke lahko izdelujejo samo člani skrbniške vloge za iskanje v podjetju.


Za dodajanje vsebine zbirki ali podajanje možnosti razčlenjevanja, indeksiranja ali iskanja v zbirki morate biti skrbnik iskanja v podjetju ali skrbnik zbirke za zbirko.

O tej nalogi

Za informacije o vrednostih, ki jih lahko podate za novo zbirko, med izdelovanjem zbirke kliknite **Pomoč**.

Postopek

Izdelovanje zbirke v pogledu Zbirke:

1. V pogledu Zbirke kliknite **Izdelava zbirke**.
2. Na strani Izdelava zbirke vnesite informacije ali opravite izbiro v naslednjih poljih:
 - **Ime zbirke**. Podajte opisno ime za vsebino ali namen zbirke.
 - **Zaščita zbirke**. Podajte ali želite omogočiti zaščito za zbirko. Ko izdelate zbirko, te nastavitve ne morete spremeniti. Če omogočite zaščito zbirke, lahko pozneje podate možnosti za uveljavitev nadzora dostopa na ravni dokumentov.
 - **Pomembnost dokumentov (model statičnega rangiranja)**. Podajte strategijo za dodeljevanje faktorja statičnega rangiranja, ki se bo uporabil za rangiranje dokumentov v rezultatih iskanja. Ko izdelate zbirko, te vrednosti ne morete več spremeniti.
 - **Vrsta kategorizacije**. Podajte ali želite, da bo mogoče iskanje dokumentov po kategorijah, katerim pripadajo.
 - **Uporabljeni jezik**. Podajte privzeti jezik za iskanje dokumentov v zbirki.
3. Sprejmite privzete vrednosti za naslednja polja ali podajte možnosti, ki jih želite uporabiti v zbirki.
 - **Opis**. Po privzetku se opis ne izdela.
 - **Pričakovano število dokumentov**. Privzeta pričakovana velikost zbirke je 1 000 000 dokumentov. Sistem uporablja to vrednost za to, da oceni velikost pomnilnika in diska, ki bo potrebna za zbirko in ne za omejitve zbirke.
 - **Nahajališče podatkov zbirke**. Privzeto nahajališče za datoteke, povezane z zbirko, je na indeksnem strežniku. Ko izdelate zbirko, te vrednosti ne morete več spremeniti.
 - **ID zbirke**. Privzet ID zbirke temelji na imenu zbirke. Ko izdelate zbirko, te vrednosti ne morete več spremeniti. Če podate ID zbirke po meri, lahko iskalne aplikacije zbirko pokličejo s tem identifikatorjem in ne s potencialno šifriranim identifikatorjem, ki ga izdela sistem.
 - **Segmentacija n-gram**. Privzeta metoda segmentacije je segmentacija na osnovi Unicode s praznimi prostori. Segmentacijo n-gram izberite samo, če zbirka vsebuje kitajske, japonske ali korejske dokumente in želite, da razčlenjevalnik za razmejitev besed uporabi n-grame. Ko izdelate zbirko, te vrednosti ne morete več spremeniti.
Več informacij o načinih konfiguriranja podpore za razčlenjevanje n-gram in izdelavi žetonov v zbirkah za iskanje v podjetju poiščite v <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088>.
4. Kliknite **Potrdi**.
V pogledu Zbirke je nova zbirka navedena z drugimi zbirkami v sistemu iskanja v podjetju. Če ste za zbirko omogočili zaščito, je poleg imena zbirke prikazana ikona  **Zaščita zbirke je omogočena**.

Zbirka je prazna, dokler vanjo ne dodate vsebine. Za dodajanje vsebine v novo zbirko izberite zbirko v pogledu Zbirke, jo uredite, izdelajte vsaj enega pajka in podajte možnosti za razčlenjevanje, indeksiranje in iskanje dokumentov.

Zagnati morate postopke za pajkanje, razčlenjevanje, indeksiranje in iskanje v zbirki. Po novi zbirki lahko iščete z vzorčno iskalno aplikacijo, dokler niste pripravljeni, da uporabite vašo iskalno aplikacijo po meri.

Urejanje zbirke

Zbirko lahko uredite, če želite podati informacije o dokumentih, ki jih želite vključiti v zbirko.

Preden začnete


Zbirko lahko urejate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

O tej nalogi

Pri urejanju zbirke podate možnosti za pajkanje virov podatkov, razčlenjevanje dokumentov, gradnjo indeksa in iskanje po indeksirani vsebini in beleženje sporočil o napakah. Ko izdelate zbirko, jo morate urediti, da ji dodate vsebino. Pozneje lahko zbirko uredite, da posodobite vsebino ali spremenite način pajkanja, razčlenjevanja, indeksiranja, iskanja ali beleženja informacij.

Postopek

Urejanje zbirke:

1. V orodjarni kliknite **Zbirke**, da odprete pogled Zbirke.
2. Na seznamu zbirk poiščite zbirko, ki jo želite urediti in kliknite  **Uredi**.
3. Napravite spremembe na katerikoli izmed naslednjih strani:

Pajkanje

Podate lahko vire podatkov, v katerih želite izvesti pajkanje, in možnosti za pajkanje vsebine. Vsaka zbirka mora vsebovati vsaj enega pajka in posamezna zbirka lahko vsebuje podatke iz več vrst virov podatkov. Za vsako vrsto vira podatkov, ki jih želite vključiti v zbirko, morate definirati vsaj enega pajka.

Razčlenjevanje

Podate lahko možnosti za razčlenjevanje in analiziranje dokumentov, ki so bili pajkani. Konfigurirate lahko kategorije, ki uporabnikom omogočajo, da iščejo po podnizih zbirke, in pravila, ki uporabnikom omogočajo iskanje določenih delov dokumentov XML in HTML. Če iskalnemu sistemu podjetja dodate motorje za analizo besedila po meri, lahko izberete enega, ki se uporabi pri analizi in razlaganju vsebine v tej zbirki. Polja lahko povežete s poviševalnimi razredi in tako vplivate na rangiranje dokumentov v rezultatih iskanja.

Indeksiranje

Podate lahko razporede za gradnjo celotnega indeksa in posodabljanje indeksa z novo in spremenjeno vsebino. Lahko tudi konfigurirate možnosti za uporabo univerzalnih znakov v poizvedbah, omejevanje pogleda zbirke na obseg URI-jev, skrčevanje rezultatov iskanja iz istega spletnega mesta in odstranjevanje URI-jev iz indeksa.

Iskanje

Podate lahko možnosti za postopke iskanja, kot sta konfiguriranje iskalnega predpomnilnika in izbiranje jezika. Konfigurirate lahko tudi hitre povezave. To je funkcija, ki zagotavlja, da poizvedba vrne vnaprej določene URI-je, kadarkoli uporabnik v poizvedbo vključi določeno besedo ali besedno zvezo. Če ste v sistem iskanja v podjetju dodali prilagojene slovarje, lahko izberete slovarje, ki jih želite uporabiti za iskanje po tej zbirki.

Dnevnik

Podate lahko vrste sporočil, ki jih želite beležiti, in možnosti za izdelavo in rotiranje datotek dnevnika. Konfigurirate lahko opozorila, tako da boste obveščeni, ko se zgodijo določeni dogodki, in podate možnosti za prejemanje

elektronskih sporočil, kadarkoli se zabeležijo določena sporočila ali določena vrsta sporočil. Podate lahko možnosti za beleženje informacij o dokumentih, ki jih izbrišete iz sistema iskanja v podjetju.

Splošno

Podate lahko splošne informacije o zbirki in si ogledate nastavitve, ki jih ne morete spremeniti. Če ste ob izdelavi zbirke omogočili zaščito, lahko konfigurirate možnosti zaščite na ravni dokumenta.

Brisanje zbirke

Brisanje zbirke popolnoma odstrani vse informacije o zbirki iz iskalnega sistema podjetja.

Preden začnete

Zbirke lahko brišejo samo člani skrbniške vloge za iskanje v podjetju.

Zbirko lahko izbrišete šele, ko zaustavite vse procese, povezane z njo.


O tej nalogi

Brisanje zbirke je lahko precej zamudno. Ko potrdite brisanje zbirke, sistem izbriše vse podatke v sistemu, ki so povezani z zbirko.

Nasvet: Morda bo sistem prikazal sporočilo, da je zahtevana operacija potekla, četudi se proces v ozadju še zmeraj izvaja. Da bi ugotovili, ali se je naloga že končala, na skrbniški konzoli kliknite možnost **Osveži** (ne kliknite **Osveži** v spletnem brskalniku). Proces brisanja je končan, ko ime zbirke ni več prikazano na seznamu zbir.

Postopek

Brisanje zbirke:

1. Kliknite **Zbirke**, da odprete pogled Zbirke.
2. Na seznamu zbir poiščite zbirko, ki jo želite izbrisati in kliknite  **Izbriši**.

Ugotavljanje ID-ja zbirke

Za večino skrbniških nalog morate poznati ID zbirke.

Predn začnete

Če si želite ogledati ID zbirke, morate biti član s skrbniško vlogo v iskanju v podjetju.

O tej nalogi

Pri izdelavi zbirke lahko bodisi podate vrednost za ID ali pa prepustite sistemu, da to stori namesto vas. Če želite ugotoviti ID zbirke po izdelavi, lahko uporabite skrbniško konzolo ali pa si ogledate konfiguracijsko datoteko.

Postopek

1. Če želite ugotoviti ID zbirke z uporabo skrbniške konzole, naredite naslednje:
 - a. Kliknite **Zbirke**, da odprete pogled Zbirke.
 - b. Na strani Splošno kliknite **Ogled nastavitv zbirke**.

Stran Nastavitve zbirke prikazuje ID zbirke, popolnoma kvalificirano pot do mesta, kjer so shranjeni podatki zbirke, in model statičnega rangiranja, uporabljen z dokumenti v tej zbirki.

2. Če želite ugotoviti ID zbirke z ogledom konfiguracijske datoteke, naredite naslednje:

a. Odprite datoteko `ES_NODE_ROOT/master_config/collections.ini`. Uredite datoteko za lažje pregledovanje. V naslednjih vzorčnih izhodnih podatkih je `col1` ID zbirke:

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

Skrbnišтво nad pajkom

Pajke konfigurirate za različne tipe podatkov, ki jih želite vključiti v zbirko. Ena zbirka lahko vsebuje poljubno število pajkov.

Nasvet: Na naslovu <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html> je na voljo spletna vadnica. Vadnica opisuje posamezne korake, potrebne za namestitev in konfiguriranje izdelka, prikazuje načine iskanja po posameznih tipih virov podatkov ter opisuje, kako s pomočjo aplikacijskega programerskega vmesnika izdelka razširite iskanje v podjetju. Čeprav se vadnica nanaša na starejšo različico izdelka OmniFind Enterprise Edition, je večina zasnov in postopkov še vedno aktualna.

Podrobnejša navodila o načinih konfiguriranja pajkov in omogočitvi zaščite v majhnih, srednjih in velikih organizacijah poiščite v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Konfiguriranje pajkov

Skrbniška konzola iskanja v podjetju se uporablja za izdelovanje, urejanje in brisanje pajkov. Običajno konfigurira pajka strokovnjak za tip podatkov, v katerih pajkate. Za nastavitev pajka za pajkanje v virih podatkov Lotus Notes mora biti skrbnik zbirke skrbnik Notes ali pa tesno sodelovati z nekom, ki se spozna na baze podatkov, v katerih boste izvajali pajkanje.

Pri izdelavi pajka vam čarovnik za tip podatkov, v katerih boste pajkali, pomaga podati lastnosti, ki krmilijo, kako pajek uporablja sistemske vire. Čarovnik vam tudi pomaga izbrati vire, ki jih želite preiskati.

Spremembe v obstoječih pajkih lahko opravite kadarkoli. Po potrebi lahko uredite lastnosti pajka ali dele prostora za pajkanje. Tudi pri teh spremembah vam pomagajo čarovniki.

Poselitev novega pajka z osnovnimi vrednostmi

Pajka lahko izdelate s pomočjo sistemskih privzetih vrednosti ali s kopiranjem vrednosti, ki so podane za drugega pajka istega tipa. Če uporabite kot osnovo za novega pajka obstoječega pajka, lahko hitro izdelate več pajkov s podobnimi lastnostmi in jih nato na primer konfigurirate za pajkanje v različnih virih podatkov ali za delovanjem po različnih razporedih pajkanja.

S kopiranjem pajka lahko porazdelite obremenitev pajkanja med več pajkov, ki uporabljajo enaka pravila za pajkanje. Tako lahko na primer prekopirate pajka Notes, ker želite z drugim strežnikom Lotus Notes uporabiti enake lastnosti in pravila za pajkanje polj. Edina razlika so lahko baze podatkov, v katerih pajka posamezen pajek in nastavitve zaščite na ravni dokumentov.

Združevanje tipov pajkov v zbirko

Pajki iskanja v podjetju so oblikovani tako, da zbirajo informacije iz različnih tipov virov podatkov. Pri konfiguriranju pajkov za zbirko se morate odločiti, kako združiti te različne tipe virov podatkov, da bodo uporabniki lahko preprosto iskali v podatkih podjetja. Če želite uporabnikom na primer omogočiti preiskovanje datotečnih sistemov Microsoft Windows in javnih map Microsoft Exchange Server z eno samo poizvedbo, izdelajte zbirko, ki vključuje pajke datotečnega sistema Windows in pajke Exchange Server.

Pri združevanju različnih tipov pajkov v eno zbirko zagotovite, da bodo lahko vsi pajki uporabljali enak način statičnega rangiranja. (Način statičnega rangiranja podate pri izdelavi zbirke.) Če na primer združite spletne vire (ki uporabljajo kot faktor rangiranja povezave dokumentov) in vire NNTP (ki kot faktor rangiranja običajno uporabljajo datum dokumentov), se lahko kakovost rezultatov iskanja zmanjša.

Konfiguriranje zaščite na ravni dokumentov

Če omogočite zaščito zbirke pri njeni izdelavi, lahko konfigurirate možnosti zaščite na ravni dokumentov. Vsak pajek lahko poveže žetone zaščite z dokumenti, v katerih pajka. Če pri konfiguriranju pajka podate, da želite uporabljati zaščito na ravni dokumentov, pajek poveže žetone zaščite, ki jih podate, z vsakim dokumentom, ti žetoni pa so dodani v indeks z dokumenti.

Če omogočite zaščito v iskalnih aplikacijah po meri, lahko uporabljajo aplikacije za overjanje uporabnikov žetone zaščite, ki so jih pajki povezali z dokumenti. Ta zmožnost omogoča omejitve dostopa do nekaterih dokumentov v zbirki in za druge dokumente omogoča, da jih preiščejo vsi uporabniki. Tako lahko na primer v eni zbirki vsem uporabnikom omogočite dostop do vseh dokumentov v javnih mapah Microsoft Exchange Server, v bazah podatkov Lotus Notes pa omogočite dostop do dokumentov samo uporabnikom s specifičnimi ID-ji uporabnikov.

Uveljavite lahko prilagojena poslovna pravila, da s kodiranjem pravil v razredu Java določite vrednost žetonov zaščite. Ko konfigurirate lastnosti pajka, podate ime dodatka, ki ga naj uporablja pajek, ko pajka dokumente. Žetoni zaščite, ki jih doda vaš dodatek, so shranjeni v indeksu in jih lahko uporabite za nadzor dostopa do dokumentov.

Pri konfiguriranju določenih tipov pajkov lahko podate dodatne krmilne elemente zaščite. Tako lahko na primer podate, da želite med obdelovanjem poizvedbe preveriti uporabnike. Če omogočite to možnost, so poverilnice uporabnika primerjane s trenutnimi krmilnimi elementi za nadzor dostopa, ki jih vzdržujejo viri podatkov, v katerih boste izvajali iskanje. To preverjanje trenutnih poverilnic lahko opravite namesto ali poleg preverjanja, ki temelji na žetonih zaščite v indeksu iskanja v podjetju.

S tem povezani pojmi

“Zaščita na ravni dokumenta” na strani 238

S tem povezana opravila

“Nadziranje pajkov” na strani 275

S tem povezane povezave

“Zahteve za nastavitve pajka za podporo zaščite” na strani 253

Izdelava pajka

Ko izdelujete pajka, podate vrsto pajka, ki ga želite izdelati. Čarovnik vam pomaga podati informacije o podatkih, ki jih želite vključiti v zbirko.

Preden začnete

Pajka lahko izdelate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

O tej nalogi

Za vsako zbirko morate izdelati vsaj enega pajka. Vrsta pajka, ki ga izdelate, je odvisna od vrste podatkov, ki jih želite vključiti v zbirko. Čarovnik za vrsto pajka, ki ga izdelujete, vam

pomaga podati možnosti za pajka. Tako vam čarovnik na primer pomaga podati možnosti za to, kako naj pajek uporablja sistemske vire. Čarovnik vam tudi pomaga izbrati vire podatkov, ki jih želite vključiti v zbirko.

Postopek

Izdelava pajka:

1. Uredite zbirko, izberite stran Pajkanje in kliknite **Izdelava pajka**.
2. Izberite vrsto pajka in osnovne vrednosti za pajka:
 - a. Izberite vrsto pajka, ki podpira vrsto podatkov, ki jih želite pajkati, kot so na primer spletne strani, baze podatkov Lotus Notes ali datotečni sistemi UNIX.
Po izbiri vrste pajka se prikažejo možnosti za način njegove izdelave.
 - b. Izberite osnovne vrednosti za pajka:

Za novega pajka uporabi privzete sistemske vrednosti

Začetne nastavitve pajka izpolni s privzetimi namestitvenimi vrednostmi.

Če izberete to možnost, kliknite **Naprej**, da začnete s konfiguriranjem novega pajka.

Za novega pajka prekopiraj vrednosti iz obstoječega pajka

Začetne nastavitve pajka izpolni z vrednostmi, ki so konfigurirane za drugega pajka te vrste.

Če izberete to možnost, se prikaže seznam pajkov, ki se ujemajo s to vrsto pajka. Izberite pajka, ki ga želite uporabiti za novega pajka, nato pa kliknite **Naprej**, da začnete s konfiguriranjem novega pajka.

Odpre se čarovnik za vrsto pajka, ki ga izdelujete. Za izdelavo pajka upoštevajte pozive čarovnika. Na katerikoli strani čarovnika lahko kliknete **Pomoč**, če želite izvedeti več o možnostih, ki jih lahko podate za to vrsto pajka.

Novi pajek je naveden na strani Pajkanje z drugimi pajki, ki pripadajo isti zbirki. Kadarkoli želite napraviti spremembe v pajku, lahko kliknete možnosti, da uredite lastnosti pajka in prostor pajkanja.

Urejanje lastnosti pajka

Spremenite lahko informacije o pajku in način, kako pajka podate. Tako lahko na primer spremenite, kako pajek uporablja vire sistema.

Preden začnete

Lastnosti pajka lahko urejate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

O tej nalogi

Med urejanjem lastnosti pajka lahko kliknete **Pomoč**, da si ogledate vrste sprememb, ki jih lahko napravite. Lastnosti, ki jih lahko uredite, so odvisne od vrste pajka.

Postopek

Urejanje lastnosti pajka:

1. Uredite zbirko, izberite stran Pajkanje, poiščite pajka, ki ga želite urediti, in kliknite

 **Lastnosti pajka.**

2. Spremenite lastnosti pajka in kliknite **Potrdi**.
3. Če želite uveljaviti spremembe, zaustavite in znova zaženite pajka. (Če ste spremenili samo opis pajka, vam ni potrebno ponovno zagnati pajka.)

Urejanje prostora za pajkanje

Spremenite lahko informacije o virih podatkov, ki jih pajek pajka. Tako lahko na primer dodate vire podatkov, odstranite vire podatkov, spremenite razpored pajkanja in spremenite pravila za pajkanje dokumentov v določenem viru podatkov.

Preden začnete


Prostor za pajkanje lahko uredite samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

O tej nalogi

Če želite izvedeti, katere spremembe lahko napravite za vrsto pajka, nad katerim izvajate skrbništvo, med urejanjem prostora za pajkanje kliknite **Pomoč**.

Postopek

Urejanje prostora za pajkanje:

1. Uredite zbirko, izberite stran Pajkanje, poiščite pajka, ki ga želite urediti, in kliknite
2.  **Prostor za pajkanje.** Spremenite prostor za pajkanje tako, da izberete možnosti, ki jih želite spremeniti. Možnosti, ki jih lahko izberete, so odvisne od vrste pajka. Za nekatere možnosti, kot je na primer dodajanje virov podatkov zbirki, se odpre čarovnik za vrsto pajka, ki vam pomaga spremeniti prostor za pajkanje.
3. Če želite uveljaviti spremembe, zaustavite in znova zaženite pajka.

Brisanje pajka

Brisanje pajka popolnoma odstrani vse informacije o pajku iz iskalnega sistema podjetja. Informacije, v katerih je predhodno pajkal pajek, pa ostanejo v indeksu do naslednje gradnje glavnega indeksa.

Preden začnete

Pajka lahko izbrišete samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.


O tej nalogi

Brisanje pajka je lahko precej zamudno. Ko potrdite brisanje pajka, sistem izbriše vse podatke v sistemu, ki so povezani s pajkom.

Nasvet: Ker to opravilo traja dlje časa, se bo morda prikazalo sporočilo o preteku čakalnega časa za zahtevano operacijo, kljub temu da proces v ozadju še vedno teče. Če želite ugotoviti, ali se je naloga dokončala, na skrbniški konzoli klikajte **Osveži** (ne klikajte možnosti **Osveži** v spletnem brskalniku). Postopek brisanja je končan, ko ime pajka ni več prikazano na seznamu pajkov.

Postopek

Brisanje pajka:

1. Uredite zbirko in izberite stran Pajkanje.
2. Poiščite pajka, ki ga želite izbrisati in kliknite  **Izbriši**.

Razporedi pajka

Pajki, ki jih izdelate za spletne vire, se izvajajo neprekinjeno. Ko spletnega pajka enkrat zaženete, ga običajno ni potrebno zaustaviti, razen v primeru, če spremenite njegovo konfiguracijo. Za vse druge tipe pajkov podate razpored pajkanja pri konfiguriranju pajka.

Za nekatere tipe virov podatkov en sam razpored krmili, kdaj bo pajek obiskal vse vire podatkov v prostoru za pajkanje. Za druge tipe virov podatkov lahko podate različne razporede za specifične vire podatkov. Tako lahko na primer podate različne razporede za pajkanje baze podatkov Lotus Notes, v kateri pajka pajek.

Pri konfiguriranju razporeda podate tip pajkanja, ki se bo izvajalo. Načrtujete lahko celotno pajkanje vseh dokumentov v prostoru za pajkanje, pajkanje, ki vključuje vse posodobitve v prostoru za pajkanje (novi dokumenti, spremenjeni dokumenti in izbrisani dokumenti) ali pajkanje, ki vključuje samo nove in spremenjene dokumente. Celotno pajkanje zahteva največ časa. Pajkanje, ki odstrani izbrisane dokumente, traja dlje časa od pajkanja, ki izbrisane dokumente prezre.

Pri urejanju prostora pajkanja podate še drug razpored pajkanja. Tako lahko na primer konfigurirate razpored za pajkanje vseh dokumentov v prostoru za pajkanje vsako soboto ponoči, drug razpored, ki se bo izvajal pogosteje, pa za pajkanje novih in spremenjenih dokumentov.

Izdelava več razporedov pajka omogoča boljši nadzor nad tem, kdaj pajek obiše ciljne vire. Če želite na primer izvesti pajkanje v bazah podatkov v različnih časovnih pasovih, lahko načrtujete pajka ob urah, ko bodo uporabniki najverjetneje zaključili svoje delo.

Pajki Content Edition

Če želite v zbirko iskanja v podjetju vključiti repozitorije IBM WebSphere Information Integrator Content Edition, morate konfigurirati pajka Content Edition.

Pajka Content Edition lahko uporabljate za pajkanje naslednjih tipov repozitorijev:

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

Pri konfiguriranju pajka podate možnosti, ki določajo, kako bo pajek pajkal v vseh repozitorijih v prostoru za pajkanje. Izberete tudi razrede postavk, v katerih želite pajkati v vsakem repozitoriju.

Nasvet:

Podrobnejše informacije o načinih konfiguriranja spojnikov in zaščiti pajka Content Edition najdete v scenariju za srednje veliko organizacijo, opisanem v publikaciji IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Konfiguriranje strežnika pajkanja

Način, na katerega boste pripravili pajkanje repozitorijev, je odvisen od tega, ali načrtujete za povezovanje s podatki za pajkanje uporabljati neposreden ali strežniški način. Če uporabite neposreden način, morate v WebSphere Information Integrator Content Edition konfigurirati spojnik. Če uporabite strežniški način, morate na strežniku pajkanja zagnati skript. Ta skript, ki je na voljo z OmniFind Enterprise Edition, omogoča pajku Content Edition komuniciranje s strežniki WebSphere Information Integrator Content Edition.

Če uporabite strežniški način, pred izdelavo pajka Content Edition opravite nalogo, ki je ustrezna za vaše okolje:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za WebSphere II Content Edition” na strani 42.
- “Konfiguriranje strežnika pajkanja v sistemu Windows za WebSphere II Content Edition” na strani 43.

Podrobna navodila o tem, kako konfigurirati sistem iskanja v podjetju za preiskovanje repozitorijev WebSphere Information Integrator Content Edition, najdete v članku IBM developerWorks, *Search WebSphere Portal Document Manager using WebSphere Information Integrator OmniFind Edition* na URL-ju <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0606lee/>.

Pregled konfiguriranja

Če želite izdelati ali spremeniti pajka Content Edition, se prijavite na skrbniško konzolo iskanja v podjetju. Biti morate član s skrbniško vlogo v iskanju v podjetju ali skrbnik zbirke, ki je lastnik pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh repozitorijih v prostoru za pajkanje.
- Podajanje, ali bo pajek za dostopanje do repozitorijev uporabljal neposredni ali strežniški način. Za strežniški način morate podati tudi informacije, ki pajku omogočajo dostop do spletnega aplikacijskega strežnika.
- Izbira repozitorijev, v katerih želite pajkati.
- Podajanje ID-jev uporabnikov in gesel, ki pajku omogočajo dostop do vsebine v izbranih repozitorijih.
- Nastavitev razporeda za pajkanje repozitorijev.
- Izbira razredov postavk, ki jih želite pajkati v vsakem repozitoriju.
- Podajanje možnosti, ki omogočajo preiskovanje lastnosti razredov postavk. Tako lahko na primer iz prostora pajkanja izključite določene tipe dokumentov ali podate, da želite pajkati v določeni različici repozitorija.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo iskalnim aplikacijam, da uveljavijo krmilne elemente zaščite na osnovi seznamov za nadzor dostopa ali žetonov zaščite.

Za razrede postavk Documentum, FileNet Panagon Content Services, Hummingbird DM, Portal Document Manager in SharePoint lahko izberete tudi možnost, s katero ob predložitvi poizvedbe preverite poverilnice uporabnika. V tem primeru sistem ne primerja poverilnic uporabnika z indeksiranimi podatki o zaščiti, pač pa s trenutnimi seznamami za nadzor dostopa, ki jih vzdržuje izviren vir podatkov. Ta vrsta preverjanja trenutnih poverilnic ni na voljo za druge tipe repozitorijev.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

Način neposrednega dostopa do repozitorijev Content Edition

Pajka Content Edition lahko konfigurirate za dostop do repozitorijev WebSphere Information Integrator Content Edition v neposrednem načinu.

O tej nalogi

V neposrednem načinu pajek uporablja spojnik WebSphere Information Integrator Content Edition, ki se namesti na strežnik pajkanja ob namestitvi OmniFind Enterprise Edition. Pajek uporablja za neposredno povezovanje z repozitoriji za pajkanje API-je za integracijo vsebine. Če se izvaja strežnik za integracijo vsebine v neposrednem načinu, ni na voljo vsa funkcionalnost strežnika za integracijo vsebine. Informacije o izvajanju strežnika za integracijo vsebine v neposrednem načinu in o tem, kako se funkcionalnost razlikuje, če se izvaja strežnik za integracijo v strežniškem načinu, najdete v dokumentaciji izdelka WebSphere Information Integrator Content Edition.

Naslednji postopek povzema korake, ki jih morate nastaviti za dostop v neposrednem načinu. Podrobnejša navodila najdete v članku IBM developerWorks, *Search WebSphere Portal Document Manager using WebSphere Information Integrator OmniFind Edition* na URL-ju <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0606lee/>.

Postopek

Postopek za konfiguriranje sistema, tako da bo pajek lahko dostopal do repozitorijev v neposrednem načinu.

1. Potrdite, da spremenljivki okolja **VBR_HOME** in **JAVA_HOME** v datoteki *iice_install_root/bin/config.sh* (v UNIX) ali v datoteki *iice_install_root\bin\config.bat* (v Microsoft Windows) podajata ustrezen imenik.
2. Če želite konfigurirati skrbniško konzolo WebSphere Information Integrator Content Edition za izvajanje v neposrednem načinu, dodajte lastnost sistema **-Dvbr.as.operationMode=direct** Java v datoteko *iice_install_root/bin/Admin.sh* (v UNIX) ali datoteko *iice_install_root\bin\Admin.bat* (v Windows). Primer:

Admin.sh file

```
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
```

Admin.bat file

```
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. Zaženite skrbniško konzolo WebSphere Information Integrator Content Edition v neposrednem načinu in konfigurirajte spojnik za strežnik pajkanja OmniFind Enterprise Edition.
4. Če uporabite za konfiguriranje pajka Content Edition skrbniško konzolo iskanja v podjetju, izberite možnost neposrednega načina.

Način strežniškega dostopa do repozitorijev WebSphere II Content Edition

Pajka Content Edition lahko konfigurirate tako, da do repozitorijev dostopi v strežniškem načinu.

Spojnik WebSphere Information Integrator Content Edition, ki ga pajek uporablja za dostop do podatkov, je v strežniškem načinu nameščen kot aplikacija podjetja na aplikacijskem strežniku WebSphere Application Server, pajek pa do repozitorijev dostopa prek strežnika. S tem pristopom lahko izkoristite prednosti okolij aplikacijskega strežnika J2EE.

Pred konfiguriranjem pajka za dostop do repozitorijev WebSphere Information Integrator Content Edition v strežniškem načinu morate na strežniku pajkanja zagnati skript. Skript, ki je na voljo skupaj z OmniFind Enterprise Edition, pajku Content Edition omogoča dostop do repozitorijev na strežniku.

Opravite nalogo, ki ustreza vašemu okolju:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za WebSphere II Content Edition”.
- “Konfiguriranje strežnika pajkanja v sistemu Windows za WebSphere II Content Edition” na strani 43.

Konfiguriranje strežnika pajkanja v sistemu UNIX za WebSphere II Content Edition

Če namestite izdelek OmniFind Enterprise Edition v računalnik, v katerem se izvaja operacijski sistem IBM AIX, Linux ali Solaris in konfigurirate pajka Content Edition tako, da pri dostopu do repozitorijev uporablja strežniški način, morate zagnati skript in konfigurirati strežnik pajkanja. Skript pajku Content Edition omogoči dostop do repozitorijev WebSphere Information Integrator Content Edition.

O tej nalogi

Pajek Content Edition uporablja knjižnice Java za WebSphere Information Integrator Content Edition kot odjemalca Java. Te knjižnice Java v strežniškem načinu zahtevajo z EJB povezane knjižnice Java za aplikacijski strežnik WebSphere Application Server. Če se želite prepričati, da je mogoče pajka Content Edition uporabljati skupaj s knjižnicami Java, morate po namestitvi aplikacijskega strežnika WebSphere Application Server zagnati namestitveni skript, ki ga na strežniku pajkanja nudi OmniFind Enterprise Edition.

WebSphere Information Integrator Content Edition je nameščen na strežniku pajkanja skupaj z izdelkom OmniFind Enterprise Edition. Če želite pajka Content Edition uporabiti v strežniškem načinu, morate prekopirati datoteko `vbr_access_services.jar` s strežnika WebSphere Information Integrator Content Edition na strežnik pajkanja.

Postopek

Če želite konfigurirati strežnik pajkanja tako, da bo mogoče z njim pajkati repozitorije WebSphere Information Integrator Content Edition, naredite naslednje:

1. Če je OmniFind Enterprise Edition nameščen v konfiguraciji z več strežniki, namestite in vežite knjižnice Java aplikacijskega strežnika WebSphere Application Server.
2. Na strežniku pajkanja zaženite namestitveni skript pajka Content Edition:
 - a. Prijavite se kot skrbnik iskanja v podjetju.
 - b. Zaženite naslednji skript, nameščen v imeniku `$ES_INSTALL_ROOT/bin`, in odgovorite na pozive:

escrvbr.sh

3. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:

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

4. Prekopirajte datoteko `vbr_access_services.jar` s strežnika WebSphere Information Integrator Content Edition na strežnik pajkanja.

Kopiranje iz:

Datoteka `vbr_access_services.jar` je na privzetem nahajališču:

```
was_install_root/installedApps/server_name/application_name
```

was_install_root je namestitveni imenik aplikacijskega strežnika WebSphere Application Server, *server_name* je ime, ki ste ga dodelili strežniku, *application_name* pa je ime aplikacije, ki ste ga podali za aplikacijo WebSphere Information Integrator Content Edition na aplikacijskem strežniku WebSphere Application Server.

Kopiranje v:

Ciljni imenik na strežniku pajkanja je *iice_install_root/lib*, pri čemer je *iice_install_root* namestitveni imenik izdelka WebSphere Information Integrator Content Edition na strežniku pajkanja.

Konfiguriranje strežnika pajkanja v sistemu Windows za WebSphere II Content Edition

Če namestite OmniFind Enterprise Edition v računalnik, v katerem se izvaja operacijski sistem Microsoft Windows, in konfigurirate pajka Content Edition tako, da pri dostopu do repozitorijev uporablja strežniški način, morate zagnati skript in konfigurirati strežnik pajkanja. Skript pajku Content Edition omogoči dostop do repozitorijev WebSphere Information Integrator Content Edition.

O tej nalogi

Pajek Content Edition uporablja knjižnice Java za WebSphere Information Integrator Content Edition kot odjemalca Java. Te knjižnice Java v strežniškem načinu zahtevajo z EJB povezane knjižnice Java za aplikacijski strežnik WebSphere Application Server. Če se želite prepričati, da je mogoče pajka Content Edition uporabljati skupaj s knjižnicami Java, morate po namestitvi aplikacijskega strežnika WebSphere Application Server zagnati namestitveni skript, ki ga na strežniku pajkanja nudi OmniFind Enterprise Edition.

WebSphere Information Integrator Content Edition je nameščen na strežniku pajkanja skupaj z izdelkom OmniFind Enterprise Edition. Če želite pajka Content Edition uporabiti v strežniškem načinu, morate prekopirati datoteko `vbr_access_services.jar` s strežnika WebSphere Information Integrator Content Edition na strežnik pajkanja.

Postopek

Če želite konfigurirati strežnik pajkanja tako, da bo mogoče z njim pajkati repozitorije WebSphere Information Integrator Content Edition, naredite naslednje:

1. Če je OmniFind Enterprise Edition nameščen v konfiguraciji z več strežniki, namestite in vežite knjižnice Java aplikacijskega strežnika WebSphere Application Server.
2. Na strežniku pajkanja zaženite namestitveni skript pajka Content Edition:
 - a. Prijavite se z ID-jem skrbnika za iskanje v podjetju (ID uporabnika ste podali ob namestitvi izdelka OmniFind Enterprise Edition).
 - b. Zaženite naslednji skript, ki je nameščen v imeniku `%ES_INSTALL_ROOT%\bin` in odgovorite na pozive:

escrvbr.vbs

3. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:

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

4. Prekopirajte datoteko `vbr_access_services.jar` s strežnika WebSphere Information Integrator Content Edition na strežnik pajkanja.

Kopiranje iz:

Datoteka `vbr_access_services.jar` je na privzetem nahajališču:

```
was_install_root\installedApps\server_name\application_name
```

was_install_root je namestitveni imenik aplikacijskega strežnika WebSphere Application Server, *server_name* je ime, ki ste ga dodelili strežniku, *application_name* pa je ime aplikacije, ki ste ga podali za aplikacijo WebSphere Information Integrator Content Edition na aplikacijskem strežniku WebSphere Application Server.

Kopiranje v:

Ciljni imenik na strežniku pajkanja je *iice_install_root*\lib, pri čemer je *iice_install_root* namestitveni imenik izdelka WebSphere Information Integrator Content Edition na strežniku pajkanja.

Pajki DB2

S pajkom DB2 lahko vključite baze podatkov IBM DB2 v zbirko.

Če z izdelkom IBM WebSphere Information Integrator izdelate tabele vzdevkov za naslednje tipe sistemov baz podatkov in jih povežete v zveze, lahko s pajkom DB2 pajkate tabele po vzdevkih:

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

Za vsak strežnik baz podatkov, ki ga želite pajkati, morate konfigurirati ločenega pajka. Pri konfiguriranju pajka podate možnosti, ki določajo, kako bo pajek pajkal v vseh bazah podatkov na istem strežniku. Izberete tudi specifične tabele, ki jih želite pajkati v vsaki bazi podatkov.

Tabele, ki jih izberete za pajkanje, morajo biti tabele baze podatkov, tabele nadomestnih imen ali pogledi. Pajek DB2 ne podpira združenih tabel.

Nasvet:

Podrobnejše primere z načini konfiguriranja zaščitenega pajka DB2 najdete v scenarijih za večjo organizacijo v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Konfiguriranje strežnika pajkanja

Preden začnete pajkati tabele baz podatkov, morate na strežnik pajkanja namestiti izdelek DB2 Administration Client. Na strežniku pajkanja morate nato zagnati skript. Ta skript, ki je na voljo z OmniFind Enterprise Edition, omogoča pajku DB2 komuniciranje s strežniki baz podatkov.

Preden za konfiguriranje pajka DB2 uporabite skrbniško konzolo iskanja v podjetju, izpolnite nalogo, ki ustreza vašemu okolju:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2” na strani 46.
- “Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2” na strani 47.

Objavljanje dogodkov

Če uporabljate WebSphere Information Integrator Event Publisher Edition in če povežete baze podatkov, ki jih želite pajkati, s preslikavami čakalne vrste objavljanja, lahko pajek DB2 uporabi preslikave za pajkanje posodobitev v tabelah baz podatkov.

Preslikava čakalne vrste objavljanja identificira čakalno vrsto WebSphere MQ, ki pri objavi posodobitev v tabeli baze podatkov sprejme sporočila XML. Pajek v čakalni vrsti spremlja informacije o teh objavljenih dogodkih in pri posodobitvi tabel posodobi prostor pajkanja (pajek pri prvem pajkanju tabele pajka v vseh dokumentih).

Objavljanje dogodkov omogoča razpoložljivost novih in spremenjenih dokumentov za preiskovanje hitreje kot dokumentov, v katerih pajka pajek po razporedu pajka.

Če so nekatere ali vse tabele konfigurirane za uporabo objavljanja dogodkov, lahko pri konfiguriranju pajka podate informacije, ki pajku omogočajo dostop do strežnika WebSphere MQ in preslikav čakalne vrste objavljanja.

Zagotoviti morate, da sta WebSphere MQ in WebSphere Information Integrator Event Publisher Edition konfigurirana na strežniku, na katerem boste pajkali, in da je modul odjemalca WebSphere MQ konfiguriran na strežniku pajkanja. Za uporabo objavljanja dogodkov s pajkom DB2 izpolnite naslednje naloge:

- “Konfiguriranje programa WebSphere MQ za pajke DB2” na strani 50.
- “Konfiguriranje programa WebSphere Information Integrator Event Publisher Edition za pajke DB2” na strani 48.

Pregled konfiguriranja

Če želite izdelati ali spremeniti pajka DB2, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh bazah podatkov na določenem strežniku baz podatkov.
- Podajanje informacij o tipih baz podatkov, v katerih želite pajkati.

Če nameravate pajkati v oddaljenih bazah podatkov, ki niso katalogizirane na lokalnem strežniku baz podatkov, morate pred uporabo pajka DB2 za pajkanje teh baz podatkov na oddaljenem strežniku zagnati skrbniški strežnik DB2. Pri konfiguriranju pajka morate podati tudi gostiteljsko ime in vrata oddaljenega strežnika baz podatkov.

- Podajanje baz podatkov, ki jih želite pajkati.
- Podajanje ID-jev uporabnikov in gesel, ki pajku omogočajo dostop do baz podatkov, ki uporabljajo krmilne elemente za dostop.
- Nastavitev razpored za pajkanje baz podatkov.
- Izbira tabel, ki jih želite pajkati v vsaki bazi podatkov.

Opozorilo: Če želite optimizirati zmogljivost procesov odkrivanja (in preprečiti prekinitve konfiguracijskega postopka pajka), izberite pajkanje vseh tabel samo v primeru, če baza podatkov ne vsebuje veliko tabel ali če tabele ne vsebujejo veliko stolpcev. Če izberete takojšnje pajkanje nekaterih tabel, lahko prostor pajkanja uredite kasneje in v zbirko dodate več tabel.

- Izbira tabel, v katerih bo izvedeno pajkanje, ko bodo posodobitve v njih objavljene v čakalni vrsti za objavljane dogodkov in podajte informacije, ki pajku omogočajo dostop do čakalne vrste za objavljane dogodkov.
- Podajanje možnosti, ki omogočajo preiskovanje stolpcev v določenih tabelah. Tako lahko na primer določene stolpce omogočite za uporabo v parametričnih poizvedbah ali podate, kateri stolpci so lahko vrnjeni v rezultatih iskanja.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali varnostnih žetonih.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2

Če OmniFind Enterprise Edition namestite na računalnik, na katerem se izvaja operacijsko okolje IBM AIX, Linux ali Solaris, morate za konfiguriranje pajka zagnati skript. Skript pajku DB2 omogoča komuniciranje s strežniki baze podatkov. Če uporabljate objavljane dogodkov, skript pajku omogoči tudi dostop do upravljalnikov vrste WebSphere MQ in vrst.

O tej nalogi

Če želite zagotoviti, da bo pajek DB2 lahko pajkal tabele baz podatkov, izvedite namestitveni skript, `escrib2.sh`, ki ga OmniFind Enterprise Edition nudi na strežniku pajkanja.

Preden zaženete skript, morate zagotoviti, da je skrbniški odjemalec DB2 nameščen na strežniku pajkanja.

Če uporabljate objavljane dogodkov, morate na strežniku pajkanja namestiti module WebSphere MQ 5.3 za Java Messaging, tako da bo pajek DB2 lahko dostopal do upravljalnikov vrst WebSphere MQ in do vrst. Ko namestite module WebSphere MQ, morate izvesti namestitveni skript `escrib2.sh`.

Postopek

Konfiguriranje strežnika pajkanja, da bo podpiral pajkanje pajkov DB2:

1. Neobvezno: Če nameravate uporabljati objavljane dogodkov, na strežniku pajkanja namestite module WebSphere MQ 5.3 za Java Messaging:
 - a. Prijavite se kot korenski uporabnik in vnesite naslednji ukaz:


```
export LD_ASSUME_KERNEL=2.4.19
```
 - b. Vstavite zgoščenko WebSphere MQ.
 - c. Spremenite imenik, kjer se nahajajo moduli MQ za Java Messaging.

- d. Vnesite naslednji ukaz, da namestite module:
`rpm -i MQSeriesJava-5.3.0-1.i386.rpm`
2. Na strežniku pajkanja zaženite namestitveni skript za pajka DB2:
 - a. Prijavite se kot skrbnik za iskanje v podjetju (ID uporabnika ste podali ob namestitvi izdelka OmniFind Enterprise Edition).
 - b. Zaženite naslednji skript, ki je nameščen v imeniku `$ES_INSTALL_ROOT/bin` in odgovorite na pozive:
`escrdb2.sh`
3. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:
`esadmin system stopall`
`esadmin system startall`

Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2

Če OmniFind Enterprise Edition namestite na računalnik Microsoft Windows, morate za konfiguriranje strežnika pajkanja zagnati skript. Skript pajku DB2 omogoča komuniciranje s strežniki baze podatkov. Če uporabljate objavljane dogodkov, skript pajku omogoči tudi dostop do upravljalnikov vrste WebSphere MQ in vrst.

O tej nalogi

Če želite zagotoviti, da bo pajek DB2 lahko pajkal tabele baz podatkov, izvedite namestitveni skript, `escrdb2.vbs`, ki ga OmniFind Enterprise Edition nudi na strežniku pajkanja.

Preden zaženete skript, morate zagotoviti, da je skrbniški odjemalec DB2 nameščen na strežniku pajkanja.

Če uporabljate objavljane dogodkov, morate na strežniku pajkanja namestiti module WebSphere MQ 5.3 za Java Messaging, tako da bo pajek DB2 lahko dostopal do upravljalnikov vrst WebSphere MQ in do vrst. Ko namestite module WebSphere MQ, morate izvesti namestitveni skript `escrdb2.vbs`.

Postopek

Konfiguriranje strežnika pajkanja, da bo podpiral pajkanje pajkov DB2:

1. Neobvezno: Če nameravate uporabljati objavljane dogodkov, na strežniku pajkanja namestite module WebSphere MQ 5.3 za Java Messaging:
 - a. Vstavite zgoščenko WebSphere MQ.
 - b. Zaženite namestitveni program WebSphere MQ.
 - c. V oknu Izbira komponent izdelka kot namestitveno možnost izberite **Java Messaging**.
2. Na strežniku pajkanja zaženite namestitveni skript za pajka DB2:
 - a. Prijavite se z ID-jem skrbnika za iskanje v podjetju (ID uporabnika ste podali ob namestitvi izdelka OmniFind Enterprise Edition).
 - b. Zaženite naslednji skript, ki je nameščen v imeniku `%ES_INSTALL_ROOT%\bin` in odgovorite na pozive:
`escrdb2.vbs`
3. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:
`esadmin system stopall`
`esadmin system startall`

Konfiguriranje programa WebSphere Information Integrator Event Publisher Edition za pajke DB2

Preden konfigurirate pajka DB2 za objavljanje dogodkov, zagotovite, da je na strežniku, ki bo pajkan, konfiguriran IBM WebSphere Information Integrator Event Publisher Edition.

O tej nalogi

Pri konfiguriranju programa WebSphere Information Integrator Event Publisher Edition za uporabo s pajkom DB2 upoštevajte naslednje smernice:

- Tako spremenjeni kot tudi nespremenjeni stolpci v izvornih tabelah morajo biti izbrani za objavljanje.
- Izbrisane vrstice v izvornih tabelah morajo biti izbrane za objavljanje.
- Čakalne vrste za objavljanje dogodkov ne more souporabljati več baz podatkov.
- Ena baza podatkov ima lahko več preslikav čakalne vrste in čakalnih vrst.
- Tabela mora imeti povezano eno publikacijo XML z eno preslikavo objavitvene čakalne vrste. (Tabela ne sme imeti povezane več kot ene publikacije XML z eno preslikavo objavitvene čakalne vrste. Tabela ima lahko več kot eno publikacijo XML, če je vsaka publikacija XML povezana z drugo preslikavo objavitvene čakalne vrste.)

Postopek

Če želite pajku DB2 omogočiti dostop do posodobitev tabele, ki so objavljene v čakalni vrsti za objavljanje dogodkov, izvedite naslednje korake, s katerimi boste konfigurirali strežnik baz podatkov. (Za pomoč si oglejte dokumentacijo za WebSphere Information Integrator Publisher Edition .)

1. Namestite WebSphere Information Integrator Event Publisher Edition na strežnik baze podatkov, ki bo pajkan.
2. Zaženite Lansirnik centra za repliciranje:

AIX, Linux ali Solaris
db2rc

Ukazni poziv Windows

Kliknite **Start** → **IBM DB2 Replication Center**.

3. Izdelava nadzornih tabel Q Capture:
 - a. Kot pogled lansirnika izberite **Objavljanje dogodkov**, nato izberite **Izdelava nadzornih tabel Q Capture** in kliknite **Naprej**.
 - b. V polju **Strežnik Q Capture** s seznama strežnikov baz podatkov, ki so na voljo, izberite strežnik, ki ga želite uporabljati kot strežnik Q Capture, in kliknite **Potrdi**.
 - c. Podajte ID uporabnika in geslo, ki ga tako pooblaščate za dostop do izbranega strežnika Q Capture. Spremenite shemo Q Capture ali sprejmite privzeto ime sheme in kliknite **Naprej**.
 - d. Podajte imena upravljalnika čakalne vrste in skrbniške čakalne vrste ter zaženite čakalno vrsto, ki ste jo podali, ko ste konfigurirali WebSphere MQ na tem strežniku baz podatkov in kliknite **Naprej**.
 - e. Kliknite **Dokončaj**. Ko se prikaže stran s sporočili in skripti SQL, kliknite **Zapri**.
 - f. Kot možnosti obdelave izberite **Zaženi zdaj** in kliknite **Potrdi**. Ko se prikaže sporočilo, ki kaže, da so skripti SQL dokončani, kliknite **Zapri**.
4. Izdelajte publikacijo XML:
 - a. V Lansirniku centra za repliciranje kot pogled lansirnika izberite **Objavljanje dogodkov**, nato izberite **Izdelava publikacije XML** in kliknite **Naprej**.
 - b. Na strani Zagon kliknite **Dokončaj**.

- c. Na strani Strežnik in preslikava čakalne vrste potrdite, da sta strežnik in shema Q Capture pravilna, nato kliknite možnost poleg polja **Preslikava objavitvene čakalne vrste** in kliknite **Nova**, da izdelate preslikavo objavitvene čakalne vrste.
 - d. Na strani Splošno vnesite ime za preslikavo čakalne vrste.
 - e. Na strani Lastnosti podajte ime pošiljalske čakalne vrste (kot je ime podatkovne čakalne vrste, ki ste jo podali med konfiguriranjem WebSphere MQ na tem strežniku), za vrsto vsebine sporočila izberite ali **Operacija v vrstici** ali **Transakcija**, počistite potrditvena polja za pošiljanje sporočil o stanju sistema in dodajanje oglavij sporočil JMS ter kliknite **Potrdi**.
 - f. Ko se prikaže stran s sporočili in skripti SQL, kliknite **Zapri**.
 - g. Kot možnosti obdelave izberite **Zaženi zdaj** in kliknite **Potrdi**. Ko se prikaže sporočilo, ki kaže, da so skripti SQL dokončani, kliknite **Zapri**.
 - h. Na strani Izbira preslikave objavitvene čakalne vrste izberite preslikavo čakalne vrste, ki ste jo izdelali in kliknite **Potrdi**.
 - i. Na strani Strežnik in preslikava čakalne vrste potrdite, da je ime preslikave čakalne vrste pravilno in kliknite **Naprej**.
 - j. Na strani Izvorna tabela kliknite **Dodaj**, nato **Prikliči vse**, izberite tabelo, ki jo želite omogočiti za objavljanje dogodkov ter kliknite **Potrdi** in **Naprej**.
 - k. Na strani Stolpci in vrstice izberite stolpce, za katere želite, da jih bo pajek DB2 pajkal (ali vse stolpce) in izberite ključne stolpce. Na strani, kjer izberete vrstice za pajkanje (ali vse vrstice), izberite možnost za objavljanje izbranih delov izvorne tabele. Ko končate s konfiguriranjem teh možnosti, kliknite **Naprej**.
 - l. Na strani Vsebina sporočila za podatke iz stolpcev izberite možnost za vključitev tako spremenjenih kot tudi nespremenjenih stolpcev ter izberite možnost za samo nove vrednosti podatkov. Prepričajte se, da je potrditveno polje za samodejen zagon publikacij XML izbrano in kliknite **Naprej**.
 - m. Na strani Pregled in dokončevanje publikacij XML kliknite **Naprej**.
 - n. Na strani Povzetek kliknite **Dokončaj**. Ko se prikaže stran s sporočili in skripti SQL, kliknite **Zapri**.
 - o. Kot možnosti obdelave izberite **Zaženi zdaj** in kliknite **Potrdi**. Ko se prikaže sporočilo, ki kaže, da so skripti SQL dokončani, kliknite **Zapri**.
5. Zaženite strežnik Q Capture:
- a. Zaprite Lansirnik centra za repliciranje in zaženite Center za repliciranje.
 - b. V drevesu objekta kliknite **Q Replication** → **Definitions** → **Q Capture Servers**.
 - c. Z desno miškino tipko kliknite ikono strežnika Q Capture, ki ste ga konfigurirali, in izberite **Omogočanje baze podatkov za Q Replication**.
 - d. Ko se prikaže opozorilno sporočilo, kliknite **Potrdi**.
 - e. Ko se prikaže stran s sporočili DB2, kliknite **Zapri**.
 - f. V drevesu objektov z desno miškino tipko kliknite ikono strežnika Q Capture in izberite **Zaženi program Q Capture**.
 - g. Kot možnost obdelave izberite **Zaženi zdaj**, podajte ime sistema, ID uporabnika in geslo za uporabnika DB2, pot do imenika, kjer so shranjeni dnevniki in ime primerka DB2 ter kliknite **Potrdi**.
 - h. Ko se prikaže sporočilo, da je bila zahteva poslana, kliknite **Zapri**.
 - i. V drevesu objektov z desno miškino tipko kliknite ikono strežnika Q Capture in izberite **Preveri status**.
- Prikaže se status strežnika Q Capture. Če je prišlo do napak, sporočilo o statusu pravi, da strežnik domnevno ne deluje. Če želite pregledati dnevnik in ugotoviti vzrok mogočih napak, v ukazno vrstico vnesite naslednji ukaz:
- ```
asnqcap Capture_Server=capture server name LOGSTDOUT=y
```

## Konfiguriranje programa WebSphere MQ za pajke DB2

Preden konfigurirate pajka DB2 za objavljanje dogodkov, zagotovite, da je na strežniku, ki ga bo pajek poslušal, konfiguriran IBM WebSphere MQ.

### Preden začnete

Prepričajte se, da sta na ciljnem strežniku baz podatkov nameščeni izdaji izdelka DB2 WebSphere Information Integrator Event Publisher Edition in WebSphere MQ.

### Omejitve

Če je ciljni strežnik baz podatkov nameščen v računalniku Linux, morajo vsi uporabniki DB2, uporabniki WebSphere MQ in uporabniki OmniFind Enterprise Edition nastaviti naslednjo spremenljivko okolja:

```
export LD_ASSUME_KERNEL=2.4.19
```

Ta spremenljivka okolja omogoča izvoz izvedb nitenja LinuxThread iz vsake lupine, kjer se izvaja namestitev, izdajajo ukazi WebSphere MQ ali izvajajo aplikacije WebSphere MQ. WebSphere MQ zahteva izvoz teh spremenljivk okolja.

Pajki DB2, ki uporabljajo vzpostavljane zveze objavljanja dogodkov za povezavo s čakalnimi vrstami WebSphere MQ z odjemalsko povezavo. Povezave z odjemalcem vzpostavite tako, da se prijavite kot skrbnik WebSphere MQ, nato pa z naslednjim ukazom nastavite CCSID upravljalnika čakalnih vrst na 819:

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

### O tej nalogi

Pajek DB2 podpira način za odjemalsko povezavo s strežnikom WebSphere MQ. Pajek posluša sporočila XML, ki so objavljena v čakalni vrsti za objavljanje dogodkov. Pajek ne more poslušati sporočil XML, ki se prenašajo preko več čakalnih vrst.

Ko konfigurirate WebSphere MQ, za pridobivanje sporočil XML iz objaviteljske čakalne vrste pajek DB2 uporablja ime upravljalnika čakalne vrste, ime čakalne vrste, ime gostitelja strežnika in ime kanala strežnika. Pajek razčleni sporočila in posodobi prostor pajkanja z informacijami o posodobljenih tabelah.

### Postopek

Če želite pajku DB2 omogočiti poslušanje čakalne vrste za objavljanje dogodkov, izvedite naslednje korake, s katerimi boste konfigurirali strežnik baz podatkov. (Za pomoč si oglejte dokumentacijo za WebSphere MQ.)

1. Prijavite se v vlogi skrbnika WebSphere MQ in vnesite naslednje ukaze, s katerimi boste izdelali upravljalnika čakalne vrste in čakalne vrste.
  - a. V ukazno vrstico vnesite naslednji ukaz:  
crtmqm QM1
  - b. Ko se prikaže sporočilo Namestitev je končana, vnesite naslednji ukaz:  
strmqm QM1
  - c. Ko se prikaže sporočilo 'QM1' je zagnan, vnesite naslednji ukaz:  
runmqsc QM1

- d. Ko se prikaže sporočilo Zagon MQSC za upravljalnika čakalne vrste QM1 vnesite naslednji ukaz, s katerim izdelate skrbniško čakalno vrsto:
 

```
DEFINE QLOCAL('ASN.QM1.ADMINQ')
```
  - e. Ko se prikaže sporočilo Izdelana je bila čakalna vrsta WebSphere MQ, vnesite naslednji ukaz, s katerim izdelate čakalno vrsto za vnovičen zagon:
 

```
DEFINE QLOCAL(' ASN.QM1.RESTARTQ')
```
  - f. Ko se ponovno prikaže sporočilo Izdelana je bila čakalna vrsta WebSphere MQ, vnesite naslednji ukaz, s katerim izdelate podatkovno čakalno vrsto:
 

```
DEFINE QLOCAL(' ASN.QM1.DATAQ')
```
  - g. Ko se ponovno prikaže sporočilo Izdelana je bila čakalna vrsta WebSphere MQ, vnesite naslednji ukaz za izhod:
 

```
end
```
2. Vnesite naslednji ukaz, s katerim zaženete Poslušalnik MQ na strežniku baz podatkov (Poslušalnik MQ se mora izvajati, ko izdelate pajka DB2, ki uporablja objavljane dogodkov). V tem primeru je 1414 številka vrat strežnika, uporabi pa se privzeti kanal SYSTEM.DEF.SVRCONN:
 

```
runmqtsr -m QM1 -t TCP -p 1414 &
```
  3. Vnesite naslednje ukaze, s katerimi uporabnika DB2 pooblastite za dostop do upravljalnika čakalne vrste in čakalnih vrst skozi vmesnik MQI (Message Queuing Interface) za objavljane dogodkov (v tem primeru je ID uporabnika 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. Vnesite naslednje ukaze za ID uporabnika, ki se uporablja za izdelavo in izvajanje pajka DB2 z objavljanim dogodkov. Ti ukazi ID uporabnika pooblastijo za dostop do upravljalnika čakalne vrste in do čakalnih vrst preko Vmesnika za postavljanje sporočil v čakalno vrsto (MQI) za objavljane dogodkov. V tem primeru je ID uporabnika 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
```

## Pajkanje baz podatkov DB2 na klasičnem strežniku vira podatkov

Pajek DB2 lahko bazo podatkov DB2 na klasičnem strežniku vira podatkov pajka preko programa WebSphere Information Integrator Classic Federation.

### O tej nalogi

Pajkanje baze podatkov DB2 na klasičnem strežniku vira podatkov je mogoče samo, če je baza podatkov federirana z bazo podatkov DB2 na neklasičnem strežniku vira podatkov z uporabo ovojnice ODBC programa WebSphere Information Integrator Classic Federation.

### Postopek

Federiranje baze podatkov (povezovanje v zvezo):

1. Namestite WebSphere Information Integrator Classic Federation na klasičen strežnik vira podatkov.
2. Namestite odjemalski modul WebSphere Information Integrator Classic Federation na neklasičen strežnik vira podatkov, ki ima bazo podatkov, ki bo federirana z bazo podatkov na klasičnem strežniku vira podatkov.

3. Konfigurirajte gonilnik ODBC za WebSphere Information Integrator Classic Federation za povezavo s klasičnim strežniku vira podatkov.
4. Povežite se z zvezno bazo podatkov in izdelajte ovojnico ODBC za federiranje z bazo podatkov na klasičnem strežniku vira podatkov.
5. Konfigurirajte pajka DB2 tako, da bo pajkal zvezno bazo podatkov na neklasičnem strežniku vira podatkov. To omogoča pajkanje baze podatkov na klasičnem strežniku vira podatkov preko programa WebSphere Information Integrator Classic Federation.

---

## Pajki DB2 Content Manager

Če želite v zbirko iskanja v podjetju vključiti tipe postavk IBM DB2 Content Manager, morate konfigurirati pajka DB2 Content Manager.

### Konfiguriranje strežnika pajkanja

Preden lahko izvedete pajkanje na strežniku DB2 Content Manager, morate na strežniku pajkanja zagnati skript. Ta skript, ki je na voljo z OmniFind Enterprise Edition, omogoča pajku DB2 Content Manager komuniciranje s strežniki DB2 Content Manager.

Preden za konfiguriranje pajka DB2 Content Manager uporabite skrbniško konzolo iskanja v podjetju, izpolnite nalogo, ki ustreza vašemu okolju:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2 Content Manager” na strani 53.
- “Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2 Content Manager” na strani 54.

#### Nasvet:

Podrobnejše primere z načini konfiguriranja zaščitene pajka DB2 Content Manager najdete v scenarijih za večjo organizacijo v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

### Pregled konfiguriranja

Pajka DB2 Content Manager lahko uporabite za pajkanje poljubnega števila strežnikov DB2 Content Manager. Pri konfiguriranju pajka podate možnosti, ki določajo, kako bo pajek pajkal na vseh strežnikih DB2 Content Manager v prostoru za pajkanje. Izberete tudi specifične tipe postavk, ki jih želite pajkati na posameznih strežnikih.

Če želite izdelati ali spremeniti pajka DB2 Content Manager, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh tipih postavk na vseh strežnikih DB2 Content Manager v prostoru za pajkanje.
- Izbira strežnikov DB2 Content Manager, na katerih želite pajkati.
- Podajanje ID-je uporabnikov in gesel, ki pajku omogočajo dostop do vsebine na strežnikih DB2 Content Manager.
- Nastavitev razporeda za pajkanje strežnikov.
- Izbira tipa postavk, ki jih želite pajkati na posameznem strežniku DB2 Content Manager.

- Podajanje možnosti, ki omogočajo preiskovanje atributov v nekaterih tipih postavk. Tako lahko na primer iz prostora za pajkanje izključite določene tipe dokumentov in podate, kateri atributi so lahko vrnjeni v rezultatih iskanja.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi seznammi za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

## Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2 Content Manager

Če OmniFind Enterprise Edition namestite na računalnik, na katerem se izvaja operacijsko okolje AIX, Linux ali Solaris, morate za konfiguriranje pajka zagnati skript. Skript pajku DB2 Content Manager omogoča komuniciranje s strežniki IBM DB2 Content Manager.

### O tej nalogi

Pajek DB2 Content Manager uporablja spojnik Java za DB2 Content Manager različice 8 za dostop do strežnikov DB2 Content Manager. Spojnik namestite tako, da na strežnik pajkanja namestite enega od naslednjih izdelkov:

- IBM DB2 Information Integrator for Content različice 8.3 za AIX, Solaris ali Linux
- IBM DB2 Information Integrator for Content različice 8.2 za AIX ali Solaris
- IBM DB2 Content Manager Toolkit različice 8.2 za Linux

Izvedite namestitveni skript, če želite zagotoviti, da bo pajek programa DB2 Content Manager lahko deloval s programom DB2 Content Manager, ki ga OmniFind Enterprise Edition nudi na strežniku pajkanja potem, ko ste namestili konektor.

### Postopek

Konfiguriranje strežnika pajkanja, da lahko pajka strežnike DB2 Content Manager:

1. Na strežnik pajkanja namestite konektor Java za DB2 Content Manager različice 8:
  - a. Na strežnik pajkanja se prijavite kot korenski uporabnik:  
su - root
  - b. Zaženite datoteko db2profile. Primer:  
./home/db2inst/sqllib/db2profile
  - c. Izvozite spremenljivko okolja JAVAHOME. Primer:  
export JAVAHOME=/usr/IBMJava2-141
  - d. Dodajte imenik Java okoljski spremenljivki PATH:  
export PATH=\$PATH:\$JAVAHOME/bin
  - e. Vstavite namestitveno zgoščenko DB2 Information Integrator for Content in zaženite namestitveni čarovnik.
  - f. V oknu Izbira komponente izvedite naslednja dejanja: (Če uporabljate izdelek DB2 Information Integrator for Content različice 8.3, boste ob izvedbi namestitve po meri zagledali okno za izbiro komponente.)

- 1) Na seznamu **Komponente** izberite **Lokalni konektorji** in nato **Konektor Content Manager V8** s seznama **Podkomponente**.
- 2) Na seznamu **Komponente** izberite **Komplet orodij in vzorci za konektor** in nato **Konektor Content Manager V8** s seznama **Podkomponente**.
- g. Podajte ime baze podatkov, uporabniško ime in geslo za knjižnico DB2 Content Manager in sprejmite privzete nastavitve za preostala okna.
2. Na strežniku pajkanja se prijavite z ID-jem uporabnika, ki se nahaja v skrbniški skupini DB2.
3. Katalogizirajte oddaljeno bazo podatkov strežnika knjižnice programa DB2 Content Manager in preverite, da se strežnik pajkanja lahko poveže s strežnikom DB2 Content Manager:

```
db2 catalog tcpip node ime_vozlišča remote gostiteljsko_ime server vrata
db2 catalog database ime_baze_podatkov as vzdevek at node ime_vozlišča
```

Kjer je:

*ime\_vozlišča*

Kratko gostiteljsko ime strežnika DB2 Content Manager (na primer ibmes).

*gostiteljsko\_ime*

Popolnoma kvalificirano gostiteljsko ime strežnika DB2 Content Manager (na primer ibmes.ibm.com).

*vrata*

Številka vrat strežnika DB2 Content Manager.

*ime\_baze\_podatkov*

Ime baze podatkov DB2 Content Manager (na primer ICMNLSDB).

*vzdevek*

Vzdevek baze podatkov DB2 Content Manager (na primer CMSVR)

4. Neobvezno: Prijavite se kot korenski uporabnik in preizkusite povezavo baze podatkov:

```
. Information_Integrator_for_Content_install_directory/bin/cmbenv81.sh
cd Information_Integrator_for_Content_install_directory/samples/java/icm
javac *.java
java SConnectDisconnectICM ICMdatabase_name CAdmin_ID CAdmin_password
```

5. Na strežniku pajkanja zaženite namestitveni skript za pajka DB2 Content Manager:

- a. Premaknite se v imenik ES\_INSTALL\_ROOT/bin:

```
cd $ES_INSTALL_ROOT/bin
```

- b. Zaženite naslednji skript in se odgovorite na pozive:

```
escrm.sh
```

6. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:

```
esadmin system stopall
```

```
esadmin system startall
```

## Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2 Content Manager

Če OmniFind Enterprise Edition namestite na računalnik Microsoft Windows, morate za konfiguriranje strežnika pajkanja zagnati skript. Skript pajku DB2 Content Manager omogoča komuniciranje s strežniki programa IBM DB2 Content Manager.

### O tej nalogi



Pajek DB2 Content Manager uporablja spojnik Java za DB2 Content Manager različice 8 za dostop do strežnikov DB2 Content Manager. Spojnik namestite tako, da na strežnik pajkanja namestite IBM DB2 Information Integrator for Content različice 8.2 ali različice 8.3 za Windows. Izvedite namestitveni skript, če želite zagotoviti, da bo pajek programa DB2 Content Manager lahko deloval s programom DB2 Content Manager, ki ga OmniFind Enterprise Edition nudi na strežniku pajkanja potem, ko ste namestili konektor.

## Postopek

Konfiguriranje strežnika pajkanja, da lahko pajka strežnike DB2 Content Manager:

1. Na strežnik pajkanja namestite konektor Java za DB2 Content Manager različice 8:
  - a. Vstavite namestitveno zgoščenko DB2 Information Integrator for Content. Namestitveni program se zažene samodejno. Odpre se namestitveni čarovnik za DB2 Content Manager Enterprise Information Portal.
  - b. V oknu Izbira komponente izvedite naslednja dejanja: (Če uporabljate izdelek DB2 Information Integrator for Content različice 8.3, boste ob izvedbi namestitve po meri zagledali okno za izbiro komponente.)
    - 1) Na seznamu **Komponente** izberite **Lokalni konektorji** in nato **Konektor Content Manager V8** s seznama **Podkomponente**.
    - 2) Na seznamu **Komponente** izberite **Komplet orodij in vzorci za konektor** in nato **Konektor Content Manager V8** s seznama **Podkomponente**.
  - c. Podajte ime baze podatkov, uporabniško ime in geslo za knjižnico DB2 Content Manager in sprejmite privzete nastavitve za preostala okna.
2. Katalogizirajte oddaljeno bazo podatkov strežnika knjižnice programa DB2 Content Manager in preverite, da se strežnik pajkanja lahko poveže s strežnikom DB2 Content Manager. V ukazni poziv na strežniku pajkanja vnesite naslednje ukaze:

```
db2 catalog tcpip node ime_vozlišča remote gostiteljsko_ime server vrata
db2 catalog database ime_baze_podatkov as vzdevek at node ime_vozlišča
```

Kjer je:

*node\_name*

Kratko gostiteljsko ime strežnika DB2 Content Manager (na primer *ibmes*).

*hostname*

Popolnoma kvalificirano gostiteljsko ime strežnika DB2 Content Manager (na primer *ibmes.ibm.com*).

*port*

Številka vrat strežnika DB2 Content Manager.

*database\_name*

Ime baze podatkov DB2 Content Manager (na primer *ICMNLSDDB*).

*alias*

Vzdevek baze podatkov DB2 Content Manager (na primer *CMSVR*)

3. Neobvezno: Preverite povezavo baze podatkov tako, da odprete ukazni poziv in vnesete naslednje ukaze:

```
cmbenv81.bat
```

```
cd Information_Integrator_for_Content_install_directory\samples\java\icm
```

```
javac *.java
```

```
java SConnectDisconnectICM ICMdatabase_name CAdmin_ID CAdmin_password
```

4. Na strežniku pajkanja zaženite namestitveni skript za pajka DB2 Content Manager:

- a. Premaknite se v imenik ES\_INSTALL\_ROOT/bin:  
`cd %ES_INSTALL_ROOT%\bin`
  - b. Zaženite naslednji skript in se odgovorite na pozive:  
`escrm.vbs`
5. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:
- ```
esadmin system stopall  
esadmin system startall
```

Pajki Domino Document Manager

Če želite v zbirko iskanja v podjetju vključiti knjižnice in kabinete Domino Document Manager, morate konfigurirati pajka Domino Document Manager.

Konfiguriranje strežnika pajkanja

Če uporablja strežnik Domino Document Manager, na katerem boste pajkali, protokol NRPC (Notes remote procedure call), morate na strežniku pajkanja zagnati namestitveni skript. Ta skript, ki je na voljo z OmniFind Enterprise Edition, omogoča pajku Domino Document Manager komuniciranje s strežniki, ki uporabljajo NRPC.

Če uporablja strežnik Domino Document Manager, na katerem boste pajkali, protokol DIIOP (Domino Internet Inter-ORB Protocol), na strežniku pajkanja ni potrebno zagnati namestitvenega skripta, vendar morate strežnik Domino Document Manager konfigurirati tako, da lahko pajek Domino Document Manager dostopa do strežnika.

Pomembno: Če strežnik Domino Document Manager uporablja DIIOP in konfigurirate pajka za uporabo HTTPS ali DIIOP prek SSL, tako da so prenosi med pajkom in strežnikom šifrirani, morate prekopirati datoteko `TrustedCerts.class` (na primer `c:\certs` ali `/data/certs`) s strežnika Domino Document Manager na strežnik pajkanja. V konfiguraciji z dvema ali štirimi strežniki morate prekopirati tudi datoteko `TrustedCerts.class` na strežnike, na katerih je nameščena iskalna komponenta. Prepričajte se, da se datoteka na strežniku pajkanja nahaja na enakem mestu kot na iskalnih strežnikih. Imeniško pot do datoteke `TrustedCerts.class` podate med konfiguriranjem pajka.

Če ste v sistem IBM AIX namestili OmniFind Enterprise Edition, morate zagotoviti, da je na strežniku pajkanja nameščen in razpoložljiv modul V/I zaključnih vrat.

Preden za konfiguriranje pajka Domino Document Manager uporabite skrbniško konzolo iskanja v podjetju, izpolnite naloge, ki ustrezajo vašemu okolju:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino” na strani 69.
- “Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino” na strani 71.
- “Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP” na strani 72.
- “Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino” na strani 73.

Zaščita na ravni dokumenta

Če je omogočena zaščita zbirke in strežnik, na katerem boste pajkali, uporablja protokol NRCP, morate na strežniku pajkanja konfigurirati overjeni strežnik Lotus Domino. Overjeni strežnik se uporablja za uveljavitev krmilnih elementov dostopa na ravni dokumentov. Preden omogočite zbirko za iskanje, izpolnite naslednje naloge:

- Konfigurirajte overjene strežnike Lotus Domino za prevetjanje poverilnic uporabnikov.

- Na aplikacijskem strežniku WebSphere Application Server omogočite globalno zaščito in konfigurirajte iskalno aplikacijo za uporabo zaščite. S tem korakom zagotovite, da bodo prijavnice poverilnice preverjene, ko bodo uporabniki poskusili uporabiti iskalno aplikacijo. Iskalni strežniki uporabljajo poverilnice za preverjanje pooblastila vsakega uporabnika za dostopanje do dokumentov Lotus Domino.

Pregled konfiguriranja

Pajka Domino Document Manager lahko uporabite za pajkanje poljubnega števila knjižnic Domino Document Manager. Pri izdelavi pajka izberete knjižnice za pajkanje z enega strežnika Domino Document Manager. Pozneje lahko pri urejanju prostora pajkanja dodate dokumente z drugega strežnika Domino Document Manager, ki jih želite vključiti v isti prostor pajkanja. Pri izdelavi ali urejanju pajka podate, ali želite izvesti pajkanje v vseh kabinetih v knjižnicah, ki jih izberete za pajkanje ali želite pajkati samo v določenih kabinetih.

Če želite izdelati ali spremeniti pajka Domino Document Manager, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh dokumentih v prostoru za pajkanje.
- Identificiranje strežnika Domino Document Manager in komunikacijskega protokola.
- Če je strežnik konfiguriran za uporabo protokola DIIOP, lahko podate, kako se bo pajek povezal z objekti Domino. Tako lahko na primer podate možnosti za uporabo HTTPS ali plasti zaščitenih vtičnic (SSL) za šifriranje komunikacij.
- Izbira knjižnic, v katerih želite pajkati.
- Nastavljanje razporeda za pajkanje knjižnic.
- Izbira dokumentov, v katerih želite pajkati. Pajek lahko pajka v vseh kabinetih v knjižnici ali pa samo v dokumentih v kabinetih, ki jih izberete.
- Podajanje možnosti, ki omogočajo preiskovanje polj v različnih knjižnicah in kabinetih. Tako lahko na primer iz prostora pajkanja izključite določena polja in podate možnosti za preiskovanje prilog.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi seznamami za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezani pojmi

“Nasveti za pajkanje baz podatkov Lotus Domino” na strani 68

“Zagotavljanje zaščite na ravni dokumenta za dokumente Lotus Domino” na strani 256

S tem povezana opravila

“Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino” na strani 69

“Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino” na strani 71

“Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP” na strani 72

“Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino” na strani 73

“Konfiguriranje overjenih strežnikov Lotus Domino za preverjanje poverilnic uporabnikov” na strani 257

Pajki strežnika Exchange Server

Če želite v zbirko iskanja v podjetju vključiti javne mape strežnika Microsoft Exchange Server, morate konfigurirati pajka Exchange Server.

Pajka Exchange Server lahko uporabite za pajkanje poljubnega števila map in podmap na strežnikih javnih map Exchange Server. Pri izdelavi pajka izberete vsebino, ki jo želite pajkati na strežniku javnih map. Pozneje lahko prostor pajkanja uredite in mu dodate vsebino z drugega strežnika javnih map.

Če želite izdelati ali spremeniti pajka Exchange Server, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh podmapah na vseh strežnikih v prostoru za pajkanje.
- Podajanje informacij o strežniku javnih map Exchange Server, na katerem želite pajkati. Podati morate ID uporabnika in geslo, da bo pajek lahko dostopal do vsebine na strežniku. Če uporablja strežnik protokol SSL (plast zaščitene vtičnice), lahko podate možnosti, ki pajku omogočajo dostopanje do datoteke skladišča ključev na strežniku pajkanja.
- Nastavitev razporeda za pajkanje strežnika javnih map.
- Izbira map in podmap za pajkanje.
- Podajanje možnosti, ki omogočajo preiskovanje dokumentov v podmapah. Tako lahko na primer iz prostora pajkanja izključite določene tipe dokumentov.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali varnostnih žetonih.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezana opravila

“Preverjanje dostopa do zaščitene dokumentov strežnika Exchange Server” na strani 256

Pajki baze podatkov JDBC

Pajka baze podatkov JDBC uporabite za vključitev baz podatkov, do katerih lahko dostopate s protokolom JDBC (Java Database Connectivity), v zbirko iskanja v podjetju.

Za vsak sistem baz podatkov, ki ga želite pajkati, morate konfigurirati ločenega pajka. Pri izdelavi pajka podate možnosti za pajkanje ene baze podatkov. Pozneje lahko v prostor pajkanja dodate baze podatkov istega tipa.

Vsaka vrstica v tabeli baze podatkov je obravnavana kot dokument, vrednosti stolpcev baze podatkov pa so razčlenjene in indeksirane kot polja z možnostjo iskanja. Pajka lahko konfigurirate za pajkanje več strukturiranih tabel tako, da pri konfiguriranju lastnosti pajka povežete vtičnik s pajkom. S tem vtičnikom se lahko vrstice iz več tabel v relacijski bazi

podatkov, ki imajo enaka polja ključev, združijo (join) in obravnavajo kot en sam dokument. Če uporabnik preišče bazo podatkov, se podatki iz združenih tabel ob prikazu dokumenta v rezultatih iskanja prikažejo kot dodatna polja.

Podprti sistemi in gonilniki baz podatkov

Za uporabo protokola JDBC za pajkanje tabel v bazi podatkov mora na strežniku pajkanja obstajati ustrezen gonilnik JDBC. Pajek baze podatkov JDBC podpira naslednje sisteme baze podatkov in gonilnike JDBC tipa 4:

Sistem baze podatkov	Ime gonilnika JDBC tipa 4	Poti razredov standardnega gonilnika JDBC
IBM DB2 Universal Database različice 8.2 in IBM DB2 Enterprise Server Edition različice 9.1 za Linux, UNIX in Windows	com.ibm.db2.jcc.DB2Driver	<i>koren_namestivte_db2</i> /java/db2jcc.jar <i>koren_namestivte_db2</i> /java/db2jcc_license_cu.jar
Oracle 9i in 10g	Oracle.jdbc.driver.OracleDriver	<i>domač_imenik_oracle</i> /jdbc/lib/ojdbc14.jar
Microsoft SQL Server 2000	com.microsoft.jdbc.sqlserver.SQLServerDriver	<i>osnova_mssql_jdbc</i> /lib/mssqlserver.jar <i>osnova_mssql_jdbc</i> /lib/msbase.jar <i>osnova_mssql_jdbc</i> /lib/msutil.jar
Microsoft SQL Server 2005	com.microsoft.sqlserver.jdbc.SQLServerDriver (gonilnik JDBC za SQL Server 2005 ni podprt v sistemih AIX.)	<i>namestivteni_imenik</i> /sqljdbc_1.0/loc/sqljdbc.jar kjer predstavlja <i>loc</i> vaše področne nastavitve, kot je na primer <i>namestivteni_imenik</i> /sqljdbc_1.0/enu/sqljdbc.jar

Pajki baze podatkov JDBC v primerjavi s pajki DB2

Če trenutno uporabljate pajka DB2, ga boste morda želeli uporabljati še naprej. Podatkov, shranjenih za pajka DB2, ne morete preseliti v pajka baze podatkov JDBC.

V naslednjih primerih uporabite namesto pajka baze podatkov JDBC pajka DB2:

- Če želite pajkati baze podatkov DB2 z gonilnikom JDBC tipa 2.
- Če želite pajkati baze podatkov Oracle in SQL Server, ki so zvezane z bazo podatkov DB2. S pajkom DB2 lahko dostopite do vseh teh tipov baz podatkov prek nadomestnega imena.
- Če želite pajkati baze podatkov DB2 for z/OS, DB2 for iSeries, Informix, Sybase, VSAM, IMS, CA-Datacom ali baze podatkov Software AG Adabas. Povežite te tipe baz podatkov v zvezo z bazo podatkov DB2 in jih obiščite s pajkom DB2 prek vzdevka.
- Za posodobitev indeksa iskanja v podjetju pri objavi posodobitev v bazi podatkov želite uporabiti objavljanje dogodkov.

Pregled konfiguriranja

Če želite izdelati ali spremeniti pajka baze podatkov JDBC, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh bazah podatkov v prostoru za pajkanje.
- Podajanje tipa baze podatkov, v kateri želite pajkati.

- Izbira baze podatkov, v kateri želite pajkati in po potrebi podajte še ID uporabnika in geslo, ki pajku omogočata dostopanje do baze podatkov.
- Nastavitev razporeda za pajkanje baze podatkov.
- Izbira tabel, v katerih želite pajkati.

Opozorilo: Če želite optimizirati učinkovitost procesov odkrivanja in preprečiti pretek čakalnega časa za postopek konfiguriranje pajka, izberite možnost pajkanja vseh tabel samo, če baza podatkov ne vsebuje več tabel ali če tabele ne vsebujejo več stolpcev. Če izberete takojšnje pajkanje nekaterih tabel, lahko prostor pajkanja uredite kasneje in v zbirko dodate več tabel.

- Podajanje možnosti, ki omogočajo preiskovanje stolpcev v določenih tabelah. Tako lahko na primer določene stolpce omogočite za uporabo v parametričnih poizvedbah ali podate, kateri stolpci so lahko vrnjeni v rezultatih iskanja.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali varnostnih žetonih.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

Preslikave razmerij za baze podatkov JDBC

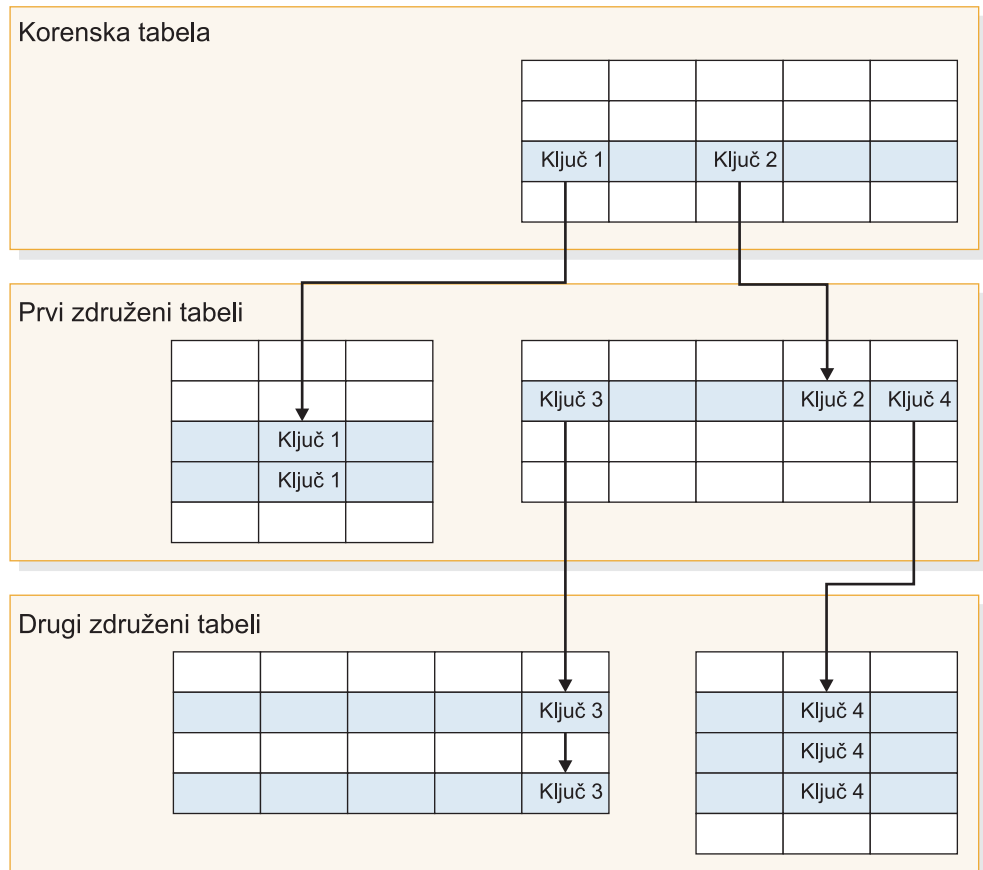
Ko izdelate pravila za vtičnik, ki pajka več strukturiranih tabel baze podatkov JDBC, podate informacije o korenski tabeli in načinu združitve nadrejenih in podrejenih tabel.

Vtičnik, ki je na voljo za iskanje v podjetju, pajku baze podatkov JDBC omogoča pridružitve v več strukturiranih tabel. Vtičnik izdelate tako, da v datoteki `ES_INSTALL_ROOT/default_config/crawler_rdb_plugin.xml` podate pravila. Po konfiguriranju pajka za uporabo vtičnika se vrstice iz tabel, ki imajo enaka ključna polja, združijo (jion) in obravnavajo kot en sam dokument. Če uporabnik preišče bazo podatkov, se podatki iz združenih tabel ob prikazu dokumenta v rezultatih iskanja prikažejo kot dodatna polja.

Združevanje tabel po stolpcih ključev

Naslednja slika prikazuje, kako je izdelana preslikava razmerij za več tabel. Pajek baze podatkov JDBC pregleda korensko tabelo v bazi podatkov. Nekateri od stolpcev tabele so polja ključev, ki jih je mogoče uporabiti za združevanje tabele z drugimi tabelami. Stolpce v združenih tabelah je mogoče uporabiti kot ključe za združitev dodatnih tabel. Vrstice v več tabelah so obravnavane kot en sam dokument v prostoru pajkanja. Vrednosti stolpcev so obravnavane kot metapodatki dokumentov. Korenska tabela je *nadrejeni element* v razmerju, tabela, pridružena na prvi ravni pa je *podrejeni element*. Podrejene tabele na prvi ravni so lahko nadrejeni elementi tabelam, ki so pridružene na drugi ravni.

V tem primeru sta stolpca Ključ 1 in Ključ 2 v korenski (nadrejeni) tabeli polji ključev, ki omogočata, da tabeli pridružite dodatne tabele, ki prav tako vsebujejo stolpca Ključ 1 in Ključ 2. Ena od tabel, pridruženih na prvi ravni, vsebuje polja ključev, Ključ 3 in Ključ 4, ki omogočata združitev tabele z dodatnimi tabelami.

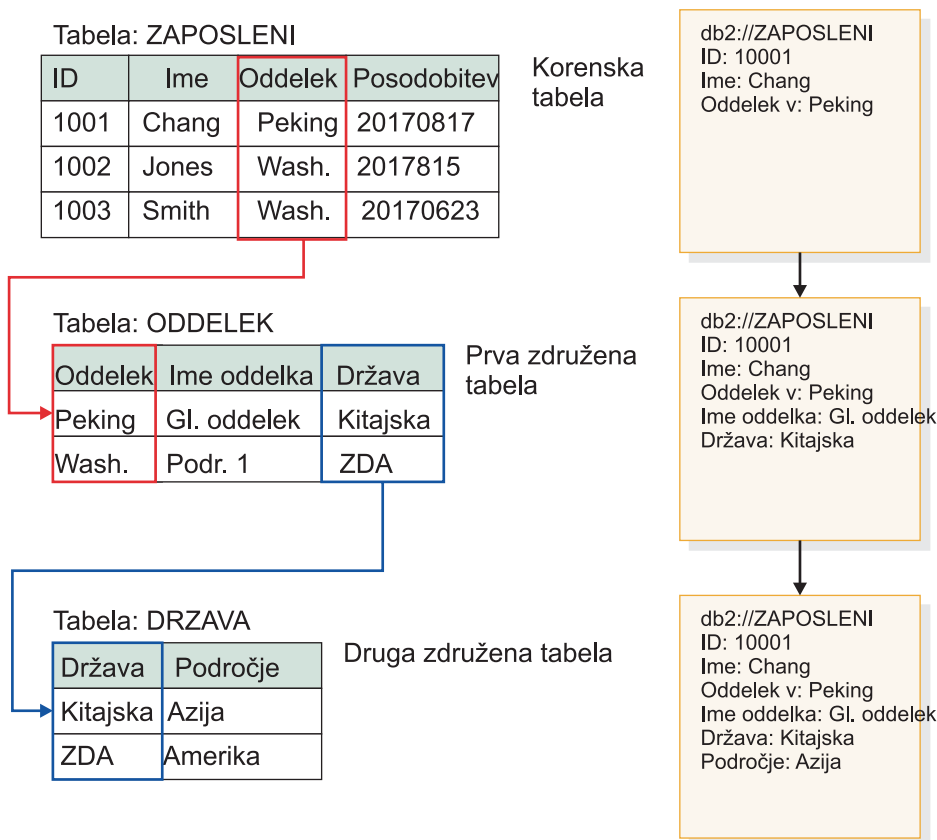


Slika 2. Združevanje tabel baze podatkov JDBC po poljih ključev

Ogledovanje rezultatov iskanja

Naslednja slika prikazuje, kako so v rezultatih iskanja podatki iz več strukturiranih tabel predstavljeni kot en sam dokument. Brez vtičnika je uporabnik, ki je preiskoval tabelo USLUŽBENCI, videl v rezultatih iskanja prikazano vrstico iz korenske tabele, v njej pa le vrednosti iz stolpcev tabele USLUŽBENCI (ID, ime in pisarna).

Z vtičnikom lahko pajek uporabi stolpec Pisarna kot ključ za združitev tabele USLUŽBENCI s tabelo PISARNA. Stolpec Država v tabeli PISARNA opravlja vlogo ključa za združitev tabele s tabelo DRŽAVA. Po združitvi tabel lahko uporabniki, ki preiskujejo tabelo USLUŽBENCI, med rezultati iskanja vidijo vrednosti stolpcev iz tabel PISARNA in DRŽAVA, pa tudi dodatna polja.



Slika 3. Vrednosti iz združenih tabel JDBC so prikazane v rezultatih iskanja

Pajkanje več strukturiranih tabel baze podatkov JDBC

Pajka baze podatkov JDBC lahko konfigurirate tako, da združi več strukturiranih tabel z istimi polji ključev.

Preden začnete

Če želite izvesti to nalogo, morate imeti pooblastila za prijavo kot skrbnik za iskanje v podjetju.

O tej nalogi

Pri konfiguriranju lastnosti pajka baze podatkov JDBC lahko podate vtičnik, ki omogoča pajkanje več strukturiranih tabel, ki so druga z drugo povezane s polji ključev. Brez tega vtičnika so vrstice v tabeli baze podatkov obravnavane kot posamezni dokumenti, po vrednostih stolpcev baze podatkov pa je mogoče iskati enako kot po posameznih poljih. S tem vtičnikom se vrstice iz več tabel v relacijski bazi podatkov, ki imajo enaka polja ključev, združijo (join) in obravnavajo kot en sam dokument. Pajek doda podatke, ki jih priključijo iz združenih tabel, v metapodatke za izvorno vrstico tabele baze podatkov. Ko uporabnik preišče bazo podatkov, se dodatni podatki ob prikazu dokumenta v rezultatih iskanja prikažejo kot dodatna polja.

Omejitve

Tipi podatkov, po katerih pajkanje ni mogoče

Pajek ne more pajkati polj v združenih tabelah, ki vsebujejo naslednje dvojiške podatkovne tipe:

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

Omejitve v obsegu prostora pajkanja

Tabele, ki jih boste združili, morajo biti v isti relacijski bazi podatkov. Združevanje tabel iz različnih baz podatkov ni mogoče.

Če je tabela v bazi podatkov konfigurirana tako, da bo združena z drugimi tabelami, je ta nastavitev univerzalna za vse pajke v zbirki, ki so omogočeni za uporabo vtičnika. Vendar pa je mogoče izdelati več zbirk in konfigurirati posebne pajke, ki pajkajo različne korenske tabele in združujejo različne tabele.

Omejitve pri uporabi drugih vtičnikov

Če konfigurirate pajka za uporabo vtičnika pri pajkanju več strukturiranih tabel, ne morete povezati drugega vtičnika s pajkom. Tako na primer ne morete podati vtičnika po meri za uveljavitev poslovnih in varnostnih pravil. Prav tako s pajkom ne morete povezati več kot enega vtičnika.

Omejitve v številu tabel, vrstic, polj in ključev

Na posamezno bazo podatkov je lahko združenih največ pet baz podatkov, vsota vrstic v teh tabelah pa mora biti manj kot milijon. Največje število polj, ki jih je mogoče prebrati iz tabele, je 10. Za združevanje tabel se uporabi par ključev. To pomeni, da ni mogoče združiti tabele z uporabo več ključev.

Zagotavljanje pajkanja sprememb v združenih tabelah

Če se vrstice v korenski tabeli med posameznimi postopki pajkanja ne spremenijo in pajek ni konfiguriran za izvedbo celotnega pajkanja, pajek prezre nespremenjene vrstice. Če se vrstice v tabeli, ki je pridružena korenski tabeli, spremenijo, četudi se korenska tabela ne, morate z enim od naslednjih dejanj zagotoviti, da bodo spremembe zaznane in pajkane:

- Korenska tabela v ciljni bazi podatkov mora imeti polje s časovnim žigom. Konfigurirajte ciljno bazo podatkov tako, da bo vsebovala polje s časovnim žigom, ki se posodobi ob vsaki spremembi vrstice v korenski tabeli ali spremembi vrstic v kateri od združenih dodatnih tabel. Če nastavite pajka baze podatkov JDBC, ne pozabite podati polja s časovnim žigom kot polja, na podlagi katerega mora pajek ugotoviti, ali je prišlo v tabelah do sprememb.
- Podajte, naj pajek ob konfiguriranju razporeda izvede celotno pajkanje. S to možnostjo zagotovite, da pajek vsakič pajka po vseh tabelah, ne glede na to, ali je prišlo do sprememb ali ne.

Postopek

Če želite pajka baze podatkov JDBC nastaviti za pajkanje več strukturiranih tabel:

1. Prijavite se na strežnik pajkanja kot skrbnik za iskanje v podjetju in prekopirajte datoteko ES_INSTALL_ROOT/default_config/crawler_rdb_plugin.xml ter izdelajte datoteko ES_NODE_ROOT/master_config/crawler_rdb_plugin.xml.
2. Z urejevalnikom besedil, ki podpira kodiranje UTF-8, uredite datoteko ES_NODE_ROOT/master_config/crawler_rdb_plugin.xml.
 - a. Uredite element <Server DBURL="jdbc:db2://db_server_url:50000/SAMPLE"> in zamenjajte jdbc:db2://db_server_url:50000/SAMPLE z URL-jem baze podatkov JDBC, ki jo nameravate pajkati. Pri konfiguriranju pajka ne pozabite podati enakega URL-ja tudi za bazo podatkov, ki jo želite pajkati.

- b. Če baza podatkov, ki jo boste pajkali, ni baza podatkov DB2, uredite element `<JDBCdriver>com.ibm.db2.jcc.DB2Driver</JDBCdriver>` in zamenjajte `com.ibm.db2.jcc.DB2Driver` z ustreznim gonilnikom JDBC. Pri konfiguriranju pajka ne pozabite podati enakega gonilnika tudi za bazo podatkov, ki jo želite pajkati.
- c. Uredite element `<User>username</User>` in zamenjajte `username` z ID-jem uporabnika, ki ima pooblastila za dostop do baze podatkov, ki jo želite pajkati.
- d. Uredite element `<Password Encryption="True">encrypted_password</Password>` in zamenjajte `encrypted_password` s šifriranim geslom za podani ID uporabnika. Šifrirano geslo lahko prekopirate iz datoteke `ES_NODE_ROOT/master_config/col_collection_name.JDBC_crawler_name/jdbccrawler.xml` in ga prilepite sem. Če šifriranje gesla ni potrebno, zamenjajte `Encryption="True"` z `Encryption="False"` ter `encrypted_password` z geslom iz čistega besedila.
- e. Če pustite element `<Delimiters Use="True">` nespremenjen, bo več izrazov v stolpcu ločenih z znakom (,), definiranim z elementom `<Delimiter>`. Nizi izrazov v tabeli so ločeni z znaki (;), definiranimi z elementom `<SecondDelimiter>`. Če nastavite `<Delimiters Use="True">` na `<Delimiters Use="False">`, se ločevalni znaki ne uporabijo, več polj z metapodatki z enakim imenom pa se doda kot metapodatki dokumenta.
- f. Če uporabite element `<Delimiters Use="True">`, uredite elementa `<Delimiter>`,`</Delimiter>` in `<SecondDelimiter>`,`</SecondDelimiter>` tako, da bosta podajala znake, ki bodo uporabljeni kot ločila vrednosti.
- g. Uredite element `<RelationMap Root="DB2INST1.TABLE_0">` in zamenjajte `DB2INST1.TABLE_0` z imenom korenske tabele, ki jo želite pajkati.
- h. Uredite element `<Relation Parent="DB2INST1.TABLE_0" ParentAlias="T0" ParentKey="ID" Child="DB2INST1.TABLE_1" ChildAlias="T1" ChildKey="ID"/>`.
 - Zamenjajte `Parent="DB2INST1.TABLE_0"` z imenom tabele, ki je v razmerju nadrejena.
 - Zamenjajte `ParentKey="T0"` z vzdevkom nadrejene tabele. Ta vzdevek mora biti unikatni, in ne sme biti vsebovan v datoteki `crawler_rdb_plugin.xml`.
 - Zamenjajte `ParentKey="ID"` z imenom stolpca, ki je v razmerju v uporabi kot polje ključev.
 - Zamenjajte `Child="DB2INST1.TABLE_1" ChildAlias="T1" ChildKey="ID"` z informacijami o podrejeni tabeli, ki jo želite pajkati.

Ta struktura določa način združitve (join) tabel. Naslednja preslikava razmerij na primer podaja, da bo pajkana korenska tabela z imenom `DB2INST1.TABLE_A`. Tabeli `DB2INST1.TABLE_B` in `DB2INST1.TABLE_C` se združita pod pogojem `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. Ponovite korak 2h, da izdelate elemente `<Relation>` za vsa razmerja, ki združujejo tabele iz korenske tabele.
- j. Uredite element `<Target TableAlias="T1">` in zamenjajte vrednost `TableAlias="T1"` z `ChildAlias`, ki ste jo definirali v koraku 2h.
- k. Uredite element `<Field Name="ID" FieldName="ID_1" Enabling="True" Searchable="True" FieldSearchable="True" IsContent="True"/>`.
 - Zamenjajte `Name="ID"` z imenom stolpca v dokumentih, ki jih želite pajkati.

- Zamenjajte `FieldName="ID_1"` z imenom polja metapodatkov v dokumentih, ki jih želite pajkati. Ta vrednost je v uporabi kot prikazno ime stolpca v konzoli za skrbništvo nad iskanjem v podjetju in rezultatih iskanja.
 - Zamenjajte `Enabling="True"` s `"False"`, če tega stolpca ne želite vključiti v metapodatke dokumenta.
 - Zamenjajte `Searchable="True"` s `"False"`, da preprečite uporabnikom preiskovanje tega stolpca s poizvedbo po prostem besedilu.
 - Zamenjajte `FieldSearchable="True"` s `"False"`, da uporabnikom preprečite iskanje tega stolpca po imenu.
 - Zamenjajte `IsContent="True"` s `"False"`, s čimer nakažete, da stolpec ne vsebuje vsebine, ki bi jo bilo mogoče iskati. Če podate `Searchable="True"` in `IsContent="True"`, se vrednost stolpca uporabi pri odkrivanju podvojenih dokumentov in postane del dinamičnega povzetka dokumentov v rezultatih iskanja.
- l. Ponovite korak 2k na strani 64, da izdelate elemente `<Field>` za vse stolpce, ki jih nameravate pajkati.
 - m. Ponovite koraka 2j na strani 64 in 2k na strani 64, da izdelate elementa `<Target>` in `<Field>` za vse podrejene tabele, na katere obstaja sklic v preslikavi razmerij (`<RelationMap>`).
 - n. Ponovite korake 2g na strani 64 do 2m, da izdelate več preslikav razmerij za več korenskih tabel.
 - o. Ponovite korake 2a na strani 63 do 2n, da konfigurirate preslikavo razmerij za drugo bazo podatkov.
3. Konfiguriranje pajka za uporabo vtičnika:
 - a. Odprite konzolo za skrbništvo nad iskanjem v podjetju, uredite zbirko in izberite stran Pajkanje.
 - b. Izdelajte pajka baze podatkov JDBC ali pa poiščite pajka, ki ga želite spremeniti, in kliknite **Lastnosti pajka**.
 - c. V polje **Ime razreda vtičnika** vnesite ime vtičnika za pajkanje več strukturiranih tabel:
com.ibm.es.plugin.rdb.RDBPlugin
 - d. V polje **Pot do razreda vtičnika** vnesite popolnoma kvalificirane poti do vtičnikov in gonilnikov JDBC, ki jih uporablja vtičnik. Pot do gonilnika JDBC za bazo podatkov DB2 v sistemu Windows je lahko na primer:
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. Kliknite bodisi **Naprej**, da nadaljujete z izdelavo pajka, ali **Potrdi**, da shranite opravljene spremembe.
 4. Če želite razmestiti datoteko `crawler_rdb_plugin.xml` v konfiguracijo sistema, znova zaženite sistem iskanja v podjetju:

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

Pajki NNTP

Če želite v zbirko iskanja v podjetju vključiti članke iz novičarskih skupin NNTP, morate konfigurirati pajka NNTP.

Pajka NNTP lahko uporabite za pajkanje poljubnega števila strežnikov NNTP. Pri konfiguriranju pajka izberete novičarske skupine, v katerih boste pajkali z enega strežnika NNTP. Pozneje lahko pri urejanju prostora za pajkanje dodate druge strežnike NNTP, v katerih želite, da pajka pajek.

Pri identificiranju novičarskih skupin za pajkanje lahko izberete skupine, ki jih boste vključili v prostor pajkanja in skupine, ki jih boste izključili iz njega. S to zasnovano lahko pajku na preprost način omogočite pajkanje v večini novičarskih skupin na strežniku in mu preprečite pajkanje v nekaj novičarskih skupinah, za katere ne želite, da jih uporabniki preiščejo.

Tako lahko na primer podate pravila za vključitev vseh novičarskih skupin na določenem strežniku NNTP in nato podate, da želite izključiti novičarske skupine na tem strežniku, če njihova imena vključujejo niz **zasebno**.

Če želite izdelati ali spremeniti pajka NNTP, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh novičarskih skupinah v prostoru za pajkanje.
- Identificiranje strežnika NNTP za pajkanje. Če je strežnik zaščiten z geslom, morate podati ID uporabnika in geslo, ki ju bo uporabil pajek za dostopanje do novičarskih skupin na strežniku.
- Nastavitev razporeda za pajkanje strežnike.
- Podajanje vzorcev, ki vključujejo novičarske skupine, in vzorcev, ki iz prostora pajkanja določene novičarske skupine izključujejo.
- Podajanje, ali naj pajek samodejno odkrije jezik in kodno stran člankov za pajkanje ali pa naj uporabi določen jezik in kodno stran.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali varnostnih žetonih.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

Pajki Notes

Če želite v zbirko iskanja v podjetju vključiti baze podatkov IBM Lotus Notes, morate konfigurirati pajka Notes.

Nasvet:

Podrobnejše primere z načini konfiguriranja zaščitenega pajka Notes najdete v scenarijih za večjo organizacijo v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Konfiguriranje strežnika pajkanja

Če uporablja strežnik Lotus Notes, na katerem boste pajkali, protokol NRPC (Notes remote procedure call), morate na strežniku pajkanja zagnati namestitveni skript. Ta skript, ki je na voljo z OmniFind Enterprise Edition, omogoča pajku Notes komuniciranje s strežniki, ki uporabljajo NRPC.

Če uporablja strežnik Lotus Notes, na katerem boste pajkali, protokol DIIOP (Domino Internet Inter-ORB Protocol), na strežniku pajkanja ni potrebno zagnati namestitvenega skripta, vendar pa morate strežnik Lotus Notes konfigurirati tako, da bo lahko pajek Notes dostopal do strežnika.

Pomembno: Če strežnik Lotus Notes uporablja DIIOP in konfigurirate pajka za uporabo HTTPS ali DIIOP prek SSL tako, da so prenosi med pajkom in strežnikom šifrirani, morate prekopirati datoteko `TrustedCerts.class` (na primer `c:\certs` ali `/data/certs`) s strežnika Lotus Notes Manager na strežnik pajkanja. V konfiguraciji z dvema ali štirimi strežniki morate prekopirati tudi datoteko `TrustedCerts.class` na strežnike, na katerih je nameščena iskalna komponenta. Prepričajte se, da se datoteka na strežniku pajkanja nahaja na enakem mestu kot na iskalnih strežnikih. Imeniško pot do datoteke `TrustedCerts.class` podate med konfiguriranjem pajka.

Če ste v sistem IBM AIX namestili OmniFind Enterprise Edition, morate zagotoviti, da je na strežniku pajkanja nameščen in razpoložljiv modul V/I zaključnih vrat.

Preden za konfiguriranje pajka Notes uporabite skrbniško konzolo iskanja v podjetju, izpolnite naloge, ki ustrezajo vašemu okolju:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino” na strani 69.
- “Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino” na strani 71.
- “Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP” na strani 72.
- “Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino” na strani 73.

Zaščita na ravni dokumenta

Če je omogočena zaščita zbirke in uporablja strežnik, na katerem boste pajkali, protokol NRPC, morate na strežniku pajkanja konfigurirati overjeni strežnik Lotus Domino. Overjeni strežnik se uporablja za uveljavitev krmilnih elementov dostopa na ravni dokumentov. Preden omogočite zbirko za iskanje, izpolnite naslednje naloge:

- Konfigurirajte overjene strežnike Lotus Domino za preverjanje poverilnic uporabnikov.
- Na aplikacijskem strežniku WebSphere Application Server omogočite globalno zaščito in konfigurirajte iskalno aplikacijo za uporabo zaščite. S tem korakom zagotovite, da bodo prijavnice poverilnice preverjene, ko bodo uporabniki poskusili uporabiti iskalno aplikacijo. Iskalni strežniki uporabljajo poverilnice za preverjanje pooblastila vsakega uporabnika za dostopanje do dokumentov Lotus Domino.

Pregled konfiguriranja

Pajka Notes lahko uporabite za pajkanje poljubnega števila standardnih baz podatkov Lotus Notes (datoteke `.nsf`). Pri izdelavi pajka izberete baze podatkov ali imenike za pajkanje z enega strežnika Lotus Notes. Pozneje lahko pri urejanju prostora pajkanja dodate dokumente z drugega strežnika Lotus Notes, ki jih želite vključiti v isti prostor pajkanja. Pri izdelavi ali urejanju pajka lahko podate, ali želite pajkati v vseh bazah podatkov ali imenikih na strežniku ali pa samo v določenih bazah podatkov, pogledih in mapah.

Če želite izdelati ali spremeniti pajka Notes, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh dokumentih v prostoru za pajkanje.
- Identificiranje gostiteljskega imena strežnika Lotus Notes, vrat in komunikacijskega protokola.

- Če je strežnik konfiguriran za uporabo protokola DIIOP, lahko podate, kako se bo pajek povezal z objekti Domino. Tako lahko na primer podate možnosti za uporabo HTTPS ali plasti zaščitene vtičnice (SSL) za šifriranje komunikacij.
- Izbira baz podatkov in imenikov, v katerih želite pajkati. Pri pajkanju imenikov lahko podate vzorce, s katerimi vključite ali izključite baze podatkov, ki vam lahko olajšajo nalogo pajkanja velikih imenikov, tako da jo razporedite med več pajkov.
- Nastavitev razporeda za pajkanje baz podatkov ali imenikov.
- Izbira dokumentov, v katerih želite pajkati. Pajkate lahko v vseh dokumentih v imeniku, v vseh dokumentih v bazi podatkov ali v dokumentih iz izbranih pogledov in map baze podatkov.
- Podajanje možnosti, ki omogočijo preiskovanje polj v različnih bazah podatkov, pogledih in mapah. Tako lahko na primer iz prostora pajkanja izključite določena polja in podate možnosti za preiskovanje prilog.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi seznamami za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezani pojmi

“Nasveti za pajkanje baz podatkov Lotus Domino”

“Zagotavljanje zaščite na ravni dokumenta za dokumente Lotus Domino” na strani 256

S tem povezana opravila

“Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino” na strani 69

“Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino” na strani 71

“Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP” na strani 72

“Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino” na strani 73

“Konfiguriranje overjenih strežnikov Lotus Domino za preverjanje poverilnic uporabnikov” na strani 257

Nasveti za pajkanje baz podatkov Lotus Domino

Preden konfigurirate pajka Notes, preglejte smernice za pajkanje baz podatkov Lotus Domino.

- Baze podatkov Notes, ki temeljijo na standardnih predlogah (kot je razpravna baza podatkov), so najboljši tip baz podatkov za pajkanje.
- Pajek Notes uveljavi naslednja pravila za preslikavo polj:
 - Imena glavnih polj iz standardnih predlog Domino so začetno registrirana.
 - Vrednosti iz polj Notes, ki so podane v tabeli s pravili za preslikavo, so uporabljene kot povzetki dokumentov v rezultatih iskanja.
 - Vrednosti iz polj Notes, ki niso podane v tabeli s pravili za preslikavo, niso uporabljene v povzetkih dokumentov.
 - Vrednosti iz polj Notes, ki so preslikane v polje Naslov, so uporabljene kot naziv dokumenta v rezultatih iskanja.

- Polja iz naslednje tabele so po privzetku preslikana v imena iskalnih polj:

Tabela 2. Privzeta pravila za preslikavo polj

Ime polja baze podatkov Notes	Ime iskalnega polja
Naslov	Naslov
NaslovDogodka	Naslov
Zadeva	Naslov
Telo	Telo
Naloga	Telo
Od	Izdelovalec
Avtor	Izdelovalec
Ključne besede	Kategorije
Kategorije	Kategorije
ImeSobeEkipe	Organizacija
ImeEkipe	Organizacija
Oddelek	Organizacija

- Pajek Notes lahko pajka v vseh tipih polj, razen v poljih, ki so izračunana za prikaz.
- Pajkanje se ne izvaja v statičnem besedilu in slikah na obrazcu Notes.
- Pri konfiguriranju pajka izberite potrditveno polje **Vse** pod možnostjo **Pajkaj**, da boste izvedli pajkanje v vseh poljih in povečali podatke polja za pajkanje (za omejitve polj za pajkanje lahko uporabite polje **Pajkaj v vseh poljih, razen**).
Če želite zmanjšati pajkanje nepotrebnih polj, počistite potrditveno polje **Pajkaj** za vsa polja, razen za tista, ki so preslikana v iskalna polja.

S tem povezani pojmi

“Pajki Domino Document Manager” na strani 56

“Pajki Notes” na strani 66

“Pajki QuickPlace” na strani 75

Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino

Če namestite OmniFind Enterprise Edition v računalnik, v katerem se izvaja operacijski sistem IBM AIX, Linux ali Solaris, in nameravate pajkati po strežnikih, ki uporabljajo protokol Notes (NRPC), morate zagnati skript in konfigurirati strežnik pajkanja. Skript pajkom Notes, QuickPlace, in Domino Document Manager omogoči komuniciranje s strežniki baze podatkov.

Omejitve

Strežnik Domino se na istem računalniku ne more izvajati hkrati s pajkom Notes, QuickPlace, ali Domino Document Manager, ki je konfiguriran za uporabo protokola NRPC. Če poskusite katerega od teh pajkov zagnati med izvajanjem strežnika Domino, sistem javi napako, pajek pa se zaustavi.

O tej nalogi

Pajki, ki uporabljajo protokol NRPC, uporabljajo knjižnice Domino kot odjemalca. Te knjižnice namestite skupaj z namestitvijo strežnika Lotus Domino Server na strežnik pajkanja.

Da bi se prepričali, ali so pajki združljivi s knjižnicami Domino, po namestitvi knjižnic Domino zaženite nastavitveni skript, ki ga na strežniku pajkanja nudi OmniFind Enterprise Edition.

Postopek

Če želite konfigurirati strežnik pajkanja tako, da bo lahko pajkal po strežnikih Lotus Notes, Lotus QuickPlace in Domino Document Manager, naredite naslednje:

1. Na strežniku pajkanja izdelajte uporabnika **server** in skupino **notes**:
 - a. Prijavite se kot korenski uporabnik:
`su - root`
 - b. Dodajte uporabnika:
`useradd server`
 - c. Vnesite geslo za tega uporabnika:
`passwd server`
Sistem od vas zahteva, da spremenite geslo.
2. Namestite Lotus Domino Server na strežnik pajkanja:
 - a. Vstavite CD strežnika Domino Server in ga nastavite. (Če nimate CD-ja, lahko prenesete sliko.)
 - b. Pomaknite se v mapo operacijskega sistema.

AIX: `cd /mnt/cdrom/aix`
Linux: `cd /mnt/cdrom/linux`
Solaris: `cd /mnt/cdrom/solaris`
 - c. Zaženite namestitveni program:
`./install`
 - d. Upoštevajte pozive in sprejmite privzete vrednosti ali podajte želene nastavitve za namestitev (na primer poti do namestitvenega in podatkovnega imenika).
Če potrebujete pomoč pri nameščanju izdelka Domino Server, preglejte dokumentacijo Domino.
 - e. Prepričajte se, da ima ID skrbnika za iskanje v podjetju dovoljenje za dostop do imenika `home/server`. ID skrbnika podate pri namestitvi OmniFind Enterprise Edition.
3. Na strežniku pajkanja zaženite nastavitveni skript, ki ga nudi OmniFind Enterprise Edition:
 - a. Prijavite se kot skrbnik za iskanje v podjetju (ID uporabnika ste podali ob namestitvi izdelka OmniFind Enterprise Edition).
 - b. Zaženite naslednji skript, ki je nameščen v imeniku `$ES_INSTALL_ROOT/bin`:
`escrnote.sh`
 - c. Odgovorite na pozive:
 - Na naslednji poziv odgovorite **D**, če je Domino Server nameščen v privzetem imeniku, in **N**, če ni:
Sistem je našel pot do imenika Lotus Notes `/opt/lotus/notes/latest/linux`.
Ali je to prava pot do imenika Lotus Notes?

Privzeta pot za sistem AIX je `/opt/lotus/notes/latest/ibmpow`.
Privzeta pot za sistem Linux je `/opt/lotus/notes/latest/linux`.
Privzeta pot za sistem Solaris je `/opt/lotus/notes/latest/sunspa`.
 - Če Domino Server ni nameščen v privzetem imeniku na strežniku pajkanja, kot odgovor na naslednji poziv podajte imenik, v katerem je nameščen Domino:

Vnesite pot do imenika Lotus Notes

Na računalniku s sistemom Linux lahko na primer podate `/opt/lotus/notes/latest/linux`.

- Na naslednji poziv odgovorite z Y, če je podatkovni imenik izdelka Domino Server nameščen v privzetem imeniku, in N, če ni:

Sistem je našel pot do podatkovnega imenika Lotus Notes `/local/notesdata`. Ali je to prava pot do podatkovnega imenika Lotus Notes?

Privzeta pot je `/local/notesdata`.

- Če podatkovni imenik Domino Server ni razmeščen na privzeto mesto na strežniku pajkanja, kot odgovor na naslednji poziv podajte podatkovno pot Domino:

Vnesite pot do podatkovnega imenika Lotus Notes.

4. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:

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

S tem povezani pojmi

“Pajki Domino Document Manager” na strani 56

“Pajki Notes” na strani 66

“Pajki QuickPlace” na strani 75

Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino

Če namestite izdelek OmniFind Enterprise Edition v računalnik s sistemom Microsoft Windows in nameravate pajkati po strežnikih, ki uporabljajo protokol NRPC Notes, morate zagnati skript in konfigurirati strežnik pajkanja. Skript pajkom Notes, QuickPlace, in Domino Document Manager omogoči komuniciranje s strežniki baze podatkov.

Omejitve

Lotus Domino Server in odjemalec Lotus Notes se na istem računalniku ne moreta izvajati hkrati, skupaj s pajkom Notes, QuickPlace, ali Domino Document Manager, konfiguriranim za uporabo protokola NRPC. Če poskusite katerega od teh pajkov zagnati med izvajanjem strežnika Domino, sistem javi napako, pajek pa se zaustavi.

O tej nalogi

Pajki, ki uporabljajo protokol NRPC, uporabljajo knjižnice odjemalca Lotus Domino. Te knjižnice namestite skupaj z namestitvijo strežnika Lotus Domino Server na strežnik pajkanja. Da bi se prepričali, ali so pajki združljivi s knjižnicami Domino, po namestitvi knjižnic Domino zaženite nastavitveni skript, ki ga na strežniku pajkanja nudi OmniFind Enterprise Edition.

Postopek

Če želite konfigurirati strežnik pajkanja tako, da bo z njim mogoče pajkati po strežnikih Lotus Notes, Lotus QuickPlace in Domino Document Manager, naredite naslednje:

1. V strežnik pajkanja se prijavite z ID-jem uporabnika, ki je član skupine skrbnikov. Prepričajte se, da ima ID uporabnika pooblastila za namestitev izdelka Lotus Notes.
2. Nameščanje izdelka Lotus Notes:
 - a. Vstavite CD strežnika Domino Server. (Če nimate CD-ja, lahko prenesete sliko.)

- b. Zaženite namestitveni program: `setup.exe`
 - c. Upoštevajte pozive in sprejmite privzete vrednosti ali podajte zelene nastavitve za namestitev (na primer poti do namestitvenega in podatkovnega imenika).
Če potrebujete pomoč, preglejte dokumentacijo Lotus Domino.
 - d. Prepričajte se, da ima ID skrbnika za iskanje v podjetju dovoljenje za dostop do podatkovnega imenika Domino. ID skrbnika podate pri namestitvi OmniFind Enterprise Edition.
3. Na strežniku pajkanja zaženite nastavitveni skript, ki ga nudi OmniFind Enterprise Edition:
- a. Prijavite se z ID-jem skrbnika za iskanje v podjetju (ID uporabnika ste podali ob namestitvi izdelka OmniFind Enterprise Edition).
 - b. Zaženite naslednji skript, ki je nameščen v imeniku `%ES_INSTALL_ROOT%\bin: escrnote.vbs`
 - c. Odgovorite na pozive:
 - Na naslednji poziv odgovorite D, če je Lotus Notes nameščen v privzetem imeniku, in N, če ni:
Sistem je našel pot do imenika Lotus Notes `c:\lotus\notes`.
Ali je to prava pot do imenika Lotus Notes?

Običajna namestitvena pot v računalniku s sistemom Windows je `c:\lotus\notes` ali `c:\lotus\domino`.

 - Če Lotus Notes ni nameščen v privzetem imeniku na strežniku pajkanja, kot odgovor na naslednji poziv podajte imenik, v katerem je nameščen Lotus Notes:
Vnesite pot do imenika Lotus Notes
 - Na naslednji poziv odgovorite z Y, če je podatkovni imenik izdelka Lotus Notes razmeščen na privzetem mestu, in z N, če ni:
Sistem je našel pot do podatkovnega imenika Lotus Notes `c:\lotus\notes\data`.
Ali je to prava pot do podatkovnega imenika Lotus Notes?

Običajna namestitvena pot v računalniku s sistemom Windows je `c:\lotus\notes\data` ali `c:\lotus\domino\data`.

 - Če podatkovni imenik Lotus Notes ni razmeščen na privzeto mesto na strežniku pajkanja, kot odgovor na naslednji poziv podajte pot do podatkovnega imenika:
Vnesite pot do podatkovnega imenika Lotus Notes.
4. Na strežniku pajkanja zaustavite sistem iskanja v podjetju, nato pa ga znova zaženite:

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

S tem povezani pojmi

“Pajki Domino Document Manager” na strani 56

“Pajki Notes” na strani 66

“Pajki QuickPlace” na strani 75

Konfiguriranje strežnikov, ki uporabljajo protokol DIIOF

Če želite pajkati strežnike, ki uporabljajo protokol DIIOF (Domino Internet Inter-ORB Protocol), morate konfigurirati strežnik tako, da bodo lahko pajki Notes, QuickPlace, in Domino Document Manager uporabljali ta protokol.

Preden začnete

Na strežniku, ki ga želite pajkati, se morajo izvajati naloge DIIOF in HTTP.

Postopek

Če želite konfigurirati strežnike, ki uporabljajo protokol DIIOP:

1. Konfiguriranje dokumenta strežnika:
 - a. Odprite dokument **strežnika** na strežnikih Lotus Notes, Lotus QuickPlace, ali Domino Document Manager, ki jih želite pajkati. Ta dokument je shranjen v imeniku Domino.
 - b. Na strani Konfiguracija razširite razdelek **strežnik**.
 - c. Na strani Zaščita v področju **Omejitve programabilnosti** podajte omejitve zaščite, ki so primerne za vaše okolje. Vnesite jih v naslednja polja:
 - **Izvajanje omejenih posrednikov Lotus Script/Java**
 - **Izvajanje omejenega Java/Javascript/COM**
 - **Izvajanje neomejenega Java/Javascript/COM**Podate lahko na primer zvezdico (*), s čimer posrednikom Lotus Script/Java dovolite neomejen dostop, nato pa podate uporabniška imena, ki so registrirana na strežniku Domino Directory za omejitve Java/Javascript/COM.

Pomembno: Pajek, ki ga konfigurirate za pajkanje tega strežnika s protokolom DIIOP, mora znati uporabljati uporabniška imena, ki jih vnesete v ta polja.
 - d. Odprite stran Internetni protokol, nato odprite stran HTTP in nastavite možnost **Dovoli odjemalcem HTTP pregledovanje baz podatkov na Da**.
2. Konfiguriranje uporabniškega dokumenta:
 - a. Na strežniku Lotus Notes, Lotus QuickPlace ali Domino Document Manager odprite uporabniški dokument, ki ga želite pajkati. Ta dokument je shranjen v imeniku Domino.
 - b. Na strani Osnove v polje **Internetno geslo** vnesite geslo.

Če s skrbniško konzolo za iskanje v podjetju konfigurirate možnosti za pajkanje tega strežnika, ID in geslo podajte na strani, kjer izberete strežnik, po katerem boste pajkali. Pajek s pomočjo danih poverilnic dostopi do strežnika.
3. Znova zaženite nalogo DIIOP na strežniku.

S tem povezani pojmi

“Pajki Domino Document Manager” na strani 56

“Pajki Notes” na strani 66

“Pajki QuickPlace” na strani 75

Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino

Preden boste lahko začeli uporabljati pajke Notes, QuickPlace in Domino Document Manager v sistemu IBM AIX, morate namestiti modul IOCP (V/I zaključna vrata) in ga konfigurirati za uporabo skupaj s pajkom.

O tej nalogi

Če ne uporabite modula IOCP, procesi odkrivanja ob poskusu izdelave pajka ne bodo uspeli. Sistem prikaže naslednje sporočilo o napaki:

```
FFQM0105E Sistem je prejel napako s strežnika -  
Sporočilo: FFQG0024E Nepričakovano izjemno stanje: odkrivanje
```

Sporočilo, ki vsebuje napako ENOEXEX, se zapiše v datoteko \$ES_NODE_ROOT/logs/system_yyyymmdd.log. (Del besedila v sporočilu je razdeljen na več vrstic, kar olajša branje sporočila.)

```

5/20/05 18:08:52.423 JST [Error] [ES_ERR_EXCEPTION_DEFAULT_MESSAGE] [] [discovery]
ies10.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)

```

Postopek

Če želite namestiti modul IOCP in se prepričati, da je pravilno nameščen na strežnik pajkanja:

Narediti morate naslednje:

1. Namestite modul IOCP (`bos.iocp.rte`) s CD-ja z izdelkom AIX na strežnik pajkanja.

Potem ko namestite modul IOCP in preden izdelate pajka Notes, QuickPlace ali Domino Document Manager, uveljavite programski popravek za modul. Navodila za to najdete v informacijah na naslednji povezavi:

<http://www.ibm.com/support/docview.wss?uid=swg21086556>

2. Vnesite naslednji ukaz, s katerim se prepričate, da je na strežniku pajkanja nameščen modul IOCP:

```
$ lsllp -l bos.iocp.rte
```

Izhodni podatki iz ukaza `lsllp` morajo biti podobni naslednjemu zgledu:

Datotečni niz	Raven	Stanje	Opis
Pot: /usr/lib/objrepos bos.iocp.rte	5.2.0.10	ODOBRENO	API V/I zaključnih vrat
Pot: /etc/objrepos bos.iocp.rte	5.2.0.10	ODOBRENO	API V/I zaključnih vrat

3. Vnesite naslednji ukaz, s katerim se prepričate, da je status vrat IOCP **Na voljo**:

```
$ lsdev -Cc iocp
```

Izhodni podatki iz ukaza `lsdev` se morajo ujemati z naslednjim zgledom:

```
iocp0 Na voljo V/I zaključna vrata
```

4. Če je status vrat IOCP **Definirana**, spremenite status na **Na voljo**:

- a. Prijavite se na strežnik pajkanja kot korenski uporabnik in izdajte naslednji ukaz:

```
# smit iocp
```

- b. Izberite možnost **Spreminjanje / Prikaz značilnosti V/I zaključnih vrat** in spremenite **STANJE, ki bo konfigurirano ob vnovičnem zagonu sistema iz Definirano v Na voljo**.
- c. Znova zaženite strežnik pajkanja.
- d. Znova vnesite ukaz `lsdev` in potrdite, da je status vrat IOCP spremenjen v **Na voljo**.

S tem povezani pojmi

“Pajki Domino Document Manager” na strani 56

“Pajki Notes” na strani 66

“Pajki QuickPlace”

Pajki QuickPlace

Če želite v zbirko iskanja v podjetju vključiti prostore in sobe Lotus QuickPlace, morate konfigurirati pajka QuickPlace.

S pajkom QuickPlace lahko pajkate tudi mesta, ki jih upravljate s storitvami Lotus Quickr za Lotus Domino. Če uporabite storitve Lotus Quickr za WebSphere Portal, s pajkom seznama semen dodajte dokumente, shranjene v knjižnicah Lotus Quickr, v zbirko za iskanje v podjetju.

Nasvet:

Podrobnejše informacije o načinih konfiguriranja zaščitene pajka QuickPlace najdete v scenariju za majhno organizacijo, opisanem v publikaciji IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Konfiguriranje strežnika pajkanja

Če strežnik, ki ga nameravate pajkati, uporablja protokol NRPC (Notes remote procedure call), morate na strežniku pajkanja zagnati namestitveni skript. Ta skript, ki je na voljo z OmniFind Enterprise Edition, omogoča pajku QuickPlace komuniciranje s strežniki, ki uporabljajo NRPC.

Če strežnik, ki ga nameravate pajkati, uporablja protokol DIIOP (Domino Internet Inter-ORB Protocol), vam na strežniku pajkanja ni treba zagnati namestitvenega skripta. Kljub temu pa morate konfigurirati ciljni strežnik in omogočiti dostop pajku QuickPlace do njega.

Če strežnik, ki ga nameravate pajkati, uporablja strežnik s protokolom LDAP (Lightweight Directory Access Protocol) LDAP, je treba ciljni strežnik konfigurirati za uporabo protokola DIIOP (pajek QuickPlace ne more uporabiti protokola NRPC za pajkanje podatkov LDAP). Konfigurirati morate tudi bazo podatkov pomoči za imenike in ciljni strežnik tako, da bo uporabljal strežnik LDAP kot sekundarni strežnik Domino.

Pomembno: Če ciljni strežnik uporablja DIIOP in konfigurirate pajka za uporabo HTTPS ali DIIOP prek SSL tako, da se prenosi med pajkom in strežnikom ne šifrirajo, morate prekopirati datoteko `TrustedCerts.class` (na primer `c:\certs` ali `/data/certs`) s ciljnega strežnika na strežnik pajkanja. V konfiguraciji z dvema ali štirimi strežniki morate prekopirati tudi datoteko `TrustedCerts.class` na strežnike, na katerih je nameščena iskalna komponenta. Prepričajte se, da se datoteka na strežniku pajkanja nahaja na enakem mestu kot na iskalnih strežnikih. Imeniško pot do datoteke `TrustedCerts.class` podate med konfiguriranjem pajka.

Potem ko konfigurirate pajka in podate ID uporabnika, ki ga bo pajek uporabljal, morate podati ID z zadostnimi pooblastili za dostop do vseh mest QuickPlace v domeni Domino. To storite tako, da dodelite ID uporabnika rezervirani skupini z imenom QuickPlaceAdministratorsSUGroup.

Če ste v sistem IBM AIX namestili OmniFind Enterprise Edition, morate zagotoviti, da je na strežniku pajkanja nameščen in razpoložljiv modul V/I zaključnih vrat.

Preden za konfiguriranje pajka QuickPlace uporabite skrbniško konzolo iskanja v podjetju, izpolnite naloge, ki ustrezajo vašemu okolju:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino” na strani 69.
- “Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino” na strani 71.
- “Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP” na strani 72.
- “Konfiguriranje strežnika QuickPlace za uporabo zaščite Lokalni uporabnik” na strani 258.
- “Konfiguriranje pomoči za imenike na strežniku QuickPlace” na strani 259.
- “Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino” na strani 73.

Zaščita na ravni dokumenta

Če je omogočena zaščita zbirke in strežnik, na katerem boste pajkali, uporablja protokol NRCP, morate na strežniku pajkanja konfigurirati overjeni strežnik Lotus Domino. Overjeni strežnik se uporablja za uveljavitev krmilnih elementov dostopa na ravni dokumentov. Preden omogočite zbirko za iskanje, izpolnite naslednje naloge:

- Konfigurirajte overjene strežnike Lotus Domino za preverjanje poverilnic uporabnikov.
- Omogočite globalno zaščito na strežniku WebSphere Application Server in konfigurirajte iskalno aplikacijo za uporabo zaščite. S tem korakom zagotovite, da bodo prijavnice poverilnice preverjene, ko bodo uporabniki poskusili uporabiti iskalno aplikacijo. Iskalni strežniki uporabljajo poverilnice za preverjanje pooblastila vsakega uporabnika za dostopanje do dokumentov Lotus Domino.

Pajkanje v prilogah

V izdelku Lotus QuickPlace lahko uvozite in objavite dokumente Microsoft Office (možnosti, ki jih lahko izberete pri uvažanju, vključujejo Uvožena stran, Stran Microsoft Word, Stran Microsoft Excel, Stran Microsoft PowerPoint in Več uvoženih strani). Pajek QuickPlace lahko po teh tipih uvoženih dokumentov pajka samo v obliki priponk, in sicer pod naslednjimi pogoji:

- Strežnik uporablja protokol DIIOP.
- Pajkanje po prilogah omogočite pri konfiguriranju možnosti pajkanja.
- Pajka konfigurirate za pajkanje polja "\$FILE" ali vseh polj.

Pregled konfiguriranja

Pajka QuickPlace lahko uporabite za pajkanje poljubnega števila prostorov QuickPlace. Pri izdelavi pajka izberete prostore za pajkanje z enega strežnika QuickPlace. Pozneje lahko pri urejanju prostora pajkanja dodate dokumente z drugega strežnika QuickPlace, ki jih želite vključiti v isti prostor pajkanja. Pri izdelavi ali urejanju pajka lahko podate, ali želite pajkati vsa mesta na strežniku ali pa le mesta, ki jih podate, in ali želite pajkati vse sobe v mestih, ki bodo pajkana, ali le sobe, ki jih podate.

Omejitev: Ko podajate ID uporabnika, ki ga bo uporabljal pajek, pazite, da boste podali ID, ki ima ustrezno pooblastilo za dostopanje do vseh prostorov QuickPlace v domeni Domino. To lahko naredite s konfiguriranjem strežnika QuickPlace in dodelitvijo ID-ja uporabnika rezervirani skupini, imenovani QuickPlaceAdministratorsSUGroup. Ne spreglejte, da ime skupine ne vsebuje vdelenih presledkov in razlikuje med velikimi in malimi črkami.

Če želite izdelati ali spremeniti pajka QuickPlace, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo, kako pajek pajka v vseh dokumentih v prostoru za pajkanje.
- Identificiranje strežnika QuickPlace in komunikacijskega protokola.
- Če je strežnik konfiguriran za uporabo protokola DIIOP, lahko podate, kako se bo pajek povezal z objekti Domino. Tako lahko na primer podate možnosti za uporabo HTTPS ali plasti zaščitenih vtičnic (SSL) za šifriranje komunikacij.
- Podajanje informacij o uporabniškem imeniku, ki je povezan s strežnikom (pajek potrebuje te informacije, da lahko uveljavi krmilne elemente dostopa, ko uporabniki preiskujejo zbirko).
- Izbira prostorov, v katerih želite pajkati.
- Nastavitev razporeda za pajkanje prostorov.
- Izbira dokumentov, v katerih želite pajkati. Pajek lahko pajka v vseh sobah v prostoru ali pa pajka samo v dokumentih v izbranih sobah.
- Podajanje možnosti, ki omogočajo preiskovanje polj v različnih prostorih in sobah. Tako lahko na primer iz prostora pajkanja izključite določena polja in podate možnosti za preiskovanje prilog.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi seznamami za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezani pojmi

“Nasveti za pajkanje baz podatkov Lotus Domino” na strani 68

“Zagotavljanje zaščite na ravni dokumenta za dokumente Lotus Domino” na strani 256

S tem povezana opravila

“Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino” na strani 69

“Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino” na strani 71

“Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP” na strani 72

“Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino” na strani 73

“Konfiguriranje overjenih strežnikov Lotus Domino za preverjanje poverilnic uporabnikov” na strani 257

“Konfiguriranje strežnika QuickPlace za uporabo zaščite Lokalni uporabnik” na strani 258

Pajki seznama semen

Če uporabljate storitve IBM Lotus Quickr za WebSphere Portal, lahko s pajkom seznama semen dodate dokumente iz knjižnic Lotus Quickr v zbirko za iskanje v podjetju.

Knjižnica Lotus Quickr je vsebnik za datoteke dokumentov. Pajek seznama semen ne podpira pajkanja spletne vsebine, kot so na primer wikiji in blogi.

Če uporabljate storitve Lotus Quickr za Lotus Domino, s pajkom QuickPlace dodajte dokumente Lotus Quickr v zbirko.

Konfiguriranje strežnika WebSphere Portal

Če namestite Lotus Quickr na strežnik WebSphere Portal različice 6, lahko s pajkom seznama semen pajkate dokumente knjižnice Lotus Quickr. Možnosti za pajkanje teh dokumentov je mogoče konfigurirati neodvisno od možnosti, ki jih podate za spletna mesta portala, po katerih pajka pajek WebSphere Portal.

Pred izdelavo pajka seznama semen morate opraviti postopek, s katerim namestite iskanje v podjetju v WebSphere Portal. Če želite vzpostaviti okolje za iskanje v podjetju, zaženite skript (`wp6_install.sh` v sistemu AIX, Linux ali Solaris oziroma `wp6_install.bat` v sistemu Windows), ki je na voljo skupaj z OmniFind Enterprise Edition na iskalnih strežnikih.

Posredniški niz uporabnika sporoča, kateri brskalnik ali robot dostopa do strežnika. Pri pajkanju strežnika Lotus Quickr pajek seznama semen uporablja posredniški niz uporabnika OmniFind SeedlistCrawler/1.0.

Pregled konfiguriranja

S pajkom seznama semen je mogoče pajkati poljubno število dokumentov Lotus Quickr. Ko konfigurirate pajka, podate tudi strežnik, ki ga želite pajkati. Pajek nato pajka vse dokumente v knjižnicah Lotus Quickr, ki se nahajajo na tem strežniku.

Dostop do dokumentov, ki bodo pajkani, mora biti mogoč z enakim ID-jem in geslom skrbnika Lotus Quickr. Če želite pajkati spletna mesta, ki uporabljajo druge poverilnice, morate konfigurirati ločenega pajka seznama semen.

Če želite izdelati ali spremeniti pajka seznama semen, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka določajo način, na katerega pajek pajka dokumente Lotus Quickr.
- Podajanje strežnika, ki bo pajkan, in informacij, ki pajku omogočajo vzpostavitev povezave s strežnikom.

Pri izdelavi ali urejanju pajka lahko preizkusite njegovo zmožnost vzpostavitve povezave z dokumenti, ki bodo pajkani. V sporočilih je tako še pred zagonom pajka navedeno, ali ta lahko dostopi do dokumentov, ki jih želite pajkati.

- Podajanje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo uporabo nadzora dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi seznamami za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Pomembno: Če želite preiskati zaščitene dokumente Lotus Quickr, morate predložiti iskanja, pri tem pa uporabiti portalski programček za iskanje, ki omogoča iskanje v podjetju s portala WebSphere Portal. Iskanja, ki jih predložite iz vzorčne iskalne aplikacije, ESSearchApplication, ne bodo imela ustreznih poverilnic, zato sistem ne bo mogel preveriti pooblastil uporabnika za dostop do dokumentov.

- Podajanje informacij, ki omogočajo komuniciranje pajka z strežnikom proxy, če z njim strežete strani.
- Če za zaščito strežnika WebSphere Portal in dokumentov Lotus Quickr uporabljate drug izdelek (na primer IBM Tivoli Access Manager WebSEAL ali CA SiteMinder SSO Agent for PeopleSoft), podajte poverilnice enotne prijave, ki pajku omogočajo dostop do dokumentov na strežniku.
- Podajanje informacij o datoteki shrambe ključev, tako da lahko pajek za vzpostavitev povezave s strežnikom uporabi protokol SSL (Secure Sockets Layer).
- Podajanje jezika in kodne strani dokumentov za pajkanje.
- Podajanje možnosti za pajkanje in iskanje metapodatkov v dokumentih Lotus Quickr.
- Podajanje razporedov za pajkanje dokumentov Lotus Quickr.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezani pojmi

“Integracija s portalom WebSphere Portal” na strani 311

S tem povezana opravila

“Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 6” na strani 318

“Nastavitev portalskega programčka za iskanje v podjetju za Lotus Quickr” na strani 323

Pajki datotečnega sistema UNIX

Če želite v zbirko iskanja v podjetju vključiti dokumente, ki so shranjeni v datotečnih sistemih AIX, Linux ali Solaris, morate konfigurirati pajka datotečnega sistema UNIX.

Pajka datotečnega sistema UNIX lahko uporabite za pajkanje poljubnega števila datotečnih sistemov. Pri konfiguriranju pajka morate izbrati lokalne in oddaljene imenike in podimenike, ki jih želite pajkati.

Če namestite strežnik pajkanja na računalnik Windows, tega strežnika ne morete uporabiti za pajkanje virov datotečnega sistema AIX, Linux ali Solaris (pajek datotečnega sistema UNIX ni na seznamu razpoložljivih tipov pajkov).

Pajek datotečnega sistema UNIX pajka v dokumentih v skladu s pooblastili za branje, ki so podana za skrbnika iskanja v podjetju.

Če želite izdelati ali spremeniti pajka datotečnega sistema UNIX, se prijavite na skrbniško konzolo iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo način pajkanja vseh podimenikov v prostoru za pajkanje.
- Nastavitev razporeda za pajkanje datotečnih sistemov.
- Izbira podimenikov in ravni podimenikov, v katerih želite, da pajka pajek.
- Podajanje možnosti za omogočitev dokumentov v podimenikih za iskanje. Tako lahko na primer iz prostora pajkanja izključite določene tipe dokumentov.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali varnostnih žetonih.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

Spletni pajki

Če želite v zbirko sistema iskanja v podjetju vključiti strani s spletnih mest, morate konfigurirati spletnega pajka.

Za pajkanje poljubnega števila strežnikov HTTP (Hypertext Transfer Protocol) in strežnikov zaščitene HTTP (HTTPS) lahko uporabite spletnega pajka. Ta obiše spletno mesto in tam prebere podatke. Nato sledi povezavam v dokumentih in tako pajka dodatne dokumente. Spletni pajek lahko pajka in povzame povezave iz posameznih strani ali *skupine okvirjev* (strani, izdelane iz okvirjev HTML).

Pajkani podatki so lahko v kateremkoli izmed splošnih formatov in so lahko iz različnih virov znotraj intraneta ali z interneta. Splošni formati vključujejo HTML, PDF, Microsoft Word, Lotus WordPro, XML (Extensible Markup Language) in tako naprej.

Nasvet:

Podrobnejše informacije o načinih konfiguriranja spletnega pajka najdete v scenariju za srednje veliko organizacijo, predstavljenem v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Če želite izdelati ali spremeniti spletnega pajka, se prijavite na skrbniško konzolo sistema iskanja v podjetju. Prav tako morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo način pajkanja vseh spletnih strani v prostoru za pajkanje.
- Podajanje pravil za dovoljene in prepovedane obiske spletnih mest. Pri specificiranju pravil za pajkanje lahko le-ta preskusite in s tem preverite, ali pajek lahko dostopi do spletnih mest, ki jih nameravate vključiti v prostor za pajkanje.
- Podajanje možnosti za vključitev določenih tipov datotek in izključitev datotek z določenimi priponami.
- Podajanje pravil, ki jih spletni pajek uporabi v primeru strani s programskimi napakami.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali varnostnih žetonih.

- Podajanje možnosti za pajkanje z geslom zaščiteneh spletnih mest (spletni strežniki, ki jih želite pajkati, morajo uporabljati osnovno overjanje HTTP ali obrazce HTML za prikaz poziva za geslo).
- Podajanje možnosti za pajkanje spletnih mest, ki jih streže strežnik proxy.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

Konfiguriranje uporabniškega posrednika

Če želite pajkati spletno mesto, ki uporablja protokol za izključevanje robotov, zagotovite, da datoteka `robots.txt` v spletnem mestu imenu uporabniškega posrednika, ki ste ga konfigurirali za spletnega pajka, omogoča dostop do spletnega mesta.

Pri zagonu sistema iskanja v podjetju bo spletni pajek naložil ime uporabniškega posrednika, ki ste ga konfigurirali. Pajek bo pred snemanjem strani s spletnega mesta, ki je prej še ni obiskal (ali je ni obiskal že nekaj časa), najprej poskusil prenesti datoteko, imenovano `robots.txt`. Ta datoteka je v korenskem imeniku spletnega mesta.

Če datoteka `robots.txt` ne obstaja, bo spletno mesto odprto za neomejeno pajkanje. Če datoteka obstaja, podaja področja spletnega mesta (imenikov), ki pajkom niso na voljo. Datoteka `robots.txt` podaja dovoljenja za pajke s tem, da identificira imena njihovih uporabniških posrednikov.

Protokol za izključevanje robotov je neobvezen, vendar mu spletni pajek sistema iskanja v podjetju poskuša ustrezati:

- Če datoteka `robots.txt` vsebuje vnos za ime uporabniškega posrednika, ki je konfigurirano za spletnega pajka, bo ta ustrezal omejitvam na tem uporabniškem posredniku.
- Če ime uporabniškega posrednika v datoteki `robots.txt` ni navedeno, zadnji vnos pa podaja `User-agent: *` (ki pomeni poljubnega uporabniškega posrednika), omejitev pa je `Disallow: /` (ki pomeni prepoved kakršnegakoli pajkanja z začetkom v korenskem imeniku spletnega mesta), spletnemu pajku pajkanje tega spletnega mesta ne bo dovoljeno.
- Če ime uporabniškega posrednika v datoteki `robots.txt` ni navedeno, zadnji vnos pa podaja `User-agent: *`, omejitev pa je `Allow: /`, bo spletni pajek lahko pajkal to spletno mesto.

Skrbniki spletnih mest pogosto podajo zadnji vnos, ki prepoveduje dostop vsem pajkom, ki jim dostop ni izrecno dovoljen. Če konfigurirate novega spletnega pajka in veste, da nekatera spletna mesta, ki jih želite pajkati, uporabljajo protokol za izključevanje robotov, se obrnite na skrbnike spletnih mest in jih prosite, naj v svoje datoteke `robots.txt` dodajo vnos za vašega pajka.

Prepričajte se, da ste v lastnostih spletnega pajka in v vseh datotekah `robots.txt`, ki pripadajo zelenim spletnim mestom, podali isto ime uporabniškega posrednika.

Če nobeno od spletnih mest, ki jih želite pajkati, ne uporablja protokola za izključevanje robotov, vrednost, ki jo podate za lastnost uporabniškega posrednika, običajno ni pomembna. Vendar pa bodo nekateri aplikacijski strežniki, JSP-ji in strežniški programčki svoje odzive prilagodili imenu uporabniškega posrednika. Za obravnavanje nezdružljivosti brskalnikov obstajajo na primer različni odzivi. Ime uporabniškega posrednika, ki ga podate za spletnega pajka, je lahko v teh situacijah pomembno, ne glede na protokol za izključevanje robotov. Če želite pajkati te tipe spletnih mest, se obrnite na njihove skrbnike in zagotovite, da ima spletni pajek dostop.

Kako spletni pajek uporablja protokol za izključevanje robotov

Spletni pajek poskuša biti usklajen s protokolom za izključevanje robotov in ne pajka spletnih mest, če tega ne dopušča datoteka robots.txt na strežniku.

Do uspešnega snemanja (download) pride, če pajek iz spletnega strežnika uspe priklicati datoteko robots.txt ali potrditi, da datoteka robots.txt ne obstaja. Snemanje je obravnavano kot neuspešno, če pajek ne more pridobiti pravil ali ne more potrditi, da datoteka robots.txt obstaja.

Uspešno snemanje ne pomeni, da ima pajek dovoljenje za pajkanje, saj lahko pravila v datoteki robots.txt prepovedujejo pajkanje. Neuspeh pri snemanju začasno prepove pajkanje, saj pajek ne more ugotoviti, kakšna so pravila.

Sledijo koraki, ki jih pajek izvede pri poskusu snemanja datoteke robots.txt:

1. Ko pajek odkrije novo spletno mesto, poskusi pridobiti naslov IP strežnika. Če ta poskus ne uspe, pajkanje ni možno.
2. Če je na voljo vsaj en naslov IP, pajek s pomočjo metode **HTTP (ali HTTPS) GET** poskusi prenesti datoteko robots.txt.
3. Če čakalni čas povezave z vtičnico poteče, je ta prekinjena ali pa pride do druge napake nižje ravni (kot je na primer težava s potrdilom SSL), pajek zabeleži težavo in poskus ponovi na vsakem naslovu, znanem za ciljni strežnik.
4. Če povezave ni mogoče vzpostaviti po poskusu vzpostavitve z vsemi naslovi, pajek počaka dve sekundi in nato še enkrat poskusi vzpostaviti povezavo z vsemi naslovi.
5. Če vzpostavi povezavo in pride do izmenjave oglavij HTTP, bo sistem preveril vrnjeni status. Če je statusna koda enaka 500 ali višja, pajek to obravnava kot neuspelo povezavo in nadaljuje s poskušanjem na drugih naslovih IP. Pri drugih statusih preneha s poskušanjem na nadomestnih naslovih IP in nadaljuje glede na statusno kodo.

Ko pajek sprejme statusno kodo HTTP, nižjo od 500, oziroma ko dvakrat preskusi vse naslove IP, nadaljuje na naslednji način:

1. Če ne prejme nobenega statusa HTTP, nižjega od 500, bo spletno mesto izključeno za ta čas.
2. Če sistem prejme status HTTP enak 400, 404 ali 410, bo spletno mesto označeno za pajkanje brez pravil.
3. Če sistem prejme status HTTP od 200 do 299, na naslednje dejanje vplivajo naslednji pogoji:
 - če je bila vsebina okrnjena, bo spletno mesto za ta čas izključeno
 - če je bila vsebina razčlenjena brez napak, bo spletno mesto označeno za pajkanje z najdenimi pravili
 - če je bila vsebina razčlenjena z napakami, bo spletno mesto označeno za pajkanje brez pravil
4. Če sistem prejme katerikoli drug status HTTP, bo spletno mesto za ta čas izločeno.

Ko pajek za spletno mesto poskusi prenesti datoteko robots.txt, zanj posodobi trajni časovni žig, imenovan robotski datum. Če je bilo spletno mesto zaradi nerazpoložljivosti informacij robots.txt izločeno, se bo števec napak trajnih robotov povečal.

Ko je dosežen interval vnovičnega poskusa, pajek za neuspelo spletno mesto znova poskusi pridobiti informacije robots.txt. Če število zaporednih neuspehov doseže največje število dovoljenih neuspehov, pajek ne bo več poskušal z iskanjem datoteke robots.txt za spletno mesto in ga bo izključil iz pajkanja.

Potem ko je stran primerna za pajkanje (pravila za preverjanje datoteke robots.txt uspejo), se števec napak nastavi na nič. Pajek bo rezultate snemanja uporabljal vse, dokler ne preteče interval za preverjanje pravil. Takrat mora spletno mesto znova označiti za pajkanje.

Nasvet:

- Če strežnik vrne vsebino, ki vsebuje skladišne napake, ali pa strežnik uporablja protokol za robote, ki ni različice 1994, ali pa vsebina vključuje nekaj drugega kot pravila za robote (na primer stran s programsko napako), bo pajek deloval, kot da strežnik nima uporabne datoteke pravil, in bo pajkal spletno mesto. To dejanje je običajno pravilno, saj skrbniki zbirke ne krmilijo vsebine spletnega mesta ali privzetega vedenja strežnika. Če skrbnik spletnega strežnika ne želi pajkanja le-tega in ne želi namestiti ustrezne datoteke pravil, lahko spletnemu pajku prepreči dostop do spletnega mesta tako, da v pravilih za pajke poda domeno, naslov IP ali predpono HTTP spletnega mesta.
- Če strežnik prejme statusno kodo 302 ali drugo preusmeritveno kodo, jo bo pajek obravnaval, kot da spletno mesto vsebuje datoteko robots.txt, ki bi jo moral uporabiti, vendar ta ni na ustreznem mestu (korenskem imeniku spletnega mesta). Skrbnik spletnega strežnika mora datoteko premakniti na pravo mesto, da bo lahko spletni pajek upošteval pravila v datoteki.
- Če pride do težav s potrdili (ko je na primer potrdilo že poteklo, služba za pooblastila ni overjena ali pa je potrdilo lastnoročno podpisano, pajek pa ni konfiguriran za sprejem lastnoročno podpisanih potrdil), bo pajek težavo obravnaval kot neuspeh pri vzpostavitvi povezave s spletnim mestom, le-to pa bo izključil. Sicer pa bi iste težave najverjetneje tudi preprečile pajkanje drugih strani tega spletnega mesta. Če želite omogočiti pajkanje spletnega mesta, mora skrbnik zbirke omogočiti lastnoročno podpisana potrdila, službo za pooblastila spletnega mesta dodati v datoteko overjene shrambe ključev ali pa prositi skrbnika spletnega strežnika, da pridobi ažurno potrdilo.
- Spletnega pajka je mogoče konfigurirati za uporabo osnovnega overjanja HTTP (vključno z osnovnim overjanjem proxy HTTP). Če ste pravilno konfigurirali sistem, bo ta za prenos datotek robots.txt od uporabnikov zahteval, da se overijo. Statusna koda 403, 407 ali drugi z overjanjem povezani odzivi uporabnika sporočijo, da je prišlo do težave pri overjanju, pajek pa izključi stran. (Podprto je samo osnovno overjanje HTTP.)
- Če datoteka robots.txt za spletno mesto preseže največjo dolžino strani robotov, lahko skrbnik zbirke poveča konfigurirano največjo vrednost (privzeta vrednost enega milijona bajtov bi morala zadostovati).

Za pomoč pri odpravljanju težav lahko med nadzorom spletnega pajka zahtevate poročilo o spletnih mestih. Izberite možnosti za prikaz vsebine datoteke robots.txt (ali želite prikazati pravila, ki spletnemu pajku prepovedujejo dostop do spletnega mesta), prikaz datuma in časa zadnjega poskusa snemanja datoteke robots.txt (pajek bo znova poskusil šele, ko preteče interval za vnovični poskus) in prikaz števila zaporednih poskusov snemanja datoteke robots.txt, ki niso uspela. Med nadzorom spletnega pajka kliknite možnost **Pomoč**, spoznajte te možnosti poročila o spletnih mestih in se naučite razlagati rezultate.

Za dodatne informacije o protokolu za izključevanje robotov preglejte naslednji URL:
<http://www.robotstxt.org/wc/exclusion.html>.

Podpora za JavaScript

Spletni pajek sistema iskanja v podjetju lahko najde nekatere povezave (URL-je), ki so vsebovani v delih JavaScript spletnih dokumentov.

Spletni pajek lahko najde relativne in absolutne povezave. Če dokument HTML vsebuje element BASE, bo pajek z njegovo pomočjo razrešil relativne povezave. V nasprotnem primeru bo uporabil lastni URL dokumenta.

Podpora za JavaScript je omejena na povzemanje povezav. Pajek ne razčlenjuje JavaScript, ne gradi DOM-a (Document Object Model) in ne prevaja ali izvaja stavkov JavaScript. Pajek poišče nize v vsebini dokumenta (vključno z, vendar ne omejeno na dele JavaScript), ki so v stavkih JavaScript najverjetneje URL-ji. To pomeni dve stvari:

- Sistem bo našel nekatere URL-je, ki jih strožji razčlenjevalnik HTML prezre. Pajek bo zavrnil vse, kar ni skladiščno veljaven URL, vendar pa so nekateri veljavni URL-ji, vrnjeni v koraku pregledovanja, vsaj malo zanimivi za iskanje.
- Vsebine dokumenta, ki jo ustvari JavaScript, ko na primer uporabnik z brskalnikom prikaže stran, ta pa izvede določeni JavaScript, spletni pajek ne more odkriti, zato ne bo indeksirana.

Ker spletni pajek ne razčlenjuje JavaScripta v datotekah HTML, URL-ji v JavaScriptu niso pajkani. Če želite omogočiti spletnega pajka za pajkanje URL-jev v JavaScriptu, lahko izvedete eno od naslednjih dejanj:

- V konzoli za skrbništvo nad iskanjem v podjetju uredite spletnega pajka, nato pa na strani Spletni prostor za pajkanje dodajte URL-je na seznam URL-jev, ki jih pajek uporablja kot začetno točko za dodajanje URL-jev v zbirko (**Začetni URL-ji**). Če želite, da spremembe stopijo v veljavo, znova zaženite spletnega pajka (ni vam treba zagnati celotnega pajkanja).
- S sidrno oznako (``) podajte URL-je kot hiperpovezave v datoteki HTML.

Pravila za omejevanje prostora za spletno pajkanje

Če želite zagotoviti, da uporabniki lahko dostopijo samo do spletnih mest, ki ste jih omogočili za iskanje, morate podati pravila za omejitev vsebin, ki jih spletni pajek lahko pajka.

Ko spletni pajek pajka spletno stran, odkrije povezave na druge strani in jih nato shrani v čakalno vrsto za nadaljnje pajkanje. Pajkanje in odkrivanje se lahko ponavlja tako dolgo, dokler to omogočajo čas in pomnilniški viri. Pri konfiguriranju spletnega pajka morate podati, kje naj pajek začne s pajkanjem. Spletni pajek lahko s teh prvotnih URL-jev (imenovanih *začetni URL-ji*) pride do vsakega dokumenta na spletu, ki je povezan prek posrednih ali neposrednih povezav.

Če želite omejiti prostor za pajkanje, spletnega pajka konfigurirajte tako, da bo temeljito pajkal samo določene URL-je, povezave, ki kažejo izven zanimivega področja, pa bo prezrl. Ker pajek po privzetku sprejme vsak URL, ki ga odkrije, morate podati pravila, ki določajo, katere URL-je naj vključi v zbirko, obenem pa izločijo preostale strani.

Vsebino, ki naj jo spletni pajek pajka, lahko podate na več načinov. Podate lahko naslednje:

- seznam začetnih URL-jev, kjer naj pajek začne s pajkanjem;
- trije tipi pravil za pajkanje: domena, naslov IP (Internet Protocol) in predpona URL;
- seznam tipov MIME za dokumente, ki jih želite vključiti;
- seznam pripon datotek dokumentov, ki jih želite izključiti;
- največje število imenikov v poti URL-ja.

Pravila za pajkanje so v naslednji obliki:

dejanje tip cilj

dejanje prepoveduje (forbid) ali dopušča (allow); tip je lahko domena, naslov IP ali predpona URL-ja (HTTP ali HTTPS); cilj pa je odvisen od vrednosti tipa. Kot univerzalni znak lahko, sicer omejeno, podate zvezdico (*) in tako podate cilj, ki se ujema z vzorcem.

Domenska pravila

Cilj domenskega pravila je ime domene DNS. Za pajkanje lahko na primer podate celotno domeno `www.ibm.com`:

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

Kot univerzalni znak lahko podate zvezdico, s tem pa bo pravilo veljalo za vsa imena gostiteljev, ki se ujemajo s preostalim delom vzorca. Podate lahko na primer, da sistem ne bo pajkal nobene domene, ki se začne s server in konča z ibm.com:

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

Ujemanje imen gostiteljev razlikuje med velikimi in malimi črkami, ne glede na to, ali podate izrecno ime domene ali vzorec imena domene. Tako se *.user.ibm.com na primer ujema z joe.user.ibm.com in mary.smith.user.ibm.com, ne pa tudi z joe.user.IBM.com.

Domenska pravila, ki ne podajajo številke vrat, veljajo za vsa vrata v tej domeni. V naslednjem zgledu bodo dovoljena vsa vrata v domeni sales:

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

Če domensko pravilo podaja številko vrat, bo veljalo samo za ta vrata. V naslednjem zgledu bodo dovoljena samo vrata 443 v domeni sales:

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

Pravila predpone

Pravilo predpone krmili pajkanje URL-jev, ki se začnejo s podanim nizom. Cilj je posamezni URL, ki običajno vsebuje eno ali več zvezdic in tako označuje vzorec. Zvezdica je na primer pogosto podana kot zadnji znak v nizu predpone.

S pomočjo pravila za predpone lahko pajkate celotno spletno mesto ali le njegov del. Podate lahko pot imenika ali vzorec in nato omogočite ali onemogočite vse v drevesu imenikov, kar je od te točke naprej. Naslednja pravila na primer delujejo skupaj, da pajku dopuščajo, da pajka vse v imeniku public na naslovu sales.ibm.com, prepovedujejo pa mu dostop do drugih strani spletnega mesta:

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

Pri podajanju pravil predpone lahko podate več kot eno zvezdico, poleg tega pa jih lahko podate kjerkoli v nizu predpone, ne samo na zadnjem mestu. Naslednje pravilo pajku na primer prepoveduje pajkanje vseh dokumentov v imeniku najvišje ravni v spletnem mestu sales.ibm.com, če se ime imenika zaključuje s fs. (Lahko imate na primer vpetja datotečnega sistema, ki ne vsebujejo informacij, ki bi bile uporabne v indeksu za iskanje.)

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

Pravila za naslove

Pravilo za naslove omogoča nadzor nad pajkanjem celotnih gostiteljev ali omrežij tako, da kot cilj poda naslov IP in omrežno masko. Primer:

```
IPv4 allow address 9.0.0.0 255.0.0.0
```

IPv6 Če zaženete iskanje v podjetju na strežniku Windows 2003 in omogočite sistem iskanja v podjetju za uporabo protokola IP različice 6 (IPv6), morate naslov zapisati v oklepajih.

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

Maska omrežja omogoča, da podate ujemanje vzorca. Če želite pravilo za naslove uveljaviti za potencialni naslov IP, se morata naslov IP v pravilu in potencialni naslov IP povsem ujemati, razen v primeru, kjer maska omrežja vsebuje ničle. Pravilo za naslove definira

vzorec, maska omrežja pa značilne bite v vzorcu naslova. Ničla v maski omrežja deluje kot univerzalni znak in označuje, da se z naslovom ujema katerakoli vrednost, ki je v naslovu podana na istem bitnem mestu.

V predhodnem primeru bo pravilo za dovoljenje veljalo za vse naslove IP, ki imajo v prvem oktetu 9 in poljubno vrednost v ostalih treh okteti.

Priporočamo, da naslednje pravilo vstavite kot končni naslov na seznam pravil za naslove. To pravilo se ujema z vsemi naslovi IP, ker so zaradi maske omrežja vsi biti nepomembni (pravilo prepoveduje vse naslove, ki niso dovoljeni s predhodnimi pravili na seznamu pravil).

IPv4 forbid address 0.0.0.0 0.0.0.0

IPv6

forbid address :: ::

Omejitve za strežnike proxy: Če nameravate pajkati spletna mesta, ki jih streže strežnik proxy, ne podajajte pravil za naslove IP. Strežnik proxy se običajno uporablja v primeru, ko uporabniški posrednik (brskalnik ali pajek) nimata neposrednega dostopa do omrežij, kjer so spletni strežniki. Strežnik proxy HTTP lahko na primer zahteve HTTP iz pajka posreduje spletnemu strežniku, odzive pa nato pošlje nazaj pajku.

Če spletni pajek uporablja strežnik proxy, je naslov IP strežnika proxy edini naslov IP, ki ga ima pajek za drugega gostitelja. Če za omejitev pajka na podmrežo naslovov IP uporabite pravila za naslove IP, bodo zaradi omejitve skoraj vsi URL-ji označeni s statusno kodo 760 (ki označuje, da jih spletni prostor prepoveduje).

Vrstni red pravil za pajkanje

Pajek uporabi pravila za pajkanje ob različnih trenutkih postopka odkrivanja in pajkanja URL-jev. Vrstni red pravil je pomemben, vendar le znotraj pravil posameznega tipa. Pomembno je, ali je pravilo za naslove pred ali za drugim pravilom za naslove, ni pa pomembno, ali je pravilo za naslove pred ali za pravilom predpone, saj pajek ne uporabi vseh pravil naenkrat.

Pajek bo znotraj niza pravil določenega tipa preveril ciljno domeno, naslov ali URL, ga primerjal s posameznim pravilom, od prvega podanega pravila do zadnjega, dokler ne najde pravila, ki ga lahko uporabi. Uporabil bo dejanje, podano za prvo pravilo, ki ga je mogoče uveljaviti.

Odvisnost vrstnega reda tako določa običajno strukturo večine pravil za pajkanje:

- Niz domenskih pravil se običajno začne s pravili za prepoved, ki iz prostora za pajkanje izločijo posamezne domene. Skrbnik zbirke lahko na primer ugotovi, da določene domene ne vsebujejo uporabnih informacij.
- Seznamu pravil za prepoved običajno sledi niz pravil za dovoljenje (z univerzalnimi znaki), ki pajku omogočajo, da obiše vse domene, ki se končajo z enim izmed imen domen višje ravni, ki definirajo intranet podjetja (kot sta na primer *.ibm.com in *.lotus.com).

Niz domenskih pravil zaključite z naslednjim privzetim pravilom, ki izloči domene, ki niso bile dovoljene s predhodnimi pravili:

```
forbid domain *
```

To zadnje pravilo je pomembno, saj preprečuje, da bi v prostor za pajkanje vključili celotni internet.

- Niz pravil za naslove se običajno začne z manjšim številom pravil za dovoljenje, ki omogočajo pajkanje omrežij višje ravni (razred A, razred B ali razred C), ki obsegajo intranet podjetja.

Preberite prejšnjo razpravo o pravilih za naslove, v kateri najdete primere za podajanje končnega pravila na seznamu pravil za naslove, s čimer pajku preprečite pajkanje spletnih mest, ki so izven omrežja podjetja.

- Niz pravil predpone je običajno največji, saj vsebuje poljubno podrobne specifikacije dovoljenih in prepovedanih področij, ki so podana kot drevesa in poddrevesa. Eden izmed dobrih pristopov je, da najprej dopustite ali prepoveste strožje lokalizirana področja, nato pa podajte nasprotno pravilo z bolj splošnim vzorcem, ki dopušča ali prepoveduje vse ostalo.

Razdelek predpon se običajno ne konča s splošnim pravilom. Priporočena končna domenska pravila in pravila za naslove zagotavljajo, da pajek izven omrežja podjetja ne pajka bolj učinkovito kot s preskušanjem predpon URL.

Pajek lahko pravila predpon uveljavi učinkoviteje, če jih združite po dejanjih (prepoved ali dovoljenje). Namesto da na primer podate kratka zaporedja pravil za dovoljenje in prepoved, ki se izmenjujejo, podajte daljše zaporedje pravil, ki uporabljajo eno dejanje, nato pa podajte daljše zaporedje pravil, ki uporabljajo drugo dejanje. Pravila za dovoljenje in prepoved lahko prepletete tako, da dosežete želeni učinek za vaš prostor za pajkanje. Vendar pa lahko z združevanjem pravil za dovoljenje in združevanjem pravil za prepoved izboljšate zmogljivost pajka.

Pripone datotek, tipi MIME in največja globina pajkanja

Te možnosti nudijo dodatne načine, s katerimi lahko podate vsebino za prostor za pajkanje. Določene tipe dokumentov lahko izključite glede na pripono datoteke dokumenta, določene tipe dokumentov pa lahko vključite glede na tip MIME dokumenta. Pri podajanju tipov MIME, ki naj jih pajek pajka, upoštevajte, da je tip MIME v spletnih dokumentih pogosto nepravilno nastavljen.

Največja globina pajkanja je število poševnic v URL-ju od korenskega imenika spletnega mesta. S to možnostjo lahko preprečite, da bi pajek zašel v rekurzivne strukture datotečnega sistema z neskončno globino. Globina pajkanja ne ustreza ravnem, ki jih pajek prečka med sledenjem povezav od enega do drugega dokumenta.

Začetni URL-ji

Začetni URL-ji so URL-ji, na katerih pajek začne s pajkanjem, poleg tega pa so ti URL-ji v pajkanje vstavljeni vsakič, ko zaženete pajka. Če so bili začetni URL-ji že odkriti, ne bodo pajkani ali znova pajkani pred ostalimi spletnimi mesti, ki ste jih dopustili v pravilih za pajkanje.

Začetni URL je pomemben prvič, ko zaženete spletnega pajka in je prostor za pajkanje prazen. Prav tako je pomemben, ko URL, ki prej še ni bil odkrit, dodate na seznam začetnih URL-jev v prostoru za pajkanje.

Začetni URL-ji morajo biti popolni URL-ji in ne samo imena domen. Podati morate protokol in, če vrata niso enaka 80, tudi številko vrat.

Naslednji URL-ji so veljavni začetni URL-ji:

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

Naslednji URL ni veljaven začetni URL:

`www.ibm.com`

Začetne URL-je morate vključiti v pravila za pajkanje. Pajek ne more na primer začeti pajkanja s podanim začetnim URL-jem, če pravila za pajkanje ne dopuščajo pajkanja tega URL-ja.

Podpora za naslove IPv6: Če zažene iskanje v podjetju na strežniku Windows 2003 in omogočite sistem iskanja v podjetju za uporabo protokola IP različice 6 (IPv6), morate začetne URL-je zapisati v oklepaju. Primer:

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

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

S tem povezana opravila

“Omogočitev podpore za protokol IPv6” na strani 25

Preskušanje povezav z URL-ji s spletnim pajkom

Ko podate URL-je, ki naj jih spletni pajek pajka, lahko preverite konfiguracijo pravil za pajkanje.

Pri podajanju domen, predpon HTTP ali naslovov IP za pajkanje lahko kliknete možnost **Preskusi** ali pa izberete stran Preskus URL-jev, kjer lahko preverite, ali pajek lahko vzpostavi povezavo z začetnimi URL-ji in s podanimi URL-ji.

Rezultati preskusa kažejo, ali je pajek uspel dostopiti do URL-jev z imenom uporabniškega posrednika, ki je podano v lastnostih pajka. Rezultati preskusa pokažejo tudi, ali je URL mogoče pajkati zaradi pravil o izključevanju (ko na primer dokumenta morda ni mogoče pajkati, ker ima pripono datoteke, ki se ujema z v prostoru za pajkanje izključeno pripono).

Ko na spletnem mestu vsaj enkrat izvedete pajkanje, lahko preskusite URL-je in tako pridobite dodatne informacije. Poročilo o preizkusu na primer lahko vsebuje najnovejšo statusno kodo HTTP (ki označuje, ali je pajkanje URL-ja uspelo), prikaže, kdaj je bil URL zadnjič pajkan in kdaj je razporejen za vnovično pajkanje, prav tako pa lahko prikaže, ali uporabniški posrednik uporablja ažurno datoteko robots.txt spletnega strežnika.

Nastavitve intervala vnovičnega pajkanja v spletnem pajku

Če želite vplivati na to, kako pogosto spletni pajek obiše URL-je, lahko podate možnosti v lastnostih spletnega pajka.

Večina drugih tipov pajkov v sistemu iskanja v podjetju teče po razporedih, ki jih poda skrbnik. Nasprotno pa spletni pajek po zagonu deluje stalno. Če želite krmiliti, kako pogosto obiše URL-je, ki jih je prej pajkal, podajte najkrajši in najdaljši interval za vnovično pajkanje.

Če za izdelavo spletnega pajka ali urejanje lastnosti le-tega uporabite skrbniško konzolo sistema iskanja v podjetju, lahko izberete možnost za konfiguriranje zahtevnejših lastnosti. Na strani Zahtevnejše lastnosti spletnega pajka lahko podate možnosti za najkrajši in najdaljši interval vnovičnega pajkanja. Spletni pajek bo s pomočjo podanih vrednosti izračunal interval za vnovično pajkanje podatkov.

Pajek ob prvem pajkanju strani uporabi datum in čas pajkanja ter povprečje podanih najkrajšega in najdaljšega intervala za vnovično pajkanje in tako nastavi datum vnovičnega pajkanja. Stran pred tem datumom ne bo znova pajkana. Čas, ko bo stran vnovič pajkana po tem datumu, je odvisen od obremenitve pajka in uravnoteževanja novih in starih URL-jev v prostoru za pajkanje.

Pajek bo ob vsakem vnovičnem pajkanju strani preveril, ali se je vsebina spremenila. Če se je spremenila, bo naslednji interval za vnovično pajkanje krajši od predhodnega, nikoli pa ne bo

krajši od podanega najkrajšega intervala za vnovično pajkanje. Če se vsebina ni spremenila, bo naslednji interval za vnovično pajkanje daljši od predhodnega, nikoli pa ne bo daljši od podanega najdaljšega intervala za vnovično pajkanje.

Možnosti za obiskovanje URL-jev s spletnim pajkom

Spletnega pajka lahko prisilite, da določene URL-je obišče takoj, ko je mogoče.

Če želite prostor za pajkanje osvežiti z informacijami iz določenih spletnih mest, lahko nadzirate pajka, izberete možnost **URL-ji za obisk ali vnovični obisk** in nato podate URL-je ali vzorce URL-jev strani, ki jih želite pajkati ali vnovično pajkati.

Če na primer oddelek Komunikacije doda spletno stran na intranet ali preuredi določeno stran in nato izvede pomembno spremembo načel, lahko podate URL nove ali spremenjene strani. Če pajek deluje, bo podani URL uvrstil v čakalno vrsto in ga pajkal ob naslednjem preverjanju strani, ki čakajo na obisk (običajno vsakih deset minut). Če pajek trenutno ne deluje, bo podani URL v čakalno vrsto uvrstil tako, da ga bo pajkal ob naslednjem zagonu.

Prepričajte se, da pravila za pajkanje vključujejo pravilo, ki pajku omogočajo obisk URL-jev, ki jih podate. Pajek lahko podane URL-je obišče prej kot bi jih sicer. Vendar pa mora za pajkanje URL-ja najprej obstajati pravilo, ki omogoča pajkanje.

Novo pajkani podatki bodo na voljo za iskanje po naslednji gradnji glavnega indeksa.

Kako spletni pajek obravnava strani s programskimi napakami

Spletnega pajka lahko konfigurirate za obravnavanje prilagojenih strani, ki jih skrbniki spletnih strežnikov izdelajo, če kot odziv na zahteve za določene strani ne želijo vrniti standardne kode napake.

Če strežnik HTTP ne more vrniti strani, ki jo zahteva odjemalec, običajno vrne odziv, sestavljen iz oglavja in statusne kode. Statusna koda označuje, kaj je težava (kot na primer napaka 404, ki označuje, da datoteke ni mogoče najti). Nekateri skrbniki spletnih mest izdelajo posebne strani, ki podrobneje pojasnjujejo težavo in strežnik HTTP konfigurirajo tako, da vrne te strani. Te prilagojene strani se imenujejo *strani s programskimi napakami*.

Strani s programskimi napakami lahko popačijo rezultate spletnega pajka. Pajek lahko na primer namesto oglavja, ki označuje težavo, prejme stran s programsko napako in statusno kodo 200, ki označuje uspešno snemanje veljavne datoteke HTML. Vendar pa ta prenesena stran s programsko napako ni povezana z zahtevanim URL-jem, temveč je njena vsebina skoraj vedno enaka vsakič, ko se prikaže namesto zahtevane strani. Te nepomembne in skoraj enake strani popačijo indeks in rezultate iskanja.

Za rešitev te situacije lahko pri konfiguriranju spletnega pajka podate možnosti za obravnavanje strani s programskimi napakami. Spletni pajek potrebuje naslednje informacije o posameznem spletnem mestu, ki vrne strani s programskimi napakami:

- Vzorec URL-ja za spletno mesto, ki uporablja strani s programskimi napakami. Ta vzorec URL-ja je sestavljen iz protokola (HTTP ali HTTPS), imena gostitelja, številke vrat (če niso običajna) in imena poti. V vzorcu za URL-je lahko uporabite zvezdico (*) kot univerzalni znak in s tem poiščete ujemanja enega ali več znakov do naslednje pojavitve neuniverzalnega znaka v vzorcu. Podani vzorec razlikuje med malimi in velikimi črkami.
- Vzorec naziva za besedilo, ki ustreza oznaki <TITLE> dokumenta HTML. V vzorcih lahko kot univerzalni znak uporabite zvezdico (*). Podani vzorec razlikuje med malimi in velikimi črkami.

- Vzorec vsebine za besedilo, ki ustreza vsebini dokumenta HTML. Vsebina ni samo vsebina oznake <BODY>, če je oznaka <BODY> prisotna. Vsebina je vse, kar v datoteki sledi oglavju HTTP. V vzorcih lahko kot univerzalni znak uporabite zvezdico (*). Podani vzorec razlikuje med malimi in velikimi črkami.
- Celó število, ki predstavlja statusno kodo za dokumente, ki se ujemajo s podanimi vzorci URL-jev, nazivov in vsebine.

Zgled

Naslednja konfiguracija spletnemu pajku pove, da mora vse veljavne strani HTML (statusna koda 200), vrnjene s spletnega mesta http://www.mojemesto.com/hr/*, primerjati s podanimi vzorci naziva in vsebine. Če se oznaka <TITLE> strani začne z "Oprostite, stran" in vsebina dokumenta vključuje karkoli (*), jo bo pajek obravnaval na enak način kot statusno kodo 404 (strani ni bilo mogoče najti).

Tabela 3. Zgled strani s programskimi napakami

Vzorec URL-ja	Vzorec naslova	Vzorec vsebine	Statusna koda HTTP
http://www.mojemesto.com/hr/*	Oprostite, stran*	*	404

Za isto spletno mesto lahko izdelate več vnosov, ki bodo obravnavali različne statusne kode. Vsaka statusna koda iz istega spletnega mesta zahteva v konfiguraciji spletnega pajka svoj vnos.

Uporaba univerzalnih znakov

Vzorci URL-jev, nazivov in vsebine niso regularni izrazi. Zvezdica predstavlja ujemanje s katerikoli znakom do naslednje ponovitve kateregakoli neuniverzalnega znaka. Primer:

*404 se ujema s *poljubni znaki*404

404: * se ujema s 404: *poljubni znaki*

http://*.mojemesto.com/* se ujema s <http://poljubni.gostitelj.mojemesto.com/poljubna.datoteka>

* se ujema s *poljubnimi znaki*

Vpliv na zmogljivost

Pri konfiguriranju možnosti za obravnavanje strani s programskimi napakami morate povečati količino časa obdelave pajka, saj mora ta preveriti vse uspešno pajkane strani. Za preverjanje ujemanj vzorcev in ugotovitev, ali naj sistem vrne stran ali nadomestno statusno kodo, je potreben dodaten čas obdelave.

Podpora za pajkanje zaščitenih spletnih mest

S podajanjem poverilnic na skrbniški konzoli sistema iskanja v podjetju lahko spletnemu pajku omogočite dostop do omejene vsebine, kot so na primer dokumenti, ki za dostop zahtevajo geslo.

Če spletni strežnik za omejevanje dostopov do spletnih mest uporablja osnovno overjanje HTTP ali overjanje na osnovi obrazcev HTML, lahko v konfiguraciji spletnega pajka podate poverilnice, s katerimi omogočite pajkanje strani v z geslom zaščitenih spletnih mestih. Prav tako lahko podate možnosti za ročno konfiguriranje datotek s piškotki.

Spletna mesta, zaščitená z osnovnim overjanjem HTTP

Če spletni strežnik za omejevanje dostopa do spletnih mest uporablja osnovno overjanje HTTP, lahko podate poverilnice za overjanje, s katerimi spletni pajek lahko dostopi do z geslom zaščitenih strani.

Večina spletnih strežnikov uporablja shemo overjanja odjemalca, imenovano osnovno overjanje HTTP, s katero ugotovijo identiteto uporabnika in tako ugotovijo, ali ima uporabnik (ali odjemalska aplikacija) dovoljenje za dostop do strani spletnega mesta. Običajno to poteka interaktivno:

- Ko uporabniški posrednik HTTP (kot je spletni brskalnik) zahteva stran, ki je zaščitena z osnovnim overjanjem HTTP, spletni strežnik odgovori s statusno kodo 401, ki označuje, da zahtevnik nima pooblastil za dostop do zahtevane strani.
- Spletni strežnik od zahtevnika tudi zahteva poverilnice, s katerimi lahko preveri, ali je uporabniku dovoljen dostop do omejene vsebine.
- Spletni strežniki uporabniku prikažejo pogovorno okno, kjer morajo vnesti ime uporabnika, geslo in druge informacije, ki so potrebne za predstavitev poverilnic uporabnika.
- Spletni brskalnik poverilnice zakodira in jih nato vključi, ko ponovi zahtevo za zaščitene strani.
- Če so poverilnice veljavne, bo spletni strežnik odgovoril s povratno kodo 200 in vsebino zahtevane strani.
- Nadaljnje zahteve za strani iz istega spletnega strežnika običajno vključujejo iste poverilnice, zaradi česar lahko pooblaščen uporabnik dostop do dodatnih omejenih vsebin brez podajanja dodatnih poverilnic.

Ko spletni strežnik ugotovi identiteto uporabnika, si z uporabniškim posrednikom HTTP izmenjata žetone, imenovane *piškotki*, s katerimi je med zahtevami HTTP mogoče ohraniti status prijave uporabnika.

Ker spletni pajek ne teče interaktivno, morate poverilnice, ki mu omogočajo pajkanje z geslom zaščitene strani, podati pred začetkom pajkanja. Pri izdelavi spletnega pajka ali urejanju prostora za pajkanje podajte informacije o posameznem zaščitenem spletnem mestu, ki ga želite pajkati.

Če želite podati te informacije, se morate obrniti na skrbnike spletnih mest ali spletnih strežnikov, ki so zaščiteni z osnovnim overjanjem HTTP. Ti vam morajo posredovati zahteve za zaščito spletnih mest, ki jih želite pajkati, vključno z vsemi informacijami, namenjenimi za overjanje identitete spletnega pajka in določanju ustreznosti dovoljenj za pajkanje omejenih strani.

Če ste pri izdelavi zbirke omogočili zaščito, lahko podate varnostne žetone, kot so ID-ji uporabnikov, ID-ji skupin ali vloge uporabnikov, ki pri konfiguriranju pajka krmilijo dostop do dokumentov. Spletni pajek bo te varnostne žetone povezal z vsakim dokumentom, ki ga pajka na drevesu datotečnega sistema za podani korenski URL. Sistem bo žetone uporabil poleg varnostnih žetonov na ravni dokumentov, ki jih konfigurirate za celotni prostor za spletno pajkanje.

Vrstni red URL-jev je pomemben. Ko dodate informacije o z geslom zaščitenem spletnem mestu, ga morate postaviti v vrstni red, v katerem naj ga pajek obdela. Najprej navedite bolj specifične URL-je, bolj splošne pa postavite na konec seznama. Spletni pajek pri ocenjevanju ciljnega URL-ja uporabi podatke overjanja, podane za prvi URL na seznamu, ki se ujema s ciljnim URL-jem.

Spletna mesta, zaščitena z overjanjem na osnovi obrazcev

Če spletni strežnik za omejevanje dostopov do spletnih mest uporablja obrazce HTML, lahko podate poverilnice za overjanje, s katerimi lahko spletni pajek dostopi do z geslom zaščitene strani.

Večina spletnih strežnikov uporablja za ugotovitev identitete uporabnika obrazce HTML, s katerimi ugotovijo, ali ima uporabnik (ali odjemalska aplikacija) dovoljenje za dostop do strani spletnega mesta. Običajno to poteka interaktivno:

- Ko uporabniški posrednik HTTP (kot je spletni brskalnik) zahteva stran, ki je zaščitena z overjanjem na osnovi obrazcev, spletni strežnik preveri, ali zahteva vključuje piškotek, ki vzpostavi identiteto uporabnika.
- Če piškotek ni prisoten, spletni strežnik uporabniku prikaže poziv za vnos varnostnih podatkov v obrazec. Ko uporabnik predloži obrazec, spletni strežnik vrne zahtevane piškotke, zahteva za z geslom zaščitene strani pa se lahko nadaljuje.
- Nadaljnje zahteve, ki vključujejo zahtevane piškotke, bodo prav tako omogočene. Pooblaščen uporabnik lahko dostopi do dodatne omejene vsebine brez vnašanja v obrazec in podajanja poverilnic za posamezno zahtevo.

Ker spletni pajek ne teče interaktivno, morate poverilnice, ki mu omogočajo pajkanje z geslom zaščitene strani, podati pred začetkom pajkanja. Pri izdelavi spletnega pajka ali urejanju prostora za pajkanje podajte informacije o posameznem zaščitenem spletnem mestu, ki ga želite pajkati.

Podana polja ustrezajo poljem, ki jih izpolni interaktivni uporabnik, ko spletni brskalnik prikaže poziv, in vsem skritim in statičnim poljem, potrebnim za uspešno prijavo.

Če želite podati te informacije, se morate obrniti na skrbnike spletnih mest ali spletnih strežnikov, ki so zaščiteni z overjanjem na osnovi obrazcev. Ti vam morajo posredovati zahteve za zaščito spletnih mest, ki jih želite pajkati, vključno z vsemi informacijami, namenjenimi za overjanje identitete spletnega pajka in določanju ustreznosti dovoljenj za pajkanje omejenih strani.

Vrstni red vzorcev URL je pomemben. Ko dodate informacije o z geslom zaščitenem spletnem mestu, ga morate postaviti v vrstni red, v katerem naj ga pajek obdela. Najprej navedite bolj specifične vzorce URL-jev, na koncu seznama pa dodajte bolj splošne vzorce URL-jev. Spletni pajek pri ocenjevanju ciljnih URL-jev uporabi podatke obrazca, podane za prvi vzorec URL-ja na seznamu, ki se ujema s ciljnim URL-jem.

Uporaba vtičnika za pajkanje zaščitene strani WebSphere Portal

Če je na strežniku WebSphere Application Server omogočena globalna zaščita in želite s spletnim pajkom pajkati zaščitena spletna mesta portala WebSphere Portal, morate izdelati vtičnik pajka, ki bo obravnaval zahteve za overjanje na podlagi obrazcev. Razpravo o overjanju na osnovi obrazcev in vzorčni program, ki ga lahko priredite za vtičnik spletnega pajka po meri, najdete v temi <http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani>.

Vtičnik je potreben, če s spletnim pajkom pajkate spletna mesta v portalu WebSphere Portal, vključno z mesti Workplace Web Content Management in mesti Lotus Quickr.

Spletna mesta, ki jih strežejo strežniki proxy

Če spletnemu pajku ni dovoljen neposredni dostop do omrežja, lahko konfigurirate pajka, ki za dostop do vsebin za pajkanje uporablja strežnik proxy HTTP.

Če v računalniku, v katerem želite zagnati spletnega pajka, ni mogoče dostopiti do omrežja TCP/IP ali pa je dostop omejen na pooblaščen procese, lahko spletnega pajka konfigurirate tako, da bo uporabil strežnik proxy HTTP. Proxy HTTP je proces, ki na določenih vratih v določenem gostitelju spremlja zahteve HTTP. Strežnik proxy te zahteve nato posreduje spletnemu strežniku, odzive iz spletnega strežnika pa pošlje nazaj odjemalcu (spletnemu pajku), ki jih je zahteval. Strežnik proxy lahko teče v istem računalniku kot spletni pajek, lahko pa teče v drugem računalniku.

Pri pajkanju brez proxyja bo zahteva za URL poslana neposredno gostitelju. Pri pajkanju s proxyjem bo zahteva poslana strežniku proxy.

Pri izdelavi spletnega pajka ali urejanju prostora za pajkanje podajte informacije o strežnikih proxy, ki jih spletni pajek uporablja v primeru, ko so strani za pajkanje v domeni strežnika proxy. Pred dodajanjem strežnika proxy v prostor za pajkanje pridobite imena domen, ki jih streže strežnik proxy, ime gostitelja ali naslov IP strežnika proxy in številko vrat, ki jo strežnik proxy uporablja.

Če strežnik proxy zahteva overjanje, pridobite tudi ime uporabnika in geslo, s katerima bo spletni pajek lahko dostopil do strani, ki jih streže strežnik proxy. Spletni pajek podpira samo osnovno overjanje proxy HTTP, kot je opisano v temi RFC2616 (<http://rfc.net/rfc2616.html>). Drugi tipi overjanj, vključno z Windows NT LAN Manager (NTLM), niso podprti.

Ko dodate strežnik proxy, ga morate izbrati in ga postaviti v vrstni red, v katerem naj ga pajek obdela. Najprej navedite bolj specifična imena domen, na koncu seznama pa dodajte bolj splošna imena domen. Spletni pajek pri ocenjevanju ciljnega URL-ja uporabi podatke strežnika proxy, podane za prvo domeno na seznamu, ki se ujema s ciljnim URL-jem. (Za URL-je, ki se ne ujemajo z nobenim pravilom strežnika proxy, sistem predpostavi, da lahko pajek do njih dostopi neposredno.)

Skrbništvo nad piškotki

Običajno se skrbništvo nad piškotki izvaja samodejno in skrbniku iskanja v podjetju ni potrebno opraviti nobenega dejanja. Če pa je potrebno, lahko ročno podate piškotke za spletno sejo pajkanja.

Piškotki so nerazumljivi žetoni, ki jih vrne spletni strežnik uporabniškemu posredniku kot del oglavja odziva HTTP. Pomen imajo samo za spletni strežnik, ki jih je izdal, in se uporabljajo za vzdrževanje stanja med zahtevami HTTP. Tako lahko na primer spletni strežnik med overjanjem odjemalca vrne piškotek, ki strežniku omogoča, da ugotovi, da je overjeni uporabnik že prijavljen. Prisotnost piškotka uporabniku omogoča izdajo dodatnih zahtev za strani na tem spletnem strežniku brez prikaza poziva za vnovično prijavo.

Spletni pajek zadrži piškotke, ki jih prejme s spletnih strežnikov in jih uporablja v času trajanja primerka pajka. Piškotke shrani v datoteko `cookies.ini`, ki jo pajek ob koncu vsake seje pajka prepiše. Ko se spletni pajek zaustavi, shrani vse nepretečene piškotke in jih ob zagonu naslednje seje znova naloži.

Če piškotke podate ročno, jih shranite v ločeno datoteko, nato pa jih po potrebi združite s piškotki v datoteki `cookies.ini`. Pajek nepretečenih piškotkov ne zavrže, toda v primeru, da težava prepreči zapis celotne zbirke piškotkov, najbrž ne želite izgubiti piškotkov, ki ste jih podali ročno. Piškotke morate združiti s piškotki, ki jih pajek vzdržuje samodejno, pred začetkom seje pajkanja.

Format piškotkov

Piškotki, ki jih želite združiti z datoteko `cookies.ini` iskanja v podjetju, morajo biti v posebnem formatu.

- Vsak piškotek mora biti v eni vrstici. Prazne vrstice in komentarji so sicer dovoljeni, vendar se v datoteki `cookies.ini` ne ohranijo.
- Vsak piškotek mora imeti naslednjo obliko:

```
CookieN(cookie_length,URL_length)cookie_text,validation_URL
```

Cookie

Zahtevana ključna beseda, ki nakazuje začetek zapisa piškotka.

Ključna beseda za piškotek ne sme vsebovati presledkov in mora imeti pripeto eno števko: 0, 1 ali 2. Številka nakazuje vrsto piškotka: version-0 (Netscape), version-1 (RFC2109) ali version-2 (RFC2965). Sezname vrat v piškotkih RFC2965 niso podprti.

cookie_length

Dolžina povezanega besedila piškotka v znakih.

URL_length

Dolžina povezanega URL-ja preverjanja v znakih.

cookie_text

Vsebina piškotka, ki bo poslana izvornemu spletnemu strežniku. Ta niz (ki predstavlja desno stran napotka Set-Cookie v oglavju HTTP zahteve) podaja ime piškotka in par vrednosti ter kakršnokoli drugo vsebino (kot je pot, nastavitve zaščite in ostalo), ki se pošlje s piškotkom. Temu nizu sledi ločilo vejica (,).

validation_URL

URL, kjer je bil odkrit ta piškotek. Ta URL se uporablja, da se ugotovi, kam naj bo piškotek poslan (na primer tako, da podaja ime domene in ime poti). URL preverjanja mora zadovoljiti omejitve izvornega spletnega strežnika, povezane z zaščito in zasebnostjo.

Naslednji zgled je prikazan v dveh vrsticah zato, da je lažje berljiv. Piškotki, ki jih boste podali vi, pa morajo biti zapisani v eni vrstici.

```
Cookie0(53,40)ASPSESSIONIDQSQTACSD=SLNSIDFNLSIDNFLSIDNFLSNL;path=/  
https://www.ibm.com:443/help/solutions/
```

Konfiguriranje piškotkov za spletnega pajka

Za sejo spletnega pajkanja lahko ročno podate piškotke in jih združite s piškotki spletnega pajka.

Preden začnete

Za ročno konfiguriranje piškotkov, ki jih bo uporabil spletni pajek, morate biti skrbnik iskanja v podjetju.

Postopek

Ročno konfiguriranje piškotkov za spletnega pajka:

1. V skrbniški konzoli iskanja v podjetju nadzirajte zbirko, ki ji želite podati piškotke, in zaustavite spletnega pajka.
2. Prijavite se na strežnik pajkanja kot skrbnik iskanja v podjetju. (Ta ID uporabnika je bil podan med namestitvijo OmniFind Enterprise Edition.)
3. Pomaknite se v podatkovni imenik za pajka, ki ga želite konfigurirati, pri čemer je *crawler_session_ID* ID, ki ga je seji pajka dodelil iskalni sistem podjetja. Primer:
`ES_NODE_ROOT/data/col_56092.WEB_88534`
4. Uredite datoteko `cookies.ini`, pripnite postavke piškotkov, ki ste jih ročno vnesli, že navedenim piškotkom ter shranite in zapustite datoteko. Prepričajte se, da vaši piškotki ne bodo preglasili že prisotnih piškotkov.
5. V skrbniški konzoli iskanja v podjetju znova zaženite pajka, ki ste ga zaustavili.

Konfiguriranje globalnega prostora za spletno pajkanje

Za spletne pajke lahko konfigurirate globalni prostor za pajkanje, ki omogoča boljši nadzor nad odstranjevanjem URL-jev iz indeksa.

Vsak spletni pajek je konfiguriran za prostor za pajkanje, ki definira URL-je, ki jih želite pajkati, in URL-je, ki jih ne želite pajkati. Odkriti URL-ji, ki so v prostoru za pajkanje, bodo ohranjeni (v bazi podatkov) za kasnejše pajkanje, URL-ji, ki niso v prostoru za pajkanje, pa bodo zavrženi. Če pajek začne s prazno bazo podatkov, bosta definicija prostora za pajkanje in baza podatkov med izvajanjem pajka ostali skladni.

Včasih se pajek zaustavi, njegov prostor za pajkanje pa se ne zmanjša (na primer zaradi novih pravil, ki prepovedujejo pajkanje strani). Ko ga znova zaženete, definicija njegovega prostora za pajkanje in baze podatkov ne bosta več skladni. Baza podatkov vsebuje URL-je (nekaj pajkanih in nekaj nepajkanih), ki niso v novem manjšem prostoru za pajkanje.

Če ima zbirka samo enega spletnega pajka, lahko ta skladnost doseže tako, da statusne kode HTTP teh URL-jev spremeni na 760 (ki podaja, da jih je potrebno izključiti), obenem pa zahteva odstranitev sedaj izključenih strani iz indeksa.

Če prostor za pajkanje razdelite med dva ali več spletnih pajkov (če želite s tem na primer zagotoviti, da so nekatere strani pajkane pogosteje kot druge), bo vsak spletni pajek vzdrževal neodvisne tabele baze podatkov (ki so na začetku prazne), poleg tega pa bo vsak pajek pajkal različne dele prostora za spletno pajkanje. Izvirni prostor za pajkanje se bo tako potem, ko drugi pajki odstranijo dele za pajkanje, zmanjšal za ustrezni delež. Do težav pride, ko izvirni pajek poskusi obnoviti skladnost tako, da iz indeksa odstrani premaknjene strani. Ker premaknjene strani sedaj pajkajo drugi pajki, morajo ostati v indeksu.

S konfiguriranjem višje ravni globalnega prostora za pajkanje lahko identificirate URL-je, ki jih izvirni pajek ne bo pajkal, vendar jih tudi ne bo odstranil iz indeksa. URL-ji, ki jih ni v nobenem prostoru za pajkanje nobenega pajka, bodo kljub temu označeni za izključitev v procesih odkrivanja in jih bo sistem iz indeksa odstranil ob naslednjem vnovičnem pajkanju.

Globalni prostor za pajkanje je definiran s konfiguracijsko datoteko, imenovano `global.rules`, ki mora obstajati v konfiguracijskem imeniku pajka (prisotnost datoteke `global.rules` omogoča delovanje globalnega prostora za pajkanje). Če ta datoteka obstaja, jo bo sistem prebral med inicializacijo pajka. Če ne obstaja, bo pajek deloval z enoravenskim prostorom za pajkanje, dokumente iz indeksa pa bo odstranil za namene vzdrževanja skladnosti med definicijo prostora za pajkanje in bazo podatkov.

Če globalni prostor za pajkanje obstaja, bo pajek URL-je na osnovi pravil izključil ali vključil tako kot prej, vendar pa bo zahteval odstranitev URL-ja iz indeksa samo v primeru, če URL ni v nobenem prostoru za spletno pajkanje.

Datoteka `global.rules` ima enako skladnjo kot lokalna datoteka `crawl.rules`, le da lahko vsebuje samo pravila za imena domen. Ta omejitev prostora za pajkanje omogoča razdelitev med pajke samo na osnovi imen gostiteljev DNS in ne na osnovi naslovov IP ali vzorcev predpon HTTP. Globalni prostor za pajkanje ne bo vplival na URL-je, ki jih v lokalnem prostoru za pajkanje izključite s pravili s predpono URL ali naslovom IP (kot je to definirano v datoteki `crawl.rules`), vendar pa bodo kljub temu izključeni.

Globalni prostor za pajkanje je namenjen samo preprečitvi odstranjevanja URL-jev, ki so v prostoru za pajkanje enega pajka iz indeksa izključeni zaradi pravila lokalne domene. Naslednja pravila veljajo v naslednjem vrstnem redu:

1. Če je URL iz baze podatkov pajka izključen zaradi pravila lokalne predpone ali pravila naslova, mu bo sistem dodelil statusno kodo 760 in ga odstranil iz indeksa. URL ne bo več pajkan.
2. Če je URL iz baze podatkov pajka izključen zaradi pravila lokalne domene, globalni prostor za pajkanje pa ne obstaja, mu bo sistem dodelil statusno kodo 760 in ga odstranil iz indeksa. URL ne bo več pajkan.

3. Če je URL iz baze podatkov pajka izključen zaradi pravila lokalne domene, vendar je izrecno dovoljen s pravilom v globalnem prostoru za pajkanje, bo URL-ju dodeljena statusna koda 761. Pajek URL-ja ne bo znova pajkal, vendar pa ga ne bo odstranil iz indeksa (predpostavljeno je, da je v lokalnem prostoru za pajkanje nekega drugega pajka).
4. Če je URL iz baze podatkov pajka izključen zaradi pravila lokalne domene, pravilo v globalnem prostoru za pajkanje pa ga izrecno ne dopušča, bo sistem URL-ju dodelil statusno kodo 760 in ga odstranil iz indeksa.

Ker je globalni prostor za pajkanje namenjen samo preprečevanju brisanja URL-jev, ki so bili že izključeni v lokalnem prostoru za pajkanje, bo privzeti rezultat v globalnem prostoru za pajkanje, če se na ciljni URL ne nanaša nobeno pravilo, prepovedano pajkanje.

Datoteka `global.rules` mora obstajati v imeniku `master_config` vseh pajkov, ki imajo v skupni rabi globalni prostor za pajkanje. Vse kopije datoteke `global.rules` in posamezne lokalne datoteke `crawl.rules` morate urediti tako, da boste zagotovili medsebojno skladnost.

Napotki za prepoved sledenja in indeksiranja

Kakovost iskanja lahko izboljšate tako, da podate napotke za spletnega pajka, ki krmilijo, ali bo ta sledil povezavam na straneh in bo indeksiral strani.

Nekatere spletne strani vsebujejo napotke za prepoved sledenja ali indeksiranja, ki robotom (kot je spletni pajek) povedo, da ne smejo slediti povezavam, ki jih najdejo na teh straneh, da vsebine teh strani ne smejo vključevati v indeks ali da ne smejo izvajati nobenega izmed teh dejanj.

S krmiljenjem teh nastavitev lahko izboljšate kakovost pajkanja. Nekatere strani imenika lahko na primer vsebujejo na tisoče povezav, vendar nobene uporabne vsebine, zato naj pajek te strani pajka, naj sledi tem povezavam, vendar pa od indeksiranja samih strani imenika ne boste imeli nobene koristi.

Včasih boste morda želeli, da pajek ne gre nižje v hierarhiji, vendar pa želene podrejene strani vsebujejo povezave, ne vsebujejo pa napotkov za prepoved sledenja. Ker so nekatere izmed teh strani samodejno ustvarjene, nimajo lastnikov, ki bi vstavili zahtevane napotke.

Če želite podati pravila za pajkanje tovrstnih strani, izdelajte ali pa uredite konfiguracijsko datoteko z imenom `followindex.rules`. Pri podajanju pravil v tej datoteki upoštevajte naslednje smernice:

- Pravila, ki jih konfigurirate, morajo podajati predpone URL (spletnih strani ni mogoče prepoznati po naslovu IP ali imenu DNS gostitelja).
- V predponah URL lahko kot univerzalni znak, ki dopušča ali prepoveduje več strani s podobnimi URL-ji, vključujejo zvezdice (*).
- Vrstni red je pomemben (pajek najprej uporabi prvo pravilo, ki se ujema s ciljnim URL-jem).
- Pravila, ki eksplicitno dopuščajo ali prepovedujejo sledenje ali indeksiranje, bodo nadomestila druge nastavitve, vključno s tistimi v ciljnim dokumentu.

Preglasitev ukazov no-follow in no-index na spletnih straneh

V konfiguracijski datoteki lahko podate pravila, ki podajajo ali naj pajek sledi povezavam na strani in indeksne strani, ki vsebujejo ukaze no-follow ali no-index.

Pređen začnete

Ukaze no-follow in no-index za spletnega pajka lahko podate samo, če ste skrbnik iskanja v podjetju. Ukazi, ki jih podate, preglasijo ukaze, ki obstajajo na straneh, ki bodo pajkane.

Postopek

Preglasitev ukazov no-follow in no-index:

1. V skrbniški konzoli iskanja v podjetju nadzirajte zbirko, za katero želite konfigurirati pravila, in zaustavite spletnega pajka.
2. Prijavite se na strežnik pajkanja kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
3. Pomaknite se v konfiguracijski imenik za pajka, ki ga želite konfigurirati, pri čemer je *crawler_session_ID* ID, ki ga je seji pajka dodelil iskalni sistem podjetja. Primer:
ES_NODE_ROOT/master_config/col_56092.WEB_88534
4. Izdelajte ali uredite datoteko z imenom followindex.rules.
5. Vnesite pravila za pajka v naslednji obliki; pri čemer so *URLprefix* začetni znaki za spletna mesta, katerim želite omogočiti, da bodo sledene ali indeksirane:

```
forbid follow URLprefix  
allow follow URLprefix  
forbid index URLprefix  
allow index URLprefix
```
6. Shranite in zapustite datoteko.
7. V skrbniški konzoli iskanja v podjetju znova zaženite pajka, ki ste ga zaustavili.

Konfiguriranje datuma, ki ga bo spletni pajek uporabljal za že pajkane dokumente

V konfiguracijski datoteki lahko podate možnost, s katero krmilite, kateri datum bo spletni pajek uporabil kot datum pajkanega dokumenta.

Preden začnete

Po privzetku spletni pajek nastavi vrednost polja Datum v pajkanih dokumentih na datum, na katerega je bil dokument pajkan. Če želite, lahko spletnega pajka konfigurirate tako, da spremeni prikaz v Nazadnje spremenjeno in prikaže datum in čas, ki ju vrne spletni strežnik.

Podatki za Nazadnje spremenjeno lahko niso na voljo za vse dokumente. Če pajka konfigurirate za uporabo te vrednosti in spletni strežnik za dokument ne vrne podatkov o zadnji spremembi (Nazadnje spremenjeno), pajek nastavi vrednost polja Datum za pajkani dokument na čas in datum, ko je bil dokument pajkan.

Tabela 4. Kako spletni pajek konfigurira polje Datum za pajkane dokumente

Konfiguriranje pajka	Podatki, ki jih vrne spletni strežnik	Metapodatkovno polje Datum za pajkane dokumente
Privzeta konfiguracija	Spletni strežnik vrne vrednost Nazadnje spremenjeno za pajkani dokument	Datum in čas pajkanja
	Spletni strežnik ne vrne vrednosti Nazadnje spremenjeno za pajkani dokument	Datum in čas pajkanja
Konfiguriran za uporabo datuma Nazadnje spremenjeno	Spletni strežnik vrne vrednost Nazadnje spremenjeno za pajkani dokument	Vrednost Nazadnje spremenjeno, ki jo vrne spletni strežnik
	Spletni strežnik ne vrne vrednosti Nazadnje spremenjeno za pajkani dokument	Datum in čas pajkanja

Če želite podati, kateri datum naj spletni pajek uporabi za pajkane dokumente, morate biti skrbnik iskanja v podjetju.

Postopek

Če želite konfigurirati spletnega pajka tako, da bo datum pajkanega dokumenta datum in čas zadnje spremembe (Nazadnje spremenjeno), ki ju vrnejo spletni strežniki, naredite naslednje:

1. V strežnik pajkanja se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
2. Uredite datoteko `crawl.properties` za spletnega pajka, ki ga želite konfigurirati, pri čemer je `crawler_session_ID` ID, ki ga je seji pajka dodelil sistem iskanja v podjetju.
`ES_NODE_ROOT/master_config/crawler_session_ID/crawl.properties`

Primer:

```
/home/esadmin/master_config/col_00112.WEB_23344/crawl.properties  
C:\Program Files\IBM\esadmin\master_config\col_55667.WEB_78899\crawl.properties
```

3. Dodajte naslednjo vrstico in shranite datoteko:
`which_date=LastModified`
4. Iz skrbniške konzole za iskanje v podjetju znova zaženite spletnega pajka.

Če spletni strežnik vrne podatke o zadnji spremembi (Nazadnje spremenjeno) za dokumente, ki jih pajka ta spletni pajek, pajek vrnjena datum in čas uporabi kot Datum za pajkane dokumente.

Pajki Web Content Management

Če želite dokumente IBM Workplace Web Content Management vključiti v zbirko sistema iskanja v podjetju, morate konfigurirati pajka Web Content Management.

Konfiguriranje strežnika WebSphere Portal

Če Workplace Web Content Management namestite na strežnik WebSphere Portal različice 6, lahko za pajkanje spletnih mest Web Content Management uporabite pajka Web Content Management. Možnosti za pajkanje teh spletnih mest lahko konfigurirate ločeno od možnosti, podane za spletna mesta portala, ki jih pajkate s pajkom WebSphere Portal.

Pred izdelavo pajka Web Content Management morate opraviti postopek, s katerim namestite iskanje v podjetju v WebSphere Portal. Če želite vzpostaviti okolje za iskanje v podjetju, zaženite skript (`wp6_install.sh` v sistemu AIX, Linux ali Solaris oziroma `wp6_install.bat` v sistemu Windows), ki je na voljo skupaj z OmniFind Enterprise Edition na iskalnih strežnikih.

Format URL

Pri podajanju URL-jev za pajkanje morate uporabiti naslednji format:

```
protokol_http://ime_gostitelja_portala:številka_vrat/predpona_portala  
/pot_programčka_streznika_iskalne_sledi_WCM/searchseed?site=ime_spletnega_mesta_WCM&lib=ime_knjižnica
```

Naslednji zgled prikazuje URL za spletno mesto v privzeti namestitveni poti Workplace Web Content Management na portalu WebSphere Portal:

```
http://portal.server.ibm.com:80/wps/wcmsearchseed/  
searchseed?site=PreskusnoSpletnoMesto01&lib=Spletna+vsebina
```

Če ime spletnega mesta ali ime knjižnice vsebuje presledke, jih morate nadomestiti z znakom plus (+). Primer, Spletna vsebina zamenjajte s Spletna+vsebina.

Pregled konfiguriranja

S pajkom Web Content Management lahko pajkate poljubno število spletnih mest Web Content Management. Pri konfiguriranju pajka morate podati URL-je spletnih mest, ki jih želite pajkati. Pajek bo nato s podanih spletnih mest prenesel povezane strani.

Spletna mesta, ki jih želite pajkati, morajo omogočati dostop z istim ID-jem in geslom skrbnika WebSphere Portal. Če želite pajkati spletna mesta, ki uporabljajo druge poverilnice, morate konfigurirati ločenega pajka Web Content Management.

Če izdelate ali spremenite pajka Web Content Management, se prijavite na skrbniško konzolo sistema iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo način pajkanja vseh dokumentov v prostoru za pajkanje.
- Podajanje URL-jev za spletna mesta, ki jih želite pajkati, in informacij, s pomočjo katerih se pajek lahko poveže s spletnimi mesti.

Pri izdelavi ali urejanju pajka lahko preskusite, ali pajek lahko vzpostavi povezavo z URL-ji za pajkanje. V sporočilih je tako še pred zagonom pajka navedeno, ali ta lahko dostopi do dokumentov, ki jih želite pajkati.

- Podajanje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo uporabo nadzora dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi seznammi za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Pomembno: Če želite iskati po zaščiteneh straneh Web Content Management, morate iskanja predložiti s pomočjo portalskega programčka Iskanje za sistem iskanja v podjetju in izdelku WebSphere Portal. Iskanja, ki jih predložite iz vzorčne iskalne aplikacije, ESSearchApplication, ne bodo imela ustreznih poverilnic, zato sistem ne bo mogel preveriti pooblastil uporabnika za dostop do dokumentov.

- Podajanje informacij, s katerimi bo pajek lahko komuniciral z strežnikom proxy, če spletna mesta Web Content Management za streženje dokumentov uporabljajo strežnik proxy.
- Če za zaščito strežnika WebSphere Portal in spletnih mest uporabljate drug izdelek (na primer IBM Tivoli Access Manager WebSEAL ali CA SiteMinder SSO Agent for PeopleSoft), podajte poverilnice za enotno prijavo, ki pajku omogočajo dostop do dokumentov na strežniku.
- Podajanje informacij o datoteki shrambe ključev, s pomočjo katerih lahko pajek za vzpostavitev povezave s spletnimi mesti Web Content Management uporabi protokol SSL (Secure Sockets Layer).
- Podajanje jezika in kodne strani dokumentov za pajkanje.
- Podajanje možnosti za pajkanje in iskanje metapodatkov v dokumentih Web Content Management.
- Podajanje razporedov za pajkanje spletnih mest Web Content Management.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezani pojmi

“Integracija s portalom WebSphere Portal” na strani 311

S tem povezana opravila

“Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 6” na strani 318

Pajki WebSphere Portal

Če želite v zbirko za iskanje vključiti strani s spletnega mesta IBM WebSphere Portal, morate konfigurirati pajka WebSphere Portal.

Konfiguriranje strežnika WebSphere Portal

Pred izdelavo pajka WebSphere Portal morate zagnati skript, s katerim nastavite okolje sistema iskanja v podjetju v izdelku WebSphere Portal. Za različne različice izdelka WebSphere Portal potrebujete različne skripte. Skripti so v iskalne strežnike nameščeni ob namestitvi izdelka OmniFind Enterprise Edition.

- Pri izdelku WebSphere Portal različice 5.1.0 ali novejši zaženite skript `wp5_install.sh`, namenjen sistemu AIX, Linux in Solaris, oziroma skript `wp5_install.bat`, namenjen sistemu Windows.
- Pri izdelku WebSphere Portal različice 6 v sistemih AIX, Linux in Solaris zaženite skript `wp6_install.sh`, v sistemu Windows pa skript `wp6_install.bat`.

Nasvet:

Podrobnejše informacije o načinih konfiguriranja zaščitene pajka WebSphere Portal poiščite v scenariju za srednje veliko organizacijo, opisanem v publikaciji IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Pregled konfiguriranja

Za pajkanje posameznega spletnega mesta WebSphere Portal lahko uporabite pajka WebSphere Portal. Pri konfiguriranju pajka morate podati URL spletnega mesta portala, ki ga želite pajkati. Pajek bo nato prenesel portalske programčke in strani, povezane s podanim URL-jem portala. Če želite pajkati drugo spletno mesto portala, izdelajte drugega pajka.

Če izdelate ali spremenite pajka WebSphere Portal, se prijavite na skrbniško konzolo sistema iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo način pajkanja vseh strani v spletnem mestu.
- Podajanje URL-ja za spletno mesto portala, ki ga želite pajkati, in informacij, s katerimi pajek lahko vzpostavi povezavo s spletnim mestom. Ker so tipi URL-jev lahko dolgi in vključujejo kodirane znake, ki niso ASCII, lahko URL prekopirate s strežnika WebSphere Portal in ga prilepite na skrbniško konzolo sistema iskanja v podjetju.

Pri izdelavi ali urejanju pajka lahko preskusite, ali pajek lahko vzpostavi povezavo z URL-jem za pajkanje. V sporočilih je tako še pred zagonom pajka navedeno, ali ta lahko dostopi do dokumentov, ki jih želite pajkati.

- Podajanje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo uporabo nadzora dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi seznammi za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Pomembno: Če želite iskati po zaščiteneh straneh WebSphere Portal, morate iskanje predložiti iz izdelka WebSphere Portal s pomočjo portalskega programčka Iskanje za sistem iskanja v podjetju. Iskanja, ki jih predložite iz vzorčne iskalne aplikacije, ESSearchApplication, ne bodo imela ustreznih poverilnic, zato sistem ne bo mogel preveriti pooblastil uporabnika za dostop do dokumentov.

- Podajanje informacij, s katerimi lahko pajek komunicira z strežnikom proxy, če spletno mesto WebSphere Portal za streženje strani uporablja strežnik proxy.
- Če za zaščito strežnika WebSphere Portal in spletnih mest uporabljate drug izdelek (na primer IBM Tivoli Access Manager WebSEAL ali CA SiteMinder SSO Agent for PeopleSoft), podajte poverilnice za enotno prijavo, ki pajku omogočajo dostop do dokumentov na strežniku.
- Podajanje informacij o datoteki shrambe ključev, s pomočjo katerih lahko pajek za vzpostavitev povezave s spletnim mestom WebSphere Portal uporabi protokol SSL (Secure Sockets Layer).
- Podajanje jezika in kodne strani dokumentov za pajkanje.
- Podajanje možnosti za pajkanje in iskanje metapodatkov v dokumentih WebSphere Portal.
- Podajanje razporedov za pajkanje dokumentov WebSphere Portal.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezani pojmi

“Integracija s portalom WebSphere Portal” na strani 311

S tem povezana opravila

“Kopiranje URL-ja za pajkanje iz programa WebSphere Portal”

“Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 6” na strani 318

“Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 5.1” na strani 313

Kopiranje URL-ja za pajkanje iz programa WebSphere Portal

Če se želite izogniti možnosti vnosa nepravilnega URL-ja, lahko URL spletnega mesta, ki ga želite pajkati, pri konfiguriranju pajka WebSphere Portal ali Web Content Management prekopirate in ga nato prilepite v ustrezno polje.

O tej nalogi

Pri izdelavi pajka WebSphere Portal ali Web Content Management morate podati URL spletnega mesta na strežniku WebSphere Portal, ki ga želite pajkati. Ker so URL-ji dolgi in običajno vsebujejo kodirane znake, ki niso ASCII, ga lahko s tem postopkom prekopirate s strežnika WebSphere Portal in ga prilepite na skrbniško konzolo sistema iskanja v podjetju.

Postopek

Če želite podati URL, ki naj ga pajek pajka, storite naslednje:

1. Ko ste pripravljeni na vnos URL-ja ali URL-jev, ki jih želite pajkati, na skrbniško konzolo sistema iskanja v podjetju, zagotovite, da je strežnik WebSphere Portal zagnan, nato pa se v program WebSphere Portal prijavite kot skrbnik.
2. Če uporabljate WebSphere Portal različice 5.1, na strežniku WebSphere Portal storite naslednje:
 - a. V zgornjem desnem vogalu kliknite možnost **Skrbništvo**.
 - b. V usmerjevalnem okvirju na levi kliknite možnost **Nastavitve portala** in nato še **Upravljanje iskanja**.
 - c. Na strani Upravljanje zbirk za iskanje kliknite možnost **PortalCollection** v področju Zbirke za iskanje. Če so na voljo druge zbirke, lahko izberete tudi drugo zbirko.
 - d. V razdelku Viri vsebin v področju Zbirka kliknite možnost **Dodaj vir vsebin**.
 - e. Za **Tip vira vsebin** kliknite **Spletno mesto portala**. V polju **Zberi dokumente, povezane s tem URL-jem** se prikaže URL spletnega mesta.
 - f. URL prekopicirajte v odložišče. URL na primer označite, pritisnite in držite tipko Ctrl, nato pa pritisnite tipko Insert.
3. Če uporabljate WebSphere Portal različice 6, na strežniku WebSphere Portal storite naslednje:
 - a. V spodnjem levem vogalu kliknite možnost **Skrbništvo**.
 - b. V usmerjevalnem področju na levi kliknite **Upravljalnik iskanja**.
 - c. Na strani Upravljanje iskanja povežavo **Zbirke za iskanje**.
 - d. Na strani Upravljanje iskanja **Privzeta storitev za iskanje po portalu** med možnostmi storitev za iskanje.
 - e. V tabeli Zbirke za iskanje kliknite zbirko **Vsebine portala**.
 - f. V tabeli Viri vsebin kliknite ikono Uredi na skrajni desni strani (ikona svinčnika), ki je poleg možnosti **Vir vsebine portala**.
 - g. Za **Tip vira vsebin** izberite **Spletno mesto portala**. V polju **Zberi dokumente, povezane s tem URL-jem** se prikaže URL spletnega mesta.
 - h. URL prekopicirajte v odložišče. URL na primer označite, pritisnite in držite tipko Ctrl, nato pa pritisnite tipko Insert.
4. Vrnite se na skrbniško konzolo sistema iskanja v podjetju in URL, ki ste ga prekopicirali, prilepite v polje z URL-jem spletnega mesta.

S tem povezani pojmi

“Pajki WebSphere Portal” na strani 100

Pajki datotečnega sistema Windows

Če želite v zbirko za iskanje vključiti dokumente, shranjene v datotečnih sistemih Microsoft Windows, morate konfigurirati pajka datotečnega sistema Windows.

Pajka datotečnega sistema Windows lahko uporabite za pajkanje poljubnega števila datotečnih sistemov Windows. Pri konfiguriranju pajka morate izbrati lokalne in oddaljene imenike in podimenike, ki jih želite pajkati.

Če namestite strežnik pajkanja v sistem AIX, Linux ali Solaris, tega strežnika ne morete uporabiti za pajkanje virov datotečnega sistema Windows (pajek datotečnega sistema Windows se ne pojavi na seznamu razpoložljivih tipov pajkov).

Nasvet:

Podrobnejše informacije o načinih konfiguriranja zaščenega pajka datotečnega sistema Windows poiščite v scenariju za majhno organizacijo, opisanem v publikaciji IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Pajkanje omrežnih imenikov v skupni rabi

Pajek datotečnega sistema Windows pajka dokumente glede na dovoljenja za branje, podana za skrbnika sistema iskanja v podjetju. Skrbnik je račun storitev OmniFind Enterprise Edition.

Za imenike, ki jih želite pajkati, lahko podate ID uporabnika in geslo. Vendar pa bosta ID uporabnika in geslo namenjena le vzpostavitvi povezave z omrežnimi imeniki v skupni rabi. Pajek bo datoteke pajkal glede na dovoljenja za branje, nastavljena za tega uporabnika v omrežnih imenikih v skupni rabi in ne za lokalne pogone.

Povezave z omrežnimi imeniki se ne prekinejo, dokler storitve OmniFind Enterprise Edition ne zaženete znova. Ko vzpostavite povezavo, se lahko zgodi, da do imenika dostopite z nepravilnim ID-jem uporabnika in geslom. Vendar pa je ta povezava dovoljena samo za odkrivanje datotečnega sistema Windows in seje pajka, ki so pod nadzorom sistema iskanja v podjetju. Če želite preprečiti morebitna varnostna tveganja, poskrbite za pravilno nastavitve overjanj za račun skrbnika za iskanje v podjetju (pod katerim se izvaja storitev OmniFind Enterprise Edition).

Če se želite izogniti težavam pri vzpostavljanju povezave z omrežnim imenikom, za isti omrežni imenik podajte isti ID uporabnika in geslo. Če podate napačna ID uporabnika in geslo, nato pa znova zaženete storitev OmniFind Enterprise Edition, pajek datotečnega sistema Windows morda ne bo uspel izvesti pajkanja, saj bo poskusil vzpostaviti povezavo z imenikom z uporabo napačnih poverilnic. Pajkanje lahko uspe, če omrežno povezavo vzpostavi drug pajek datotečnega sistema Windows, ki uporablja pravilni ID uporabnika in geslo.

Pregled konfiguriranja

Če izdelate ali spremenite pajka datotečnega sistema Windows, se prijavite na skrbniško konzolo sistema iskanja v podjetju. Za to morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke, ki ima v lasti pajka.

Čarovnik vam pri izdelavi pajka pomaga pri naslednjih nalogah:

- Podajanje lastnosti, ki krmilijo delovanje pajka in njegovo uporabo sistemskih virov. Lastnosti pajka krmilijo način pajkanja vseh podimenikov v prostoru za pajkanje.
- Nastavitve razporeda za pajkanje datotečnih sistemov.
- Izbira podimenikov za pajkanje.
Podate lahko število ravni podimenikov, ki naj jih pajek pajka. Če želite pajkati oddaljene datotečne sisteme, lahko podate tudi ID uporabnika in geslo, s katerima bo pajek dostopil do podatkov.
- Podajanje možnosti za omogočitev dokumentov v podimenikih za iskanje. Iz prostora za pajkanje lahko na primer izključite določene tipe dokumentov ali pa podate ID uporabnika in geslo, s katerima lahko pajek dostopi do datotek v določenem podimeniku.
- Konfiguriranje možnosti zaščite na ravni dokumentov. Če ste pri izdelavi zbirke omogočili zaščito, lahko pajek podatke zaščite poveže z dokumenti v indeksu. Ti podatki omogočajo, da iskalne aplikacije uporabijo nadzor dostopov, ki temelji na shranjenih seznamih za nadzor dostopov ali na varnostnih žetonih.

Tako lahko izberete možnost za preverjanje poverilnic uporabnika v času predložitve poizvedbe. V tem primeru sistem ne bo primerjal poverilnic uporabnika z indeksiranimi podatki zaščite, temveč bo poverilnice primerjal s trenutnimi sezname za nadzor dostopov, s katerimi upravlja izvorni vir podatkov.

Če želite uporabiti zaščito na ravni dokumentov, morate zagotoviti, da so na strežniku pajkanja pravilno konfigurirane informacije o računih uporabnikov in domen.

Med izdelavo pajka kliknite možnost **Pomoč**, da se seznanite s polji čarovnika in spoznate, kako v pajka vnesete informacije, potrebne za pajkanje podatkov.

S tem povezani pojmi

“Zagotavljanje zaščite na ravni dokumenta za dokumente datotečnega sistema Windows” na strani 259

“Zaščiteno iskanje po overjenih domenah Windows” na strani 261

Konfiguriranje podpore za aplikacije poslušalnika podatkov

Iskanje v podjetju lahko razširite s pomočjo API-ja poslušalnika podatkov in izdelate zunanjega pajka. Vaše aplikacije poslušalnika podatkov po meri lahko dodajo podatke v zbirko, odstranijo podatke iz zbirke ali spletnemu pajku ukažejo obisk oz. vnovični obisk URL-jev.

Preden začnete

Če želite konfigurirati aplikacije poslušalnika podatkov, morate imeti vlogo skrbnika za iskanje v podjetju.

Pomembno: Poslušalnik podatkov v naslednjih izdajah ne bo podprt. Za razvoj odjemalskih aplikacij za iskanje v podjetju namesto API-jev poslušalnika podatkov uporabite raje API-je za iskanje in indeksiranje (SI-API). Naslednje informacije so namenjene uporabnikom, ki so predhodno izdelali aplikacije poslušalnika podatkov.

O tej nalogi


Odjemalska aplikacija poslušalnika podatkov omogoča pajkanje po tipih virov podatkov, po katerih privzeti pajki iskanja v podjetju ne morejo izvajati pajkanja. Preden lahko uporabite aplikacijo poslušalnika podatkov, morate konfigurirati poverilnice, ki aplikaciji omogočijo dostop do zbirk in njihovo posodabljanje.



Ko odjemalska aplikacija poslušalnika podatkov vzpostavi povezavo s poslušalnikom podatkov, morate posredovati ID in geslo odjemalske aplikacije ter ID zbirke, ki jo želite posodobiti. Vnesene informacije se morajo ujemati z informacijami, ki jih za aplikacijo konfigurirate na skrbniški konzoli.

Poslušalnik podatkov se samodejno zažene ob zagonu sistema iskanja v podjetju. Če po konfiguriranju aplikacije na skrbniški konzoli spremenite številko vrat, morate znova zagnati poslušalnik podatkov.

Postopek

Če želite konfigurirati aplikacije poslušalnika podatkov, naredite naslednje:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
3. Na strani Poslušalnik podatkov kliknite možnost **Konfiguriranje aplikacij poslušalnika podatkov**.

4. Na strani Aplikacije poslušalnika podatkov podajte število niti, ki jih lahko izdela poslušalnik podatkov in s katerimi nato obdeluje zahteve odjemalskih aplikacij, ter številko vrat, na katerih poslušalnik podatkov spremlja morebitne zahteve. Podajte tudi največje število dokumentov iz posamezne zbirke, ki jih lahko sistem zadrži včasem pomnilniku, dokler jih razčlenjevalnik ne začne razčlenjevati.
 5. Kliknite **Dodaj aplikacijo poslušalnika podatkov**, da dodate informacije o odjemalski aplikaciji.
 6. Na strani Dodajanje aplikacije poslušalnika podatkov podajte informacije o overjanju, ki odjemalskim aplikacijam poslušalnika podatkov omogočajo dostop do zbirk za iskanje v podjetju. ID-ji odjemalca poslušalnika podatkov morajo biti v sistemu iskanja v podjetju unikatni.
 7. Izberite zbirke, ki jih lahko posodobi aplikacija poslušalnika podatkov:
 - Kliknite **Vse zbirke**, če želite, da aplikacija posodobi vse zbirke.
 - Kliknite **Specifične zbirke**, če želite, da aplikacija posodobi samo zbirke, ki jih podate. Ko izberete to možnost, se prikaže seznam z imeni zbirk. Izberite potrditveno polje **Izberi** pred vsako zbirko, ki jo lahko aplikacija posodobi.
 8. Kliknite **Potrdi**.
 9. Če ste spremenili številko vrat poslušalnika podatkov ali število dokumentov, ki jih je mogoče zadržati včasem pomnilniku, znova zaženite poslušalnik podatkov:
 - a. Kliknite  **Nadzor**, da preklopite v pogled za nadzor sistema.
 - b. Na strani Poslušalnik podatkov kliknite  **Zaženi znova**.
- S tem povezana opravila**
“Nadziranje poslušalnika podatkov” na strani 288

Prilagojeni vtičniki za pajke

Pri konfiguriranju lastnosti za pajke lahko podate razred Java, s katerim lahko izsilite nadzor dostopov na ravni dokumentov. Z razredom Java pa lahko dodate, spremenite ali odstranite metapodatke in vsebino dokumentov ter tako posodobite indeks. Če napišete vtičnik, lahko razširite sposobnosti pajka na pajkanje arhivskih datotek.

Vtičnik vsebuje razred Java, ki ga sistem pokliče za vsak dokument, ki ga pajek pajka. Sistem razredu Java nato iz indeksa, varnostnih žetonov, metapodatkov in vsebine dokumenta posreduje identifikator dokumenta (URI). Razred lahko vrne nove ali spremenjene varnostne žetone, metapodatke in vsebino, lahko pa jih tudi odstrani.

Ko sistem enkrat pajka vse dokumente v prostoru za pajkanje, bo vtičnik poklical samo za nove ali spremenjene dokumente. Če želite spremeniti varnostne žetone, metapodatke ali vsebino dokumentov, ki so v indeksu sistema iskanja v podjetju, vendar v izvornem viru podatkov še niso bili posodobljeni, zaženite popolno pajkanje vseh dokumentov v prostoru za pajkanje in nato znova zgradite glavni indeks.

Vtičniki za izsiljeno zaščito

Zaščito na ravni dokumentov izsilite tako, da enega ali več varnostnih žetonov (z vejico ločenega niza) povežete s posameznim dokumentom, ki ga pajek pajka. Kot varnostni žetoni se pogosto uporabljajo identifikatorji skupin.

Po privzetku je vsakemu dokumentu dodeljen javni žeton, ki poskrbi, da je dokument na voljo vsem. Javni žeton lahko nadomestite z vrednostjo, ki jo poda skrbnik, ali z vrednostjo, povzeto iz polja v pajkanem dokumentu.

Vtičnik omogoča, da za določanje vrednosti varnostnih žetonov pajkanih dokumentov uporabite svoja poslovna pravila. Varnostni žetoni, ki jih povežete s posameznim dokumentom, bodo shranjeni v indeks. Z njimi lahko filtrirate dokumente, ki se ujemajo z varnostnimi žetoni, poleg tega pa zagotovite, da bo sistem v rezultatih iskanja vrnil samo dokumente, za katere ima uporabnik pooblastilo za prikaz.

Vtičniki za dodajanje, spreminjanje ali odstranjevanje metapodatkov

Metapodatke dokumenta, kot so datum zadnjega spreminjanja dokumenta, bo sistem izdelal za vse pajkane dokumente. Z vtičnikom pajka lahko za določitev vrednosti metapodatkov, ki jih želite indeksirati za posamezni dokument, uporabite svoja poslovna pravila.

Sistem bo metapodatke izdelal kot par imena in vrednosti. Uporabniki lahko iščejo po metapodatkih s poizvedbo s prostim besedilom ali s poizvedbo, ki podaja ime polja metapodatkov.

Vtičniki za dodajanje, spreminjanje ali odstranjevanje vsebine dokumentov

Vsebina dokumentov je sestavljena iz delov, ki vsebujejo vsebino, po kateri je mogoče iskati, in vsebino, ki lahko postane del dinamičnega povzetka dokumenta v rezultatih iskanja. Z vtičnikom pajka lahko za določitev vsebine, ki jo želite indeksirati za posamezni dokument, uporabite svoja poslovna pravila.

Vtičniki spletnega pajka

Z aplikacijskimi programerskimi vmesniki za spletnega pajka lahko krmilite, kako bo sistem pajkal dokumente in kako jih bo pripravil na razčlenjevanje. V oglavje zahteve HTTP lahko na primer dodate polja, ki jih sistem uporabi, ko pajek zahteva dokument. Po pajkanju dokumenta in pred njegovo razčlenitvijo ter dodelitvijo žetonov lahko spremenite vsebino, varnostne žetone in metapodatke. Prav tako pa lahko prekinete pošiljanje dokumenta razčlenjevalniku.

Razpravo o overjanju na osnovi obrazcev in vzorčni program, ki ga lahko priredite za vtičnik spletnega pajka po meri, najdete v temi <http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani>.

Vtičniki za arhivske datoteke

Če napišete vtičnik, lahko razširite sposobnosti pajka in omogočite podporo za pajkanje formatov arhivskih datotek, ki niso ZIP ali TAR. Napišete lahko na primer vtičnik, s katerim omogočite podporo za pajkanje dokumentov v formatu LZH.

Neograjen način

Pri konfiguriranju nespletnega pajka lahko izberete možnost za izvajanje vtičnika v neograjenem načinu. V tem načinu se procesi vtičnika izvajajo v okviru procesov pajka, s čimer se lahko izboljša učinkovitost vtičnika.

Pomembno: Če vtičnik naleti na težavo, ki je ne more odpraviti med izvajanjem v tem načinu, se lahko proces pajka zaustavi.

S tem povezani pojmi



Vtičniki pajka

Podpora za pajkanje arhivskih datotek

Pajki iskanja v podjetju lahko povzamejo datoteke iz arhivske datoteke (kot je na primer datoteka ZIP ali TAR), da je mogoče indeksirati in preiskati posamezne datoteke v arhivu.

Podprti formati arhivskih datotek

Podprti so naslednji formati arhivskih datotek:

Tabela 5. Formati arhivskih datotek, ki jih podpirajo pajki iskanja v podjetju

Pripona datoteke	tip MIME	Tip podatkov	Opombe
.zip, .ZIP	application/zip	zip	<ul style="list-style-type: none">• odvisno od zmožnosti paketa java.util.zip• podpira upadajoče (način 8) stiskanje:<ul style="list-style-type: none">– brez podpore za šifrirane datoteke– brez podpore za zip64
.tar	application/tar	tar	podprti formati tar: <ul style="list-style-type: none">• GNU tar 1.13• POSIX 1003.1-1998 (ustar)• POSIX 1003.1-2001 (pax)
.tar, .gz, .tgz	ni uporabno	tgz	odvisno od zmožnosti paketa java.util.zip

Omejitve in smernice

Samodejno odkrivanje kodne strani ni na voljo za datoteke, ki so povzete iz arhivske datoteke. Pri povzemanju datotek uporablja pajek nastavitve kodne strani, ki je konfigurirana za uporabo s čistim besedilom in neznanimi tipi dokumentov. Če za konfiguriranje nastavitve jezika in kodne strani za pajek uporabite skrbniško konzolo iskanja v podjetju, podate kodno stran, ki jo bo uporabljal pajek za dokumente čistega besedila in za dokumente, katerih kodne strani ni mogoče odkriti samodejno.

Da ugotovi, kdaj je potrebno izvesti vnovično pajkanje datotek v arhivski datoteki, uporablja pajek popravljen datum v podatkih oglavja arhivskega vnosa za vsako datoteko. Pri nadzoru pajka vključujejo prikazani statistični podatki za pajkane dokumente, vključno s statističnimi podatki za vstavljene, posodobljene in izbrisane dokumente, informacije o datotekah, ki so bile povzete iz arhivskih datotek.

Če želite pajkom omogočiti pajkanje po arhivskih datotekah v drugih formatih, kot so na primer datoteke LZH, morate najprej napisati vtičnik za pajka, nato pa pajka konfigurirati za njegovo uporabo.

Selitev

Če želite za pajkanje arhivskih datotek omogočiti pajke, ki so obstajali pred namestitvijo OmniFind Enterprise Edition različice 8.4, morate urediti prostor pajkanja. Pazite, da nastavitve za tipe MIME za izključitev in pripone datotek za izključitev ne bodo vsebovale referenc na datoteke .zip, .tar, .tgz ali .gz.

Pomembno: Če spremenite te nastavitve za pajka Content Edition ali za pajka DB2 Content Manager, morate za uveljavitev sprememb znova izvesti pajkanje v vseh dokumentih.

Formati URI v indeksu iskanja v podjetju

URI (uniform resource identifier) posameznega dokumenta v indeksu iskanja v podjetju nakazuje tip pajka, ki je dodal dokument v zbirko.

URI-je ali vzorce URI lahko podate pri konfiguriranju kategorij, območij in hitrih povezav za zbirko. URI lahko podate tudi, ko morate odstraniti dokumente iz indeksa, ali ko želite prikazati podrobne informacije o statusu določenega URI-ja.

Preiščite zbirko, da ugotovite URI-je ali vzorce URI za dokument. V rezultatih iskanja lahko s klikom na URI-je priključite dokumente, ki vas zanimajo. Prekopirate lahko URI iz rezultatov iskanja in ga uporabite na skrbniški konzoli za iskanje v podjetju. Podate lahko na primer vzorec URI in samodejno povežete dokumente, ki se ujemajo z vzorcem URI, s kategorijo za iskanje v podjetju.

Arhivske datoteke

Oblika zapisa URI-ja za dokumente, ki so ekstrahirani iz arhivske datoteke (kot je na primer datoteka .zip ali .tar) in nato pajkani, je:

```
Original_URI(?|&)&ArchiveEntry=Entry_Name(&ArchiveEntry=Entry_Name)
```

Parametri

Original_URI

Nahajališče arhivske datoteke v viru podatkov.

Entry_Name

Z URL kodirano ime arhivske postavke v arhivski datoteki.

Zgledi

```
file:///d:/Archive1.zip
```

```
file:///d:/Archive1.zip?ArchiveEntry=Folder1/PowerPoint.ppt
```

```
file:///d:/Archive1.zip?ArchiveEntry=Folder2/Text.txt
```

Pajki Content Edition

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek Content Edition v načinu za dostop do strežnika, je:

```
vbr://Server_Name/Repository_System_ID/Repository_Persistent_ID  
/Item_ID/Version_ID  
/Item_Type/?[Page=Page_Number&] JNDI_properties
```

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek Content Edition v načinu neposrednega dostopa, je:

```
vbr:///Repository_System_ID/Repository_Persistent_ID  
/Item_ID/Version_ID  
/Item_Type/[?Page=Page_Number]
```

Parametri

Kodiranje URL se uveljavi za vsa polja.

Server_Name

Ime strežnika WebSphere Information Integrator Content Edition.

Repository_System_ID

Sistemski ID repozitorija.

Repository_Persistent_ID

Trajni ID repozitorija.

Item_ID

ID postavke.

Version_ID

ID različice. Če je ID različice prazen, ta vrednost podaja najnovejšo različico dokumenta.

Item_Type

Tip postavke (CONTENT ali FOLDER).

Page_Number

Številka strani.

JNDI_properties

Lastnosti JNDI za odjemalca aplikacije J2EE. Obstajata dva tipa lastnosti:

java.naming.factory.initial

Ime razreda aplikacijskega strežnika, ki je v uporabi za izdelavo ročke EJB.

java.naming.provider.url

URL poimenovalne storitve aplikacijskega strežnika, ki je v uporabi za zahtevo po ročki EJB.

Zgledi Dokument:

```
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
```

Pajki DB2

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek DB2, je:

```
db2://Database_Name/Table_Name
/Unique_Identifier_Column_Name1/Unique_Identifier_Value1
[/Unique_Identifier_Column_Name2/Unique_Identifier_Value2/...
/Unique_Identifier_Column_NameN/Unique_Identifier_ValueN]
```

Parametri:

Kodiranje URL se uveljavi za vsa polja.

Database_Name

Notranje ime baze podatkov ali njen vzdevek.

Table_Name

Ime ciljne tabele, vključno z imenom sheme.

Unique_Identifier_Column_Name1

Ime prvega stolpca unikatnega identifikatorja v tabeli.

Unique_Identifier_Value1

Vrednost prvega stolpca unikatnega identifikatorja.

Unique_Identifier_Column_NameN

Ime *n*-tega stolpca unikatnega identifikatorja v tabeli.

Unique_Identifier_ValueN

Vrednost *n*-tega stolpca unikatnega identifikatorja.

Zgledi Lokalna, katalogizirana baza podatkov:

```
db2://LOCALDB/SCHEMA1.TABLE1/MODEL/ThinkPadA20
```

Oddaljena, nekatalogizirana baza podatkov:

```
db2://myserver.mycompany.com:50001/REMOTEDB/SCHEMA2.TABLE2/NAME/DAVID
```

Pajki DB2 Content Manager

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek DB2 Content Manager, je:

cm://*Server_Name/Item_Type_Name/PID*

Parametri

Kodiranje URL se uveljavi za parameter *PID*.

Server_Name

Ime strežnika knjižnic IBM DB2 Content Manager.

Item_Type_Name

Ime tipa ciljne postavke.

PID Trajni identifikator DB2 Content Manager.

Zgled

cm://cmsrvctg/ITEMTYPE1/92+3+ICM8+icm1sdb12+ITEMTYPE159+26+A1001001A
03F27B94411D1831718+A03F27B+94411D183171+14+1018

Pajki Domino Document Manager

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek Domino Document Manager, je:

dominodoc://*Server_Name:Port_Number/Database_Replica_ID/Database_Path_and_Name
/View_Universal_ID/Document_Universal_ID
/?AttNo=Attachment_Number&AttName=Attachment_File_Name*

Parametri

Kodiranje URL se uveljavi za vsa polja.

Server_Name

Ime strežnika Domino Document Manager.

Port_Number

Neobvezno: Številka vrat strežnika Domino Document Manager.

Database_Replica_ID

Identifikator dvojnika baze podatkov.

Database_Path_and_Name

Datotečno ime in pot do baze podatkov dokumenta NSF na ciljnim strežniku Domino Document Manager.

View_Universal_ID

Univerzalni ID pogleda, ki je v uporabi za pajkanje dokumentov Domino Document Manager.

Document_Universal_ID

Univerzalni ID dokumenta, ki je definiran v pajkanem dokumentu.

Attachment_Number

Neobvezno: Zaporedna številka (začenši z nič) posamezne priloge.

Attachment_File_Name

Neobvezno: Izvirno ime datoteke priloge.

Zgledi Dokument Domino Document Manager:

dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf
/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226

Priloga dokumenta:

dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf
/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
?AttNo=0&AttName=AttachedFile.doc

Pajki Exchange Server

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek Exchange Server, je:
exchange://OWA_path[?useSSL=true]

Parametri

OWA_Path

Pot OWA (Outlook Web Access) brez protokola.

useSSL=true

Dodano, ko je protokol izvirne poti OWA HTTPS.

Zgledi Telo dokumenta:

exchange://exchangesvr.ibm.com/public/RootFolder1/Folder1/Document.EML

Priloga dokumenta:

exchange://exchangesvr.ibm.com/public/RootFolder1/Folder1/Document.EML/
AttachedFile.doc

Omogočeno za SSL:

exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/Document.EML
?useSSL=true

Pajki baze podatkov JDBC

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek baze podatkov JDBC, je:

```
jdbc://DB_URL/Table_Name  
/Unique_Identifier_Column_Name1/Unique_Identifier_Value1  
/[Unique_Identifier_Column_Name2/Unique_Identifier_Value2  
/.../Unique_Identifier_Column_NameN/Unique_Identifier_ValueN]
```

Parametri

Kodiranje URL se uveljavi za vsa polja.

DB_URL

URL baze podatkov.

Table_Name

Ime ciljne tabele, vključno z imenom sheme.

Unique_Identifier_Column_Name1

Ime prvega stolpca unikatnega identifikatorja v tabeli.

Unique_Identifier_Value1

Vrednost prvega stolpca unikatnega identifikatorja.

Unique_Identifier_Column_NameN

Ime *n*-tega stolpca unikatnega identifikatorja v tabeli.

Unique_Identifier_ValueN

Vrednost *n*-tega stolpca unikatnega identifikatorja.

Zgledi:

Baza podatkov DB2:

jdbc:db2://host01.svl.ibm.com:50000/SAMPLE/DB2INST1.ORG/DEPTNUMB/51

Baza podatkov Oracle:

jdbc:oracle:thin:@/host01.svl.ibm.com:1521:ora/SCOTT.EMP/EMPNO/7934

Baza podatkov MS SQL Server 2000:

jdbc:microsoft:sqlserver://host01.svl.ibm.com:1433;
DatabaseName=Northwind/dbo.Region/RegionID/100

Baza podatkov MS SQL Server 2005:
jdbc:sqlserver://host01.svl.ibm.com:1433;
DatabaseName=Northwind/dbo.Region/RegionID/100

Pajki Notes

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek Notes, je:

```
domino://Server_Name[:Port_Number]/Database_Replica_ID/Database_Path_and_Name  
/[View_Universal_ID]/Document_Universal_ID  
[?AttNo=Attachment_Number&AttName=Attachment_File_Name]
```

Parametri

Kodiranje URL se uveljavi za vsa polja.

Server_Name

Ime strežnika Lotus Notes.

Port_Number

Številka vrat za strežnik Lotus Notes. Številka vrat ni obvezna.

Database_Replica_ID

Identifikator dvojnika baze podatkov.

Database_Path_and_Name

Datotečno ime in pot do baze podatkov NSF na ciljnim strežniku Lotus Notes.

View_Universal_ID

Univerzalni ID pogleda, ki je definiran v ciljni bazi podatkov. Ta ID je podan samo, če je dokument izbran iz pogleda ali mape. Če ne določite pogleda ali mape za pajkanje (če na primer podate, da želite pajkati vse dokumente v bazi podatkov), univerzalni ID pogleda ni podan.

Document_Universal_ID

Univerzalni ID dokumenta, definiran v dokumentu, ki ga pajka pajka.

Attachment_Number

Zaporedna številka (začenši z nič) posameznega dokumenta. Številka priloge ni obvezna.

Attachment_File_Name

Izvirno ime datoteke priloge. Ime datoteke priloge ni obvezno.

Zgledi Dokument, izbran za pajkanje po pogledu ali mapi:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf/  
8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
```

Dokument, ki ni bil izbran za pajkanje po pogledu ali mapi:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf//  
0205F44FA3F45A9049256DB20042D226
```

Priloga dokumenta:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf//  
0205F44FA3F45A9049256DB20042D226?AttNo=0&AttName=AttachedFile.doc
```

Pajki QuickPlace

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek QuickPlace, je:

```
quickplace://Server_Name:Port_Number/Database_Replica_ID/Database_Path_and_Name  
/View_Universal_ID/Document_Universal_ID  
/?AttNo=Attachment_Number&AttName=Attachment_File_Name
```

Parametri

Kodiranje URL se uveljavi za vsa polja.

Server_Name

Ime strežnika Lotus QuickPlace.

Port_Number

Neobvezno: Številka vrat strežnika QuickPlace.

Database_Replica_ID

Identifikator dvojnika baze podatkov.

Database_Path_and_Name

Datotečno ime in pot do baze podatkov dokumenta NSF na ciljnim strežniku QuickPlace.

View_Universal_ID

Univerzalni ID pogleda, ki je v uporabi za pajkanje dokumentov QuickPlace.

Document_Universal_ID

Univerzalni ID dokumenta, ki je definiran v pajkanem dokumentu.

Attachment_Number

Neobvezno: Zaporedna številka (začeni z nič) posamezne priloge.

Attachment_File_Name

Neobvezno: Izvirno ime datoteke priloge.

Zgledi Dokument:

```
quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace%5CPageLibrary4925704300021490.nsf/A7986FD2A9CD47090525670800167225/2B02B1DE3A82B2CE49257043001C2498
```

Priloga strani:

```
quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace%5CPageLibrary4925704300021490.nsf/A7986FD2A9CD47090525670800167225/2B02B1DE3A82B2CE49257043001C2498?AttNo=0&AttName==QPCons3.ppt
```

Pajki seznama semen

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek seznama semen, je:

```
seedlist://Page_URL?pageID=Page_ID[&useSSL;=true]
```

Parametri

Kodiranje URL se uveljavi za vsa polja.

Page_URL

URL dokumenta (unikaten za vsak dokument).

Page_ID

Identifikator objekta za dokument.

useSSL

Če je protokol HTTPS, se v URI doda `&useSSL;=true`. V nasprotnem primeru je `useSSL` izpuščen.

Zgled Protokol HTTPS:

```
seedlist://quickrserver.ibm.com:10035/1otus/my poc?uri=dm:bec6090046f1cd52bc5cfc06e9f4550&verb;=view&pageID;=N1FSZUR1MkJQNjZSMdZQMUMwM1FPNjZCQzY2SUw2SUhPNk1RQ0M2Uk80Nk9PNjVCRUM2UUs2TDFDMA==&useSSL;=true
```

Pajki datotečnega sistema UNIX

Oblika zapisa URI-ja za dokumente, ki jih pajka pajek datotečnega sistema UNIX, je:

`file:///Directory_Name/File_Name`

Parametri

Kodiranje URL se uveljavi za vsa polja.

Directory_Name

Ime absolutne poti do imenika.

File_Name

Ime datoteke.

Zgled

`file:///home/user/test.doc`

Pajki Web Content Management

Oblika zapisa URI-ja za dokumente WebSphere Content Management, ki jih pajka pajek Web Content Management, je:

`wcm://Page_URL?pageID=Page_ID[&useSSL=true]`

Parametri

Kodiranje URL se uveljavi za vsa polja.

Page_URL

URL dokumenta (unikaten za vsak dokument).

Page_ID

Identifikator strani.

useSSL

Če je protokol HTTPS, se v URI doda `useSSL=true`. V nasprotnem primeru je `useSSL` izpuščen.

Zgledi

Protokol HTTP:

`wcm://wp6server.ibm.com:9081/wps/wcm/myconnect/Web+Content/Site01/SiteArea01/ContentTest01?pageID=6QReDeJ9DI3R0663E03Q06L1E2MR47MHOC3Q862RD6J0863B0GJS86J9E0`

Protokol HTTPS:

`wcm://wp6server.ibm.com:9444/wps/wcm/myconnect/Web+Content/Site01/SiteArea01/ContentTest01?pageID=6QReDeJ9DI3R0663E03Q06L1E2MR47MHOC3Q862RD6J0863B0GJS86J9E0&useSSL=true`

Pajki WebSphere Portal: WebSphere Portal različice 5

Oblika zapisa URI-ja za dokumente WebSphere Portal različice 5, ki jih pajka pajek WebSphere Portal, je:

`wps://Page_URL?portletDefID=Portlet_Def_ID&portletID=Portlet_ID&pageID=Page_ID[&useSSL=true]`

Parametri

Kodiranje URL se uveljavi za vsa polja.

Page_URL

URL dokumenta (unikaten za vsak dokument).

Portlet_Def_ID

Identifikator definicije portalskega programčka.

Portlet_ID

Identifikator portalskega programčka.

Page_ID

Identifikator strani.

useSSL

Če je protokol HTTPS, se v URI doda `useSSL=true`. V nasprotnem primeru je `useSSL` izpuščen.

Zgledi Telo dokumenta:

```
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
```

Zgledi Omogočeno za 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
```

Pajki WebSphere Portal: WebSphere Portal različice 6

Oblika zapisa URI-ja za dokumente WebSphere Portal različice 6, ki jih pajka pajek WebSphere Portal, je:

```
wp6://Page_URL?portletURL=Portlet_URL?portletDefID=Portlet_Def_ID&pageID=Page_ID[&useSSL=true]
```

Parametri

Kodiranje URL se uveljavi za vsa polja.

Page_URL

URL dokumenta (unikaten za vsak dokument).

Portlet_URL

Unikatni URL dokumenta.

Portlet_Def_ID

Identifikator definicije portalskega programčka.

Page_ID

Identifikator strani.

useSSL

Če je protokol HTTPS, se v URI doda `useSSL=true`. V nasprotnem primeru je `useSSL` izpuščen.

Zgledi Protokol HTTP:

```
wp6://wp6server.ibm.com:9081/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGhgYWjobuhmaGDka0Bu5uHqFRpkaAAAEisaBQ!!?portletURL=/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BiUZnBwMqsyabykYGBmS2tcDoDE40MebyRiU-woYGFoaG7oZuhgZGjgbuZh6uXY4ihAYOMGQ7Z0CDBAwDcXPkM&portletDefID=6_24LS10811G1F102A0G6HEJUR10&pageID=3_24LS108110R6B02A0GGU94LN00
```

Protokol HTTPS:

```
wp6://wp6server.ibm.com:9444/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGh
```

```
gYWjobuhmaGDka0Bu5uHq5RtqaAAA50L41Q!!  
?portletUrl=/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59  
gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAdzjIPrz6_Tzyc1P1C3IjDHQdF  
RUBTu-saA!/d12/d0/Y2BiUZnBwMqsyabykYGBmS2tcDoDE40MebyRiU-woYGFOaG7oZu  
hgZGjgbuZh6uXWZihAYOMGQ5Z31BDaAwAk73P2  
&portletDefID=6_24LS10811G1F102A0G6HEJMU10  
&pageID=3_24LS108110R6B02A0GGU94T410&useSSL=true
```

Pajki datotečnega sistema Windows

Oblike zapisa URI-ja za dokumente, ki jih pajka pajek datotečnega sistema Windows, so:

file:///Directory_Name/File_Name

file:///Network_Folder_Name/Directory_Name/File_Name

Parametri

Kodiranje URL se uveljavi za vsa polja.

Directory_Name

Ime absolutne poti do imenika.

File_Name

Ime datoteke.

Network_Folder_Name

Samo za dokumente na oddaljenih strežnikih je to ime mape v skupni rabi v omrežju a Windows.

Zgledi Lokalni datotečni sistem:

file:///d:/directory/test.doc

Omrežni datotečni sistem:

file:///filesvr.ibm.com/directory/file.doc

Skrbnišvo nad razčlenjevalnikom

Če želite izboljšati berljivost dokumentov, lahko podate možnosti za razčlenjevanje, analizo in kategoriziranje dokumentov in metapodatkov, preden jih sistem doda v indeks sistema iskanja v podjetju.

Možnosti, ki jih podate za razčlenjevanje vsebine dokumentov in optimiziranje berljivosti informacij, vključujejo naslednje:

Konfiguriranje možnosti za razčlenjevanje kitajskih, japonskih in korejskih dokumentov

Tukaj lahko podate možnosti za uporabo segmentacije n-gram za razčlenjevanje dokumentov, napisanih v kitajskem, japonskem ali korejskem jeziku. Prav tako lahko iz praznih prostorov v kitajskih in japonskih dokumentih odstranite znake za novo vrstico.

Omogočitev iskanja po izvornem XML-ju

Če zbirka vključuje dokumente XML, lahko omogočite iskanje s pomočjo skladnje poizvedbe za izvorni XML, kot so XPath in fragmenti XML. Iskanje po izvornem XML-ju omogoča uporabnikom, da poizvedbe podajo na osnovi razmerij med različnimi elementi XML.

Konfiguriranje kategorij

Dokumente, ki imajo podoben vzorec URI ali pa vsebujejo specifične besede, lahko združite v kategorije. Uporabniki lahko tako pri iskanju po zbirki rezultate iskanja omejijo na dokumente, ki pripadajo specifičnim kategorijam.

Konfiguriranje iskalnih polj

Elemente v dokumentih XML lahko preslikate v iskalna polja v indeksu. Prav tako pa lahko v iskalna polja preslikate elemente metapodatkov v dokumentih HTML. Z izdelavo iskalnih polj v indeksu sistema iskanja v podjetju boste uporabnikom omogočili, da izdajajo poizvedbe za specifične dele dokumentov XML in HTML in tako povečajo točnost rezultatov iskanja.





Konfiguriranje možnosti za obdelavo besedila

Če ste v sistem iskanja v podjetju dodali prilagojene motorje za analizo besedila, lahko izberete tistega, ki ga želite uporabiti za zbirko. Ko motor za analizo povežete z zbirko, lahko podate možnosti za preslikavo vsebine, da jo bo sistem jezikovno analiziral in razložil. Prav tako lahko podate, kako naj sistem rezultate analize preslika v indeks sistema iskanja v podjetju ali v tabele baze podatkov JDBC.

Preslikava polj v poviševalne razrede

Tukaj lahko podate, da naj sistem dokumente s polji, ki se ujemajo z izrazi poizvedbe, v rezultatih iskanja rangira višje kot druge dokumente, ki se ne ujemajo z izrazi poizvedbe. Ko polja preslikate v poviševalne razrede, morate podati, katera polja vsebine in metapodatkov želite pospešiti. Prav tako lahko konfigurirate ocene, ki jih posamezni poviševalni razred uporablja za rangiranje dokumentov.

S tem povezani pojmi

-  Linguistic support for semantic search
-  Text analysis included in enterprise search
-  Basic concepts used in text analysis processing
-  Semantic search applications

 Semantic search query terms

“Jezikovna podpora in podpora za kodno stran” na strani 151

“Odkrivanje formata dokumentov” na strani 137

Delo s kategorijami

Kategorije omogočajo združevanje dokumentov s podobnimi značilnostmi ter iskanje in pridobivanje samo tistih dokumentov, ki ustrezajo kriteriju za članstvo v tej skupini.

Če povežete dokumente s kategorijami in vaše iskalne aplikacije podpirajo to zmožnost, lahko uporabniki preiščejo podniz zbirke tako, da podajo ime kategorije. Če uporabniki preiščejo celotno zbirko, lahko izboljšajo rezultate iskanja in preiščejo samo tiste dokumente, ki so v isti kategoriji kot eden od nastalih dokumentov.

Pri konfiguriranju kategorije podate pravila, ki dajo pajku navodila za povezavo dokumentov s kategorijo. Združite lahko dokumente, ki souporabljajo vzorec URI ali dokumente, ki vsebujejo specifično vsebino (na primer dokumente, ki vsebujejo ali izključujejo določene besede in fraze).

Za izdelavo kategorij in izvajanje skrbništva nad njimi lahko uporabite skrbniško konzolo iskanja v podjetju:

- Tip kategorizacije izberete pri izdelavi zbirke. Izberete lahko, da kategorij ne boste uporabljali ali uporabite kategorije, ki temeljijo na pravilih.
- Pri konfiguriranju razčlenjevalnih pravil za zbirko lahko po potrebi spremenite tip kategorizacije. Če spremenite tip kategorizacije, ko so dokumenti že pajkani in indeksirani, se kakovost iskanja zmanjša, dokler ne izvedete vnovičnega pajkanja vseh dokumentov in znova zgradite glavni indeks.
- Če izberete kategorije, ki temeljijo na pravilih, uporabite skrbniško konzolo za izvajanje skrbništva nad drevesom kategorij, kategorijami in pravili kategorij. Če spremenite kategorije ali pravila kategorij za tem, ko so dokumenti že pajkani in indeksirani, se kakovost iskanja zmanjša, dokler ne izvedete vnovičnega pajkanja vseh dokumentov in znova zgradite glavni indeks.

Kategorije, ki temeljijo na pravilih

Konfigurirate lahko pravila, ki krmilijo, kateri dokumenti bodo povezani s kategorijami v zbirki iskanja v podjetju.

Za zbirke, ki jih izdelate v iskanju v podjetju in za kategorije, temelječe na pravilih, ki jih preselite iz zbirke portala IBM WebSphere Portal, lahko izdelate pravila kategorij. Za konfiguriranje pravil za kategoriziranje dokumentov morate pri izdelavi zbirke ali pri podajanju razčlenjevalnih možnosti za zbirko podati, da želite uporabiti kategorije, temelječe na pravilih.

Razčlenjevalnik uporablja pravila, ki jih podate za povezovanje dokumentov z eno ali več kategorijami:

- Če dokument ustreza vsaj enemu pravilu v kategoriji, razčlenjevalnik poveže dokument s kategorijo.
- Če dokument ustreza vsaj enemu pravilu v več kategorijah, razčlenjevalnik poveže dokument z vsemi kategorijami.
- Če dokument ne ustreza nobenemu pravilu za kategorijo, razčlenjevalnik ne poveže dokumenta s kategorijo. Uporabniki lahko pri preiskovanju zbirke poiščejo ta dokument in ga pridobijo, vendar ne morejo preiskati kategorije in pričakovati, da bodo dokument pridobili.

Pri izvajanju skrbništva nad drevesom kategorij (ali sistematike) za zbirko se odločite, kam v hierarhijo kategorij želite dodati novo kategorijo. Drevo kategorij se uporablja tudi za izbiro kategorije, ki jo želite urediti in nato za dodajanje pravil za kategoriziranje dokumentov, brisanje pravil ali spreminjanje vsebine posameznih pravil.

Pri konfiguriranju pravila za kategoriziranje dokumentov izberete, ali bo iskanje v podjetju za odločitev, ali dokument pripada v kategorijo, uporabil URI dokumenta ali vsebino dokumenta.

Pravila vzorcev URI

Pravilo URI velja za URI dokumenta. URI podate delno (vzorec) in dokumenti, ki vsebujejo podan vzorec v svojih URI-jih, ustrezajo pravilu.

Če na primer podate, da je besedilo pravila `/hr/`, prvi URI spodaj ustreza pravilu, drugi URI pa ne:

```
file:///corporate/hr/medicalform.doc  
http://company.com/human_resources/medicalform.htm
```

Ker so vsi URI-ji obravnavani kot vzorci, sistem prezre vse zvezdice, ki jih podate kot univerzalne znake na začetku ali na koncu vzorca. Tako na primer `*hr/*` in `/hr/` ustrezata enakemu nizu URI-jev.

Pravila vzorcev URI ne razlikujejo med velikimi in malimi črkami. Če vsebuje URI presledke, mora vzorec URI upoštevati pravila iskanja v podjetju za kodiranje URI-jev. Naslednji primer prikazuje pravi in napačen način podajanja URI-ja za pot datotečnega sistema Windows:

Napačen
URI: `file:///c:/program files/`
Pravilen URI: `file:///c:/program+files/`

Pravila vsebine dokumentov

Pravila vsebine dokumentov izrazite v enakem formatu kot poizvedbo. Če je dokument veljaven za poizvedbo, ustreza pravilu. Pri konfiguriranju pravila podate besede in fraze, ki jih mora dokument vsebovati ali izključevati, ter izberete jezik za uveljavljanje pravil korenjenja besed.

Naslednjo pravilo na primer podaja, da dokument ustreza pravilu, če vsebuje besedo `hr` ali frazo `human resources`:

`hr "human resources"`

Naslednje pravilo pa na primer podaja, da dokument ustreza pravilu, če vsebuje njegov naziv besedo `"health"`, ne pa tudi fraze `"employee benefits"`:

`+title:health -title:"employee benefits"`

Za pravila vsebine velja enaka lingvistična normalizacija kot za poizvedbe API-ja iskanja in indeksiranja (SI-API). Toda skladnja za pravila vsebine podpira podniz operacij, ki so na voljo v skladnji poizvedbe SI-API. V pravilih vsebine so dovoljeni samo naslednji operatorji poizvedbe:

- +** Pred izraz vpišite znak plus, če želite nakazati, da se mora izraz pojaviti v dokumentu.

- Pred izraz vpišite znak minus, če želite nakazati, da se izraz ne sme pojaviti v dokumentu.
- " " Dva ali več izrazov vključite v narekovaje, če želite nakazati, da se mora v dokumentu pojaviti natančna fraza.

ime_polja:

Pred izraz ali frazo vpišite ime polja, če želite nakazati, da se mora izraz ali fraza (ali ne sme) pojaviti v podanem polju dokumenta.

Podprta so vsa polja vsebine in metapodatkov, ki so konfigurirana tako, da so polja z možnostjo preiskovanja. Naslednje ključne besede polja SI-API in tipi polj niso podprti:

- site:
- url:
- link:
- docID:
- samegroupas:
- parametrična polja
- žetoni zaščite
- atributi (kot so \$source, \$language, \$doctype itd)

Drevesa kategorij

Drevo kategorij omogoča prikaz vseh kategorij, temelječih na pravilih, v zbirki. Drevo kategorij lahko uporabite za izdelavo kategorij, brisanje kategorij in urejanje pravil, ki povezujejo dokumente s kategorijami.

Drevo kategorij, imenovano tudi sistematika, je urejeno v hierarhijo. Drevo se začne s korenko kategorijo, iz katere izhajajo vse druge kategorije. Ugnedite lahko neomejeno število kategorij in podkategorij, ki uporabnikom nudijo različne izbire za pregledovanje in iskanje dokumentov.

Če na primer dokument ustreza pravilom v več kategorijah, je povezan z vsemi temi kategorijami. Če uporabniki pri pregledovanju rezultatov iskanja preiskujejo kategorijo ali pregledujejo dokumente, ki spadajo v kategorijo, dejstvo, da dokument spada v več kategorij, poveča verjetnost, da ga bodo uporabniki našli.

Pri izvajanju skrbništva nad drevesom kategorij lahko krmilite, kateri dokumenti spadajo v eno ali več kategorij tako, da pod obstoječe kategorije ugnedite nove. Pri izdelavi kategorije podate, ali bo izdelana na korenski ravni ali kot podkategorija druge kategorije. Drevo kategorij se uporablja tudi za brisanje kategorij iz zbirke in za spreminjanje pravil za povezovanje dokumentov s kategorijami. Pri urejanju kategorije lahko preimenujete kategorijo, dodate ali zbrisate pravila kategoriziranja ali spremenite vsebino posameznih pravil.

Pri izvajanju skrbništva nad drevesom kategorij kot smernico uporabite naslednje opise vedenja iskanja in pregledovanja:

- Če uporabnik preiskuje kategorijo visoke ravni, bo iskanje dokumentov, ki ustrezajo kriteriju iskanja, izvedeno v tej kategoriji in v vseh njenih podkategorijah. Če uporabnik preišče zbirko, ki nima dodatnih podkategorij, bo preiskana samo ta podkategorija.
- Če uporabnik preiskuje rezultate iskanja in izbere možnost za pregledovanje dokumentov, ki spadajo v določeno kategorijo, bodo prikazani samo dokumenti v tej kategoriji. V rezultatih iskanja so prikazana tudi imena vseh podkategorij, tako da se lahko uporabnik pomika med kategorijami in si ogleda posamezne podnize dokumentov.

Izbira vrste kategorizacije

Ko izberete vrsto kategorizacije, lahko podate način, s katerim želite v zbirki povezati dokumente s kategorijami.

Preden začnete

Vrsto kategorizacije lahko spremenite samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke za zbirko, ki jo spreminjate.

O tej nalogi

Vrsta kategorizacije je podana ob izdelavi zbirke. Če je potrebno, lahko spremenite način kategoriziranja dokumentov v zbirki. Uporabite lahko kategorije na osnovi pravil, ki jih konfigurirate posebej za zbirko ali pa ne uporabite kategorij.

Pomembno: Če vrsto kategorizacije spremenite po pajkanju podatkov in izdelavi indeksa za zbirko, indeks postane neskladen. Če želite zagotoviti točnost rezultatov iskanja, znova pajkajte vse dokumente v zbirki in nato znova zgradite glavni indeks.

Postopek

Izbira vrste kategorizacije:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Izbira vrste kategorizacije**.
2. Na strani Izbira vrste kategorizacije izberite eno izmed naslednjih možnosti:

Brez To možnost izberite, če dokumentov v tej zbirki ne želite kategorizirati.

Na osnovi pravil

To možnost izberite, če želite kategorizirati dokumente po pravilih, ki ste jih konfigurirana posebej za to zbirko.

Konfiguriranje kategorij

Za zbirko lahko izdelate poljubno število pajkov, poleg tega lahko vsaka kategorija vsebuje poljubno število pravil. Pravila določajo, kateri dokumenti so samodejno povezani s kategorijo.

Preden začnete

Kategorije lahko konfigurirate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke za zbirko, kateri pripadajo kategorije.

Kot vrsta kategorizacije mora biti izbrana možnost uporabe kategorij na osnovi pravil.

Če si želite ogledati, kako podati pravila za povezane dokumente med izdelavo ali urejanjem kategorije kliknite **Pomoč**.

O tej nalogi

Če vaše iskalne aplikacije podpirajo podporo za kategorije, lahko uporabniki iščejo v podnizu zbirke tako, da podajo ime kategorije. Prav tako pa lahko uporabniki v rezultatih iskanja izberejo kategorijo in pregledujejo samo tiste dokumente, ki pripadajo izbrani kategoriji.

Pomembno: Če kategorizacije ali pravila kategorije spremenite po pajkanju podatkov in izdelavi indeksa za zbirko, indeks ne bo več skladen. Če želite zagotoviti točnost rezultatov iskanja, znova pajkajte vse dokumente v zbirki in znova zgradite glavni indeks.

Postopek

Konfiguriranje kategorije:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Konfiguriranje drevesa kategorij**.
2. Na strani Drevo kategorij izberite nahajališče v drevesu, kjer želite dodati kategorijo in kliknite **Izdelaj kategorijo**.

Če izberete koren, se nova kategorija izdela na korenski ravni. Če izberete ime kategorije, se nova kategorija ugnezdi pod izbrano kategorijo na drevesu kategorij.

Odpre se čarovnik, ki vam pomaga pri podajanju pravil za povezovanje dokumentov s kategorijo:

 - a. Na strani Izdelava kategorije vnesite opisno ime za kategorijo in nato kliknite **Naprej**.
 - b. Na strani Izdelava pravil kategorije kliknite **Dodaj pravilo**.
 - c. Na strani Izdelava pravila kategorije v polje **Ime pravila** vnesite unikatno ime za pravilo. To ime mora biti unikatno v vseh kategorijah v zbirki.
 - d. Podajte pravilo, ki ga želite uporabiti za dokumente, povezane s to kategorijo in nato kliknite **Potrdi**.
 - Če želite pripadnost dokumenta tej kategoriji ugotoviti z URI-jem dokumenta, kliknite **vzorec URI-jev** in nato podajte vzorec URI-jev.

Če podano besedilo obstaja v URI-ju, bo dokument povezan s kategorijo.

Primer: file:///c:/program+files/finance
 - Če želite pripadnost dokumenta tej kategoriji ugotoviti s poizvedbo po vsebini, kjer je mogoče iskanje, kliknite **Vsebina dokumenta**, izberite jezik dokumentov in podajte besede in besedne zveze, ki se morajo ali ne smejo pojaviti v dokumentu.

Pravilo izrazite v isti obliki kot poizvedbo, vendar so dovoljeni samo operatorji vključi (+), izključi (-), besedna zveza (" ") in ime polja (*field_name*:). Pravila vsebine ne podpirajo segmentacije n-gram.

Če dokument vključuje ali ne vključuje besed, ki jih podate, bo povezan s kategorijo.

Primer: +finance -računovodstvo +naslov:"davčno leto"
 - e. Kliknite **Dokončaj**.

Vaša nova kategorija je navedena na strani Drevo kategorij skupaj z drugimi kategorijami, ki spadajo v to zbirko.

S tem povezane povezave

“Formati URI v indeksu iskanja v podjetju” na strani 107

Delo z iskalnimi polji XML

Če želite uporabnikom omogočiti iskanje po določenih delih dokumentov XML, preslikajte elemente XML v iskalna polja.

Elemente XML preslikate v iskalna polja s pomočjo skrbniške konzole za iskanje v podjetju.

Ponavadi so indeksirani vsi besedilni podatki v dokumentu XML. Po preslikavi elementov XML v iskalna polja lahko s pomočjo informacij o zgradbi elementov XML nudite podporo natančnejšim poizvedbam. Iskanje po imenu polja in vrnitev v rezultatih iskanja za podatke v elementu XML omogočite tako, da preslikate ime elementa XML v ime polja. Natančnost preslikave izboljšate tako, da podate attribute elementa XML (in vrednosti teh atributov) kot kriterije elementa, da ta postane iskalno polje.

Iskalna polja XML

Iskalna polja XML uporabnikom omogočajo poizvedovanje po določenih delih dokumentov XML.

Dokumenti XML postajajo čedalje pogostejši, saj vsebujejo tako polstrukturirano kot nestrukturirano besedilo. Struktura elementov XML je enkapsulirana in uporablja kontekst, izrecno definiran z elementi, ki obdajajo besedilo. Avtorjevo ime je lahko na primer prikazano na naslednji način:

```
<author>Janez Novak</author>
```

V tem kontekstu besedilo **Janez Novak** označuje avtorja dokumenta XML.

Pri iskanju v podjetju lahko elemente XML povežete ali preslikate v imena iskalnih polj. Ko za zbirko konfigurirate možnosti razčlenjevanja, podate, kateri elementi XML bodo preslikani v katera imena iskalnih polj. S preslikavo elementov XML v iskalna polja uporabnikom omogočite iskanje po vrednostih teh elementov tako, da v poizvedbah podajo imena polj. Poizvedbe, ki iščejo po imenovanih poljih, lahko nudijo natančnejše rezultate iskanja kot poizvedbe prostega besedila, ki iščejo po celotni vsebini dokumenta.

Če vaša zbirka vsebuje na primer dokumente XML in podate, da naj sistem polji `<title>` in `<author>` označi kot iskalni polji v indeksu, lahko uporabniki poizvedujejo neposredno po teh elementih. Iskanje za `author:Novak` najde dokumente XML, ki vsebuje dokumente, v katerih je **Novak** preslikan v polje z imenom `author`.

V drugem primeru utegne element XML z imenom `<summary>` vsebovati informacije, ki jih je koristno prikazati med rezultati iskanja. Če preslikate element `<summary>` v iskalno polje in podate, da naj se vrednost tega elementa prikaže v rezultatih iskanja, vsebina elementa postane del dobljenega dokumenta.

Če preslikate element XML v ime polja, je iskanje po besedilu znotraj elementa mogoče tudi po imenu polja, ki ga podate. Če element XML vsebuje attribute, pa sistem vrednosti atributov ne indeksira neposredno, zato neposredno iskanje po njih ni mogoče. Če želite izdati poizvedbo po vrednostih atributov, morate v konzoli za skrbništvo nad iskanjem v podjetju konfigurirati možnosti razčlenjevanja in omogočiti iskanje po izvirnem XML. Potem ko omogočite iskanje po izvirnem XML, omogočite strukturo dokumenta XML v indeksu in lahko poizvedujete po dokumentu tako, da podate omejitve poizvedbe XPath. Primer:

```
@xpath: '//elementName[@attributeName="attributeValue"]
```

S tem povezani pojmi

 [Semantic search query term](#)

S tem povezana opravila

[“Omogočanje podpore za iskanje po izvirnem XML” na strani 135](#)

Preslikovanje elementov XML v iskalna polja

Ko preslikate element XML v iskalno polje, podate, katere elemente XML lahko iščejo uporabniki tako, da v poizvedbo vnesejo ime polja.

Preden začnete

Če želite preslikati elemente XML v iskalna polja, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke, kateri pripadajo dokumenti XML.

Omejitve

V posamezni zbirki je lahko več preslikav polj XML, a le ena preslikava korenskega elementa XML. Korenski element dokumenta XML ustrezno uveljavi preslikave.

O tej nalogi

Ko izdelate preslikavo polja XML oziroma dodate, spremenite ali zbrisete polja obstoječe preslikave polja XML, spremembe stopijo v veljavo šele zatem, ko znova zaženete razčlenjevalnik. Nove in spremenjene preslikave stopijo v veljavo za nove, razčlenjene podatke, potem ko znova zaženete razčlenjevalnik in nimajo vpliva na podatke, ki so že v indeksu. Če želite posodobiti dokumente, ki so že v indeksu, morate dokumente znova prepajkati in indeksirati.

Ta naloga uporablja naslednji vzorčni dokument XML, s pomočjo katerega prikaže, kako lahko preslikate dosje osebj in uporabnikom omogočite izvajanje neposrednih poizvedb po določenih elementih.

```
<?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>
```

Postopek

Če želite preslikati elemente XML v tem primeru za iskanje polj, naredite naslednje:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Preslikaj elemente XML v polja**.
2. Na strani Preslikave polja XML kliknite **Izdelaj preslikavo XML**. Odpre se stran Izdelava preslikave polja XML.
3. V polje **Ime korenskega elementa XML** vnesite ime korenskega elementa: **personnel**. Prepričajte se, da se tukaj podano ime ujema s korenskim elementom v dokumentih XML, ki jih želite poiskati. Iskanje v podjetju pri razčlenjevanju in indeksiranju dokumentov XML na podlagi imena korenskega elementa izbere, katero preslikavo bo uporabilo.
4. V polje **Ime preslikave XML** vnesite ime za ta niz pravil za preslikavo polja XML. Potem ko izdelate niz pravil za preslikavo XML, se to ime prikaže na strani Preslikave polj XML, izberete pa ga, ko želite dodati, izbrisati ali spremeniti pravila za preslikave.
5. Preslikajte element XML **jobrole** v iskalno polje z imenom **jobrole**:
 - a. V polje **Ime polja** vnesite **jobrole**.

Nasvet: Kliknite **Pomoč**, da pridete do informacij o znakih ASCII in metaznakih, ki niso podprti v imenih polj.
 - b. V polje **Ime elementa XML** vnesite **jobrole**.
 - c. Če želite uporabnikom omogočiti poizvedovanje po polju **jobrole**, izberite potrditveno polje **Iskanje po poljih**.

- d. Če želite zagotoviti, da bo zadetek prikazan samo, če se iskalni izraz ujema s celotno vrednostjo polja **jobrole**, (če torej v polju ni drugih besed) izberite potrditveno polje **Popolno ujemanje**.
 - e. Če želite uporabnikom omogočiti razvrščanje rezultatov glede na vrednosti v polju **jobrole**, izberite potrditveno polje **Možnost razvrščanja**.
Možnost razvrščanja polja je prikazana samo, če to funkcijo podpira tudi iskalna aplikacija. Vzorčne iskalne aplikacije za iskanje v podjetju na seznamu polj, s katerimi lahko razvrščate rezultate iskanja, ne vključujejo polj, ki so preslikana iz elementov XML.
 - f. Če želite uporabnikom omogočiti, da v rezultatih iskanja prikažejo vrednost polja **jobrole**, izberite potrditveno polje **Rezultati iskanja**.
6. Preslikajte element XML **jobroles** v isto iskalno polje:
- a. Kliknite **Dodaj polje**, da na seznam pravil za preslikavo polj dodate prazno vrstico.
 - b. V polje **Ime polja** vnesite **jobrole**.
 - c. V polje **Ime elementa XML** vnesite **jobroles**.
- Nasvet:** Ni treba, da se imena elementov XML ujemajo z imeni iskalnih polj, prav tako pa se lahko več elementov XML preslika v isto iskalno polje.
- d. Če želite uporabnikom omogočiti, da poizvedujejo po polju **jobrole** (podajte, da se morajo iskalni izrazi povsem ujemati z vrednostjo polja), s pomočjo polja razvrstijo rezultate iskanja, nato pa si v rezultatih iskanja ogledajo polje, izberite ustrezna potrditvena polja.
7. Preslikajte element XML **section**, ki vsebuje atribut z imenom **ID**, v katerem je vrednost **expertise**, v iskalno polje z imenom **expertise**:
- a. Kliknite **Dodaj polje**, da na seznam pravil za preslikavo polj dodate prazno vrstico.
 - b. V polje **Ime polja** vnesite **expertise**.
 - c. V polje **Ime elementa XML** vnesite **section**.
 - d. V polje **Ime atributa XML** vnesite **id**.
 - e. V polje **Vrednost atributa XML** vnesite **expertise**.
 - f. Če želite uporabnikom omogočiti poizvedovanje po polju **expertise**, izberite potrditveno polje **Iskanje po poljih**.
 - g. Če želite zagotoviti, da bo zadetek prikazan samo, če se iskalni izraz ujema s celotno vrednostjo polja **expertise**, (če torej v polju ni drugih besed) izberite potrditveno polje **Popolno ujemanje**.
 - h. Če želite uporabnikom omogočiti razvrščanje rezultatov glede na vrednosti v polju **expertise** in če iskalna aplikacija podpira to funkcijo, izberite potrditveno polje **Možnost razvrščanja**.
 - i. Če želite uporabnikom omogočiti, da v rezultatih iskanja prikažejo vrednost polja **expertise**, izberite potrditveno polje **Rezultati iskanja**.

Zgledi:

Če želite poiskati vse osebe v organizaciji, ki delajo na iskalnih izdelkih, podajte naslednjo poizvedbo:

`jobrole:search`

Če želite poiskati vse osebe v organizaciji, ki imajo izkušnje v jezikoslovju, podajte naslednjo poizvedbo:

`expertise:linguistics`

Delo z iskalnimi polji HTML

Elemente metapodatkov HTML lahko preslikate v iskalna polja, če želite uporabnikom omogočiti, da iščejo po specifičnih razdelkih metapodatkov dokumentov HTML.

Za preslikavo elementov metapodatkov HTML v iskalna polja uporabite skrbniško konzolo sistema iskanja v podjetju.

S preslikavo elementov metapodatkov HTML v iskalna polja boste uporabnikom omogočili iskanje po dokumentih HTML z natančnejšimi poizvedbami.

Iskalna polja HTML

Iskalna polja HTML omogočajo uporabnikom izdajanje poizvedb za attribute v dokumentih HTML.

Elementi metapodatkov v dokumentih HTML so podobni atributom dokumentov, saj vsebujejo informacije o dokumentu, kako je ta oblikovan in kako je dostopen prek spleta.

Primer:

```
<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="Ta tema opisuje izdelek IBM." />
<meta name="format" content="XHTML" />
```

Sistem iskanja v podjetju lahko imena elementov metapodatkov HTML poveže ali preslika v imena iskalnih polj. Pri konfiguriranju možnosti razčlenjevanja za zbirko lahko podate, kateri elementi metapodatkov HTML bodo preslikani v katera imena iskalnih polj. S preslikavo elementov metapodatkov HTML v iskalna polja boste uporabnikom omogočili, da z vnosom imen iskalnih polj v poizvedbo poiščejo dokumente s temi elementi. Poizvedbe, ki iščejo po specifičnih poljih, nudijo natančnejše rezultate iskanja kot poizvedbe s prostim besedilom, ki iščejo po celotni vsebini dokumentov.

Če vaša zbirka na primer vključuje dokumente HTML in podate, da naj sistem elemente metapodatkov za avtorske pravice in izvleček indeksira kot iskalna polja, lahko uporabniki izdajo poizvedbo za te specifične elemente. Če izvedete iskanje po `copyright:IBM`, bo sistem našel dokumente HTML, ki v metapodatkih `copyright` vsebujejo `IBM`.

Če elemente metapodatkov HTML preslikate v iskalna polja, morate podati, ali želite preslikati vse elemente metapodatkov HTML, samo elemente, ki pripadajo nizu elementov metapodatkov Dublin Core, ali samo elemente metapodatkov HTML, ki jih podajate. Za opis elementov v nizu elementov metapodatkov Dublin Core preglejte spletno stran Dublin Core Metadata Initiative:

<http://dublincore.org/documents/dcmi-terms/#H2>

Če izberete izdelavo preslikav za vse elemente metapodatkov HTML ali vse elemente metapodatkov Dublin Core, se bo privzeto ime iskalnega polja ujemalo z imenom elementa metapodatkov. Če želite privzeto ime iskalnega polja zamenjati in za specifične elemente podati druge možnosti iskanja, dodajte elemente na seznam elementov, ki jih želite konfigurirati posebej.

Preslikava elementov metapodatkov HTML v iskalna polja

Ko preslikate element metapodatkov HTML v iskalno polje, podate, kateri uporabniki elementov metapodatkov HTML lahko iščejo s podajanjem imena polja v poizvedbi.

Preden začnete

Elemente metapodatkov HTML lahko preslikavajo v iskalna polja samo člani skrbniške vloge za iskanje v podjetju ali skrbniki zbirke za zbirko, kateri pripadajo kategorije.

O tej nalogi

Ko izdelate preslikavo polj HTML, oziroma polja v obstoječi preslikavi polj HTML dodate, spremenite ali zbrisate, bo ta sprememba stopila v veljavo po vnovičnem zagonu razčlenjevalnika. Nove in spremenjene preslikave stopijo v veljavo za nove, razčlenjene podatke, potem ko znova zaženete razčlenjevalnik in nimajo vpliva na podatke, ki so že v indeksu. Če želite posodobiti dokumente, ki so že v indeksu, morate dokumente znova prepajkati in indeksirati.

Postopek

Preslikava elementov metapodatkov HTML v iskalna polja:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Preslikava metapodatkov HTML v polja**.
2. Na strani Preslikava metapodatkov HTML podajte, katere elementi metapodatkov HTML želite preslikati v iskalna polja:

- Če preslikate vse elemente HTML ali vse elemente, ki pripadajo nizu elementov Dublin Core, morate podati ali želite, da bodo uporabniki lahko iskali v poljih po imenu polja in ali naj bodo polja prikazana v rezultatih iskanja. Podati morate tudi ali morajo iskalni izrazi uporabnikov popolnoma ustrezati celotni vrednosti elementa metapodatkov (to pomeni, da so v polju samo besede, ki ustrezajo iskalnim izrazom uporabnikov) in ali uporabniki lahko urejajo rezultate iskanja glede na to polje. Sistem kot imena polj samodejno uporabi imena elementov metapodatkov. Če želite preglasiti privzeta imena polj za določene elemente ali z določenimi elementi uporabiti drugačne možnosti iskanja, lahko preslikate ime elementa v unikatno ime polja in potem podate možnosti iskanja, ki jih želite uporabiti s tem elementom.
- Preslikava določenih elementov metapodatkov v iskalna polja:
 - a. Kliknite **Dodaj polje**, da na seznam pravil za preslikavo polj dodate prazno vrstico.
 - b. Vnesite ime, ki ga želite povezati z elementom metapodatkov HTML, ki ga preslikavate. Uporabniki lahko to ime polja podajo pri izvajanju poizvedb za dokumente HTML v tej zbirki.

Nasvet: Kliknite **Pomoč**, da pridete do informacij o znakih ASCII in metaznakih, ki niso podprti v imenih polj.

- c. Vnesite ime elementa metapodatkov, ki ga želite preslikati.
- d. Če želite uporabnikom omogočiti poizvedbe v tem polju, izberite potrditveno polje **Iskanje po poljih**.

Nasvet: Če je izbira tega potrditvenega polja razveljavljena, po tem polju ni mogoče iskati s poizvedbo po poljih ali s poizvedbo po prostem besedilu. To dejanje je lahko na primer koristno, če ste izbrali možnost za vključitev vseh elementov podatkov HTML ali elementov podatkov Dublin Core, vendar želite preprečiti iskanje po določenih poljih.

- e. Če želite uporabnikom omogočiti poizvedbe po tem polju in podati, da se dokument ujema samo, če se poizvedbeni izrazi ujemajo s celotno vrednostjo polja (to pomeni, da v polju ne obstajajo nobene druge besede), izberite potrditveno polje **Popolno ujemanje**.

- f. Če želite uporabnikom omogočiti razvrščanje rezultatov v abecednem vrstnem redu (razvrstitev v nizu) v tem polju, izberite potrditveno polje **Poljubna razvrstitev**. Možnost razvrščanja polja je prikazana samo, če to funkcijo podpira tudi iskalna aplikacija. Vzorčne iskalne aplikacije za iskanje v podjetju na seznamu polj, s katerimi lahko razvrščate rezultate iskanja, ne vključujejo polj, ki so preslikana iz elementov metapodatkov HTML.
- g. Če je podatkovni tip tega polja DECIMAL, DOUBLE, INTEGER, SHORT, TIME ali TIMESTAMP in želite uporabnikom omogočiti, da pri preiskovanju tega polja podajo parametrične poizvedbe ali razvrstijo rezultate po številkah, glede na vrednost tega polja, izberite potrditveno polje **Parametrično iskanje**.
- h. Če želite uporabnikom omogočiti ogled tega polja v rezultatih iskanja, izberite potrditveno polje **Rezultati iskanja**.

Zgled:

Uporabniki lahko poizvedujejo po preslikanih imenih polj, da najdejo dokumente HTML z določenimi metapodatki. Če ste na primer preslikali element metapodatkov HTML z imenom opis v iskalno polje z imenom izvleček, lahko uporabniki vnesejo poizvedbo, podobno naslednji, če iščejo dokumente HTML, ki govorijo o računalnikih Thinkpad:

izvleček:thinkpad

Prilagojena obdelava besedila

Kakovost in natančnost rezultatov iskanja lahko izboljšate tako, da integrirate algoritme za obdelavo besedila po meri v zbirke za iskanje v podjetju.

OmniFind Enterprise Edition podpira arhitekturo IBM Unstructured Information Management Architecture (UIMA), ki predstavlja sestav za izdelavo, odkrivanje, oblikovanje in razmeščanje funkcij za analizo besedila. Razvijalci aplikacij izdelajo in preizkusijo algoritme za analizo vsebine, ki bo preiskana, nato pa izdelajo arhiv motorja za obdelavo (datoteka .pear), ki vključuje vse vire, potrebne za uporabo arhiva pri iskanju v podjetju. Če želite preiskovati zbirke s pomočjo algoritmov za analizo po meri, morate dodati arhiv (v katerem je motor za analizo besedila) v sistem iskanja v podjetju.

Komponenta logike analize v motorju za analizo besedila se imenuje *razlagalnik*. Vsak razlagalnik izvaja specifične naloge za jezikoslovno analizo. Motor za obdelavo besedila lahko vsebuje vrsto razlagalnikov ali pa je sestavljen iz več motorjev za analizo besedila, od katerih vsaka vsebuje lastne razlagalnike po meri.

Informacije, ki jih ustvarijo razlagalniki, se imenujejo *rezultati analize*. Rezultati analize, ki ustrezajo informacijam, ki jih želite poiskati, se zapišejo v podatkovno strukturo, imenovano *struktura splošne analize*.

Konfiguriranje možnosti za obdelavo besedila zbirke zajema naslednje naloge:

- Izberite motor za analizo besedila, s katerim želite razložiti dokumente v zbirki.
- Če vaša zbirka vsebuje dokumente XML z označevanjem s pomenom in želite to označevanje uporabiti v svoji analizi besedila po meri, lahko datoteke za preslikavo povežete z zbirko in preslikate izhodne podatke preslikave XML v strukturo splošne analize.

Preslikati je mogoče na primer vsebino elementov <naslovljenca> in <stranke> v opombe o osebah v strukturi splošne analize. Do teh opomb lahko nato dostopajo vaši razlagalniki po meri, ki lahko odkrijejo dodatne informacije (odkrijejo lahko na primer spol osebe).

Opombe o osebi lahko preslikate tudi v indeks za iskanje v podjetju, s čimer uporabnikom omogočite iskanje po osebah, ne da bi morali za to poznati izvirne elemente XML.







Če želite uporabnikom omogočiti podajanje izvernih elementov XML v poizvedbah, vam ni treba konfigurirati preslikav XML. Namesto tega morate konfigurirati možnosti razčlenjevanja in omogočiti iskanje izvirnega XML za zbirko.

- Preslikajte strukturo splošne analize v indeks za iskanje v podjetju, s čimer omogočite iskanje po razloženih dokumentih s semantičnim iskanjem.

Glede na entitete in povezavo, ki jih odkrijejo razlagalniki, lahko uporabniki iščejo zasnove, ki se pojavijo v istem stavku (na primer specifična imena in ime konkurence) ali ključne besede in zasnovo (na primer ime Janko in njegova telefonska številka).

- Preslikajte strukturo splošne analize v relacijsko bazo podatkov. Podatek lahko preslikate v tabele IBM DB2 ali tabele Oracle. Tovrstna preslikava omogoča, da rezultate analize uporabite v aplikacijah baz podatkov, kot so na primer aplikacije za rudarjenje podatkov. Prav tako omogoča uporabo poizvedb SQL za iskanje podatkov, ki so izven dosega iskanja v podjetju.

S tem povezani pojmi

-  Custom text analysis integration
-  Basic concepts used in text analysis processing
-  Workflow for custom analysis integration
-  Text analysis algorithms
-  Semantic search applications
-  Semantic search query term

Dodajanje motorjev za analizo besedila v sistem

Če izdelate motor za analizo besedila po meri, ga morate dodati v sistem, preden ga lahko začnete uporabljati v iskanjih v sistemu. Zbirke lahko s pomočjo motorja za analizo razčlenijo dokumente in povečajo natančnost rezultatov iskanja.

Preden začnete

Če želite v sistem dodati motorje za analizo besedila, morate imeti vlogo skrbnika za iskanje v podjetju.

O tej nalogi


Razvijalci aplikacij lahko izdelajo arhiv motorja za obdelavo (.pear), ki se drži sestava UIMA za analizo besedila. Arhiv vključuje vse vire, potrebne za preiskovanje zbirk za iskanje v podjetju. Če želite preiskovati zbirke s pomočjo algoritmov za analizo po meri, morate dodati arhiv (v katerem je motor za analizo besedila) v sistem iskanja v podjetju.

Potem ko dodate motor za analizo besedila v sistem, lahko spremenite njegovo prikazno ime in izberete možnost za prikaz izvorne kode XML. Izvorna koda XML prikaže, katere informacije izdela motor.





Če je motor za analizo besedila povezan z zbirko, ga ne boste mogli odstraniti iz sistema.

Postopek


Če želite dodati motor za analizo po meri v sistem iskanja v podjetju, naredite naslednje:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
3. Na strani Razčlenjevanje kliknite **Konfiguriraj motorje za analizo besedila**.
4. Na strani Motorji za analizo besedila kliknite **Dodaj motorje za analizo besedila**.
5. Na strani Dodajanje motorja za analizo besedila vnesite opisno ime novega motorja. Sistem na podlagi tega prikaznega imena prepozna motor za analizo besedila na skrbniški konzoli.
6. Podajte nahajališče datoteke `.pear`. Če je datoteka manjša kot 8 MB, je lahko datoteka v lokalnem računalniku, poiščete pa jo lahko s pregledovanjem. Če je datoteka večja kot 8 MB, se mora nahajati na indeksnem strežniku, vnesti pa morate popolnoma kvalificirano pot do nje.
7. Kliknite **Potrdi**. Vaš motor za analizo besedila je naveden na strani Motorji za analizo besedila.

S tem povezani pojmi

-  [Workflow for custom analysis integration](#)
-  [Custom text analysis integration](#)
-  [Basic concepts used in text analysis processing](#)
-  [XML markup in analysis and search](#)

S tem povezana opravila

-  [Creating an XML elements to the common analysis structure mapping file](#)

Povezovanje motorja za analizo besedila z zbirko

Če so motorji za analizo besedila po meri povezani s sistemom iskanja v podjetju, lahko izberete tistega, ki ga želite uporabiti skupaj z zbirko. Uporabniki lahko nato pri preiskovanju zbirke podajo semantične poizvedbe ter izboljšajo kakovost in natančnost rezultatov iskanja.

Preden začnete

Če želite povezati motor za analizo besedila z zbirko, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

O tej nalogi

Če je motor za analizo besedila že povezan s to zbirko, se ob povezavi drugega motorja zgodijo naslednja dejanja:

- Če izberete možnost **Brez analize po meri**, se vse preslikave analize besedila, ki ste jih pred tem definirali za zbirko, nastavijo na novo. Zbiranje se začne z uporabo privzetih sistemskih vrednosti.
- Če izberete ime drugega motorja za analizo besedila po meri, sistem zadrži vse preslikave analize besedila, ki ste jih pred tem definirali za zbirko. Če na primer spremenite `engine_1` v `engine_2`, `engine_2` podeduje datoteke za preslikavo XML, ki ste jih konfigurirali za `engine_1`.




Postopek

Če želite povezati motor za analizo besedila z zbirko, naredite naslednje:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Konfiguriraj možnosti za obdelavo besedila**.

2. Kliknite **Izberi motor za analizo besedil**. Če v sistem iskanja v podjetju niste dodali motorjev za analizo besedila ali če zbirka za analizo uporablja privzete algoritme, je ime motorja **Privzetek**.
3. Na strani Izbira motorja za analizo besedila za to zbirko, izberite ime motorja, ki ga želite uporabiti skupaj s to zbirko. Če ni na voljo nobenega motorja za analizo besedila ali če izberete **Brez analize po meri**, razčlenjevalnik pri dodajanju opomb v dokumente in pripravi dokumentov za indeksiranje uveljavi privzeta pravila za analizo besedila.

S tem povezani pojmi

-  Workflow for custom analysis integration
-  Custom text analysis integration
-  Basic concepts used in text analysis processing

Preslikovanje elementov XML v strukturo splošne analize

Če vaša zbirka vsebuje dokumente XML z označevanjem s pomenom in želite s tem označevanjem uporabnikom omogočiti semantično iskanje po indeksu za iskanje v podjetju ali tabelah relacijske baze podatkov, lahko elemente XML preslikate v strukturo splošne analize.

Preden začnete

Če želite preslikati elemente XML v strukturo splošne analize, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

Omejitve

Datoteka za preslikavo je lahko velika največ 8 MB.

O tej nalogi

Če želite procesom analize besedil po meri omogočiti dostop do specifičnih elementov v dokumentih XML ali preslikati več elementov XML v skupni tip, ki ga boste uporabili v semantičnem iskanju, lahko izdelate datoteke za preslikavo po meri. Datoteke za preslikavo se morajo držati sestava UIMA za analizo besedil.

Ko dodate datoteke za preslikavo v zbirko, ki uporablja motor za analizo besedila po meri, v izvornih dokumentih, ki bodo preslikani v opombe v strukturi splošne analize. Te opombe lahko nato uporabi motor za analizo besedila po meri. Strukturo splošne analize je mogoče preslikati v indeks in uporabnikom pri preiskovanju zbirke s semantičnim iskanjem omogočiti poizvedovanje po opombah.

Preslikati je mogoče na primer vsebino naslovljenca in elemente stranke v opombe o osebah v strukturi splošne analize. Do teh opomb lahko nato dostopajo vaši razlagalniki po meri, ki lahko odkrijejo dodatne informacije (odkrijejo lahko na primer spol osebe). Opombe o osebi lahko preslikate tudi v indeks za iskanje v podjetju, s čimer uporabnikom omogočite iskanje po osebah, ne da bi morali za to poznati izvirne elemente XML.

Če želite uporabnikom omogočiti podajanje izvornih elementov XML v poizvedbah, vam ni treba konfigurirati datotek preslikav. Namesto tega morate konfigurirati možnosti razčlenjevanja in omogočiti iskanje izvirnega XML za zbirko.

Postopek

Če želite preslikati elemente XML v strukturo splošne analize, naredite naslednje:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Konfiguriraj možnosti za obdelavo besedila**.
2. V področju **Preslikava elementov XML v strukturo splošne analize** kliknite **Dodaj preslikavo**.
3. Na strani Preslikava elementov XML v strukturo splošne analize vnesite opisno prikazno ime datoteke za preslikavo.
4. Podajte nahajališče datoteke. Če se datoteka za preslikavo nahaja v lokalnem sistemu, se lahko do nje pomaknete s pregledovanjem. Če pa je datoteka za preslikavo na indeksnem strežniku, morate vnesti popolnoma kvalificirano pot do nje.
5. Kliknite **Potrdi**. Vaša nova datoteka za preslikavo je dodana na stran Možnosti za obdelavo besedila.

S tem povezani pojmi



Workflow for custom analysis integration



Custom text analysis integration



Basic concepts used in text analysis processing



XML markup in analysis and search

S tem povezana opravila



Creating an XML elements to the common analysis structure mapping file

Preslikava strukture splošne analize v indeks

Podate lahko, kateri rezultati analize besedila iz strukture splošne analize bodo preslikani v indeks in bodo posledično na voljo uporabnikom, ki poizvedujejo po zbirki z uporabo semantičnega iskanja.

Preden začnete

Če želite preslikati strukturo splošne analize v indeks, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

Omejitve

Datoteka za preslikavo je lahko velika največ 8 MB.

O tej nalogi

S preslikavo strukture splošne analize v indeks iskanja v podjetju, uporabnikom omogočite, da podajo semantično natančne poizvedbe in izboljšajo kakovost rezultatov iskanja.

Glede na entitete in povezavo, ki jo odkrijejo razlagalniki, lahko uporabniki iščejo zasnove, ki se pojavijo v istem stavku (na primer specifična imena in ime konkurence) ali ključne besede in zasnovo (na primer ime Janko in njegova telefonska številka).





Postopek

Če želite preslikati strukturo splošne analize v indeks, naredite naslednje:


1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Konfiguriraj možnosti za obdelavo besedila**.

2. V področju **Preslikava strukture splošne analize v indeks** kliknite **Izberi datoteko za preslikavo**.
3. Na strani **Izbira datoteke za preslikavo** za to zbirko izberite datoteko za preslikavo, ki jo želite uporabiti z indeksom za iskanje v podjetju:
 - Če želite uporabiti privzeta pravila za preslikavo z indeksom za iskanje v podjetju, izberite možnost **Privzetek**.
 - Če želite preslikati strukturo analize po meri v indeks, podajte nahajališče datoteke za preslikavo. Če se datoteka nahaja v lokalnem sistemu, jo lahko poiščete s pregledovanjem. Če pa je datoteka na indeksnem strežniku, vnesite popolnoma kvalificirano pot do nje.
4. Kliknite **Potrdi**. Datoteka za preslikavo, ki ste jo podali, je prikazana na strani **Možnosti za obdelavo besedila**.

S tem povezani pojmi

-  [Workflow for custom analysis integration](#)
-  [Custom text analysis integration](#)
-  [Basic concepts used in text analysis processing](#)
-  [Index mapping for custom analysis results](#)

S tem povezana opravila

-  [Creating the common analysis structure to index mapping file](#)

Preslikava strukture splošne analize v relacijsko bazo podatkov

Podate lahko, kateri rezultati analize besedila iz strukture splošne analize bodo preslikani v relacijsko bazo podatkov, ki bo uporabljena v aplikacijah baze podatkov.

Preden začnete

Če želite preslikati strukturo splošne analize v relacijsko bazo podatkov, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

Omejitve

Datoteka za preslikavo je lahko velika največ 8 MB.

O tej nalogi

S preslikavo strukture splošne analize v relacijsko bazo podatkov aplikacijam baze podatkov omogočite uporabo podatkov. Uporabniki lahko na primer podajo poizvedbe SQL izven iskanja v podjetju in z njimi iščejo opombe, ki jih je dodala analiza besedila. Dobljene informacije lahko uporabite tudi za dodatno obdelavo besedila, na primer v aplikacijah za rudarjenje podatkov.

Za preslikavo strukture splošne analize v relacijsko bazo podatkov lahko naenkrat uporabite po eno datoteko za preslikavo. Ko naložite novo datoteko za preslikavo, prepisete trenutno datoteko.

Postopek

Če želite preslikati strukturo splošne analize v relacijsko bazo podatkov, naredite naslednje:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Konfiguriraj možnosti za obdelavo besedila**.
2. V področju **Preslikava strukture splošne analize v relacijsko bazo podatkov** kliknite **Dodaj preslikavo**.
3. Na strani Preslikava strukture splošne analize v relacijsko bazo podatkov vnesite opisno prikazno ime datoteke za preslikavo, s pomočjo katere želite preslikati informacije v relacijsko bazo podatkov.
4. Podajte nahajališče datoteke za preslikavo. Če se datoteka nahaja v lokalnem sistemu, jo lahko poiščete s pregledovanjem. Če pa je datoteka na indeksnem strežniku, vnesite popolnoma kvalificirano pot do nje.
5. Kliknite **Potrdi**. Prikazno ime datoteke za preslikavo je prikazano na strani Možnosti za obdelavo besedila.

S tem povezani pojmi



Workflow for custom analysis integration



Custom text analysis integration



Basic concepts used in text analysis processing



Database mapping for selected analysis results

S tem povezana opravila



Creating the common analysis structure to database mapping file

Konfiguriranje niti za storitev razčlenjevalnika

Če imate dovolj pomnilniških virov, lahko povečate število niti, ki so razčlenjevalniku na voljo za razčlenjevanje dokumentov.

Praden začnete

Če imate veliko število zbirk, povečajte število niti razčlenjevalnika. Zagotovite, da ima sistem dovolj pomnilnika, da lahko podpira dodatne niti. Razčlenjevalnik z eno nitjo zahteva 200 MB pomnilnika. Za vsako dodatno nit je potrebnih dodatnih 50 MB pomnilnika.

Če želite konfigurirati število niti, zagnanih za razčlenjevalnik, morate biti član z vlogo skrbnika sistema iskanja v podjetju ali skrbnik zbirke.

Postopek

Če želite konfigurirati število niti razčlenjevalnika, storite naslednje:

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Konfiguriraj možnosti razčlenjevanja**.
2. Podajte največje število niti razčlenjevalnika, ki naj jih sistem zažene ob zagonu razčlenjevalnika, in nato kliknite **Potrdi**.
3. Razčlenjevalnik znova zaženite.

Omogočanje zahtevnejše analize zloženek

Kakovost iskanja lahko izboljšate tako, da za razčlenjevalnik omogočite izboljšano analizo zloženek. Zahtevnejša analiza razčleni zloženke, tako da je mogoče vsak del obravnavati kot celoto.

Preden začnete

Če želite podati možnosti za razčlenjevanje zloženek, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

O tej nalogi

Nekateri jeziki združijo več polnopomenskih izrazov v eno samo besedo, mednje pa ne uvrstijo presledkov (*zloženke*). Zahtevnejša analiza in razčlenjevanje zloženek je koristno za iskanje v jezikih, kot je na primer nemščina, ter kritično pri iskanju v jezikih, kot je denimo korejščina.

Če za zloženke omogočite zahtevnejšo analizo, lahko uporabniki iščejo polnopomenske besede, ki tvorijo zloženke, ne da bi pri tem morali uporabiti univerzalne znake. Iskanje nemške besede **Organ** (organ) na primer vrne tudi dokumente, ki vsebujejo besedo **Organspender** (darovalec organov), ne vrne pa dokumentov, ki vsebujejo besedo **Organisation** (organizacija). V nasprotju s poizvedbo z univerzalnim znakom **Organ***, ki lahko vrne vse nize, ki sledijo besedi **Organ**, zahtevnejše iskanje vrne samo polnopomenske besede znotraj nadrejene zloženke.

Beseda **Mustermann** se na primer razdeli v dva žetona (muster in mann), ki se vsak posebej shranita v indeks. Če vnesete poizvedbo z univerzalnim znakom **Musterma***, iskalni postopki ne prepoznajo dela **Musterma** kot predpone razstavljenе besede. Posledično iskanje ne najde dokumentov z izrazom **Mustermann**. Če želite uporabnikom omogočiti vnos poizvedb z univerzalnimi znaki pri iskanju zloženek, ne omogočite zahtevnejše analize zloženek.

Uporabniško definirana besedišča, kot so sinonimi in poviševalne besede, veljajo tudi za komponente zloženke, ki so v poizvedbi uporabljene kot neodvisne besede.

Postopek

Če želite omogočiti zahtevnejšo analizo zloženek, naredite naslednje:

1. Uredite zbirko, izberite možnost Razčlenjevanje in kliknite **Konfiguriraj možnosti razčlenjevanja**.
2. Izberite potrditveno polje **Omogoči zahtevnejšo analizo zloženek** in kliknite **Potrdi**.

S tem povezani pojmi

“Univerzalni znaki v poizvedbah” na strani 168



Linguistic support for semantic search



Text analysis included in enterprise search

Omogočanje podpore za iskanje po izvornem XML

Če zbirka vsebuje dokumente XML, lahko uporabnikom omogočite, da pri iskanju dokumentov uporabijo označevanje XML, in sicer tako, da za zbirko omogočite iskanje po izvornem XML.

Preden začnete

Če želite omogočiti podporo za iskanje dokumentov XML s pomočjo iskanja po izvirnem XML, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

O tej nalogi

Iskanje po izvirnem XML, kot je na primer XPath ali fragmenti XML, lahko obrodi natančnejše rezultate iskanja tako, da uporabi označevanje XML v dokumentih znotraj poizvedbe. Uporabniki lahko podajo, da se mora iskalni izraz pojaviti znotraj določenega elementa ali atributa XML.





Računi prodajalca računalniške opreme, izdani v formatu XML utegnejo vsebovati postavke o <order> (naročilu), ki vsebujejo elementa <company> (podjetje) in <computertype> (vrsta računalnika). Če želite priklicati račune, ki vsebujejo naročila za IBM-ove prenosnike, iskanje s ključnima besedama IBM in prenosnik prikličejo tudi dokumente, ki vključujejo prenosne računalnike proizvajalca Dell in IBM-ove namizne računalnike. Z iskanjem XML lahko podate, da se mora znotraj elementa <company> pojaviti IBM, da mora element <computertype> vsebovati besedo prenosnik ter da se morata oba elementa pojaviti v istem elementu <order>. S tem prikličete račune, ki obravnavajo zgolj IBM-ove prenosnike.

Postopek

Če želite uporabnikom omogočiti iskanje po zbirki z uporabo iskanja po izvirnem XML.

1. Uredite zbirko, izberite stran Razčlenjevanje in kliknite **Konfiguriraj možnosti razčlenjevanja**.
2. Izberite možnost **Omogoči uporabnikom iskanje dokumentov XML z iskanjem po izvirnem XML**.
3. Kliknite **Potrdi**.

S tem povezani pojmi

-  Linguistic support for semantic search
-  Text analysis included in enterprise search
-  Semantic search applications
-  Semantic search query term

“Iskalna polja XML” na strani 123

Odkrivanje formata dokumentov

Sistem iskanja v podjetju s pomočjo privzete preslikave pripone URL in tipa MIME ugotovi tip dokumenta in tip razčlenjevalnika, ki naj ga uporabi za posamezni dokument.

Z urejanjem konfiguracijske datoteke `parserTypes.cfg` lahko nadomestite in razširite privzete preslikave pripon URL in tipov MIME v tipe razčlenjevalnika. Datoteka `parserTypes.cfg` definira pravila za preslikavo pripon datotek ali tipov MIME v tipe razčlenjevalnika. Preslikate lahko na primer pripono datoteke, kot je na primer `.content`, in podate, da naj dokumente tega tipa razčleni razčlenjevalnik HTML.

Različni formati dokumentov imajo različne notranje predstavitve. Sistem iskanja v podjetju uporabi za razčlenjevanje dokumentov notranje filtre in filtre drugih proizvajalcev, večino dokumentov pa razčleni s storitvami razčlenjevalnika, ki so specializirane za specifični format.

Zaznavanje formata dokumentov in dodeljevanje razčlenjevalniku poteka na naslednji način:

1. Algoritem za zaznavanje formata dokumentov preveri pripono URL-ja obdelanega dokumenta.
2. Sistem preveri tip MIME dokumenta, ki je del metapodatkov, ki jih nastavi pajek.
3. Sistem poskusi vsakemu dokumentu dodeliti ustrezn tip razčlenjevalnika. Za dokumente HTML, besedilo (TXT) in dokumente XML sistem dodeli tip razčlenjevalnika, ki je specifičen za posamezen format dokumentov.

Pri drugih formatih dokumentov bo sistem uporabil razčlenjevalnik Stellent. Ne pozabite, da je tehnologija filtriranja dokumentov Stellent sedaj v lasti družbe Oracle. Sklici na Stellent v tej dokumentaciji so zato sopomenke sklicev na tehnologijo Oracle Outside In Content Access.

Razčlenjevalnik Stellent podpira več sto formatov dokumentov, vendar je v sistemu iskanja v podjetju omogočen le podniz filtrov za dokumente. Vendar pa lahko konfiguracijske datoteke uredite tako, da bo razčlenjevalnik Stellent razčlenjeval tudi druge tipe dokumentov.

Pomembno: Filtri za dokumente, ki jih dodate, ne pripadajo podnizu filtrov za dokumente, ki so omogočeni za iskanje v podjetju v privzeti sistemski konfiguraciji, niso preizkušeni, zato tudi niso podprti.

4. Če datoteka `parserTypes.cfg` ni na voljo, bo sistem za ugotovitev tipa dokumenta in uporabljenega razčlenjevalnika uporabil privzeto preslikavo. Sistem bo za ugotovitev tipa dokumenta storil naslednje:
 - a. pripono URL bo primerjal z uporabniško definiranimi pravili za pripone v datoteki `parserTypes.cfg`;
 - b. tip MIME bo primerjal z uporabniško definiranimi pravili za tipe MIME v datoteki `parserTypes.cfg`;
 - c. pripono URL bo primerjal s privzetimi pravili sistema iskanja v podjetju;
 - d. tip MIME bo primerjal s privzetimi pravili za tipe MIME sistema iskanja v podjetju.
5. Če sistem ne more identificirati formata dokumenta, ga bo zavrnil. Morda se bo prikazalo sporočilo o napaki, ki označuje, da tip dokumenta ni podprt.

Če ste kot tip razčlenjevalnika dodelili Stellent, se lahko sporočilo o napaki prikaže v primeru, ko Stellent ne more prepoznati formata dokumenta. Do napake lahko pride v naslednjih primerih:

- Dokument je pokvarjen.
- Dokument ni v formatu, ki ga Stellent podpira. Če želite odpraviti to težavo, morate v datoteko `stellentTypes.cfg` dodati zavrnjene formate datotek. Prav tako morate posodobiti datoteko `parserTypes.cfg` in podati, da naj sistem tip MIME ali pripono zavrjenih formatov dokumentov poveže z razčlenjevalnikom Stellent.

Privzeti podprti tipi dokumentov

Sistem bo pri zaznavanju formata dokumenta ovrednotil samo nekatere tipe dokumentov.

Spodnji formati dokumentov so izvorni tipi, ki jih samodejno odkrijejo in razčlenijo vgrajene storitve razčlenjevalnika zbirke:

HTML
čisto besedilo
XML

Po privzetku razčlenjevalnik Stellent razčleni naslednje formate dokumentov:

Adobe Portable Document Format (PDF)
Lotus 1-2-3
Lotus Freelance Graphics
Lotus Word Pro
Just System Ichitaro
Microsoft Excel (različice do 2007)
Microsoft PowerPoint (različice do 2007)
Microsoft Visio
Microsoft Word (različice do 2007)
Obogateno besedilo (RTF)
StarOffice/OpenOffice Calc
StarOffice/OpenOffice Impress
StarOffice/OpenOffice Draw
StarOffice/OpenOffice Writer

Pri odprtih oblikah zapisa datotek XML v Microsoft Office in oblikah zapisa OpenDocument v OpenOffice ni potrebno spreminjati konfiguracijskih datotek.

Če želite razčleniti druge tipe dokumentov, morate posodobiti konfiguracijske datoteke (`parserTypes.cfg` in `stellenttypes.cfg`) ter podati pravila za preslikavo določenih tipov dokumentov v storitev razčlenjevalnika zbirke ali filter Stellent.

Omejitev: Obdelava dvosmernega besedila v datotekah PDF, da bi ustrezalo logičnemu zaporedju branja besedila, je izven obsega tehnologije prikazovalnika Stellent. Razčlenjevalnik Stellent ne zagotavlja pravilnega zaporedja besedila, ki ga ekstrahira iz datotek PDF. Zaradi dvosmernega besedila v datotekah PDF zaporedje, v katerem je besedilo razčlenjeno, morda ne bo enako logičnemu zaporedju branja besedila. Zaradi te omejitve pride do težave pri obdelavi datotek PDF, ki so napisane v jezikih, kot sta hebrejščina ali arabščina, ki se povečini pišejo od desne proti levi (dvosmerno).

Tipi dokumentov, povezani z razčlenjevalniki zbirke in razčlenjevalniki Stellent

Če želite zagotoviti, da bodo dokumenti v prostoru za pajkanje točno in učinkovito razčlenjeni, lahko izdelate konfiguracijske datoteke, ki podajajo, katere tipe dokumentov naj razčleni razčlenjevalnik zbirke in katere filtri za dokumente Stellent.

V zbirki sistema iskanja v podjetju bodo večino formatov dokumentov obdelali vgrajeni razčlenjevalniki HTML ali XML. Določenih tipov dokumentov (kot so dokumenti Postscript) sistem običajno ne bo razčlenjeval, druge tipe dokumentov (kot so tipi dokumentov Microsoft Word, Microsoft Excel, Microsoft PowerPoint, Lotus Freelance, Lotus 123, PDF, RT in Ichartar) pa bodo obravnavale funkcije za razčlenjevanje Stellent.

Ker so metapodatki lahko zavajajoči, lahko sistem razčlenjevalniku Stellent pošlje napačne dokumente s čistim besedilom in HTML, ta pa jih nato pošlje nazaj enemu izmed vgrajenih razčlenjevalnikov, kar vpliva na zmogljivost. Pri drugih dokumentih sistem ne more zaznati tipa dokumenta, zato dokumente preskoči. Če se želite izogniti tej situaciji, lahko izdelate konfiguracijske datoteke, ki krmilijo, kdo in kako bo razčlenil različne tipe dokumentov.

Povezovanje tipov dokumentov z razčlenjevalnikom zbirke in razčlenjevalnikom Stellent vključuje naslednje naloge:

1. Konfiguriranje tipov dokumentov za razčlenjevalnik zbirke. Ta korak vključuje izdelavo konfiguracijske datoteke, ki preslika tipe dokumentov v razčlenjevalnik, ki ga uporablja zbirka. Za vsako zbirko lahko izdelate eno izmed teh konfiguracijskih datotek.
2. Konfiguriranje tipov dokumentov za razčlenjevalnik Stellent. Ta korak vključuje izdelavo konfiguracijske datoteke, ki preslika tipe dokumentov v filtre za dokumente Stellent, ki jih uporablja zbirka. Za vsako zbirko lahko izdelate eno izmed teh konfiguracijskih datotek.
3. Zaustavitev in vnovični zagon razčlenjevalnika. Če želite, da spremembe stopijo v veljavo, s pomočjo skrbniške konzole sistema iskanja v podjetju zaženite nadzornika zbirke, za katero ste konfigurirali tipe dokumentov, nato pa razčlenjevalnik zaustavite in znova zaženite.

Povezovanje tipov dokumentov z razčlenjevalnikom zbirke

Če želite povezati določene tipe dokumentov z razčlenjevalnikom zbirke, izdelajte konfiguracijsko datoteko `parserTypes.cfg`. Skrbniška konzola za iskanje v podjetju ne nudi podpore za to nalogo.

Preden začnete

Če želite izvesti to nalogo, se morate prijaviti kot skrbnik za iskanje v podjetju.

O tej nalogi

Če konfiguracijska datoteka ne obstaja, razčlenjevalnik zbirke uporabi privzeta pravila za storitve razčlenjevalnika. Če konfiguracijska datoteka obstaja, pravila v datoteki podajajo:

- V kateri tip razčlenjevalnika bosta preslikana določena pripona URL in določen tip MIME.
- Kako razčlenjevati dokumente, katerih tip zaradi nepopolnih metapodatkov ni znan.

Format datoteke `parserTypes.cfg` je zaporedje vrstic, v katerih je vsaka vrstica eno od naslednjih pravil:

EXTENSION *extension parser*

Vse dokumenti, katerih URL se konča s podano pripono, bo obdelal podani razčlenjevalnik. Priponi ne dodajajte pike. Primerjava ne loči med malimi in velikimi črkami.

CONTENTTYPE *type/subtype parser*

Vse dokumente, katerih tip vsebine se ujema s podanim tipom/podtipom, bo obdelal podani razčlenjevalnik. Če predpostavimo, da je tip vsebine dokumenta `t/s`, pride do ujemanja, če je `t` enak tipu in `s` enak podtipu ali pa je podtip univerzalni znak (zvezdica, `*`).

UNKNOWN parser

Vse dokumente, katerih pripona in tip vsebine nista znana (to pomeni, da ju pajek ne ponudi na razpolago), bo obdelal podani razčlenjevalnik.

DEFAULT parser

Vse dokumente, ki jih ne pokrivajo druga pravila, bo obdelal podani razčlenjevalnik.

V vseh primerih mora *parser* podajati `text`, `html`, `xml`, `stellent` ali `none`, pri čemer `none` pomeni, da dokumenti tega tipa ne bodo razčlenjeni.

Če za dokument velja več pravil, prevlada tisto, ki je bolj specifično, ne glede na zaporedje, v katerem se pojavijo pravila:

- Pravilo `EXTENSION` je bolj specifično kot pravilo `CONTENTTYPE`.
- Pravilo `CONTENTTYPE`, ki vključuje podtip, je bolj specifično kot pravilo z univerzalnim znakom. Pravilo za tip vsebine `application/postscript` ima prednost pred pravilom za `application/*`.
- Izognite se uporabi dveh pravil za isto pripono ali tip vsebine. V tem primeru bo implementacija opredelila, katero pravilo bo imelo prednost.

Postopek

Če želite povezati tipe dokumentov z razčlenjevalnikom zbirke, naredite naslednje:

1. V indeksni strežnik se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
2. Upoštevajoč navodila izdelajte konfiguracijsko datoteko, pri čemer *collection_ID* označuje zbirko, ki jo boste konfigurirali:
`ES_NODE_ROOT/master_config/collection_ID.parserdriver/parserTypes.cfg`
3. Z urejevalnikom besedil uredite datoteko in podajte pravila za storitve razčlenjevalnika, nato pa shranite in zapustite datoteko.
4. Spremembe stopijo v veljavo šele, ko s skrbniško konzolo za iskanje v podjetju nadzirate razčlenjevalnik zbirke, ga zaustavite, nato pa znova zaženete.

Primer

V tem primeru vgrajeni razčlenjevalnik HTML obdela vse dokumente s pripono `txt`, `htm` ali `html`, s tipom vsebine, ki se začne na `text/`, ali z neznano pripono in tipom vsebine. Vgrajeni razčlenjevalnik XML obdela vse dokumente s pripono `xml` ali s tipom vsebine `text/xml`. Vsi drugi dokumenti, vključno z dokumenti tipa vsebine, ki se začne na `application/`, so poslani razčlenjevalniku 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
```

S tem povezana opravila

“Spreminjanje pravil zamenjave za nekatere oznake HTML” na strani 142

“Povezovanje tipov dokumentov z razčlenjevalnikom Stellent” na strani 144

S tem povezane povezave

“Privzeta pravila za storitev razčlenjevalnika zbirke” na strani 141

Privzeta pravila za storitev razčlenjevalnika zbirke

Če za preslikavo tipov datotek in tipov vsebin v razčlenjevalnik zbirke ne izdelate konfiguracijske datoteke, bodo za razčlenjevanje dokumentov uporabljena privzeta pravila.

Sledijo privzeta pravila, ki jih uporablja razčlenjevalnik zbirke:

```
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 jttd stellent
EXTENSION jtdd stellent
EXTENSION jtdd 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
```

S tem povezana opravila

“Povezovanje tipov dokumentov z razčlenjevalnikom zbirke” na strani 139

Razčlenjevanje neznanih tipov dokumentov

Če je tip dokumenta neznan (če na primer dokument nima datotečne pripone ali če ni tipa MIME, ki bi bil povezan z dokumentom), lahko konfigurirate pravila, s katerimi preprečite, da bi razčlenjevalnik izbrisal dokument.

O tej nalogi

Če razčlenjevalnik ne prepozna datotečnega formata, poskusi razčleniti dokument s privzetim razčlenjevalnikom HTML. Če vsebina ni v formatu HTML, razčlenjevalnik zavrne dokument. Zapis vseh zavrženih dokumentov se hrani na indeksnem strežniku na naslednji lokaciji:

ES_NODE_ROOT/data/collection_ID/dropped_doc_logs/dropped_docs_pd_date.log

Postopek

Če se želite temu izogniti, konfigurirajte razčlenjevalnik tako, da neznane tipe dokumentov razčleni z razčlenjevalnikom ASCII:

1. V indeksni strežnik se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
2. Izdelajte ali uredite naslednjo konfiguracijsko datoteko razčlenjevalnika, pri čemer *collection_ID* označuje zbirko, ki jo želite konfigurirati:
ES_NODE_ROOT/master_config/collection_ID.parserdriver/parserTypes.cfg
3. Dodajte naslednje pravilo v datoteko:
UNKNOWN text
4. Spremembe stopijo v veljavo šele, ko s skrbniško konzolo za iskanje v podjetju krmilite razčlenjevalnik zbirke, ga zaustavite, nato pa znova zaženete.

Spreminjanje pravil zamenjave za nekatere oznake HTML

Spremenite lahko pravila zamenjave za oznake HTML, ki jih razčlenjevalnik uporablja pri razčlenjevanju dokumentov HTML. Skrbniška konzola za iskanje v podjetju ne nudi podpore za to nalogo.

Praden začnete

Če želite izvesti to nalogo, se morate prijaviti kot skrbnik za iskanje v podjetju.

O tej nalogi

Da bi indeksu za iskanje v podjetju nudil besedilne informacije, razčlenjevalnik HTML zamenja informacije o oznakah (oznake HTML) z drugimi znaki, ki izražajo pomen oznak na enak način, kot jih obravnava spletni brskalnik. Oznaka za odstavek (<p>) na primer pri strukturi splošne analize (CAS) in v indeksu povzroči ločilo odstavka. Pravila zamenjave določajo, kateri deli besedilnih informacij se prikažejo v istem odstavku.

Če želite pomen nekaterih oznak HTML približati standardu HTML, lahko posodobite konfiguracijsko datoteko razčlenjevalnika za zbirko in spremenite pravila zamenjave.

Nasvet: Če se želite seznaniti z drugimi načini za nadzor obravnavanja oznak HTML v dokumentih HTML, preberite <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011251>.

Postopek

Če želite spremeniti pravila zamenjave za oznake HTML, naredite naslednje:

1. V indeksni strežnik se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
2. Uredite datoteko ES_NODE_ROOT/master_config/collection_ID.parserdriver/collection.properties, pri čemer *collection_ID* označuje zbirko, ki jo želite konfigurirati.
3. Nastavite naslednji parameter na vrednost true:
treview.tokenizer.newHtmlTagReplacement=true

Pravila zamenjave za naslednje oznake HTML se spremenijo v vrednosti, prikazane spodaj. Vsa druga pravila zamenjave za oznake HTML še naprej uporabljajo privzeta pravila.


```
dfn    EMPTY
div    NEW LINE
q      BLANK
span   EMPTY
```

4. Spremembe stopijo v veljavo šele, ko s skrbniško konzolo za iskanje v podjetju krmilite razčlenjevalnik zbirke, ga zaustavite, nato pa znova zaženete.

S tem povezana opravila

“Povezovanje tipov dokumentov z razčlenjevalnikom zbirke” na strani 139

S tem povezane povezave

“Privzeta pravila zamenjave za HTML”

Privzeta pravila zamenjave za HTML

Razčlenjevalnik HTML interpretira in zamenja informacije oznak v dokumentih HTML ter pridobi besedilne informacije za indeks za iskanje v podjetju.

Razčlenjevalnik HTML za iskanje v podjetju pri razčlenjevanju oznak HTML uporablja naslednja pravila zamenjave. Prvi stolpec prikazuje ime oznake HTML. Drugi stolpec prikazuje nadomestno vrednost.

```
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

S tem povezana opravila

“Spreminjanje pravil zamenjave za nekatere oznake HTML” na strani 142

Povezovanje tipov dokumentov z razčlenjevalnikom Stellent

Če želite podati, katere tipe dokumentov bodo razčlenili filtri dokumentov Stellent, izdelajte konfiguracijsko datoteko `stellenttypes.cfg`. Skrbniška konzola za iskanje v podjetju ne nudi podpore za to nalogo.

Tehnologija Oracle Outside In Technology:

Tehnologija filtriranja dokumentov Stellent je sedaj v lasti družbe Oracle. Sklici na Stellent v tej dokumentaciji so zato sopomenke sklicev na tehnologijo Oracle Outside In Content Access.

Omejitve

OmniFind Enterprise Edition podpira filtre Stellent za naslednje tipe dokumentov:

- Adobe Portable Document Format (PDF)
- Lotus 1-2-3
- Lotus Freelance Graphics
- Lotus Word Pro
- Ichitaro
- Microsoft Excel
- Microsoft PowerPoint
- Microsoft Visio
- Microsoft Word
- Rich Text Format (RTF)

Če želite v indeks za iskanje v podjetju vključiti dodatne tipe dokumentov, Stellent filter pa podpira formate teh dokumentov, lahko v konfiguracijski datoteki `stellenttypes.cfg` konfigurirate pravila za razčlenjevanje.

Pomembno: Formati dodanih dokumentov, ki se ne ujemajo s tipi dokumentov na seznamu, niso bili preizkušeni za iskanje v podjetju in niso podprti.

Celoten seznam formatov dokumentov Stellent najdete v temi *Outside In Technology: podprti datotečni formati* na naslovu http://www.oracle.com/technology/products/content-management/oit/ds_oitFiles.pdf.

Preden začnete

Če želite izvesti to nalogo, se morate prijaviti kot skrbnik za iskanje v podjetju.

O tej nalogi

Konfiguracijska datoteka `stellenttypes.cfg` podaja naslednje:

- Pravila za sprejetje za datotečne tipe, ki jih bo razčlenjeval razčlenjevalnik Stellent. Tip datoteke ustreza enemu od tipov datotek iz knjižnice Stellent.

- Izvirna pravila za datotečne tipe, ki bodo vrnjeni razčlenjevalniku zbirke v obdelavo z enim od vgrajenih razčlenjevalnikov. To dejanje je potrebno, ker lahko razčlenjevalnik zbirke pomotoma, zaradi zavajajočih metapodatkov pošlje dokument razčlenjevalniku Stellent.
- Pravila za zavrnitev za datotečne tipe, ki bodo zavrnjeni, ker jih iskanje v podjetju ne podpira.

Če je konfiguracijska datoteka podana, a ne obstaja, zagon razčlenjevalnika ne bo uspel. Če za lastnost `OutsideInSupportedTypes` v datoteki `stellent.properties` niste podali konfiguracijske datoteke, sistem uporabi privzeta pravila za razčlenjevalnike Stellent.

Konfiguracijska datoteka navaja tipe dokumentov in podaja, kako bodo obravnavani. Format datoteke je zaporedje vrstic, kjer vsaka vrstica predstavlja pravilo, ki ustreza enemu od naslednjih formatov:

```
accept DEFAULT
accept ALL doc_type
accept stellent_type doc_type
native DEFAULT
native stellent_type doctype
reject stellent_type
```

Kjer je:

doc_type

Vrednost, ki bo uporabljena za žeton poizvedbe `doctype`. Dokumente je mogoče iskati po tipu dokumenta. Uporabnik lahko poda na primer `$doctype::pdf` za iskanje dokumentov PDF.

stellent_type

Je ena od vrednosti za tip filtra v knjižnici Stellent, na primer `FI_123R1`.

DEFAULT

Pomeni, da seznam sprejetih ali izvirnih tipov, odvisen od tipa pravila, vključuje vsa privzeta pravila. Ta možnost vam omogoča, da razširite privzeto konfiguracijo, namesto da jo nadomestite.

All

Pomeni, da so vsi tipi, ki niso izrecno navedeni, sprejeti s podanim žetonom `doctype`.

Pravila v konfiguracijski datoteki se obdelajo na naslednji način:

1. Če obstaja pravilo `reject` za *stellent_type*, sistem ne sprejme dokumenta.
2. Če obstaja pravilo `native` za *stellent_type* (vključno s privzetimi pravili za razčlenjevanje, če je podan `native DEFAULT`), se dokument vrne vgrajenemu razčlenjevalniku, skupaj z vrednostjo za žeton *doc_type*, ki ga podaja to pravilo. Vrednost *doc_type* mora biti `txt`, `htm` ali `xml` in označevati golo besedilo, HTML ali XML.
3. Če obstaja pravilo `accept` za *stellent_type* (vključno s privzetim seznamom, če je podan `accept DEFAULT`), sistem sprejme dokument.
4. Če je podan `accept ALL`, sistem prav tako sprejme dokument.
5. V nasprotnem primeru sistem zavrne dokument in ga ne razčleni.

Če sistem sprejme tip dokumenta, uporabi vrednost *doc_type*, ki je bila podana v pravilu. To vrednost nato poleg razčlenjene vsebine pošlje nazaj v razčlenjevalnik zbirke.

Postopek

Če želite povezati tipe dokumentov z razčlenjevalnikom Stellent, naredite naslednje:

1. V indeksni strežnik se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.

2. Uredite datoteko `ES_NODE_ROOT/master_config/collection_ID.stellent/stellent.properties`, pri čemer `collection_ID` označuje zbirko, ki jo želite konfigurirati.
3. Za lastnost `OutsideInSupportedTypes` podajte absolutno pot do konfiguracijske datoteke, ki jo izdelujete.

Za posamezno zbirko lahko na primer izdelate naslednjo konfiguracijsko zbirko in jo shranite z drugimi datotekami, specifičnimi za zbirko:

```
ES_NODE_ROOT/master_config/collection_ID.stellent/stellenttypes.cfg
```

Kot drug zgled lahko izdelate naslednjo konfiguracijsko datoteko, tako da lahko uporabite iste nastavitve za vse zbirke in to datoteko shranite skupaj z drugimi datotekami na ravni sistema. Če uporabite ta pristop, ne pozabite podati te poti v datoteki `stellent.properties` za vsako zbirko, kot je podano v koraku 2.

```
ES_INSTALL_ROOT/default_config/stellent/stellenttypes.cfg
```

4. Z urejevalnikom besedil izdelajte konfiguracijsko datoteko in podajte pravila za razčlenjevanje Stellent, nato pa shranite datoteko in jo zapustite.
5. Spremembe stopijo v veljavo šele, ko s skrbniško konzolo za iskanje v podjetju nadzirate razčlenjevalnik zbirke, ga zaustavite, nato pa znova zaženete.

Primeri

V naslednji konfiguracijski datoteki seja Stellent poleg dokumentov s privzetega seznama podprtih tipov sprejme tudi dokumente v formatu Microsoft Visio.

```
accept DEFAULT
accept FI_VISIO3 visio
accept FI_VISIO4 visio
accept FI_VISIO5 visio
accept FI_VISIO6 visio
```

V naslednji konfiguracijski datoteki sistem sprejme dokumente Postscript in omogoči iskanje za dokumente tipa ps, dokumente v formatu X pixmap (XPM) vrne v vgrajeni besedilni razčlenjevalnik, dokumente v slikovnem formatu PNG zavrne, vse druge tipe pa sprejme, omogoči iskanje po njih ter jim določi tip dokumenta drugo.

```
accept DEFAULT
accept FI_POSTSCRIPT ps
native FI_XPIXMAP txt
reject FI_PNG
accept ALL other
```

S tem povezana opravila

“Povezovanje tipov dokumentov z razčlenjevalnikom zbirke” na strani 139

S tem povezane povezave

“Privzeta pravila za razčlenjevanje za razčlenjevalnike Stellent”

Privzeta pravila za razčlenjevanje za razčlenjevalnike Stellent

Če ne izdelate konfiguracijske datoteke za preslikavo datotečnih tipov v filtre dokumentov Stellent, razčlenjevalnik razčleni dokumente s pomočjo privzetih pravil.

Razčlenjevalnik Stellent sprejme in razčleni naslednje tipe dokumentov:

```
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

Razčlenjevalnik Stellent vrne naslednje tipe dokumentov razčlenjevalniku zbirk v obdelavo z enim od vgrajenih razčlenjevalnikov:

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_ASM0708	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

S tem povezana opravila

“Povezovanje tipov dokumentov z razčlenjevalnikom Stellant” na strani 144

Jezikovna podpora in podpora za kodno stran

Razčlenjevalnik in iskalni strežniki različno obravnavajo jezikovno obdelavo sistema iskanja v podjetju.

Razčlenjevalnik zaradi jezikovne obdelave ne razlikuje med jeziki in področnimi nastavitvami. Če pa uporabnik išče po zbirki, ki vključuje dokumente v več jezikih, iskalni strežniki omogočajo, da rezultate iskanja omejite na specifični jezik ali področne nastavitve.

Če na primer metapodatki v angleškem dokumentu za področne nastavitve vsebujejo `en_US`, bo dokument indeksiran kot dokument v angleškem jeziku (`en`) in kot dokument, ki za angleščino uporablja področne nastavitve Združenih držav Amerike (`en_US`). Ta tip indeksiranja omogoča pravilno predstavitev za področne nastavitve specifičnih informacij, kot so na primer številke, datumi in časi. Uporabniki lahko tako pri iskanju po zbirki najdejo dokument, ne glede na to, ali iščejo po dokumentih `en` ali `en_US`.

Če je dokument indeksiran samo po kodi jezika, kot je na primer `en`, ne bo indeksiran po področnih nastavitvah. Če uporabniki v zbirki iščejo, na primer dokumente `en_US`, sistem ne bo našel ničesar.

Sistem iskanja v podjetju nudi jezikovno podporo za naslednje jezike in dvomestne kode jezikov, kot so dokumentirane v standardu ISO 639:

Jeziki preprostih besedil:

- en=angleščina
- sq=albanščina
- az=azerbajdžanščina - latinica
- bg=bolgarščina
- be=beloruščina
- ca=katalonščina
- hr=hrvaščina
- cs=češčina
- da=danščina
- nl=nizozemščina
- et=estonsščina
- fi=finščina
- fr=francoščina
- de=nemščina
- el=grščina
- hu=madžarščina
- is=islandščina
- id=indonezijsščina
- in=indonezijsščina
- it=italijanščina
- kk=kazahstanščina
- lv=latvijščina
- lt=litvanščina
- lo=laoščina
- mk=makedonščina
- ms=malajščina
- mt=malteščina
- no=norveščina
- nb=norveščina (Bokmal)
- pl=poljščina
- pt=portugalščina
- ro=romunščina
- ru=ruščina
- sr=srbščina (cirilica)
- sh=srbščina (latinica)

sk=slovaščina
sl=slovenščina
es=španščina
sv=švedščina
tr=turščina
uk=ukrajinščina
cy=valijanščina

Pojmovni jeziki:

Pri poenostavljeni in tradicionalni kitajščini so namesto dvomestnih kod uporabljene razširjene kode jezikov.

zh-CN=kitajščina (poenostavljena)
zh-TW=kitajščina (tradicionalna)
ja=japonščina
ko=korejščina

Kompleksni jeziki besedil:

ar=arabščina
as=asamščina
bn=bengalščina
gu=Gujarati
iw=hebrejščina
he=hebrejščina
hi=hindujščina
kn=Kannada
ml=Malayalam
mr=Marathi
or=Oriya
pa=Punjabi
ta=tamilščina
te=Telugu
th=tajščina
ur=urdu
vi=vietnamščina

Sistem iskanja v podjetju lahko večino teh jezikov zazna samodejno, prav tako pa lahko samodejno zazna kodno stran, uporabljeno v dokumentih s čistim besedilom. Pri konfiguriranju pajka lahko onemogočite samodejno zaznavanje jezika in kodne strani, če želite podati izrecni jezik ali kodno stran, ki naj jo sistem uporabi.

Samodejno zaznavanje jezika

Sistem iskanja v podjetju lahko obdela dokumente v skoraj vseh jezikih.

Če je dokument v enem izmed naslednjih jezikov, lahko sistem samodejno zazna njegov jezik. Če poznate jezik svojih dokumentov, lahko pri konfiguriranju pajka podate jezik, ki naj ga sistem uporabi, in mu tako preprečite, da bi ga poskusil zaznati samodejno.

arabščina
bolgarščina
češčina
kitajščina, poenostavljena
kitajščina, tradicionalna
danščina
nizozemščina
angleščina
finščina
francoščina, kanadska
francoščina, državna
nemščina, državna
nemščina, švicarska
grščina
hebrejščina
madžarščina
islandščina

italijanščina
japonščina
korejščina
norveščina, Bokmal
poljščina
portugalščina, brazilska
portugalščina, državna
romunščina
ruščina
španščina
švedščina
tajščina
turščina

Samodejno zaznavanje kodne strani

Sistem iskanja v podjetju podpira dokumente v različnih kodnih straneh.

Sistem lahko pri besedilnih datotekah samodejno zazna naslednje kodne strani. Pri drugih formatih dokumentov sistem kodno stran zazna s pomočjo metapodatkov v dokumentu, kot so na primer elementi metapodatkov HTML. Če poznate kodno stran dokumentov, lahko pri konfiguriranju pajka podate kodno stran, ki naj jo sistem uporabi, in mu tako preprečite, da bi jo poskusil zaznati samodejno.

Oblike kodiranja Unicode:

UTF-8
UTF-16BE
UTF-16LE

Oblike večbajtnega kodiranja:

Shift-JIS
ISO-2022-CN
ISO-2022-JP
ISO-2022-KR
GB18030
EUC-JP
EUC-KR

Oblike enobajtnega kodiranja:

ISO-8859-1: danščina, nizozemščina, nemščina, angleščina, francoščina, italijanščina
italijanščina, norveščina, portugalščina, španščina, švedščina
ISO-8859-2: češčina, madžarščina, poljščina, romunščina
ISO-8859-5: ruščina
ISO-8859-6: arabščina
ISO-8859-7: grščina
ISO-8859-8: hebrejščina, hebrejščina v vidnem vrstnem redu
ISO-8859-9: turščina
Windows-1250: češčina, madžarščina, poljščina, romunščina
Windows-1251: ruščina
Windows-1252: danščina, nizozemščina, nemščina, angleščina, francoščina,
italijanščina, norveščina, portugalščina, španščina, švedščina
Windows-1253: grščina
Windows-1254: turščina
Windows-1255: hebrejščina
Windows-1256: arabščina
KOI8-R: ruščina

Zaznavanje nabora znakov ni natančna operacija. Proces zaznavanja kodne strani poskusi identificirati nabor znakov, ki se najboljše ujema z značilnostmi bajtnih podatkov, vendar je proces deloma statistične narave, zato ni mogoče zagotoviti točnih rezultatov.

Vhodni podatki naj bodo zaradi večje natančnosti predvsem v enem jeziku. Sistem prav tako zahteva najmanj nekaj sto bajtov čistega besedila v danem jeziku.

Če pride do navzkrižja med zaznanim kodiranjem in podprtimi kodiranjmi, bo sistem za zbirko uporabil privzeto kodno stran.

Jezikovna analiza za kitajske, japonske in korejske dokumente

Če želite izboljšati berljivost dokumentov, napisanih v kitajskem, japonskem in korejskem jeziku, lahko podate možnosti za jezikovno analizo.

Pri kitajskih, japonskih in korejskih dokumentih lahko podate, da naj razčlenjevalnik za slovarsko analizo uporabi metodo segmentacije n-gram. Pri kitajskih in japonskih dokumentih lahko razčlenjevalnik konfigurirate tudi tako, da iz praznih prostorov odstrani znake za novo vrstico.

Segmentacija n-gram

Pri izdelavi zbirke morate izbrati tip slovarske analize, ki ga želite uporabiti za razčlenjevanje dokumentov, napisanih v jezikih, ki za ločevanje besed ne uporabljajo praznega prostora.

Segmentacija s praznimi prostori na osnovi Unicode uporablja za ločevanje besed prazen prostor. Segmentacija n-gram pa upošteva kot posamezno besedo prekrivajoča se zaporedja s poljubnim številom znakov. Pri jezikih kot so kitajščina, japonščina in korejščina, ki za ločevanje besed ne uporabljajo praznega prostora, lahko segmentacija n-gram vrne boljše rezultate iskanja kot segmentacija s praznimi prostori na osnovi Unicode.

Metodo segmentacije, ki jo želite uporabiti pri razčlenjevanju dokumentov, izberite pri izdelavi zbirke. Po izdelavi zbirke lahko to nastavitvev prikažete v možnostih razčlenjevanja, ne morete pa je spremeniti.

Informacije o tem, kako konfigurirati podporo za celotno razčlenjevanje n-gram in izdelavo žetonov v zbirkah za iskanje v podjetju, in informacije o načinu obravnave znakov v zbirkah, ki so konfigurirane za podporo celotnemu razčlenjevanju n-gram, poiščite v temi <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088>.

Odstranjevanje presledkov iz besedila

Razčlenjevalnik lahko konfigurirate tako, da iz besedila odstrani presledke.

Preden začnete

Če želite izvesti to nalogo, se morate prijaviti kot skrbnik za iskanje v podjetju.

O tej nalogi

Če omogočite to možnost za zbirko, razčlenjevalnik odstrani zaporedja presledkov, ki ločujejo znake, sestavljene iz dveh črk. Morda boste želeli presledke odstraniti, če so vaši dokumenti napisani v jeziku, mej besed (na primer kitajščina ali japonščina) ne ločuje s presledki.

Ko konfigurirate razčlenjevalnik za odstranitev znakov za presledek, lahko podate tudi, ali želite odstraniti samo presledke, ki se pojavijo med znaki DBCS ali pa bi radi odstranili vse presledke, ne glede na kontekst. Morda boste slednjo možnost želeli uporabiti, če ste v japonski dokument vnesli besedilo v angleščini, in želite tudi iz njega odstraniti presledke.

Razčlenjevalnik odstrani naslednje znake:

- Tab (0x09)
- LF ali pomik v novo vrstico (0x0A)
- CR ali pomik na začetek vrstice (0x0D)

Spremembe stopijo v veljavo šele zatem, ko zaustavite razčlenjevalnik in ga znova zaženete. Če želite uveljaviti spremembe za dokumente, ki so že shranjeni v indeksu, znova prepakajte dokumente in znova zgradite glavni indeks.

Postopek

Če želite odstraniti znake za presledek iz besedila, naredite naslednje:

1. V indeksni strežnik se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
2. Z urejevalnikom besedil uredite naslednjo datoteko, kjer je *collection_ID* ID, ki ste ga podali za zbirko (ali ki ga je dodelil sistem) pri njeni izdelavi:

```
ES_NODE_ROOT/master_config/collection_ID.parserdriver/collection.properties
```

3. Podajte, kako nameravate odstraniti znake za presledek:

- Če želite odstraniti presledke med znaki DBCS, nastavite vrednost lastnosti `removeCjNewlineChars` na `true`:
`removeCjNewlineChars=true`
- Če želite odstraniti presledke iz celotnega dokumenta, nastavite vrednost lastnosti `removeCjNewlineChars` na `true` in vrednost lastnosti `removeCjNewlineCharsMode` na `all`:

```
removeCjNewlineChars=true  
removeCjNewlineCharsMode=all
```

Skrbnišтво nad indeksi

Sistem iskanja v podjetju izdelava za vsako zbirko indeks, ki ga vzdržuje tako, da občasno posodablja njegovo vsebino, s tem pa zagotavlja, da bodo uporabniki vedno imeli dostop do najnovejših informacij.

Če želite, da bo možno iskanje po podatkih, ki jih zberejo pajki, morate izdelati indekse. Ob prvi izdelavi zbirke bo sistem iskanja v podjetju izdelal indeks za vse podatke, ki so bili na začetku pajkani. Če pajek pajka nove in spremenjene vire podatkov, se za novo vsebino izdelava posodobite (imenovana *delta indeks*). Na koncu bo sistem moral posodobitve združiti z osnovnim indeksom. Ta postopek združevanja se imenuje gradnja *glavnega indeksa*. Ob vsaki gradnji delta ali glavnega indeksa sistem prekopira novo vsebino v iskalne strežnike in omogoči iskanje po njej.

Pajki zbirajo podatke stalno ali občasno po razporedu. Če indekse pogosto posodabljate, bodo uporabniki lahko iskali po ažurnih podatkih. Vendar pa morate indeks, ki ga sistem vedno posodablja, znova zgraditi. Indeks se tako povečuje in porablja vedno več sistemskih virov. Če želite ohraniti optimalno zmogljivost, redno izvajajte gradnjo glavnega indeksa.

Pogostost gradnje glavnega indeksa je odvisna od naslednjega:

- od sistemskih virov (prostor v datotečnem sistemu, hitrost procesorja in pomnilnik)
- koliko dokumentov mora sistem pajkati in znova pajkati
- tip podatkov za pajkanje
- kako pogosto spremenite pravila kategorij (spremembe bodo stopile v veljavo šele po gradnji glavnega indeksa)
- kako pogosto namesto razporejenega izvajanja izsilite zagon pajka

Pri zbirkah z več milijoni dokumentov, zgrajenih večinoma iz spletnih dokumentov, gradnjo glavnega indeksa izvedite približno enkrat na dan, indeks pa posodobite vsako uro ali vsaki dve uri.

Če želite ohraniti ažurni indeks, ki omogoča iskanje, izvedite naslednje naloge:

- specifikiranje razporedov za gradnjo indeksa
- spreminjanje razporeda indeksiranja
- omogočitev in onemogočitev razporeda indeksiranja
- konfiguriranje sočasne gradnje indeksov

Če želite podati možnosti, ki vplivajo na prikaz indeksa uporabniku, lahko izvedete naslednje naloge:

- konfiguriranje podpore za univerzalne znake v poizvedbah
- konfiguriranje območij za omejitev obsega dokumentov, po katerih lahko uporabniki iščejo
- krčenje dokumentov iz istega vira v rezultatih iskanja
- odstranjevanje URI-jev iz indeksa

S tem povezana opravila

“Nadziranje dejavnosti indeksa za zbirko” na strani 285

“Nadziranje čakalne vrste indeksov za iskanje v podjetju” na strani 285

Razpored gradenj indeksov

Za gradnjo glavnega indeksa in posodabljanje indeksa z novo vsebino je mogoče podati razporede.

Preden začnete

Če želite določiti razpored izgradnje indeksa, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

O tej nalogi

Če želite uporabnikom zagotoviti neprekinjen dostop do najnovejših informacij v virih, ki jih preiskujejo, načrtujte redno gradnjo indeksa. Med gradnjo glavnega indeksa sistem znova zgradi celoten indeks. Procesi indeksiranja preberejo vse podatke, ki so jih zbrali pajki in jih je razčlenil razčlenjevalnik. Med gradnjo delta indeksa sistem omogoči iskanje po informacijah, v katerih je bilo od zadnje gradnje glavnega indeksa izvedeno pajkanje.

Po privzetku je izbrana možnost za načrtovanje gradenj indeksa. Ta možnost sporoči procesu razporejanja, naj načrtuje naloge za gradnjo glavnega in delta indeksov ob zagonu sistema iskanja v podjetju. Če želite preprečiti izvedbo načrtovane gradnje indeksa, lahko kadarkoli počistite izbiro potrditvenega polja **Omogoči ob zagonu sistema**. To boste morda morali storiti, da bi odpravili morebitne težave.

Da bi prihranil sistemske vire in izboljšal zmogljivost, sistem samodejno preveri, ali je treba za izdelani indeks uveljaviti spremembe. Če spremembe niso potrebne, zavrže zahtevo za načrtovano gradnjo.

Postopek

Če želite načrtovati gradnje indeksa, naredite naslednje:

1. Uredite zbirko, izberite stran Indeksiranje in kliknite **Načrtuj gradnje indeksa**.
2. Če želite podati, kako pogosto bo v indeks vnesena nova vsebina, na strani Razpored gradenj indeksov v področju **Podajte razpored za gradnjo delta indeksa** podajte naslednje možnosti:
 - a. V področju **Začni na** lahko v polja **Leto**, **Mesec**, **Dan**, **Ura** in **Minuta** vnesete, kdaj želite začeti graditi prvi delta indeks.
 - b. V področju **Interval posodabljanja** lahko v poljih **dni**, **ur** in **minut** podate, kako pogosto želite graditi delta indekse.
Praviloma je treba delta indekse graditi precej pogosto, na primer vsako uro ali dve. Glede na to, kako pogosto se spreminja izvorna vsebina, je lahko podani interval pogostejši ali redkejši. Podate lahko na primer vsako uro (0 dni in 1 ura) ali vsakih 12 ur (0 dni in 12 ur).
3. Če želite podati, kako pogosto bo indeks zgrajen povsem na novo, v področju **Podajte razpored za gradnjo glavnega indeksa** podajte naslednje možnosti:
 - a. V področju **Začni na** lahko v poljih **Leto**, **Mesec**, **Dan**, **Ura** in **Minuta** podate, kdaj želite prvič zgraditi glavni indeks.
 - b. V področju **Interval posodabljanja** lahko v poljih **dni**, **ur** in **minut** podate, kako pogosto želite graditi glavni indeks.
Praviloma je treba glavni indeks graditi redno, na primer vsakih 24 ur. Glede na to, kako pogosto se spreminja izvorna vsebina, je lahko podani interval pogostejši ali redkejši. Podate lahko na primer vsakih 12 ur (0 dni in 12 ur) ali vsak poltretji dan (2 dneva in 12 ur).

4. Kliknite **Potrdi**.

Spreminjanje razporeda indeksiranja

Spremenite lahko razpored za gradnjo indeksa.

Praden začnete

Če želite spremeniti razpored indeksiranja, morate biti član z vlogo skrbnika sistema iskanja v podjetju ali pa morate biti skrbnik te zbirke.

Postopek

Če želite spremeniti razpored indeksiranja, storite naslednje:

1. Uredite zbirko, izberite stran Indeksiranje in spremenite ustrezne vrednosti v poljih **Mesec, Dan, Leto** in **Ura**. Podajte, kako pogosto naj sistem posodobi indeks in kako pogosto naj zgradi glavni indeks.
2. Kliknite možnost **Uveljavi**.

Omogočanje in onemogočanje razporedov indeksiranja

Omogočite ali onemogočite lahko razporede za gradnjo indeksa.

Praden začnete



Če želite omogočiti ali onemogočiti razpored indeksiranja, morate biti član z vlogo skrbnika sistema iskanja v podjetju ali pa morate biti skrbnik te zbirke.


O tej nalogi

Razpored za indeks lahko onemogočite, če želite preprečiti zagon razporejene gradnje indeksa. Če želite na primer odpraviti težave, lahko razpored onemogočite in tako preprečite, da bi sistem ob razporejenem datumu in času začel z gradnjo indeksa.

Razpored lahko omogočite ali onemogočite med urejanjem zbirke, prav tako pa ga lahko omogočite ali onemogočite med nadzorom zbirke.

Postopek

1. Če želite razpored za indeks omogočiti ali onemogočiti med urejanjem zbirke, storite naslednje:
 - a. Uredite zbirko, ki jo želite spremeniti.
 - b. Na strani Indeksiranje izberite potrditveno polje **Omogoči ob zagonu sistema** ali pa razveljavite izbiro le-tega in s tem omogočite ali onemogočite razpored za posodabljanje indeksa.
 - c. Izberite potrditveno polje **Omogoči ob zagonu sistema** ali razveljavite izbiro le-tega in s tem omogočite ali onemogočite razpored za gradnjo glavnega indeksa.
 - d. Kliknite možnost **Uveljavi**.
2. Če želite razpored za indeks omogočiti ali onemogočiti med nadzorom zbirke, storite naslednje:
 - a. Zaženite nadzornika zbirke, ki jo želite spremeniti.
 - b. Če ste razporedili indeksiranje in ne želite, da se gradnja zažene ob razporejenem datumu in času, na strani Indeksiranje kliknite možnost  **Onemogoči razpored**. Sistem indeksa ne bo zgradil, dokler razporeda ne omogočite ali kliknete možnost  **Zaženi** ter s tem zaženete proces gradnje indeksa.

- c. Če je indeksiranje razporejeno, razpored za gradnjo pa je onemogočen, kliknite možnost  **Omogoči razpored**.
Gradnja indeksa se bo izvedla na datum in čas, ki ste ga podali v razporedu indeksiranja.

Konfiguriranje sočasnih gradenj indeksa

Uporabo virov indeksiranja kontrolirate s podajanjem, koliko zbirkam se lahko sočasno obdeluje zahteve za gradnjo indeksov. Če imate zadostne systemske vire, lahko izboljšate kvaliteto iskanja tako, da omogočite posodabljanje indeksov istočasno, ko se gradi glavni indeks.

Preden začnete

Možnosti gradnje indeksa za sistem lahko podajajo samo člani skrbniške vloge za iskanje v podjetju.

O tej nalogi

Iskanje v podjetju lahko sočasno gradi več indeksov hkrati tako, da souporablja vire med zbirkami, kar omogoča paralelno obdelavo zahtev po izgradnji indeksov za več zbirk. S souporabo postopkov zagotovite, da reorganiziranje zelo velikega indeksa ne onemogoči razpoložljivosti drugih indeksov, ki v čakalni vrsti čakajo na gradnjo.

Če zahtevate gradnjo indeksa ali pa jo razporejate, bo ta uvrščena v čakalno vrsto za indeksiranje in bo čakala, da pride na vrsto za obdelavo. Ker ima vsaka zbirka svoj indeks, je lahko istočasno v čakalni vrsti za indeksiranje več zahtev za gradnjo indeksa različnih zbirk. Ko konfigurirate možnosti indeksiranja za sistem, podate, koliko zbirk lahko hkrati souporablja vire za indeksiranje in zanje poteka paralelna obdelava zahtev.


Podate lahko tudi, da se zahteve za posodabljanje indeksa obdelujejo istočasno s potekom gradnja glavnega indeksa za zbirko. Če omogočite to možnost, bo sistem iskalne strežnike osvežil z najnovejšimi dokumenti (preko delta indeksa), medtem ko bo sistem obdeloval gradnjo glavnega indeksa, ki se izvaja počasneje. Vendar je gradnja indeksa postopek, ki zahteva veliko virov. Med gradnjo indeksa je v uporabi velika količina pomnilnika sistema in prostora na disku. Če omogočite to možnost in nimate dovolj prostora na disku ali pomnilnika, se lahko zgodi, da se bo zmanjšala splošna zmogljivost sistema.

Če povečate število sočasnih gradenj indeksov, se zahteve za gradnjo indeksa, ki so že v vrsti ne zaženejo samodejno. Vaše spremembe vplivajo na nove gradnje indeksov, ki pridejo v vrsto, ko ste spremenili to vrednost.

Če zmanjšate število sočasnih gradenj indeksov, se trenutne gradnje indeksov ne zaustavijo samodejno. Vaše spremembe stopijo v veljavo potem, ko se trenutne gradnje indeksov zaustavijo, kar gradnjam indeksov v vrsti omogoči, da se zaženejo.

Postopek

Podajanje možnosti gradnje indeksov za sistem:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
3. Na strani Indeksiranje kliknite **Konfiguriranje možnosti indeksiranja**.
4. Na strani Možnosti indeksiranja na ravni sistema vnesite število zbirk, ki lahko souporablja vire sistema in za katere vzporedno poteka obdelava zahtev za gradnjo indeksa.

Število zbirk, ki souporablajo vire za indeksiranje, ne more biti večje od števila zbirk v sistemu iskanja v podjetju. Če imate na primer pet zbirk, morate vnesti število, ki je manjše ali enako pet.

5. Če imate zadostne sistemske vire za podporo več sočasnih gradenj indeksov za posamezne zbirke, lahko izberete možnost, ki omogoča gradnjo delta indeksa sočasno z gradnjo glavnega indeksa.

Omejitev: Če izberete to možnost, medtem ko poteka gradnja indeksa za katero od zbirk, gradnja indeksa morda ne bo uspela. Preden omogočite sočasno gradnjo delta in glavnega indeksa, onemogočite vse načrtovane gradnje indeksa in počakajte, da se hkratni gradnji indeksa zaustavita (ali pa ju zaustavite pred koncem). Potem ko omogočite to možnost, omogočite še načrtovane gradnje indeksa, ki ste jih poprej onemogočili.

Izgradnja indeksov samo v primeru odkritih sprememb

V primeru načrtovane gradnje indeksa do gradnje pride samo, če sistem odkrije spremembe, ki jih je treba uveljaviti za indeks. Z ukazom **startIndexBuild** lahko zaženete gradnjo indeksa in izberete, ali naj sistem preverja spremembe.

Preden začnete

Če želite ročno podati, naj do gradnje indeksa pride samo v primeru odkritih sprememb, se morate prijaviti kot skrbnik iskanja v podjetju.

Omejitve

Sistem ne preverja sprememb, ki utegnejo vplivati na indeks, dokler se glavni indeks ne zgradi vsaj dvakrat. Takšna izvedba je namenjena spremembam, ki jih ni mogoče odkriti, dokler se glavni indeks ne zgradi vsaj enkrat, na primer spremembe v definiciji obsega ali pravila za skrčenje rezultatov iskanja.

O tej nalogi

Gradnje indeksa porabljajo sistemske vire, kot so CPE, pomnilnik in V/I pasovna širina. Če želite prihraniti sistemske vire in izboljšati zmogljivost, lahko sistem za iskanje v podjetju odkrije, ali je prišlo do sprememb, ki jih je treba uveljaviti za indeks in izgradi indeks samo, kadar je treba uveljaviti spremembe. Sistem lahko odkrije naslednje dogodke za spremembo indeksa:

Samo gradnje glavnega indeksa

- Odstranjevanje URI-jev po vzorcu, s čimer se dokumenti odstranijo iz indeksa.
- Odstranjevanje pajka iz sistema, s čimer se dokumenti odstranijo iz indeksa.

Gradnje glavnega in delta indeksa

- Spreminjanje definicij obsega pravil za razvrščanje rezultatov iskanja.
- Dodajanje ali odstranjevanje dokumentov neposredno prek nastavitev pajka ali API-ja.
- Spreminjanje faktorjev povišanja z namenom vplivanja na statično rangiranje dokumentov v indeksu.
- Spreminjanje pravil za uporabo univerzalnih znakov v poizvedbah.

Za vse zahteve za načrtovane gradnje indeksa velja odkrivanje sprememb. Gradnja indeksa se izvede samo, če obstajajo spremembe, ki jih je treba uveljaviti v indeksu. Če sistem ne odkrije sprememb, zavrže zahtevo za izgradnjo indeksa.

Če želite vsiliti izgradnjo indeksa, ne glede na to, ali je prišlo do sprememb, jo lahko zaženete iz konzole za skrbništvo nad iskanjem v podjetju. Z ukazom **startIndexBuild** lahko ročno zaženete izgradnjo indeksa in podate, ali naj sistem pred obdelavo zahteve preveri morebitne spremembe.

Postopek

Če želite zagnati gradnjo indeksa iz ukazne vrstice in zahtevati, naj se gradnja nadaljuje samo, če sistem odkrije spremembe, naredite naslednje:

1. Prijavite se kot skrbnik iskanja v podjetju. V konfiguraciji z več strežniki se prijavite v indeksni strežnik.
2. Vnesite naslednji ukaz:

```
esadmin controller startIndexBuild -options
```

Možnosti:

-cid *collection_ID*

ID zbirke, za katero želite zagnati gradnjo indeksa.

-buildType *build_type*

Podaja tip indeksa, ki bo zgrajen. Veljavni vrednosti sta *main* in *delta*.

-detectChanges

Neobvezno. Povzroči, da se za zahteve po gradnji indeksa izvedejo preizkusi za odkrivanje sprememb. Gradnja indeksa se nadaljuje samo, če sistem odkrije spremembe, ki jih je treba uveljaviti za indeks.

V naslednjem primeru se zahteva za zagon gradnje glavnega indeksa za zbirko *col_1* izvede samo, če sistem odkrije spremembe, ki jih je treba uveljaviti za indeks:

```
esadmin controller startIndexBuild -cid col_1 -buildType main -detectChanges
```

Zaustavljanje gradnje indeksov

Gradnjo glavnih in delta indeksov lahko zaustavite z ukazom **stopIndex** namesto s skrbniško konzolo iskanja v podjetju.

Postopek

Če želite zaustaviti gradnjo indeksa z ukazom namesto s skrbniško konzolo:

1. Prijavite se kot skrbnik iskanja v podjetju. Pri namestitvi na več strežnikov se prijavite na indeksni strežnik.
2. Vnesite naslednji ukaz:

```
esadmin stopIndex -cid collection_id -buildType type kjer je:
```

-cid *collection_id*

ID zbirke, ki je lastnica indeksa,

type

pa je vrsta gradnje indeksa, ki jo želite zaustaviti. Dovoljeni vrednosti sta *main* (glavni) in *delta*.

Primer

```
esadmin stopIndex -cid col1 -buildType delta
```

Možnosti, ki vplivajo na iskalni pogled indeksa

Ko so dokumenti indeksirani, lahko podate možnosti, ki vplivajo na to, kako uporabniki lahko iščejo dokumente in jih prikazujejo v rezultatih iskanja.

Če želite podati možnosti, ki vplivajo na uporabnikov prikaz indeksa, lahko izvedete naslednja opravila:

- Konfigurirajte podporo za univerzalne znake v poizvedbenih izrazih. Podporo za univerzalne znake lahko vgradite v indeks ali pa podate možnosti za razširitev poizvedbenih izrazov med obdelavo poizvedbe.
- Konfigurirate območja, da omejite obseg dokumentov, v katerih uporabniki lahko iščejo. Ko uporabniki iščejo po zbirki, iščejo samo po dokumentih, ki spadajo v območje, ne pa po celotnem indeksu.
- V rezultatih iskanja skrite dokumente iz istega vira. Dokumente, ki ustrezajo URI-ju ali vzorcu URI-jev, lahko v indeksu združite v skupino in v rezultatih iskanja prikažete samo najvišje najdene dokumente (uporabniki lahko podajo možnosti za ogled skrčenih najdenih dokumentov).
- Odstranite URI-je iz indeksa. Možno je, da boste morali uporabnikom začasno preprečiti iskanje po posameznih dokumentih v indeksu.

Nekateri tipi pajkov in zbirke, za katere zaščita ni omogočena, uporabljajo odkrivanje dokumentov, s čimer preprečijo hkraten prikaz več enakih ali skoraj enakih dokumentov v rezultatih iskanja.

Možnosti indeksiranja za iskanje dokumentov

Pri konfiguriranju možnosti za iskanje pajkanih podatkov ali pri preslikavi elementov metapodatkov XML in HTML v iskalna polja morate podati način iskanja po dokumentih in prikaza le-teh v rezultatih iskanja.

Možnosti iskanja, ki jih podate, bodo skupaj z dokumenti shranjene v indeks. Tako lahko omejite, za kaj uporabniki izdajajo poizvedbe in kaj vidijo v rezultatih iskanja.

Možnosti pajka:

Pri konfiguriranju pajka za pajkanje virov podatkov, ki vsebujejo polja, lahko podate naslednje možnosti, ki krmilijo, ali je po polju mogoče iskati, kako je mogoče po poljih iskati in ali jih je mogoče vrniti v rezultatih iskanja:

- Prosto besedilno iskanje
- Iskanje po poljih
- popolno ujemanje
- Poljubna razvrstitev
- Parametrično iskanje
- Rezultati iskanja
- Vsebina dokumenta

Možnosti za preslikavo polj XML in HTML:

Ko konfigurirate razčlenjevalnik in podate, da želite elemente XML in elemente metapodatkov HTML preslikati v iskalna polja v indeksu, morate podati naslednje možnosti:

- iskanje po poljih
- popolno ujemanje
- razvrstitveno
- rezultati iskanja

Če konfigurirate možnosti iskanja za specifične elemente metapodatkov HTML, lahko v nasprotju z ostalimi elementi ali elementi, ki so v nizu elementov metapodatkov Dublin Core, podate, da je po poljih, ki vsebujejo številske vrednosti, mogoče iskati s parametričnimi poizvedbami.

Prosto besedilno iskanje

Indeks iskanja v podjetju je popolni besedilni indeks z vsebino iz različnih virov podatkov. Po vsebini lahko iščete tako, da podate preprosto poizvedbo v izvirnem jeziku. Proces iskanja iščejo po poljih in vsebini dokumentov in tako poiščejo dokumente, ki so pomembni za poizvedbo.

Če želite omogočiti iskanje po poljih s prosto besedilno poizvedbo, pri konfiguriranju pajka izberite potrditveno polje **Prosto besedilno iskanje**. Če želite iskati po poljih nazivov, ključnih besed in opisov, izberite to potrditveno polje in potrditveno polje **Iskanje po poljih**.

Zgled 1:

Prosto besedilno iskanje je lahko preprosto kot naslednja poizvedba:

veriga kolesa

Če želite označiti, katere besede se morajo ali se ne smejo prikazati v dokumentu, lahko vključite posebne označbe. Če na primer pred besedo vnesete znak plus (+), boste s tem podali, da mora dokument vsebovati to besedo, da ga bo sistem označil kot ujemanje. Če pred besedo vnesete znak minus (-), boste dokumente, ki vsebujejo to besedo, izključili iz rezultatov iskanja. Če dve ali več besed ogradite v navednice ("), boste s tem iskali besedno zvezo.

Zgled 2:

V naslednji prosto besedilni poizvedbi bo do ujemanja prišlo le v primeru, če dokument vsebuje točno besedno zvezo znanstvena fantastika in ne vsebuje besede robot:

+ "znanstvena fantastika" -robot

Iskanje po poljih

Z iskanjem po poljih lahko objekt poizvedbe omejite na specifična podatkovna polja in polja metapodatkov v dokumentu. Podate lahko na primer, da morajo v nazivu dokumenta obstajati določene besede.

Če želite omogočiti iskanje po imenih polj, pri konfiguriranju pajka ali pri konfiguriranju možnosti za preslikavo polj za elemente XML in HTML izberite potrditveno polje **Iskanje po poljih**. Če želite iskati po poljih Naziv, Ključne besede in Opis, izberite to potrditveno polje in potrditveno polje **Prosto besedilno iskanje**.

Zgled:

Če želite v sistemu iskanja v podjetju podati iskanje po poljih, v poizvedbo vključite ime polja in besedo ali besedno zvezo, ki morata obstajati v tem polju.

Naslednja poizvedba poišče dokumente, ki morajo v polju naziva vsebovati besedo ibm in besedno zvezo iskanje v podjetju:

naziv:ibm naziv:"iskanje v podjetju"

Popolno ujemanje

Pri iskanju s popolnim ujemanjem lahko povečate kakovost rezultatov iskanja, saj lahko podate natančnejše poizvedbe. Pri iskanju s popolnim ujemanjem lahko izdajate poizvedbe za polja in elemente XML ter priključite samo tiste dokumente, v katerih se z izrazom poizvedbe ujema celotna vrednost polja ali elementa XML. Če vrednost polja ali elementa vsebuje manj besedila ali dodatno besedilo, do ujemanja ne bo prišlo.

Nasvet: Sistem pri ugotavljanju, ali je prišlo do popolnega ujemanja, izraze poizvedbe pretvori v male črke, v nizu poizvedbe odstrani odvečne presledke in opravi ujemanje z vzorcem univerzalnih znakov. Vendar pa ne pride do razstavljanja ali iskanja sopomenk, sistem pa tudi ne odstrani izkjučitvenih besed. Popolno ujemanje elementa XML zahteva, da se ime elementa, brez ugnazdenih elementov, in celotna vrednost elementa točno ujemata z izrazi poizvedbe.

Po poljih lahko popolna ujemanja iščete s skladno poizvedbe sistema iskanja v podjetju (SI-API) ali skladno poizvedbe XMLFrag2 (poizvedbe XPath niso podprte). Enačaj (=) pred izrazi poizvedbe označuje, da mora sistem izvesti iskanje s popolnim ujemanjem.

Če želite omogočiti iskanje po poljih s popolnim ujemanjem, pri konfiguriranju pajka ali možnosti za preslikavo polj za elemente XML in HTML izberite potrditveno polje **Popolno ujemanje**. Če pri konfiguriranju možnosti razčlenjevanja za zbirko uporabnikom omogočite iskanje po dokumentih XML z iskanjem po izvirnem XML-ju, lahko popolna ujemanja z iskalnimi izrazi iščete po vseh elementih XML.

Zgled 1:

Polje z imenom `barva` vsebuje vrednost `temno modra`.

- Pri naslednji poizvedbi s popolnim ujemanjem bo prišlo do ujemanja, ker ne vsebuje drugih izrazov:
`barva:"temno modra"`
- Pri naslednji poizvedbi s popolnim ujemanjem ne bo prišlo do ujemanja, ker ta poleg besede `temno modro` vključuje tudi besedo `krilo`:
`barva:"temno modro krilo"`
- Naslednja poizvedba s popolnim ujemanjem se ne ujema, ker polje `barva` vsebuje tudi besedo `temno`:
`barva:=modra`

Zgled 2:

Če ne uporabite popolnega ujemanja, bo naslednja poizvedba XMLFrag2 vrnila dokumente, v katerih so izrazi `<diagnoza>intraduktalni karcinom tipa ogrca</diagnoza>` ali druge izraze, ki se točno ne ujemajo z izrazi poizvedbe:

```
@xmlf2::'<diagnoza>intraduktalni karcinom</diagnoza>'
```

Pri popolnem ujemanju lahko z naslednjo poizvedbo XMLFrag2 zagotovite, da bo sistem vrnil samo tiste dokumente, v katerih se z izrazi poizvedbe ujema celotna vsebina vrednosti elementa XML:

```
@xmlf2::'=<diagnoza>intraduktalni karcinom</diagnoza>'
```

Poljubna razvrstitev

Če vir podatkov vključuje polja, oziroma uporabniki iščejo dokumente XML ali HTML, lahko omogočite razvrščanje rezultatov po vrednostih v določenem polju. V vzorčni iskalni

aplikaciji sistema iskanja v podjetju so navedena imena vseh polj, ki so konfigurirana kot razvrstitvena polja. Uporabniki lahko namesto prikaza rezultatov po relevantnosti ali datumu izberejo abecedno razvrstitev rezultatov (po vrsti niza) tako, da izberejo eno od navedenih polj. Uporabniki lahko prav tako izberejo, ali bodo dokumenti razvrščeni v rastočem ali padajočem vrstnem redu.

Dokumenti, ki ne vsebujejo polja za razvrščanje, bodo v rezultatih iskanja navedeni na koncu. Dokumenti, ki vsebujejo razvrstitveno polje, vendar so bili indeksirani, preden jih je skrbnik konfiguriral kot razvrstitvena polja, bodo prav tako navedeni na koncu rezultatov iskanja.

Če želite uporabnikom omogočiti abecedno razvrščanje rezultatov iskanja po vrednostih v polju, pri konfiguriranju pajka ali konfiguriranju možnosti preslikave polj za elemente XML in HTML izberite potrditveno polje **Poljubna razvrstitev**. Če polje vsebuje numerične vrednosti, izberite potrditveno polje **Parametrično iskanje**, s čimer podate, da je mogoče s pomočjo vrednosti polj numerično razvrstiti rezultate iskanja.

Parametrično iskanje

Parametrično iskanje je tip iskanja po poljih, s katerim lahko izdajate primerjalne ali ocenjevalne poizvedbe za številiska in datumska polja in metapodatke. Poiščete na primer lahko dokumente določene velikosti ali dokumente, napisane na določen datum. Prav tako lahko po dokumentih iščete z atributi, kot so večje, manjše ali enako podani vrednosti.

Za numerično razvrstitev glede na vrednost polja morate polje omogočiti za parametrično iskanje.

Če želite iskati po polju s pomočjo parametrične poizvedbe ali imeti možnost numeričnega razvrščanja rezultatov, pri konfiguriranju pajka ali možnosti preslikave polj za določene elemente metapodatkov HTML izberite potrditveno polje **Parametrično iskanje**.

Zgled 1:

Naslednja poizvedba poišče postavke, ki stanejo točno 50 dolarjev (oziroma katerekoli druge valute, ki je indeksirana za polje *cena*):

```
#cena:=50
```

Zgled 2:

Naslednja poizvedba poišče dokumente, katere velikost datoteke je večja od 1024, vendar manjša ali enaka 2048:

```
#velikostdatoteke::>1024<=2048
```

Rezultati iskanja

Morda boste želeli iskati po nekaterih poljih, vendar jih ne boste želeli prikazati v rezultatih iskanja ali pa boste želeli polje prikazati v rezultatih iskanja, čeprav ga niste iskali. Morda želite na primer poizvedeti po finančnih podatkih, da bi pridobili pomembno poročilo, vendar v rezultatih ne želite prikazati plač zaposlenih, pri katerih so navedena tudi imena zaposlenih.

Če želite omogočiti prikaz polja v rezultatih iskanja, pri konfiguriranju pajka ali možnosti za preslikavo polj za elemente XML in HTML izberite potrditveno polje **Rezultati iskanja**.

Vsebina dokumenta

Pri določenih tipih dokumentov, kot so spletni dokumenti, je kot vsebina obravnavan celotni dokument. Pri drugih tipi dokumentov, kot so dokumenti s polji, pa lahko podate, katera polja, v nasprotju z metapodatki, vsebujejo uporabno vsebino.

Če želite podati, da je polje del vsebine dokumenta, pri konfiguriranju pajka izberite potrditveno polje **Vsebina dokumenta**. Če izberete obe potrditveni polji, **Vsebina dokumenta** in **Prosto besedilno iskanje**, bo sistem pri zaznavanju dvojnikov dokumentov uporabil vrednost polja, ta pa bo nato v rezultatih iskanja postala del področja z dinamičnim povzetkom dokumenta.

S tem povezani pojmi

 Skladnja poizvedbe

Odkrivanje podvojenih dokumentov

Odkrivanje podvojenih dokumentov je tehnika, s katero se prepreči, da bi rezultati iskanja vsebovali več dokumentov z enako ali skoraj enako vsebino.

Kakovost iskanja se lahko zmanjša, če rezultati iskanja vsebujejo več kopij enakega (ali skorajda enakega) dokumenta. Analiza podvojenih dokumentov se izvede samo, če držita naslednja pogoja:

- Zbirka uporablja model rangiranja na osnovi povezav. Ta model velja za pajke, ki pajkajo spletna mesta, na primer spletni pajek ali pajek WebSphere Portal.
- Zaščita zbirke je onemogočena.

Med globalno analizo procesi indeksiranja odkrijejo dvojnike tako, da pregledajo vsebino vsakega dokumenta. Če najdejo dva dokumenta z enako vsebino, ju obravnavajo kot dvojnika.

Če želite, da analiza dvojnikov upošteva tudi metapodatke dokumentov, morate pri konfiguriranju pajkov za zbirko označiti potrditveno polje **Vsebina dokumentov** in podati možnosti za pajkanje metapodatkov. V tem primeru pajek pajka metapodatkovna polja kot vsebino dokumentov in vključi metapodatke v analizo vsebine pri iskanju podvojenih dokumentov. Podobna analiza se izvede, če konfigurirate možnosti za razčlenjevanje dokumentov HTML in XML in izberete potrditveno polje **Vsebina dokumenta**.

Če podate, da polje ali metapodatkovno polje sestavlja vsebino dokumenta, se vsebina teh polj doda v dinamični povzetek dokumenta v rezultatih iskanja, kar lahko vpliva na to, ali se dokument prikaže v rezultatih iskanja. Če je v iskalni aplikaciji omogočeno odkrivanje skoraj enakih dokumentov (lastnost `NearDuplicateDetection` v metodi `setProperty` je nastavljena na `Yes`), se dokumenti s podobnimi naslovi in povzetki ne prikažejo v rezultatih iskanja. Uporabniki lahko kliknejo povezavo in si ogledajo skoraj enake, neprikazane dokumente.

V skupini podvojenih dokumentov je eden od dokumentov glavni dokument, drugi pa so dvojniki. Vsi dokumenti iz skupine dvojnikov imajo enako predpisano predstavitev vsebine. Med indeksiranjem se indeksira vsebina (žetoni) glavnega dokumenta. Za dvojnike se indeksirajo samo žetoni metapodatkov. Če izbrišete glavni dokument iz indeksa, postane naslednji dvojniki po hierarhiji glavni dokument. Ko uporabniki preiskujejo zbirko, se v rezultatih iskanja prikaže samo glavni dokument.

S tem povezani pojmi

“Analiza podvojenih dokumentov in zaščita zbirke” na strani 237

Univerzalni znaki v poizvedbah

Uporabnikom lahko omogočite, da v izraze poizvedb vnesejo univerzalne znake in iščejo besede, ki se ujemajo s podanim vzorcem.

Izraz poizvedbe z univerzalnim znakom je izraz, ki vsebuje zvezdico (*). Če uporabnik predloži poizvedbo, v kateri je med drugimi tudi univerzalni znak, rezultati iskanja vključujejo vse dokumente v indeksu, ki se ujemajo z iskanim izrazom, ter vse dokumente v indeksu, ki se ujemajo z vzorcem, podanim s pomočjo univerzalnega znaka. Končni univerzalni znak v izrazu poizvedbe naj* lahko ustreza besedam najdi, najbolje in najem.

Ko konfigurirate možnosti za univerzalne znake v indeksu, izberete, ali želite uporabnikom omogočiti, da v poizvedbah podajo univerzalne znake, in če, na kakšen način nameravate zagotoviti to podporo:

- Omogočite lahko iskanje v vseh delih dokumenta za besedami, ki se ujemajo z vzorcem, podanim s pomočjo univerzalnega znaka, ali pa omejite ujemanje vzorca na polja.
- Za vsa polja lahko omogočite podporo za poizvedbe, ki vsebujejo univerzalne znake, ali pa omejite ujemanje vzorca na podana polja.
- Univerzalne znake lahko omejite na končni znak v izrazu poizvedbe (končni univerzalni znak) ali pa dovolite, da se univerzalni znak pojavi kjerkoli v izrazu poizvedbe. (Imena polj ne smejo vsebovati univerzalnih znakov).
- Glede na to, na katerih mestih se lahko pojavijo univerzalni znaki, lahko izberete način razširjanja izrazov poizvedbe (izrazi poizvedbe, ki vsebujejo univerzalne znake, se razširijo na vse izraze v indeksu, ki se ujemajo). V indeksu je mogoče shraniti vse možne razširitve izrazov ali procese iskanja, ki med obdelavo poizvedbe razširijo izraze.

Vse spremembe, ki jih opravite v nastavitvah za univerzalne znake, stopijo v veljavo šele ob naslednji gradnji glavnega indeksa.

Razširitev indeksa

Če želite v indeks dodati razširitve izrazov, morate podati, koliko začetnih znakov besede se mora ujemati z vzorcem, podanim s pomočjo univerzalnih znakov, v izrazu poizvedbe, da pride do ujemanja. Samo izrazi poizvedb, ki vsebujejo vsaj toliko znakov (* ni všteta) vrnejo rezultate. Če na primer podate 4, mora izraz poizvedbe vsebovati vsaj štiri znake, da pride do ujemanja.

Če podate 4, se beseda tehnologija ujema z izrazom poizvedbe tehn* in izrazom poizvedbe tehn*, ne ujema pa se z izrazom poizvedbe te*.

Med gradnjo delta ali glavnega indeksa se poleg izvornih izrazov indeksirajo tudi vse možne razširitve posameznih izrazov v dokumentu. Prednost te metode je v tem, da razširitev izrazov med obdelavo poizvedb ne zahteva dodatnega časa. Vendar pa ta metoda poveča indeks, kar pomeni, da morate imeti na razpolago dovolj sistemskih virov za povečan indeks.

Ta metoda pride v poštev predvsem, če je zbirka dokaj majhna ali če sta prostor in čas, potrebna za gradnjo indeksa, manj pomembna kot odzivni čas poizvedbe. S to metodo lahko na primer iščete po katalogu ali registru zaposlenih.

Opisana metoda je na voljo samo, če omogočite podporo za končne univerzalne znake. Če omogočite podporo za univerzalne znake, ki se pojavijo kjerkoli v izrazu poizvedbe, ne boste mogli izbrati možnosti, ki omogoča vključitev razširitev izrazov v indeks.

Razširitev poizvedb

Če želite razširiti poizvedbe in uveljaviti pravila za ujemanje vzorcev, ko uporabniki predložijo poizvedbe z univerzalnimi znaki, morate podati, koliko izvedb izraza poizvedbe sestavlja ujemanje. Če na primer podate 50, je za ujemanje z izrazom poizvedbe možnih do 50 izvedb izraza poizvedbe.

V zgornjem primeru se torej izraz poizvedbe `tehn*` ujema z besedami tehnično, tehnika, tehnologija in do 50 drugimi besedami, ki se začnejo na `tehn`.

Čeprav ima razširitev izrazov zgolj majhen vpliv na velikost indeksa, lahko precej poslabša učinkovitost poizvedb. Iskalni procesi morajo ponoviti vse možne razširitve izraza poizvedbe z univerzalnimi znaki, vse do omejitve, ki jo podate v nastavitvah za univerzalne znake.

Ta metoda je najprimernejša, kadar je zbirka dokaj velika in je treba kar se da zmanjšati velikost in skrajšati čas, potreben za gradnjo indeksa. Opisano metodo lahko izberete v primeru repozitorijev elektronske pošte, kjer se mora indeks prilagajati hitro spreminjajočim se dokumentom, odzivni čas pa ni tako pomemben.

Ta metoda je na voljo, ne glede na to, ali omogočite podporo za končne univerzalne znake ali podporo za univerzalne znake, ki se pojavijo kjerkoli v izrazu poizvedbe.

Podpora za univerzalne znake v poizvedbah

Niz razširitev izraza poizvedbe z univerzalnimi znaki vsebuje vse izraze v indeksu, ki jih je mogoče pridobiti tako, da univerzalni znak zamenjate s poljubnim zaporedjem znakov. Niz je določen takole:

- Če zbirka podpira univerzalne znake, ki se lahko pojavijo kjerkoli v izrazu poizvedbe, sistem vse izraze, ki vsebujejo zvezdico, smatra za izraz z univerzalnim znakom.
- Niz v skrajnem primeru vsebuje največje število razširitev, ki ga konfigurira skrbnik iskanja v podjetju. Če indeks vsebuje več razširitev, jih sistem prezre. (Iskalni rezultati nakazujejo, ali je sistem prezrl razširitve univerzalnih znakov.)
- Če je podpora za univerzalne znake omejena zgolj na niz polj, niz vsebuje samo izraze, ki se pojavijo v enem od podanih polj. Izraz se mora pojaviti samo v enem od polj v vsaj enem dokumentu, ki je v indeksu.
- Če je izraz poizvedbe izraz iz polja, mora biti univerzalni znak podan za označevalcem polja (na primer `fieldname:*sphere`). Ime polja ne sme vsebovati dvopičja (:).
- Če je podpora za univerzalne znake omejena zgolj na niz polj, mora biti ime polja v izrazu poizvedbe z univerzalnim znakom eno od polj, ki je podano na skrbniški konzoli za iskanje v podjetju. V nasprotnem primeru sistem ne najde razširitev za podani izraz.
- Univerzalni znaki so podprti samo v izrazih iz čistega besedila, ne pa tudi v imenih elementov XML, imenih atributov ali vrednostih atributov. Izraz, ki je sestavljen izključno iz univerzalnih znakov, ni podprt.

Vpliv univerzalnih znakov na indeks

Podpora za univerzalne znake, ki temelji na razširitvi indeksa, poveča indeks in podaljša čas, potreben za njegovo izgradnjo.

Med razširitvijo indeksa se indeksira vsaka predpona izraza in tudi izraz sam. Za izraz podpora sistem na primer indeksira naslednje izraze:

p po pod podp podpo podpor podpora

Število v indeksu shranjenih izrazov raste sorazmerno s faktorjem povprečne dolžine besede. Stiskanje indeksa sicer zmanjša velikost indeksa, vendar ne bistveno. Čas, ki je potreben za gradnjo indeksa, se povečuje sorazmerno s povprečno dolžino besede.

Indeks angleških dokumentov raste s faktorjem približno štiri, saj je povprečna dolžina angleške besede pet do šest črk. Indeks se z izdelovanjem žetonov n-gram poveča približno za dvakrat, saj vsak n-gram vsebuje dva znaka.

Razširitev indeksa z namenom podpore za univerzalne znake je priporočena v naslednjih primerih:

- Zbirka je dovolj majhna, tako da prostor in čas, ki ju porablja razširjeni indeks, ne povzročata težave z zmogljivostjo.
- Vse možne razširitve univerzalnih znakov so vključene v rezultate iskanja in izpolnjujejo zahteve uporabnika (ali podjetja).

Vpliv na indeks lahko omilite tako, da podate minimalno dolžino predpone za razširitev univerzalnih znakov. Če je minimalna dolžina predpone na primer nastavljena na tri, se predponi **p** in **po** v besedi **podpora** ne indeksirajo, angleški indeks pa se namesto za faktor štiri poveča za faktor tri.

Če za omogočanje podpore za univerzalne znake uporabite metodo razširitve poizvedbe, se predpone ne zapišejo v indeks. Izrazi se razširijo ob predložitvi poizvedbe, indeks se poveča zgolj za majhno podatkovno strukturo, ki je potrebna za podporo razširitvi. Ponavadi je indeks z razširitvijo poizvedbe za med 10% in 20% večji od indeksa brez podpore za univerzalne znake, za gradnjo indeksa pa je potrebnega manj kot 10% več časa. Konfiguracija največjega števila razširitev ne vpliva na velikost indeksa ali časa, potrebnega za njegovo izgradnjo.

Konfiguriranje možnosti za univerzalne znake v poizvedbah

Ko konfigurirate možnosti indeksiranja za zbirko za iskanje v podjetju, lahko podate, ali želite uporabnikom omogočiti, da v izraze poizvedb vključijo univerzalne znake.

Preden začnete

Če želite konfigurirati možnosti za univerzalne znake, morate imeti vlogo skrbnika za iskanje v sistemu ali biti skrbnik zbirke, ki ji pripada indeks.

O tej nalogi

Potem ko podate možnosti za univerzalne znake, vaše spremembe stopijo v veljavo šele ob naslednji gradnji glavnega indeksa.

Postopek

Če želite konfigurirati podporo za univerzalne znake v poizvedbah, naredite naslednje:

1. Uredite zbirko, izberite stran Indeksiranje in kliknite **Konfigurirajte možnosti za univerzalne znake**.
2. Na strani Možnosti za univerzalne znake izberite potrditveno polje **Podpora univerzalnemu znakom v poizvedbah**.
3. Neobvezno: Podate lahko, ali želite nuditi podporo za univerzalne znake v poizvedbah, ki iščejo po prostem besedilu. Poizvedba po prostem besedilu **teh***, ki ne išče po poimenovanem polju, vrne razširjene rezultate (kot sta tehnologija ali tehnika) samo, če je izbrano to potrditveno polje.
4. Podajte, katera polja podpirajo univerzalne znake:

- Če želite podati, da univerzalnih znakov ni mogoče obdelati v poizvedbah, ki iščejo po poljih, izberite možnost **Brez polj**.
 - Če želite vsem poljem v dokumentu omogočiti podporo za poizvedbe z univerzalnimi znaki, izberite možnost **Vsa polja**.
 - Če želite omejiti podporo za univerzalne znake na zgolj nekaj polj, izberite možnost **Določena polja** in vnesite imena teh polj. Razširjeni rezultati bodo vrnjeni samo za polja, ki jih podate. Poizvedba `author:john*` na primer vrne razširjene rezultate samo, če podate, da polje `author` podpira univerzalne znake.
5. Podajte, ali se mora univerzalni znak v izrazu poizvedbe pojaviti v končnem položaju (končni univerzalni znak) ali pa ni omejen in se lahko pojavi kjerkoli v izrazu poizvedbe. Ko izberete položaj in vrsto univerzalnega znaka, morate podati tudi, kako želite omogočiti podporo za univerzalne znake. Do podrobnosti o tem pridete tako, da na skrbniški konzoli kliknete **Pomoč**.

Območja

Če želite uporabnikom omejiti pogled zbirke, konfigurirajte območje.

Območje je skupina povezanih URI-jev v indeksu. Če konfigurirate območje, s tem omejite dokumente, ki jih uporabniki lahko vidijo v zbirki. Ko uporabniki iščejo po zbirki, iščejo samo po dokumentih v območju, ne pa po celotnem indeksu. Če želite uporabiti to funkcijo, morajo imeti vaše iskalne aplikacije podporo za območja iskanja.

Ko izdelate območje, podate obseg URI-jev v indeksu, po katerih uporabniki lahko iščejo. Z omejitvijo območja dokumentov, po katerih uporabniki lahko iščejo, boste zagotovili, da so dokumenti v rezultatih iskanja specifični za informacije, ki jih uporabniki iščejo.

Tako lahko na primer izdelate eno območje, ki vključuje URI-je za oddelek tehnične podpore in drugo območje, ki vključuje URI-je za oddelek človeških virov. Če vaša iskalna aplikacija podpira območja, lahko uporabniki v oddelku tehnične podpore prikličejo dokumente iz območja Tehnična podpora, uporabniki iz oddelka človeških virov pa dokumente iz območja Človeški viri.

Izdelate jih lahko poljubno število, vendar pa lahko preveliko število območij vpliva na zmogljivost. Območja konfigurirajte tako, da bo večina zahtev za iskanje omejena na eno ali dve območji. Ker lahko območja vsebujejo celotne URI-je ali vzorce URI-jev, lahko isti dokument pripada več območjem.

Pri konfiguriranju območij boste morda morali glavni indeks zgraditi dvakrat, da bodo spremembe stopile v veljavo. Če območja konfigurirate pred gradnjo prvega glavnega indeksa zbirke, bodo uporabniki sicer lahko iskali po zbirki, vendar v rezultatih iskanja ne bodo mogli videti podatkov o območju. Če želite zagotoviti, da bodo rezultati iskanja odražali obseg URI-jev v območju, znova zgradite glavni indeks.

Če območja konfigurirate po gradnji glavnega indeksa, spremembe stopijo v veljavo ob naslednji gradnji glavnega indeksa.

Konfiguriranje območij

Ko konfigurirate območje za zbirko za iskanje v podjetju, podate URI-je ali vzorce URI-jev za vrsto dokumentov v indeksu, po katerih lahko iščejo uporabniki.

Preden začnete

Če želite konfigurirati območja, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke, ki ji območje pripada.

O tej nalogi

Če iskalne aplikacije omogočijo podporo za območja, lahko uporabniki iščejo samo po dokumentih, ki se ujemajo z URI-ji, ki definirajo meje območja pri iskanju po zbirki.

Pri konfiguriranju območij boste morda morali glavni indeks zgraditi dvakrat, da bodo spremembe stopile v veljavo. Če območja konfigurirate pred prvo gradnjo indeksa, bodo lahko uporabniki sicer iskali po zbirki, ne bodo pa mogli videti podatkov območja v rezultatih iskanja. Če želite zagotoviti, da bodo rezultati iskanja odražali obseg URI-jev v območju, znova zgradite glavni indeks.

Če območja konfigurirate po gradnji glavnega indeksa, spremembe stopijo v veljavo ob naslednji gradnji glavnega indeksa.

Postopek

Če želite konfigurirati območje, naredite naslednje:

1. Uredite zbirko, izberite stran Indeksiranje in kliknite **Konfiguriraj območja**.
2. Na strani Območja kliknite **Izdelaj območje**.
3. Podajte ime območja in URI-je ter vzorce, ki definirajo meje območja. Podate lahko tudi URI-je in vzorce URI-jev, ki jih želite izvzeti iz območja.
4. Kliknite **Potrdi**.

Vaše novo območje je navedeno na strani Območja, skupaj z drugimi območji, ki pripadajo tej zbirki.

S tem povezane povezave

“Formati URI v indeksu iskanja v podjetju” na strani 107

Skrčeni URI-ji

Iskanje v podjetju lahko organizira rezultate iskanja tako, da lahko dokumente iz virov z isto predpono URI-jev v rezultatih iskanja skrčite.

Ko so rezultati skrčeni, je dokument z največjo pomembnostjo običajno prikazan pomaknjen v levo. Pod najvišjim rezultatom so združeni in zamaknjeni rezultati z manjšo pomembnostjo.

Če želite skrčiti dokumente z različnimi predponami URI-jev kot eno skupino, lahko predpone URI-jev povežete z imenom skupine, ki jo izdelate. Če imate na primer tri strežnike za upravljanje finančnih podatkov, lahko v rezultatih iskanja dokumente iz vseh treh strežnikov združite v skupino in skrčite nižje uvrščene rezultate pod najvišjim najdenih dokumentom.

Iskalne aplikacije lahko uporabijo za skrčenje dokumentov v rezultatih iskanja predpono URI-ja ali ime skupine. V vzorčni iskalni aplikaciji za iskanje v podjetju sta prikazana zgornja dva dokumenta rezultatov iskanja. Če iskanje vrne več kot dva najdena dokumenta z enako predpono URI-ja, (ali dokumente, ki spadajo v isto skupino URI-jev), lahko izberete možnost za prikaz skrčenih rezultatov.

Uporabniki lahko uporabijo skladno poizvedb iskanja v podjetju (*samegroupas:URI prefix*) za iskanje po vseh dokumentih, ki so v isti skupini kot predpona URI-ja, ki je podana v poizvedbi.

Nasveti za organizacijo predpon URI-jev in imen skupin

Če za konfiguriranje pravil za skrčevanje rezultatov iskanja uporabljate skrbniško konzolo, podate predpone URI-jev dokumentov, ki jih želite skrčiti in po želji povezati predpone URI-jev z imenom skupine.

Vrstni red predpon URI-jev, ki jih konfigurirate, je pomemben. Indeksni strežnik pri izračunavanju vrednosti vsakega URI-ja v zbirki uporabi vrstni red predpon URI-jev. Za vsak URI:

1. Indeksni strežnik zaporedoma pregleduje predpone URI-jev v pravilih za skrčevanje rezultatov iskanja.
2. Ko najde prvo predpono URI-ja, ki se ujema s predpono dokumenta v indeksu, poveže ime skupine (ali predpono URI-ja, če pravilo ne podaja ime skupine) kot dodaten iskalni izraz za dokument.

Če spletnega dokumenta ni mogoče povezati s predpono URI-ja, indeksni strežnik kot predpono URI-ja uporabi ime gostitelja URL-ja. Če dokumenta protokola NNTP ni mogoče povezati s predpono URI-ja, indeksni strežnik kot predpono URI-ja uporabi prvi ID sporočila v vrednosti referenčnega oglavja.

Ko na seznam skrčenih predpon URI-ja dodate predpono URI-ja, ji morate določiti mesto v vrstnem redu, po katerem želite, da jo pregleda indeksni strežnik, in jo po želji povezati z dodatnim iskalnim izrazom z dokumenti v indeksu:

- Če dodate predpono URI-ja in jo ne povežete z imenom skupine, lahko izberete posamezno predpono URI-ja in jo na seznamu premaknete navzgor ali navzdol.
- Če dodate predpono URI-ja in jo povežete z imenom skupine, vedno ko na seznamu premaknete predpono URI-ja navzgor ali navzdol, premaknete celotno skupino predpon URI-jev, ki pripadajo isti skupini. Vrstni red predpon URI-jev v skupini ni pomemben. Vedno, ko izberete posamezno predpono URI-ja, s tem samodejno izberete celotno skupino.

Skrčenje URI-jev v rezultatih iskanja

Podate lahko možnosti za grupiranje in skrčenje dokumentov, dobljenih iz virov z istim predznakom URI. Izdelate lahko tudi ime skupine, ki omogoča skupno skrčevanje dobljenih dokumentov z različnimi predznaki URI.

Preden začnete

Če želite podati možnosti za skrčenje rezultatov iskanja, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

O tej nalogi



Spremembe, ki jih opravite za skrčevanje rezultatov iskanja, stopijo v veljavo šele ob naslednji gradnji glavnega indeksa.

Postopek

Če želite podati možnosti za skrčenje rezultatov iskanja, naredite naslednje:

1. Uredite zbirko, izberite stran Indeksiranje in kliknite **Skrči rezultate iskanja**.
2. Na strani Skrčenje rezultatov iskanja kliknite **Dodaj predpono URI**.
3. Na strani Dodajanje predpone URI za skrčenje rezultatov vnesite predpono URI za dokumente, ki jih želite v rezultatih iskanja prikazati skrčeno. Primer:

```
http://finance/ROI/  
http://server1.com/finance/  
db2://LOCALDB/SCHEMA1.TABLE1/  
exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/
```

4. Vnesete lahko opisno ime skupine, ki ga želite povezati s to predpono URI-ja. Če želite skrčiti dokumente, dobljene iz različnih virov, kot eno samo skupino, pri dodajanju predpon URI vnesite isto ime skupine.
5. Kliknite **Potrdi**.
6. Na strani Skrčenje rezultatov iskanja postavite novo pravilo v položaj, kjer ga naj pregleda indeksni strežnik:
 - Če ste dodali predpono URI in je niste povezali z imenom skupine, se nova predpona URI pojavi na dnu seznama. S smernimi tipkami jo premaknete v zeleni položaj.
 - Če ste novo predpono URI povezali z imenom skupine, se nova predpona URI pojavi na dnu niza predpon URI, ki pripadajo isti skupini. S smernimi tipkami lahko premaknete celotno skupino predpon URI v zeleni položaj.
7. Če želite spremeniti predpono URI ali ime skupine, izberite predpono URI in kliknite  **Uredi**.
8. Če želite odstraniti predpono URI s seznama, jo izberite in kliknite  **Odstrani**.

Odstranitev URI-jev iz indeksa

Če želite uporabnikom preprečiti iskanje po dokumentih v zbirki, lahko iz indeksa odstranite URI-je teh dokumentov.

Preden začnete

Če želite odstraniti URI-je iz indeksa, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

O tej nalogi

Ko podate popolnoma kvalificiran URI, ga uporabniki ne vidijo več med rezultati iskanja. Če pa uporabnik predloži isto poizvedbo in se dokumenti, ki jih prikliče ta poizvedba, nahajajo v iskalnem predpomnilniku, se predpomnjena stran za URI, ki ste ga odstranili, še naprej pojavlja med rezultati iskanja. Iskalni predpomnilnik se osveži in URI odstrani iz indeksa šele ob naslednji gradnji glavnega ali delta indeksa.

Če podate vzorec URI za odstranitev več URI-jev, uporabniki med rezultati iskanja vidijo URI-je, ki ustrezajo podanemu vzorcu, vse do naslednje gradnje glavnega indeksa.

Ko odstranite URI iz indeksa, ga ne odstranite iz prostora za pajkanje. Ko pajek naslednjič pajka dokument, se URI vdela v indeks in je odtlej znova na voljo za iskanje. Če želite odstraniti URI iz prostora za pajkanje, morate posodobiti pravila za pajkanje in iz njih izvzeti dokument, nato pa zaustaviti in znova zagnati pajka.

Postopek

Če želite odstraniti URI-je iz določenih dokumentov v indeksu, naredite naslednje:

1. Uredite zbirko, izberite stran Indeksiranje in kliknite **Odstrani URI-je iz indeksa**.
2. Na strani Odstranite URI-jev iz indeksa vnesite URI-je (ali vzorce URI), ki jih želite odstraniti iz indeksa.

Primer:

```
http://domain.org/hr/*  
db2://knowledgeManagement/ROI*  
cm://enterprise/finance*
```


S tem povezane povezave

“Formati URI v indeksu iskanja v podjetju” na strani 107

Skrbnništvo nad iskalnim strežnikom

Med možnostmi, ki jih lahko podate za iskalne strežnike, sodijo uporaba predpomnilniškega prostora za dobljene rezultate iskanja, krmiljenje največje dolžine prikazanih povzetkov dokumentov v rezultatih iskanja, povezovanje slovarjev po meri, ki izboljšajo kakovost iskanja, in vračanje vnaprej definiranih URI-jev v rezultatih iskanja, kadar se v poizvedbi pojavijo določeni izrazi.

Ko uporabnik predloži poizvedbo, iskalni strežniki za hitro iskanje ustreznih dokumentov uporabijo indeks. Iskalni strežniki uporabijo za pridobivanje metapodatkov za ustrezne dokumente podatkovno shrambo iskanja v podjetju, ki vsebuje razčlenjene in z žetonizirane podatke. Metapodatki lahko vključujejo URI dokumenta, naziv, opis, datum, tip podatkov in drugo, vendar niso omejeni na njih.

Pri konfiguriranju iskalnih strežnikov za zbirko podate možnosti, ki vplivajo na način obdelave poizvedb, vključno z možnostmi, ki lahko vplivajo na zmogljivost poizvedbe:

Konfiguriranje iskalnega predpomnilnika

Da bi optimizirali učinkovitost poizvedb, lahko podate možnost, s katero rezultate iskanja (odgovori na poizvedbe) shranite v predpomnilnik, nato pa konfigurirate količino prostora, ki ga želite dodeliti predpomnjenim rezultatom iskanja.

Konfiguriranje največje dolžine za prikaz povzetkov dokumentov

Večina nastalih dokumentov prikaže povzetek vsebine dokumenta, ki uporabnikom pomaga pri odločitvi, ali je dokument tisti, ki ga želijo pridobiti. Podate lahko, koliko prostora bo uporabljenega v rezultatih iskanja za prikaz teh informacij povzetka.

Podajanje drugega privzetega jezika

Privzet jezik za preiskovanje dokumentov v zbirki je podan pri izdelavi zbirke, vendar lahko po potrebi podate drug jezik.

Povezovanje slovarjev po meri

Če so razvijalci aplikacij izdelali slovarje po meri za sopomenke, zaustavitvene besede ali poviševalne besede, lahko podate slovarje, ki bodo uporabljeni, ko bodo uporabniki preiskovali zbirko.

Konfiguriranje hitrih povezav

Če želite, lahko vnaprej določite URI-je, ki bodo vrnjeni za določene ključne besede in fraze. Če uporabniki v poizvedbi podajo ključne besede ali fraze, je vnaprej definiran URI vrnjen z rezultati iskanja. URI-ji hitrih povezav so vrnjeni poleg URI-jev, ki jih vrnejo iskalni strežniki s preiskovanjem indeksa.

S tem povezani pojmi

“Rangiranje dokumentov” na strani 189

“Prilagojeni slovarji poviševalnih besed” na strani 192

Iskalni predpomnilniki

Če je obremenitev iskalnih strežnikov relativno visoka, lahko izboljšate zmogljivost tako, da rezultate iskanja shranite v predpomnilnik.

Ko iskalni strežniki obdelujejo zahteve za iskanje, najprej preverijo, ali rezultati za isto poizvedbo že obstajajo v predpomnilniku. Če najdejo ustrezen odziv na poizvedbo, lahko uporabniku hitro vrnejo rezultate iskanja. Če iskalni strežniki ne najdejo ustreznega odziva na poizvedbo, preiščejo indeks.

Ko se iskalni predpomnilnik zapolni, sistem najprej zbrši najstarejše rezultate iskanja in rezultate redkih poizvedb, s čimer naredi prostor za nove rezultate iskanja.

Iz skrbniške konzole iskanja v podjetju lahko omogočite predpomnjenje iskanj in podate zmogljivost predpomnilnika (število odzivov na poizvedbe, ki so lahko istočasno zapisani v predpomnilniku).

Če spremenite nastavitve iskalnega predpomnilnika, morate za uveljavitev sprememb znova zagnati iskalne strežnike.

Konfiguriranje iskalnega predpomnilnika

Iskalni predpomnilnik za zbirko lahko omogočite ali onemogočite. Podate lahko tudi možnosti za nadzor velikosti iskalnega predpomnilnika.

Preden začnete

Iskalni predpomnilnik za zbirko lahko konfigurirate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

Postopek

Konfiguriranje iskalnega predpomnilnika:

1. Uredite zbirko, izberite stran Iskanje in kliknite **Konfiguriraj možnosti iskalnega strežnika**.
2. Na strani Možnosti iskalnega strežnika izberite potrditveno polje **Uporaba iskalnega predpomnilnika**.
3. V polje **Največje število vnosov v predpomnilnik** vnesite največje število odgovorov na poizvedbe, ki jih je mogoče zapisati v predpomnilnik.
4. Kliknite **Potrdi**.
5. Če želite uveljaviti spremembe, zaženite nadzornika iskalnih strežnikov in znova zaženite postopke strežnikov.

Prilagojeni slovarji sopomenk

Kakovost rezultatov iskanja lahko izboljšate tako, da uporabnikom omogočite, da med iskanjem po zbirki iščejo tudi sopomenke poizvedbenih izrazov.

Če izdelate slovar sopomenk in ga dodate iskalnemu sistemu podjetja ter ga povežete z zbirko, lahko uporabniki med iskanjem po zbirkah iščejo dokumente, ki vsebujejo sopomenke poizvedbenih izrazov. Z razširitvijo poizvedb na ta način je bolj verjetno, da bodo uporabniki našli vse dokumente, ki jih zanimajo, ne samo dokumente, ki se ujemajo s točno njihovimi poizvedbenimi izrazi. Ker med izdelavo slovarja sopomenk definirate, katere besede so sopomenke, zagotovite, da uporabniki najdejo koristne dokumente, ne da bi morali podati vse različice poizvedbenega izraza.

Tako lahko na primer vaša organizacija za oddelke, opremo ali drugo uporablja kratice in okrajšave ali pa morda dokumenti v vaših zbirkah vsebujejo besedišče, specifično za vašo dejavnost. Z izdelavo slovarja sopomenk zagotovite, da poizvedbe, ki vsebujejo akronime (na primer CPE) vrnejo dokumente z razširitvami teh akronimov (CPE-je, centralna procesna enota, procesne enote in tako naprej).

Jezik poizvedbe iskanja v podjetju podpira sopomenke tako, da uporabnikom omogoča, da poizvedbenemu izrazu predpripnejo tildni operator. Tako lahko na primer poizvedba `~WAS`

vrne dokumente o strežniku WebSphere Application Server. Razvijalci aplikacij lahko omogočijo podporo za sopomenke tudi v lastnostih poizvedb, ki ne potrebujejo posebne skladnje.

Slovarji sopomenk vsebujejo različice besed in imajo naslednje lastnosti:

- Besede niso specifične za jezik, ampak se jih lahko uporabi v različnih jezikih. Za vsako zbirko je na voljo samo en slovar sopomenk.
- Besede niso sklanjane ali spregane. Vse spregatve in sklanjatve besed je potrebno dodati na seznam sopomenk. Tako so med izpeljanimi oblikami edninske in množinske besede (na primer CPE in CPE-e).

Večina izrazov, ki jih dodate v slovar sopomenk so točne pomenske ustreznice. To pomeni, da če je izraz A sopomenka izraza B, potem je tudi B sopomenka A-ja. Vsakič, ko se v poizvedbi uporabi A, je mogoče uporabiti tudi B in obratno.

Vendar lahko dodate tudi izraze, ki ustrezajo različnim uporabam določenega izraza, vključno s splošnimi ali bolj specifičnimi različicami izraza. Tako imate lahko na primer eno skupino sopomenk, ki vključuje besedi gradnja in hiša ter drugo skupino, ki vključuje besede banka, doma in hranilnica.

Manj tesno povezani izrazi bodo dali več rezultatov iskanja, vendar se lahko zgodi, da nekateri izmed teh rezultatov ne bodo uporabni za poizvedbo. API za iskanje in indeks nudi metode, ki uporabnikom omogočajo, da izberejo ustrezne sopomenke, ko predložijo iskalno zahtevo, in metode, ki uporabnikom kažejo iskalne izraze in njihove sopomenke.

Izdelava slovarja sopomenk poteka tako, da najprej strokovnjak s področja snovi zbirke izdela seznam sopomenk v obliki XML ali pa v sodelovanju z razvijalcem aplikacije izdela datoteko XML. Z orodjem za iskanje v podjetju **essyndictbuilder** se datoteka XML pretvori v binarno datoteko (.dic).

Skrbnik iskanja v podjetju binarno datoteko naloži na sistem in ji dodeli prikazno ime. Skrbniki zbirke lahko izberejo, kateri slovar sopomenk naj se uporabi za iskanje po dokumentih v zbirki, ko konfigurirajo možnosti iskalnega strežnika za zbirko.

Omejitev: Potem ko slovar sopomenk po meri dodate sistemu, ga ne morete več urejati. Sopomenke, ki so na voljo določeni zbirki, popravite tako, da:

1. Posodobite izvorno datoteko XML.
2. Pretvorite vir XML v novo slovarsko datoteko.
3. Odstranite stari slovar sopomenk iz zbirk, ki ga uporabljajo.
4. Izbrišite stari slovar sopomenk iz sistema.
5. Dodajte novi slovar sopomenk v sistem.
6. Povežite novi slovar sopomenk z zbirkami, ki ga bodo uporabljale.

Izdelate lahko skript, ki vključuje te korake, nato pa s pomočjo tega skripta znova razmestite slovar v sistem iskanja v podjetju.

S tem povezani pojmi

 [Synonym support in search applications](#)

S tem povezana opravila

 [Creating an XML file for synonyms](#)

 [Creating a synonym dictionary](#)

Dodajanje slovarjev sopomenk v sistem

Če izdelate slovarje sopomenk po meri za iskanje dokumentov v zbirki, morate slovarje povezati s sistemom iskanja v podjetju. Pozneje lahko izberete, kateri slovar sopomenk boste uporabili za preiskovanje zbirke.

Praden začnete


Če želite dodati slovarje sopomenk po meri in jih uporabljati v poizvedbah za iskanje v sistemu, morate imeti vlogo skrbnika za iskanje v podjetju.

Omejitve

Slovar sopomenk je lahko velik največ 8 MB.

Postopek

Če želite povezati sopomenke s sistemom iskanja v podjetju, naredite naslednje:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
3. Na strani Iskanje kliknite **Konfiguriraj slovarje sopomenk**.
4. Na strani Konfiguriranje slovarjev sopomenk kliknite **Dodaj slovar sopomenk**.
5. Na strani Dodajanje slovarja sopomenk vnesite unikatno prikazno ime za slovar sopomenk in, če želite, podajte še opis.
6. Podajte nahajališče datoteke .dic. Če se datoteka nahaja v lokalnem sistemu, jo lahko poiščete s pregledovanjem. Če pa je datoteka na indeksnem strežniku, vnesite popolnoma kvalificirano pot do nje.
7. Kliknite **Potrdi**. Vaš slovar sopomenk po meri je dodan v sistem iskanja v podjetju in je odslej na voljo za iskanje zbirke.

Povezovanje slovarja sopomenk z zbirko

Če so slovarji sopomenk povezani s sistemom iskanja v podjetju, lahko izberete enega od njih med iskanjem zbirke. Če se izraz iz poizvedbe ujema z izrazom iz slovarja, sistem med rezultati iskanja prikaže tudi dobljene dokumente, ki vsebujejo sopomenke tega izraza.

Praden začnete

Če želite izbrati slovar sopomenk za zbirko, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

Postopek

Če želite povezati slovar sopomenk z zbirko, naredite naslednje:

1. Uredite zbirko, izberite stran Iskanje in kliknite **Konfiguriraj možnosti iskalnega strežnika**.
2. V polju **Ime slovarja sopomenk** na strani Možnosti iskalnega strežnika izberite slovar sopomenk, ki ga želite uporabiti, ko bodo uporabniki poizvedovali po tej zbirki.
Seznam razpoložljivih slovarjev sopomenk vključuje vse slovarje sopomenk, ki so dodani v sistem iskanja v podjetju.
3. Kliknite **Potrdi**.

Prilagojeni slovarji zaustavitvenih besed

Kakovost rezultatov iskanja lahko izboljšate tako, da podate, naj se med obdelavo poizvedbe določene besede samodejno odstranijo iz poizvedbenih izrazov.

Slovar zaustavitvenih besed vsebuje izraze, specifične za podjetje, ki se pogosto uporabljajo in so zato kot poizvedbeni izrazi neuporabni. S tem ko iz poizvedb izločite te izraze, zagotovite, da uporabnikov ne preplavijo najdeni dokumenti, ki so precej nekoristni (iskanje vrne samo dokumente, ki ustrezajo drugim izrazom v poizvedbi). Med obdelavo poizvedbe iskalni strežniki odstranijo zaustavitvene besede iz poizvedbe. Odstranjene besede vključujejo zaustavitvene besede v vašem slovarju po meri in zaustavitvene besede, ki so preddefinirane za iskanje v podjetju (kot so pogosti predlogi in členi).

Pri iskanju v podjetju se po privzetku izvaja jezikovno specifično prepoznavanje zaustavitvenih besed. Ta postopek iz poizvedbe odstrani pogoste besede kot na primer **a** in **the**. Slovar zaustavitvenih besed po meri morate podati samo za zaustavitvene besede, ki so specifične v podjetju ali domeno.

Med obdelavo poizvedbe se zaustavitvene besede odstranijo, preden program poda predloge drugačnega črkovanja. Če so vse besede v poizvedbi zaustavitvene besede, se med obdelavo poizvedbe ne odstrani nobenih zaustavitvenih besed. Odstranjevanje zaustavitvenih besed je onemogočeno, če so vsi poizvedbeni izrazi zaustavitvene besede, saj se tako zagotovi, da bodo obstajali rezultati iskanja. Če je na primer beseda **avto** zaustavitvena beseda in iščete **avto**, potem rezultati iskanja vsebujejo dokumente, ki ustrezajo besedi **avto**. Če pa iščete **avto volvo**, potem rezultati iskanja vsebujejo samo dokumente, ki ustrezajo besedi **volvo**.

Izdelava slovarja zaustavitvenih besed poteka tako, da najprej strokovnjak s področja snovi zbirke izdela seznam zaustavitvenih besed v obliki XML ali pa v sodelovanju z razvijalcem aplikacije izdela datoteko XML. Z orodjem za iskanje v podjetju **esstopworddictbuilder** se datoteka XML pretvori v binarno datoteko (.dic).

Skrbnik iskanja v podjetju binarno datoteko naloži na sistem in ji dodeli prikazno ime. Skrbniki zbirke lahko izberejo, kateri slovar zaustavitvenih besed naj se uporabi za iskanje po dokumentih v zbirki, ko konfigurirajo možnosti iskalnega strežnika za zbirko.

Omejitev: Potem ko slovar zaustavitvenih besed po meri dodate sistemu, ga ne morete več urejati. Če želite popraviti zaustavitvene besede, ki so na voljo za obdelavo poizvedb:

1. Posodobite izvorno datoteko XML.
2. Pretvorite vir XML v novo slovarsko datoteko.
3. Odstranite stari slovar zaustavitvenih besed iz zbirk, ki ga uporabljajo.
4. Izbrišite stari slovar zaustavitvenih besed iz sistema.
5. Dodajte novi slovar zaustavitvenih besed v sistem.
6. Povežite novi slovar zaustavitvenih besed z zbirkami, ki ga bodo uporabljale.

Izdelate lahko skript, ki vključuje te korake, nato pa s pomočjo tega skripta znova razmestite slovar v sistem iskanja v podjetju.

S tem povezani pojmi

 [Custom stop word dictionaries](#)

S tem povezana opravila

 [Creating an XML file for stop words](#)

 [Creating a stop word dictionary](#)

Dodajanje slovarjev zaustavitvenih besed sistemu

Če za odstranjevanje besed iz poizvedb izdelate slovar zaustavitvenih besed po meri, morate slovarje dodati iskalnemu sistemu podjetja. Pozneje lahko izberete, kateri slovar zaustavitvenih besed želite uporabiti za iskanje po zbirki.

Praden začnete


Slovarje zaustavitvenih besed po meri lahko sistemu dodajajo samo člani skrbniške vloge za iskanje v podjetju.

Omejitve

Največja velikost slovarja zaustavitvenih besed je 8 MB.

Postopek

Povezovanje zaustavitvenih besed po meri z iskalnim sistemom podjetja:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
3. Na strani Iskanje kliknite **Konfiguriranje slovarjev zaustavitvenih besed**.
4. Na strani Konfiguriranje slovarjev zaustavitvenih besed kliknite **Dodajanje slovarja zaustavitvenih besed**.
5. Na strani Dodajanje slovarja zaustavitvenih besed vnesite unikatno prikazno ime za slovar.
6. Podajte nahajališče datoteke .dic. Če se datoteka nahaja v lokalnem sistemu, jo lahko poiščete s pregledovanjem. Če pa je datoteka na indeksnem strežniku, vnesite popolnoma kvalificirano pot do nje.
7. Kliknite **Potrdi**. Vaš slovar zaustavitvenih besed po meri je dodan iskalnemu sistemu podjetja in postane na voljo za iskanje po zbirkah.

Povezovanje slovarja zaustavitvenih besed z zbirko

Če so vaši slovarji zaustavitvenih besed povezani z iskalnim sistemom podjetja, lahko izberete enega, ki se uporabi pri iskanju v zbirki. Če se izraz poizvedbe ujema z izrazom v slovarju, potem se izraz odstrani iz poizvedbe pred njegovo obdelavo.

Praden začnete

Slovar zaustavitvenih besed za zbirko lahko izberete samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

Postopek

Za povezovanje slovarja zaustavitvenih besed z zbirko naredite naslednje:

1. Uredite zbirko, izberite stran Iskanje in kliknite **Konfiguriraj možnosti iskalnega strežnika**.
2. V polju **Ime slovarja zaustavitvenih besed** na strani Možnosti iskalnega strežnika izberite slovar zaustavitvenih besed, ki ga želite uporabiti, ko uporabniki poizvedujejo po tej zbirki.

Seznam slovarjev, ki so na voljo, vsebuje vse slovarje zaustavitvenih besed, ki so bili dodani iskalnemu sistemu podjetja.

3. Kliknite **Potrdi**.

Vnovična razmestitev slovarjev po meri

S konzolo za skrbništvo nad iskanjem v podjetju ne morete spremeniti slovarja sopomenk, slovarja izključitvenih besed ali slovarja poviševalnih besed. Kljub temu pa lahko dodate korake v skripte, s katerimi nato znova razmestite slovar.

Preden začnete

Če želite znova razmestiti slovar po meri, se morate prijaviti kot skrbnik iskanja v podjetju.

Postopek

Če želite znova razmestiti slovar po meri za iskanje v podjetju, naredite naslednje:

1. Na indeksnem strežniku odprite eno od naslednjih datotek in poiščite ime slovarja, ki ga želite znova razmestiti:
 - Če želite poiskati ime slovarja sopomenk, odprite `ES_NODE_ROOT/master_config/SynonymConfiguration.xml`.
 - Če želite poiskati ime slovarja izključitvenih besed, odprite `ES_NODE_ROOT/master_config/StopWordDictionaryConfiguration.xml`
 - Če želite poiskati ime slovarja poviševalnih besed, odprite `ES_NODE_ROOT/master_config/BoostingWordDictionaryConfiguration.xml`.

Naslednji primer prikazuje datoteko `SynonymConfiguration.xml` z dvema slovarjema sopomenk, ki sta bila naložena pod imeni `hello` in `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>
```

Sistem naloženim slovarjem dodeli unikatne datotečne ID-je. V prejšnjem primeru sta ID-ja `synonym_hello1.dic` in `synonym_goodby2.dic`. Absolutna pot do teh datotek na indeksnem strežniku je naslednja:

- Pot do slovarja sopomenk je `ES_NODE_ROOT/data/custom_dictionary/synonym_*.dic`.
- Pot do slovarja izključitvenih besed je `ES_NODE_ROOT/data/custom_dictionary/stopwordDictionary_*.dic`.
- Pot do slovarja poviševalnih besed je `ES_NODE_ROOT/data/custom_dictionary/boostingwordDictionary_*.dic`.

V sistemih AIX, Linux in Solaris lahko z ukazom `ls` prikažete seznam razpoložljivih slovarjev. Primer:

```
% 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. Potem ko identificirate slovar, ki ga želite posodobiti, ustavite sistem iskanja v podjetju.

3. Prepišite (zamenjajte) slovar, ki ga želite posodobiti, z novo datoteko slovarja. Prepišite na primer `synonym_hello1.dic` ali `synonym_goodbye2.dic`.
4. Če se sistem iskanja v podjetju izvaja na dveh ali štirih strežnikih, ročno prepišite tudi datoteko slovarja na teh iskalnih strežnikih. Datoteke slovarja se nahajajo na enaki poti kot na indeksnem strežniku (`ES_NODE_ROOT/data/custom_dictionary/`).
5. Znova zaženite sistem iskanja v podjetju.

S tem povezani pojmi

“Prilagojeni slovarji sopomenk” na strani 178

“Prilagojeni slovarji zaustavitvenih besed” na strani 181

“Prilagojeni slovarji poviševalnih besed” na strani 192

Dinamično povzemanje

Dinamično povzemanje je tehnika, ki ugotavlja, katere besedne zveze v dokumentih rezultatov iskanja najbolje predstavljajo pojme, ki jih uporabnik išče.

Dinamično povzemanje poskuša v sistemu iskanja v podjetju zajeti stavke v dokumentih, ki vsebujejo več različnih iskalnih izrazov. Sistem nato izbere nekaj stavkov ali delov stavkov in jih prikaže v rezultatih iskanja. Iskalni izrazi bodo v rezultatih iskanja označeni s pomočjo upodabljanja HTML.

Pri konfiguriranju možnosti iskalnega strežnika za zbirko lahko podate največjo prikazano dolžino povzetkov dokumentov v rezultatih iskanja. Ker povzetki vsebujejo označene znake, bo vmesni pomnilnik, vrnjen iskalni aplikaciji, večji od podane največje vrednosti. Vendar pa prikazana dolžina ne bo preseгла največje podane vrednosti, kljub temu da bo povzetek morda krajši (odvisno od podatkov povzetka, povzetih iz izvornega dokumenta).

Prilagajanje povzetkov dokumentov na skrbniški konzoli

Količino informacij, ki je prikazana v povzetkih dokumentov, lahko prilagodite tako, da podate možnosti za iskalni strežnik na skrbniški konzoli za iskanje v podjetju.

Preden začnete

Prikazno dolžino povzetkov za zbirko lahko krmilijo samo člani skrbniške vloge za iskanje v podjetju ali skrbnik zbirke za zbirko.

O tej nalogi

Vrednost, ki jo podate za največjo prikazno dolžino povzetkov dokumentov, deluje skupaj z vrednostjo, ki jo podate za število stavkov, ki jih lahko vsebuje vsak povzetek. Prednost pa ima vrednost, ki izdela najkrajši povzetek dokumenta.

Če na primer podate omejitev štirih stavkov, bo povzetek dokumenta vseboval samo štiri stavke kljub temu, da prikazana dolžina omogoča prikaz več znakov kot je skupno število znakov v teh štirih stavkih. Če pa na primer število stavkov omejite na 10, obenem pa za prikazano dolžino podate omejitev 500 znakov, lahko sistem prikaže povzetek dokumenta, ki vsebuje manj kot 10 stavkov.

Postopek

Konfiguriranje prikazne dolžine za povzetke dokumentov:

1. Uredite zbirko, izberite stran Iskanje in kliknite **Konfiguriraj možnosti iskalnega strežnika**.

2. Na strani Možnosti iskalnih strežnikov podajte največjo prikazno dolžino za povzetke dokumentov. Ko bodo uporabniki prikazali rezultate iskanja, povzetki dokumentov ne bodo presegali vrednosti, ki jo podate.
3. Podajte koliko stavkov lahko vsebuje vsak povzetek dokumenta (povzetki lahko vsebujejo do 10 stavkov).
4. Kliknite **Potrdi**.
5. Če želite uveljaviti spremembe, zaženite nadzornika iskalnih strežnikov in znova zaženite postopke strežnikov.

Prilagajanje povzetkov dokumentov z urejanjem lastnosti

Vsak dokument, dobljen s poizvedbo iskanja v podjetju, vključuje tudi povzetek. Količino informacij v posameznem povzetku je mogoče prilagoditi z urejanjem datoteke lastnosti.

O tej nalogi

Opise rezultatov iskanja lahko prilagodite tako, da spremenite vrednosti za naslednje lastnosti v datoteki `ES_NODE_ROOT/master_config/collection_ID.runtime.node1/runtime-generic.properties`:

MinWordsPerSentence

Najmanjše število besed v vsakem stavku povzetka. Krajši stavki se vključijo v povzetek, če ni dovolj stavkov, ki vsebujejo več besed, kot je nakazano v vrednosti `MinWordsPerSentence`. Privzeta vrednost je 4.

MaxWordsPerSentence

Največje število besed v vseh stavkih, ki bodo vključeni v povzetek. Če ima stavek več besed, kot jih dovoljuje ta omejitev, je v povzetek vključen le del stavka (tisti, ki vsebuje poizvedbene izraze, največ do vrednosti `MaxWordsPerSentence`). Preostanek stavka ni vključen. Privzeta vrednost je 20.

Stavki za povzetke dokumentov so izbrani na podlagi lastniškega, notranjega algoritma, ki določa pomembnost vseh stavkov, ki vsebujejo iskalne izraze. Izbiranje po pomembnosti se izvede pred filtriranjem stavkov glede na dolžino.

NumberOfReturnedSentences

Število stavkov, ki sestavljajo opis dokumenta. Privzeta vrednost je 5.

MaxSentencesPerDocument

Največje število stavkov v dokumentu, ki bodo obravnavani kot kandidati v postopku izdelave opisa. Privzeta vrednost je 1000.

Postopek

Če želite prilagoditi povzetke dokumentov v rezultatih iskanja, naredite naslednje:

1. V iskalne strežnike se prijavite kot skrbnik za iskanje v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
2. Z urejevalnikom besedil uredite naslednjo datoteko, kjer je `coll_ID` ID, ki ste ga podali za zbirko (ali ki ga je dodelil sistem) pri njeni izdelavi:

`ES_NODE_ROOT/master_config/coll_ID.runtime.node1/runtime-generic.properties`

Nasvet: Če želite določiti preslikavo med imenom zbirke in njenim ID-jem, si oglejte datoteko `ES_NODE_ROOT/master_config/collections.ini`.

3. Spremenite lastnosti, ki jih želite prilagoditi, nato pa shranite in zapustite datoteko.
4. Opravljene spremembe uveljavite tako, da zaustavite iskalne strežnike in jih znova zaženete.

Delo s hitrimi povezavami

Hitre povezave so dokumenti, ki jih sistem vrne v rezultatih iskanja, ko uporabnik predloži poizvedbo, ki vključuje specifične besede ali besedne zveze.

Za konfiguriranje hitrih povezav zbirke uporabite skrbniško konzolo sistema iskanja v podjetju.

Hitre povezave

Hitre povezave nudijo uporabnikom povezave na dokumente, ki so vnaprej določeni kot pomembni za izraze poizvedbe.

Hitra povezava je URI, ki ga sistem iskanja v podjetju samodejno vključi v rezultate iskanja, ko poizvedba vsebuje določene besede ali besedne zveze. Običajno se URI-ji hitrih povezav prikažejo na vrhu seznama rezultatov, zaradi česar bodo uporabniki videli dokumente, za katere ste vnaprej določili, da so pomembni za poizvedbo.

Sistem vrne hitre povezave poleg drugih rezultatov iskanja. Procesi iskanja v indeksu poiščejo dokumente, ki se ujemajo z izrazi poizvedbe, in poleg URI-jev hitrih povezav vrnejo še URI-je hitrih povezav za te dokumente.

Pri konfiguriranju hitre povezave lahko podate opisni naziv in povzetek URI-ja, ki je uporabnikom v pomoč pri prepoznavanju dokumenta, prav tako pa lahko ti hitro ugotovijo, ali je to dokument, ki ga želijo priklicati.

Za URI <http://www.ibm.com/education/us/> lahko na primer podate naziv kot je IBM-ovo izobraževanje v Združenih državah Amerike in povzetek Rešitve, izdelki in viri za strokovnjake, učitelje in študente v Združenih državah Amerike.

Če želite hitre povezave uporabiti v zbirkah sistema iskanja v podjetju, mora biti v iskalni aplikaciji na voljo možnost za prikaz hitrih povezav. V nekaterih iskalnih aplikacijah imajo uporabniki pri iskanju po zbirki možnost omogočiti ali onemogočiti prikaz hitrih povezav.

Konfiguriranje hitrih povezav

Če želite izdelati hitro povezavo do zbirke za iskanje v podjetju, morate povezati URI dokumenta s ključnimi besedami, ki poskrbijo za njegovo vključitev v rezultate iskanja.

Preden začnete

Če želite konfigurirati hitre povezave, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke, ki vsebuje hitro povezavo.

O tej nalogi

Do primerov, ki prikazujejo, kako podati ključne besede, in URI-jev za hitre povezave pridete tako, da med izdelovanjem ali urejanjem hitre povezave kliknete **Pomoč**.

Da bi opravljene spremembe stopile v veljavo, vam ni treba znova zagnati iskalnih strežnikov.

Postopek

Če želite konfigurirati hitro povezavo, naredite naslednje:

1. Uredite zbirko, izberite stran Iskanje in kliknite **Konfiguriraj hitre povezave**.
2. Na strani Hitre povezave kliknite **Izdelaj hitro povezavo**.

3. Podajte ključne besede in fraze, po zaslugi katerih bo ta hitra povezava prikazana med rezultati iskanja, URI dokumenta, za katerega ste ugotovili, da je relevanten za to poizvedbo, in druge možnosti za hitro povezavo.

V eno vrstico lahko podate eno ključno besedo, več ključnih besed ali eno frazo (dve ali več besed, zapisanih v narekovajih). Besede je treba ločiti s presledkom (ključnih besed ni mogoče razmejiti z vejico). Pritisnite tipko Enter, da se pomaknete v novo vrstico.

4. Kliknite **Potrdi**.

Vaša nova hitra povezava je prikazana na strani Hitre povezave, skupaj z drugimi hitrimi povezavami, ki sodijo v isto zbirko.

S tem povezane povezave

“Formati URI v indeksu iskanja v podjetju” na strani 107

Rangiranje dokumentov

Ko uporabnik išče po zbirki, iskalni procesi vrnejo najpomembnejše rezultate za izraze in pogoje poizvedbe.

Iskalni strežniki podpirajo skladno obogatene poizvedbe in za prikaz najpomembnejših rezultatov iskanja uporabljajo več različnih metod, na primer ocenjevanje na osnovi besedila in statično rangiranje. Privzeto vedenje rangiranja lahko razširite tako, da konfigurirate možnosti, ki vplivajo na pomembnost dokumentov v rezultatih iskanja:

- Izdelate lahko prilagojene slovarje poviševalnih besed, ki vplivajo na način rangiranja dokumentov, ki vsebujejo podane poviševalne besede, v rezultatih iskanja.
- Vplivate lahko na ocene dokumentov, ki se ujemajo s podanim vzorcem URI-ja.
- Vplivate lahko na ocene dokumentov, ki vsebujejo polja, preslikana v poviševalne razrede.

S tem povezani pojmi

“Rangiranje dokumentov, ki temelji na poviševalnih razredih” na strani 196

“Prilagojeni slovarji poviševalnih besed” na strani 192

Ocenjevanje na osnovi besedila

Sistem iskanja v podjetju za vsak dokument, ki se ujema z izrazi v poizvedbi, dinamično izračuna oceno.

Besedilni rezultat dokumenta predstavlja pomembnost poizvedbenih izrazov v dokumentu. Sistem iskanja v podjetju pri izračunavanju besedilne ocene posameznega dokumenta, ki se ujema s poizvedbo, upošteva več dejavnikov, kot so:

- Ali se na podlagi izrazov dokument loči od drugih dokumentov. Če se na primer poizvedbeni izrazi pojavijo v enem dokumentu, ne pa tudi v drugih, to pomeni, da so ti izrazi pomembni za ta dokument in da se dokument zaradi njih razlikuje od drugih dokumentov. Poizvedbeni izrazi, ki se pojavijo v večini dokumentov, prispevajo manj k rezultatu dokumenta kot poizvedbeni izrazi, ki se pojavijo v manjšem številu dokumentov.
- Število pojavitev poizvedbenih izrazov v dokumentu. Rezultat izraza se izračuna na podlagi vsake pojavitve tega izraza v dokumentu. Več kot je pojavitev poizvedbenega izraza v dokumentu, višji je rezultat dokumenta. Če na primer iščete izraz *thinkpad*, se dokument, v katerem se ta izraz pogosto pojavi, v rezultatih iskanja uvrsti više kot dokumenti, v katerih se izraz pojavi le nekajkrat.
- Pri dokumentih HTML se upoštevajo tudi atributi poizvedbenega izraza (torej kontekst izraza). Rezultat izraza se izračuna na podlagi atributov tega izraza (na primer lokacije, ali je besedilo krepko, ležeče, ali vsebuje sidra in podobno). V splošnem imajo izrazi, ki se pojavijo v naslovu dokumenta, višji rezultat kot izrazi, ki se pojavijo v odstavku. Poudarjeni izrazi (na primer izrazi, zapisani krepko) imajo višji rezultat kot izrazi, zapisani v običajnem besedilu. Pomembnost atributov lahko konfigurirate v nastavitvah razčlenjevalnika.
- Pomembnost poizvedbenih izrazov. Če želite prilagoditi pomembnost izrazov v dokumentu, lahko zanje konfigurirate poviševalne vrednosti. V tem primeru povežete slovar poviševalnih besed (ki vsebuje besede in poviševalne vrednosti) z zbirko. Sistem izdelani slovar uporablja med iskanjem, poviševalne vrednosti pa vplivajo na rezultat dokumenta. Višja kot je poviševalna vrednost, več prispeva izraz k rezultatu dokumenta.
- Medsebojna bližina poizvedbenih izrazov v dokumentu. Če se poizvedbeni izrazi v dokumentu pojavijo blizu drug drugemu, se za izračun rezultata besedila uporabi njihova leksikalna afiniteta. Predpostavimo, da imate dva dokumenta. Eden od njih opisuje

avtomobilsko hišo v mestu (besedi avtomobilska in hiša se v besedilu pojavita blizu druga drugi). Drugi dokument pa obisuje avtomobilsko razstavo v mestu, v bližini stare hiše (besedi avtomobilska in hiša se ne nahajata blizu). Če iščete izraz **avtomobilska hiša**, sistem zaradi bližine izrazov v prvem dokumentu ta dokument uvrsti više kot drugega.

- Dejavnika za določanje rezultata dokumenta sta tudi dolžina dokumenta in pestrost besedišča (število unikatnih besed).

Statično rangiranje

Z določenimi vrstami dokumentov lahko povežete dejavnik statičnega rangiranja, ki poveča pomembnost teh dokumentov v rezultatih iskanja.

Pri izdelavi zbirke podate možnosti **Pomembnosti dokumentov**. Tip pomembnosti dokumentov, ki ga izberete, določa, ali je faktor statičnega rangiranja povezan z dokumenti v zbirki. Če uporabniki preiskujejo zbirko, ki uporablja statično rangiranje, dejavnik statičnega rangiranja vpliva na rangiranje dokumentov, vrnjenih v rezultatih iskanja.

Faktor statičnega rangiranja za spletno vsebino temelji na povezavah. Število povezav na dokument iz drugih dokumentov in vir teh povezav lahko povečajo relevantnost tega dokumenta v rezultatih iskanja.

Za dokumente, ki vključujejo podatkovna polja ali metapodatke datuma, faktor statičnega rangiranja temelji na datumu dokumentov. Datumsko polje dokumenta, ki ga nudi pajek, je lahko datum spremembe ali datum zadnjega pajkanja dokumenta, odvisno od konfiguracije pajka.

Datum dokumenta lahko poviša njegovo relevantnost. Tako so na primer nedavni članki v novičarskih skupinah NNTP verjetno bolj koristni kot starejši članki. Če vir podatkov vključuje več datumskih vrednosti, lahko pri konfiguriranju pajka izberete tisto, ki je najpomembnejša za določitev relevantnosti dokumentov.

Če v zbirki uporabljate statično rangiranje, zagotovite, da v isti zbirki nimate virov podatkov, ki uporabljajo različne vrste rangiranja. Če na primer želite kot statični dejavnik rangiranja uporabiti povezave z dokumentom, zagotovite, da zbirka vsebuje samo spletne dokumente. Rangiranje dokumentov je manj natančno, če združite vire z različnimi modeli rangiranja v isti zbirki. Prav tako razvrstitev rezultatov iskanja morda ne bo v skladu s pričakovanji.

Prepričajte se tudi, da dokumenti v zbirki vsebujejo polja in vrednosti, ki omogočajo uveljavitev statičnega rangiranja. Zamislite si na primer zbirko, ki je konfigurirana za uporabo statičnega rangiranja na osnovi datumov dokumentov, in pajka v zbirki, ki je konfiguriran za uporabo specifičnega polja kot datum dokumenta. Če dokument ne vsebuje tega polja, morda ne bo pravilno razvrščen po pomembnosti, zato razvrstitev v rezultatih iskanja ne bo v skladu s pričakovanji.

Implikacije rangiranja na osnovi povezav

Statično rangiranje skupaj z dejavniki, kot je na primer dodelitev rezultata za povišanje vzorcev URI, prispeva k statičnemu rezultatu dokumenta in vpliva na njegovo pomembnost. Model rangiranja na podlagi povezav se ponavadi uveljavi za spletne zbirke, saj ta model izračuna statični rang dokumenta na osnovi števila povezav nanj. Dokument, na katerega vodijo povezave iz večjega števila drugih dokumentov, so obravnavane kot relevantnejše.

Če konfigurirate ta model za nespletno ali mešano zbirko (zbirko, ki vsebuje spletne in nespletne dokumente), se lahko kakovost iskanja zniža, saj nespletne dokumenti nimajo koncepta povezovanja.

Če je omogočeno rangiranje na osnovi povezav, je omogočeno tudi odkrivanje podvojenih dokumentov. Podvojeni dokumenti imajo enak statični rang kot glavni dokument. Če faktorji povišanja vzorcev URI niso konfigurirani za dokumente v skupini dvojnikov, imajo vsi podvojeni dokumenti enak statični rezultat.

Obnavljanje privzetih vrednosti za statično rangiranje dokumentov

Če konfigurirate možnost statičnega rangiranja dokumentov, ko izdelate zbirko, lahko lastnosti nastavite spet na privzete vrednosti tako, da uredite datoteke `runtime.properties` za zbirko.

Preden začnete

Za obnovo privzetih vrednosti rangiranja dokumentov za zbirko morate biti skrbnik iskanja v podjetju.

O tej nalogi

Če želite obnoviti privzete vrednosti rangiranja dokumentov za zbirko, morate posodobiti datoteke `runtime.properties` za to zbirko in vse iskalne strežnike v iskalnem sistemu vašega podjetja. V konfiguraciji z več strežniki se datoteka `runtime.properties` nahaja na indeksnem strežniku v imeniku `ES_NODE_ROOT/master_config/collection_ID.runtime.node_ID`, pri čemer je `collection_ID` ID zbirke in `node_ID` ID za iskalne strežnike.

Če želite na primer posodobiti zbirko `col1` v sistemu iskanja v podjetju z več strežniki, posodobite datoteko `runtime.properties` za to zbirko in oba iskalna strežnika (`node3` in `node4`):

```
ES_NODE_ROOT/master_config/col1.runtime.node3/runtime.properties
ES_NODE_ROOT/master_config/col1.runtime.node4/runtime.properties
```

Postopek

Če želite za vrednosti rangiranja dokumentov obnoviti privzete vrednosti za zbirko:

1. Se prijavite na indeksni strežnik kot skrbnik iskanja v podjetju.
2. Ugotovite ID zbirke za zbirko, ki ji želite obnoviti privzete vrednosti rangiranja. ID zbirke se nahaja v datoteki `ES_NODE_ROOT/master_config/collections.ini`. Za lažji prikaz uredite to datoteko. V naslednjem primeru je `col1` ID zbirke:

```
% 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. Uredite datoteko `runtime.properties` za zbirko, ki jo želite obnoviti, in naredite naslednje spremembe:

- a. Izbrišite naslednje lastnosti:

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- b. Če datoteka `runtime.properties` podaja `trevi.sourcetype=1`, ki nakazuje, da so dokumenti rangirani po povezavah, uredite datoteko `ES_INSTALL_ROOT/default_config/runtime.1/runtime.properties` ter kopirajte in prilepite v datoteko `runtime.properties` naslednje privzete lastnosti:
- ```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```
- c. Če datoteka `runtime.properties` podaja `trevi.sourcetype=2`, ki nakazuje, da so dokumenti rangirani po povezavah, uredite datoteko `ES_INSTALL_ROOT/default_config/runtime.2/runtime.properties` ter kopirajte in prilepite v datoteko `runtime.properties` naslednje privzete lastnosti:
- ```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```
- d. Če datoteka `runtime.properties` podaja `trevi.sourcetype=3`, ki nakazuje, da se za rangiranje dokumentov v zbirki statični faktor rangiranja ne uporablja, uredite datoteko `ES_INSTALL_ROOT/default_config/runtime.0/runtime.properties` ter kopirajte in prilepite v datoteko `runtime.properties` naslednje privzete lastnosti:
- ```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```
4. V konfiguraciji z več strežniki ponovite korak 3 na strani 191, da posodobite datoteko `runtime.properties` za isto zbirko in drugi iskalni strežnik.
5. V skrbniški konzoli nadzirajte stran Iskanje in znova zaženite postopke iskanja za to zbirko.

Po potrebi ponovite te korake za vsako zbirko, kateri želite obnoviti privzete vrednosti rangiranja dokumentov.

---

## Prilagojeni slovarji poviševalnih besed

Če želite izboljšati kakovost rezultatov iskanja, lahko z izdelavo slovarja poviševalnih besed po meri vplivate na to, kako so dokumentih uvrščenih v rezultatih iskanja.

Če podaja poizvedba besedo, ki je v slovarju poviševalnih besed, bo pomembnost dokumentov, ki vsebujejo to besedo, zvišana ali znižana glede na poviševalni faktor, ki je konfiguriran za besedo v slovarju.

S slovarjem poviševalnih besed lahko zagotovite, da se v primeru, ko uporabnik poda določene iskalne izraze, vrnejo določeni dokumenti. Predpostavimo, da imate zbirko, v kateri so dokumenti na temo avtomobilov. Za takšne dokumente lahko domnevate, da so nekatere ključne besede, povezane z modeli avtomobilov, imenom proizvajalca in podobno, zelo pomembne. Na rangiranje rezultatov iskanja lahko vplivate tako, da povežete poviševalno vrednost s ključnimi besedami (`model`, `proizvajalec` in podobno) v slovarju poviševalnih besed. Ko uporabniki preiščejo zbirko in podajo poizvedbo, v kateri je katera od ključnih besed, so dokumenti o avtomobilih v rezultatih iskanja razvrščeni više kot drugi dokumenti.

Območje poviševalnih faktorjev je od -10 do 10. Med obdelavo poizvedbe iskalni strežniki povečajo pomembnost dokumentov, ki vsebujejo besede s pozitivnimi poviševalnimi faktorji in zmanjša pomembnost dokumentov, ki vsebujejo besede z negativnimi poviševalnimi faktorji.

Tako je na primer dokument, ki se ujema z izrazi poizvedbe z visokimi poviševalnimi faktorji uvrščen višje kot bi bil, če ne bi uporabili poviševalnega faktorja. (Pospeševalni faktor je samo eden od faktorjev, ki prispeva k uvrstitvi dokumenta.)

Pri izdelavi slovarja lahko dodelite enak poviševalni faktor kakršnemukoli številu besed. Slovar lahko vsebuje izraz, sestavljen iz ene besede ali izraz, sestavljen iz več besed. Izrazi, sestavljeni iz več besed, so primerjani kot fraza.

Če je beseda utežena s poviševalno vrednostjo, ki je podana v poizvedbi, ki uporablja operator OR (na primer `to | ono`), je za izraze poizvedbe izračunano uteženo povprečje. Nastali združeni rezultat se uporablja za vse pojavitve operandov poizvedbe OR. Za različne operande poizvedbe OR niso izračunani različni rezultati.

Pospeševanje, ki temelji na slovarjih poviševalnih besed, ni podprto z izrazi poizvedbe, ki temeljijo na poljih. Pri razčlenjevanju izrazov poizvedbe je za izračun uvrstitve dokumenta uporabljeno samo besedilo poizvedbe, ne pa tudi ime polja. Za uveljavitev poviševalnih faktorjev v izrazih poizvedbe, ki se pojavijo v poljih, lahko preslikate ta imena polj v poviševalne razrede.

Za izdelavo slovarja poviševalnih besed mora strokovnjak za predmet zbirke izdelati seznam poviševalnih besed v formatu XML ali sodelovati pri razvijalcu aplikacije za izdelavo datoteke XML. Za pretvorbo datoteke XML (.dic) morate uporabiti orodje za iskanje v podjetju **esboosttermdictbuilder**.

Skrbnik iskanja v podjetju binarno datoteko naloži na sistem in ji dodeli prikazno ime. Skrbniki zbirke lahko izberejo slovar poviševalnih besed, ki bodo uporabljene za preiskovanje dokumentov v zbirki, ko konfigurirajo možnosti iskalnega strežnika za zbirko.

**Omejitev:** Ko dodate slovar poviševalnih besed po meri v sistem, ga ne morete več urejati. Če želite pregledati in popraviti poviševalne besede, ki so na voljo za obdelavo poizvedbe, morate narediti naslednje:

1. Posodobite izvorno datoteko XML.
2. Pretvorite vir XML v novo slovarsko datoteko.
3. Odstranite star slovar poviševalnih besed iz zbirk, ki ga uporabljajo.
4. Izbršite star slovar poviševalnih besed iz sistema.
5. V sistem dodajte nov slovar poviševalnih besed.
6. Povežite nov slovar poviševalnih besed z zbirkami, ki ga bodo uporabljale.

Izdelate lahko skript, ki vključuje te korake, nato pa s pomočjo tega skripta znova razmestite slovar v sistem iskanja v podjetju.

#### **S tem povezani pojmi**

“Rangiranje dokumentov” na strani 189



Custom boost word dictionaries

#### **S tem povezana opravila**



Creating an XML file for boost words



Creating a boost word dictionary

## Dodajanje slovarjev poviševalnih besed v sistem

Če izdelate slovarje poviševalnih besed po meri, jih morate povezati s sistemom iskanja v podjetju. Pozneje lahko izberete, kateri slovar poviševalnih besed želite uporabiti za preiskovanje zbirke.

### Praden začnete


Če želite v sistem dodati slovarje poviševalnih besed po meri, morate biti član s skrbniško vlogo v iskanju v podjetju.

### Omejitve

Največja dovoljena velikost slovarja poviševalnih besed je 8 Mb.

### Postopek

Postopek za povezavo poviševalnih besed po meri s sistemom iskanja v podjetju:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Kliknite ikono  **Uredi**, da preklopite v prikaz za urejanje sistema.
3. Na strani Iskanje kliknite **Konfiguriraj slovarje poviševalnih besed**.
4. Na strani Konfiguriranje slovarjev poviševalnih besed kliknite **Dodaj slovar poviševalnih besed po meri**.
5. Na strani Dodajanje slovarja poviševalnih besed vpišite unikatno ime slovarja za prikaz in po želji vnesite opis.
6. Podajte nahajališče datoteke .dic. Če se datoteka nahaja v lokalnem sistemu, jo lahko poiščete s pregledovanjem. Če pa je datoteka na indeksnem strežniku, vnesite popolnoma kvalificirano pot do nje.
7. Kliknite **Potrdi**. Slovar poviševalnih besed po meri bo dodan v sistem iskanja v podjetju in postane na voljo za preiskovanje zbirk.

## Povezovanje slovarja poviševalnih besed z zbirko

Če so s sistemom iskanja v podjetju povezani slovarji poviševalnih besed, lahko izberete enega med njimi, ki bo uporabljen med preiskovanjem zbirke. Če se izraz poizvedbe ujema z izrazom v slovarju, bo pomembnost dokumentov, ki vsebujejo ta izraz, zvišana ali znižana glede na poviševalni faktor, ki je dodeljen izrazu v slovarju.

### Praden začnete

Če želite izbrati slovar poviševalnih besed za zbirko, morate biti član s skrbniško vlogo v iskanju v podjetju ali skrbnik zbirke.

### Postopek

Postopek za povezavo slovarja poviševalnih besed z zbirko:

1. Uredite zbirko, izberite stran Iskanje in kliknite **Konfiguriraj možnosti iskalnega strežnika**.
2. V polju **Ime slovarja poviševalnih besed** na strani Možnosti iskalnega strežnika izberite slovar poviševalnih besed, ki jih želite uporabiti, če uporabniki poizvedujejo v tej zbirki.  
Seznam razpoložljivih slovarjev vključuje vse slovarje poviševalnih besed, ki so bili dodani v sistem iskanja v podjetju.

### 3. Kliknite **Potrdi**.

---

## Rangiranje dokumentov, ki temelji na vzorcih URI

Pomembnost dokumentov lahko povečate ali zmanjšate tako, da vzorcem URI-jev dodelite poviševalne faktorje.

Sistem vsem dokumentom, ko jih doda v indeks, dodeli privzeto oceno statičnega rangiranja. Privzeta ocena je odvisna od tega, ali je bilo statično rangiranje za zbirko omogočeno, in če je bilo, je odvisna tudi od tipa statičnega rangiranja (po datumu dokumentov ali pri spletnih dokumentih po številu drugih dokumentov, ki se na ta dokument navezujejo).

Na relativno pomembnost dokumenta lahko vplivate tako, da vzorcem URI-jev dodelite poviševalne faktorje. Sistem bo s pomočjo poviševalnega faktorja in privzete statične ocene rangiranja ter drugih faktorjev ocenil končno statično oceno dokumenta.

Pri tem je pomemben vrstni red vzorcev URI-jev, ki jih konfigurirate. Indeksni strežnik pri izračunavanju vrednosti vsakega dokumenta v zbirki ovrednoti vzorce URI-jev po vrstnem redu, v katerem so navedeni. Za vsak URI:

1. Indeksni strežnik zaporedoma pregleduje vzorce URI-jev.
2. Ko najde prvi vzorec URI-jev, ki ustreza dokumentu v indeksu, za dokument uporabi poviševalni faktor, ki je konfiguriran za ta vzorec URI-jev.
3. Če se dokument ne ujema z nobenim vzorcem URI-jev, potem sistem uporabi privzeto statično oceno rangiranja.

Ko konfigurirate poviševalni faktor za vzorec URI-jev, morate vzorec URI-jev postaviti v vrstni red, v katerem naj ga indeksni strežnik pregleda.

## Vplivanje na ocene dokumentov, ki ustrezajo vzorcem URI-jev

Pomembnost dokumentov, ki se ujemajo z vzorcem URI-jev, lahko povečate ali zmanjšate tako, da privzeti oceni statičnega rangiranja dodelite poviševalni faktor.

### **Preden začnete**

Na pomembnost dokumentov, ki se ujemajo z vzorcem URI-jev, lahko vplivate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

### **O tej nalogi**

Pospeševalni faktor, ki ga konfigurirate, se uporablja s privzeto oceno statičnega rangiranja za izračun nove statične ocene za vse dokumente, ki ustrezajo podanemu vzorcem URI-jev.

Pospeševalni faktorji zvišujejo samo statične ocene, poleg tega so faktorji samo eden izmed prispevkov pri izračunavanju končnega ranga dokumenta. Če na primer dokument vsebuje veliko število povezav (zaradi česar dobi visoko začetno oceno), bo dokument, ki ne vsebuje nobene povezave, vedno rangiran nižje.

### **Postopek**

Če želite vplivati na oceno dokumentov, ki se ujemajo z vzorcem URI-jev, naredite naslednje:



1. Uredite zbirko, izberite stran Stvarno kazalo in kliknite **Vpliv na ocene z ujemanjem z vzorci URI-jev**.
2. Na strani Vpliv na ocene z ujemanjem z vzorci URI-jev kliknite **Dodaj vzorec URI-jev**.

3. Vnesite vzorec URI-jev za dokumente, ki jim želite povečati ali zmanjšati pomembnost v rezultatih iskanja. Primer:  

```
http://domain.org/hr/*
db2://*ROI*
/afs/
```
4. Vnesite vrednost med -10 in 10 za poviševalni faktor. Končne statične ocene za vse dokumente, ki se ujemajo z vzorcem URI-jev, bo izračunane na osnovi tega poviševalnega faktorja.
5. Kliknite **Potrdi**.
6. Na strani Vpliv na ocene z ujemanjem z vzorci URI-jev postavite nov vzorec URI-jev v vrstni red, kakor želite, da ga indeksni strežnik prebere.  
 Indeksni strežnik izračuna ocene statičnega rangiranja po vrstnem redu, kot so navedeni URI-ji. Boljše rezultate boste dosegli, če najprej navedete bolj specifične URI-je. V naslednjem zgledu se podimenik /forms ujema z vzorcem URI-jev  

```
http://www.ibm.com/hr/*
```

 Pravilen izračun ocen za dokumente v podimeniku /forms zagotovite tako, da najprej navedete vzorec URI-jev za podimenik /forms:  

```
http://www.ibm.com/hr/forms/* 8
http://www.ibm.com/hr/* -2
```
7. Če želite spremeniti vzorec URI-jev ali poviševalni faktor, izberite vzorec URI-jev in kliknite  **Uredi**.
8. Če želite s seznama odstraniti vzorec URI-jev, ga izberite in kliknite  **Odstrani**.
9. Če želite uveljaviti poviševalni faktor za dokumente, ki so bili predhodno indeksirani, znova zgradite stvarno kazalo.

---

## Rangiranje dokumentov, ki temelji na poviševalnih razredih

S preslikavo polj v poviševalne razrede lahko vplivate na to, kako so dokumenti uvrščeni v rezultatih iskanja.

Pri razčlenjevanju dokumentov dodeli razčlenjevalnik žetonom *poviševalne razrede* glede na polja, ki jim pripadajo žetoni. Ti poviševalni razredi so vključeni v indeks in se uporabljajo med vrednotenjem poizvedbe za izračun uvrstitev, ki prispevajo k temu, kako so uvrščeni dokumenti.

Za vpliv na to, kako so izračunane uvrstitve, lahko konfigurirate številske poviševalne faktorje za poviševalne razrede. Če se izraz poizvedbe ujema z žetonom v polju, ki je preslikano v poviševalni razred, vpliva prispevek te pojavitve žetona na skupno uvrstitev dokumenta. Uvrstitev je izračunana z uveljavitvijo poviševalnega faktorja, ki je konfiguriran v poviševalnem razredu.

Za primer vzemimo, da želite pospešiti uvrstitev polj z nazivi. Če se v nazivu pojavi izraz poizvedbe, ima pojavitve visok prispevek v uvrstitvi dokumenta in pomaga, da bo dokument uvrščen višje v rezultatih iskanja.

Za vpliv na uvrstitev dokumentov uporabite skrbniško konzolo iskanja v podjetju, na kateri podajte poviševalne faktorje za poviševalne razrede in preslikavo polj v poviševalne razrede. Za iskanje v podjetju je vnaprej konfiguriranih šestnajst poviševalnih razredov. Osem poviševalnih razredov je oblikovanih za uporabo s polji vsebine, drugih osem pa za polja metapodatkov. Uvrstitve, ki so povezane s privzetimi poviševalnimi razredi, lahko uredite in s poviševalnimi razredi povežete druga ali dodatna polja.

Če spremenite preslikave polj, morate dokumente znova pajkati in razčleniti, da bodo spremembe lahko uveljavljene v predhodno indeksiranih dokumentih. Če spremenite faktorje,

ki so podani za poviševalni razred, za uveljavitev sprememb nadzirajte iskalne strežnike ter zaustavite in znova zaženite postopke iskalnega strežnika.

## Odkrivanje dvojnikov dokumentov in povzetki dokumentov

Če preslikate polje v poviševalni razred, morate podati, ali bo polje uporabljeno za odkrivanje dvojnikov dokumentov in ali je vsebino polja mogoče vključiti v povzetke dokumentov v rezultatih iskanja.

- Če je polje uporabljeno za odkrivanje dvojnikov dokumentov, je ocenjeno kot polje vsebine, in so za izbiro na voljo samo poviševalni razredi, ki so oblikovani za polja vsebine. Vsebino teh tipov polj lahko uporabite v dinamičnih povzetkih dokumentov v rezultatih iskanja.
- Če polje ni uporabljeno za odkrivanje dvojnikov dokumentov, je ocenjeno kot polje metapodatkov, in so za izbiro na voljo samo poviševalni razredi, ki so oblikovani za polja metapodatkov. Če sta v tem primeru dva dokumenta popolnoma enaka v vsakem pogledu, vendar je podano polje ocenjeno kot dvojniki, polje ne bo uporabljeno v dinamičnih povzetkih dokumentov.

## Visoke in nizke vrednosti priklica

Pri vrednotenju poizvedbe oceni postopek iskanja število nastalih dokumentov, ki bodo vrnjeni. Pragovi določajo, ali bo poizvedba upoštevana kot poizvedba z nizko vrednostjo priklica, z visoko vrednostjo priklica ali z vrednostjo, ki je med nizko in visoko vrednostjo:

### Nizka vrednost priklica

Če je pričakovano število nastalih dokumentov pod nizkim pragom, je poizvedba ocenjena kot poizvedba z nizko vrednostjo priklica.

### Visoka vrednost priklica

Če je pričakovano število nastalih dokumentov nad visokim pragom, je poizvedba ocenjena kot poizvedba z visoko vrednostjo priklica.

### Mešana vrednost priklica

Če je pričakovano število dokumentov med dvema pragoma, je vrednost priklica poizvedbe mešanica dveh pragov.

Vsak poviševalni razred podaja poviševalne faktorje, ki so med obdelavo poizvedbe povezani s poizvedbami z nizko vrednostjo priklica in z visoko vrednostjo priklica. Nizki poviševalni faktor vpliva na relativno pomembnost poizvedb z nizko vrednostjo priklica, visok poviševalni faktor pa na relativno pomembnost poizvedb z visoko vrednostjo priklica. Mešanica dveh poviševalnih faktorjev vpliva na relativno pomembnost poizvedb z mešano vrednostjo priklica.

Vrednosti poviševalnih faktorjev krmilijo relativno pomembnost vsake pojavitve izraza poizvedbe v dokumentu. Vsaka pojavitve izraza poizvedbe v dokumentu se šteje v skladu z ustreznim poviševalnim faktorjem.

Če konfigurirate poviševalne razrede za zbirko, lahko uredite privzete poviševalne faktorje. Tako lahko na primer podate poviševalne faktorje, ki zagotavljajo, da bodo izrazi poizvedbe, ki se pojavijo v poljih z nazivi, šteli petkrat več kot izrazi poizvedbe, ki se pojavijo v običajnem besedilu.

### S tem povezani pojmi

“Rangiranje dokumentov” na strani 189

## Preslikava polj v poviševalne razrede

Na relativno pomembnost polj lahko vplivate s preslikavo imen polj v poviševalne razrede.

### Preden začnete

Če želite preslikati polja v poviševalne razrede, morate biti član s skrbniško vlogo v iskanju v podjetju ali skrbnik zbirke.



### O tej nalogi

Sistem uporablja poviševalni faktor za vplivanje na uvrstitev dokumentov, ki vsebujejo izraze poizvedbe v poljih, ki so preslikana v poviševalne razrede.

Iskanje v podjetju rezervira nekatere preslikave za notranja polja in običajno besedilo, ki nimajo drugih definirajočih značilnosti. Druga polja lahko preslikate v poviševalne razrede, ki jih uporabljajo rezervirana polja, vendar rezerviranih polj ne morete urejati ali izbrisati.

### Postopek

Postopek za preslikavo polj v poviševalne razrede:

1. Uredite zbirko, izberite stran Razčlenitev in kliknite **Preslikaj polja v poviševalne razrede**.
2. Na strani Preslikava polj v poviševalne razrede kliknite **Dodaj polje**.
3. Na strani Dodajanje polja v poviševalni razred vpišite ime polja, ki ga želite preslikati v poviševalni razred.  
Podate lahko ime polja, ki obstaja v pajkanem viru ali v zunanjem viru, ime polja, ki je preslikano iz elementa XML, ime polja, ki je preslikano iz elementa metapodatkov HTML ali eno od vnaprej definiranih imen polj.
4. Podajte, ali bo polje uporabljeno za odkrivanje dvojnikov dokumentov. Če izberete potrditveno polje, vsebuje seznam razpoložljivih poviševalnih razredov razrede, ki veljajo za polja vsebine. Če je v rezultatih iskanja vrnjen dokument s tem poljem, bo vsebina polja prikazana v področju s povzetkom dokumenta.  
Če počistite potrditveno polje, vsebuje seznam razpoložljivih poviševalnih razredov razrede, ki veljajo za polja metapodatkov. Vsebina polja ne bo prikazana v področju s povzetkom dokumenta v rezultatih iskanja.
5. Izberite poviševalni razred in kliknite **Potrdi**.  
Polje, ki ste ga dodali, je prikazano na strani Preslikava polj v poviševalne razrede. Izberete lahko možnost za urejanje poviševalnega razreda ali konfiguriranje različnih poviševalnih faktorjev za določanje uvrstitev dokumentov, ki vsebujejo to polje.
6. Če želite spremeniti, ali bo polje uporabljeno za odkrivanje dvojnikov dokumentov ali za preslikavo polja v drug poviševalni razred, kliknite  **Uredi**. (Polj, rezerviranih za uporabo iskanja v podjetju, ne morete urejati.)
7. Za odstranitev polja iz poviševalnega razreda kliknite  **Odstrani**. (Polj, rezerviranih za uporabo iskanja v podjetju, ne morete odstraniti.)
8. Za uveljavitev sprememb v dokumentih, ki so bili predhodno indeksirani, znova pajkajte in indeksirajte dokumente.

### S tem povezani pojmi

“Rangiranje dokumentov” na strani 189

## Konfiguriranje poviševalnih faktorjev za poviševalne razrede

Pospeševalni faktorji, ki jih konfigurirate za poviševalne razrede, predstavljajo vašo oceno tega, kako ustrezna je prisotnost določenih polj v nastalih dokumentih glede na poizvedbo.



Pospeševalni razredi z visokimi poviševalnimi faktorji lahko povečajo pomembnost nastalih dokumentov, ki vsebujejo polja, preslikana v poviševalni razred.

### **Preden začnete**


Če želite konfigurirati poviševalne faktorje za poviševalne razrede, morati biti član s skrbniško vlogo v iskanju v podjetju ali skrbnik zbirke.

### **O tej nalogi**

Sistem uporablja poviševalne faktorje, ki so konfigurirani za poviševalni razred, privzeta uvrstitev statičnega rangiranja in druge faktorje za izračun nove uvrstitve za nastale dokumente, ki vsebujejo polja, preslikana v poviševalni razred.

### **Postopek**

Postopek za konfiguriranje poviševalnih faktorjev za poviševalne razrede:

1. Uredite zbirko, izberite stran Razčlenitev in kliknite **Preslikaj polja v poviševalne razrede**.
2. Na strani Preslikava polj v poviševalne razrede kliknite **Urejanje poviševalnih razredov**.
3. Na strani Pospeševalni razredi poiščite poviševalni razred, ki ga želite spremeniti, in kliknite  **Uredi**.
4. Na strani Urejanje poviševalnega razreda podajte nove vrednosti za visok in nizek poviševalni faktor. Za oba faktorja lahko vpišete isto vrednost.
5. Kliknite **Potrdi**.
6. Za uveljavitev sprememb nadzirajte iskalne strežnike in izberite ikone za zaustavitev in vnovični zagon iskalnih procesov. Ko uporabniki predložijo poizvedbe, bo relativna pomembnost nastalih dokumentov, ki vsebujejo polja, preslikana v ta poviševalni razred, določena z novimi poviševalnimi faktorji.

#### **S tem povezani pojmi**

“Rangiranje dokumentov” na strani 189

## **Privzete vrednosti poviševalnega razreda**

Iskanje v podjetju nudi 16 poviševalnih razredov, ki jih lahko uporabite za vpliv na to, kako so dokumenti uvrščeni v rezultatih iskanja.

Za izračun uvrstitev za polja in besedilo, ki nimajo nobenih drugih definirajočih značilnosti, so za uporabo iskanja v podjetju rezervirana naslednja polja:

```
es_special_field.regular_text
es_special_field.default_field
es_special_field.default_metadata_field
```

Druga polja lahko preslikate v poviševalne razrede, ki jih uporabljajo rezervirana polja, vendar rezerviranih polj ne morete urejati ali izbrisati.

Za vsa druga polja lahko uredite poviševalne faktorje, ki jih uporablja sistem za izračun ranga dokumenta. Kakršnokoli število polj lahko preslikate v katerikoli poviševalni razred, vključno s poviševalnimi razredi, ki jih uporabljajo rezervirana polja.

V naslednji tabeli navajamo imena poviševalnih razredov, privzete poviševalne faktorje za poizvedbe z nizko vrednostjo priklica, privzete poviševalne faktorje za razrede z visoko vrednostjo priklica in imena vnaprej definiranih polj, ki so preslikana v poviševalne razrede v privzeti konfiguraciji.

Privzeti poviševalni faktorji se spreminjajo glede na način statičnega rangiranja, ki ste ga izbrali za zbirko pri njeni izdelavi. Možnosti vključujejo neuporabo statičnega rangiranja, rangiranje, ki je določeno s številom povezav na dokument (za spletne vire) ali rangiranje, ki je določeno z datumom dokumenta.

Tabela 6. Privzete vrednosti poviševalnega razreda

| Privzeti nizki in visoki poviševalni faktorji |                            |                        |                       |                                                                                                                              |
|-----------------------------------------------|----------------------------|------------------------|-----------------------|------------------------------------------------------------------------------------------------------------------------------|
| Ime poviševalnega razreda                     | Brez statičnega rangiranja | Povezave dokumenta     | Datum dokumenta       | Vnaprej definirane preslikave polj                                                                                           |
| Razred vsebine A                              | Nizka: 4<br>Visoka: 2      | Nizka: 6<br>Visoka: 1  | Nizka: 4<br>Visoka: 2 | es_special_field.regular_text                                                                                                |
| Razred vsebine B                              | Nizka: 5<br>Visoka: 4      | Nizka: 7<br>Visoka: 3  | Nizka: 5<br>Visoka: 4 | es_special_field.html_emphasized_text<br><br>Vključuje naslednje elemente HTML: b, big, caption, dfn, em, h4, h5, h6, strong |
| Razred vsebine C                              | Nizka: 7<br>Visoka: 4      | Nizka: 9<br>Visoka: 3  | Nizka: 7<br>Visoka: 4 | es_special_field.html_headers<br><br>Vključuje naslednje elemente HTML: h1, h2, h3                                           |
| Razred vsebine D                              | Nizka: 2<br>Visoka: 5      | Nizka: 1<br>Visoka: 5  | Nizka: 2<br>Visoka: 5 | naziv                                                                                                                        |
| Razred vsebine E                              | Nizka: 1<br>Visoka: 1      | Nizka: 5<br>Visoka: 10 | Nizka: 1<br>Visoka: 1 | es_special_field.anchor                                                                                                      |
| Razred vsebine F                              | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | es_special_field.anchor_same_dir                                                                                             |
| Razred vsebine G                              | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | es_special_field.anchor_same_host                                                                                            |
| Razred vsebine H                              | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | es_special_field.default_field                                                                                               |
| Razred metapodatkov A                         | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | es_special_field.default_metadata_field                                                                                      |
| Razred metapodatkov B                         | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 |                                                                                                                              |
| Razred metapodatkov C                         | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 |                                                                                                                              |
| Razred metapodatkov D                         | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 |                                                                                                                              |
| Razred metapodatkov E                         | Nizka: 1<br>Visoka: 1      | Nizka: 5<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | ključne besede                                                                                                               |
| Razred metapodatkov F                         | Nizka: 1<br>Visoka: 1      | Nizka: 3<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | es_special_field.urlhost                                                                                                     |
| Razred metapodatkov G                         | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | es_special_field.urlpath                                                                                                     |
| Razred metapodatkov H                         | Nizka: 1<br>Visoka: 1      | Nizka: 1<br>Visoka: 1  | Nizka: 1<br>Visoka: 1 | opis                                                                                                                         |

### S tem povezani pojmi

“Rangiranje dokumentov” na strani 189

---

## Iskalne aplikacije za iskanje v podjetju

Iskalna aplikacija vam omogoča, da v sistemu iskanja v podjetju iščete po zbirkah in zunanjih virih . Izdelate lahko neomejeno število iskalnih aplikacij in vsaka iskalna aplikacija lahko išče po neomejenem številu zbirk in zunanjih virov.

### Vzorec iskalne aplikacije

Vzorec iskalne aplikacije prikazuje številne funkcije za iskanje in priklic, ki so na voljo za iskanje v podjetju. Vzorčna aplikacija je tudi delovni zgled, ki prikazuje kako lahko uporabite API za iskanje in stvarno kazalo (SI-API) IBM za izgradnjo interaktivni iskalnih aplikacij po meri, ki odražajo cilje vašega podjetja.

Vzorčna iskalna aplikacija za iskanje v podjetju omogoča, da preiščete vse zbirke in zunanje vire v sistemu iskanja v podjetju, razen če spremenite lastnosti v privzeti konfiguracijski datoteki. Z vzorčno iskalno aplikacijo lahko preskusite nove zbirke in zunanje vire in jih šele nato omogočite za uporabnike.

Vzorčna iskalna aplikacija je samodejno povezana z vsemi zbirkami in zunanjimi viri. V delovnem okolju skrbniki za iskanje v podjetju krmilijo, katere iskalne aplikacije lahko preiščejo različne zbirke.

### Iskalna aplikacija po meri

Iskalne aplikacije, ki jih izdelate lahko izvajate kot samostojne spletne aplikacije v okolju aplikacijskega strežnika WebSphere Application Server ali pa jih zaženete kot portalske programčke v okolju IBM WebSphere Portal. Z API-jem za iskanje in stvarno kazalo lahko izdelate iskalne aplikacije, ki brez težav delujejo kot vzorčne iskalne aplikacije v obeh okoljih.

Prilagajanje iskalnih aplikacij si lahko olajšate s Prilagojevalnikom iskalnih aplikacij. Ta aplikacija vam omogoča, da izbiro opravite v grafičnem vmesniku in si sproti ogledate učinek sprememb. Ko shranite spremembe, posodobite konfiguracijske datoteke iskalne aplikacije.

#### Nasvet:

Podrobnejše informacije o možnostih uporabe Prilagojevalnika iskalnih aplikacij in načinih namestitve iskalne aplikacije kot portalski programček v portalu WebSphere Portal poiščite v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

#### S tem povezani pojmi

“Možnosti indeksiranja za iskanje dokumentov” na strani 163

“Identifikatorji iskalne aplikacije” na strani 237

“Zaščita na ravni dokumenta” na strani 238



Pregled API-ja za iskanje in indeksiranje



Skladnja poizvedbe



Nastavljanje lastnosti poizvedbe

---

## Povezovanje iskalnih aplikacij z zbirkami

Preden lahko novo iskalno aplikacijo začnete uporabljati, jo morate povezati z zbirko po kateri lahko išče.

### Preden začnete

Zbirke lahko z iskalnimi aplikacijami povezujejo samo člani skrbniške vloge za iskanje v podjetju.

### Postopek

Povezovanje iskalne aplikacije z eno ali več zbirkami:

1. V orodni vrstici skrbniške konzole kliknite **Zaščita**.
2. Na strani Iskanje kliknite **Konfiguriranje iskalnih aplikacij**.
3. Na strani Konfiguriranje iskalnih aplikacij kliknite **Dodajanje iskalne aplikacije**.
4. Vnesite ime iskalne aplikacije.
5. Izberite zbirke, po katerih aplikacija lahko išče.
  - Kliknite **Vse zbirke in zunanji viri**, če želite, da bo iskalna aplikacija imela dostop do vseh zbirk, ki ste jih dodali sistemu.
  - Kliknite **Specifične zbirke in zunanji viri**, če želite, da bo iskalna aplikacija imela dostop samo do zbirk, ki jih podate.

Ko izberete to možnost, se prikaže seznam z imeni zbirk in zunanjih virov. Za vsako zbirko, po kateri zbirka lahko išče, izberite potrditveno polje **Izberi**.
6. Kliknite **Potrdi**.

---

## Funkcije vzorčne iskalne aplikacije

Vzorčna iskalna aplikacija za iskanje v podjetju prikazuje večino iskalnih funkcij, ki jih lahko vgradite v iskalne aplikacije po meri.

Z vzorčno iskalno aplikacijo lahko istočasno preiskujete vse zbirke in zunanje vire. S to aplikacijo lahko preiskujete vse zbirke in zunanje vire v sistemu iskanja v podjetju, razen če so bile spremenjene privzete lastnosti aplikacije.

### Poizvedbene funkcije

Te funkcije vam omogočajo naslednje:

- Podajanje preprostih poizvedb s prostim besedilom.
- Podajanje bolj zapletenih poizvedb, ki omogočajo bolj natančne rezultate iskanja. Tako lahko na primer iščete v specifičnih poljih ali elementih XML ali uporabite poizvedbeno skladnjo, če želite poiskati dokumente, ki vsebujejo oziroma ne vsebujejo določenih besed ali besednih zvez.
- Podajanje zbirk in zunanjih virov v katerih želite iskati.
- Iskanje v specifičnih vrstah virov ali v vseh virih.
- Iskanje v specifičnih vrstah dokumentov. Tako lahko na primer iščete samo v dokumentih Microsoft Word ali samo v dokumentih PDF (prenosljivi format dokumentov).
- Podajanje v katerem jeziku so vaši poizvedbeni izrazi. Podate lahko tudi jezik dokumentov, ki jih želite preiskati.
- Iskanje v specifičnih podnizih zbirke. Tako lahko na primer iskalna aplikacija omeji prikazane dokumente na preddefiniran obseg dokumentov (območje) ali pa lahko predložite poizvedbo, ki preišče samo dokumente, ki pripadajo imenovani kategoriji.

- Razširitev poizvedbe tako, da vključuje sopomenke poizvedbenih izrazov. Če je z zbirko povezan slovar sopomenk, so dokumenti, ki vsebujejo sopomenke izrazov poizvedbe, vrnjeni v rezultatih iskanja.

## Funkcije rezultatov iskanja

Te funkcije vam omogočajo naslednje:

- Prikaz rezultatov iskanja, ki ustrezajo vašim poizvedbenim izrazom.
- Nadzor nad tem koliko najdenih dokumentov je prikazanih na vsaki strani in pregledovanje niza rezultatov s funkcijama naprej in nazaj.
- Skrivanje in prikaz podrobnosti najdenih dokumentov. Tako lahko na primer v vsakem najdenem dokumentu prikažete kratke opise dokumentov ali podrobnosti, kot so imena polj.
- Skrčevanje dokumentov iz istega vira. Če je na primer iz enega vira vrnjenih 100 dokumentov, bosta v nizu rezultatov prikazana samo dva najbolj pomembna dokumenta, ki sta prikazana skupaj. Ostalih 98 dokumentov si lahko ogledate tako, da izberete možnost, ki vam omogoča prikaz več dokumentov iz istega vira.
- Razvrščanje dokumentov po pomembnosti, datumu ali vrednostih v posameznem polju. Pri razvrščanju po datumu ali polju lahko podate ali želite rezultate prikazati v naraščajočem ali padajočem vrstnem redu.
- Dajanje predlogov za popravke črkovanja, če program najde v vašem poizvedbenem nizu besede, ki bi lahko bile narobe črkovane.
- Prikaz informacij o kategorijah, katerim pripada najdeni dokument (če zbirka uporablja kategorije) in pregledovanje samo dokumentov, ki spadajo v specifično kategorijo.
- Podajanje dodatnih poizvedbenih izrazov za iskanje v rezultatih iskanja.

## Funkcije za priklic dokumentov

Te funkcije vam omogočajo naslednje:

- Priklic dokumentov s klikom na URI dokumenta in odpiranje dokumenta v spletnem brskalniku. Če je pajek Notes ali pajek Domino Document Manager konfiguriran za uporabo protokola DIIOP, potem je mogoče dokumente, ki jih pajkajo ti pajki, prikazati v odjemalski prikazovalni aplikaciji Lotus Notes, namesto v spletnem brskalniku.  
Če je za pajka konfigurirana zaščita za raven dokumentov, potem lahko dokumente priključijo samo uporabniki, ki so pooblašeni za dostop do zaščitene vsebine.
- Priklic dokumentov s klikom na hitre povezave. Hitra povezava poveže ključne besede z URI-ji. Če poizvedba vsebuje podane ključne besede, se povezani URI-ji (ki so bili vnaprej določeni kot posebej pomembni za te ključne besede) pojavijo na vrhu rezultatov iskanja.

---

## Lastnosti iskalnih aplikacij

Konfiguracijsko datoteko za iskalne aplikacije lahko uredite tako, da podate možnosti za svoje okolje, spremenite videz aplikacije in krmilite možnosti, ki so uporabnikom na voljo po zagonu iskalne aplikacije.

Lastnosti lahko uredite tudi s Prilagojevalnikom iskalnih aplikacij. Pri urejanju lastnosti s Prilagojevalnikom lahko takoj vidite učinek sprememb. Ko ste zadovoljni z možnostmi, ki ste jih podali za iskanje po zbirkah in prikaz rezultatov iskanja, lahko možnosti shranite in s tem posodobite konfiguracijsko datoteko za iskalno aplikacijo.

**Pomembno:** Če zaženete iskalno aplikacijo kot portalski programček s portala WebSphere Portal, ne morete interaktivno spreminjati iskalne aplikacije s prilagojevalnikom iskalnih

aplikacij. Urediti morate lastnosti in konfigurirati primerek portalskega programčka z vmesnikom za skrbništvo nad portalom WebSphere Portal.

Konfiguracijska datoteka za vzorčno iskalno aplikacijo sistema iskanja v podjetju je datoteka `config.properties`. Ta tema pojasnjuje lastnosti v tej datoteki in opisuje privzete lastnosti. Če izdelate konfiguracijske datoteke za svoje prilagojene iskalne aplikacije, bodo lastnosti v teh datotekah in vrednosti, podane za te lastnosti, morda drugačne.

## Parametri okolja

Podate lahko možnosti, ki krmilijo delovanje iskalnih aplikacij.

### **applicationName**

Podaja ime veljavne iskalne aplikacije. Privzeta vrednost je `Default`.

Privzeto vrednost spremenite, če želite kot privzeto aplikacijo uporabiti drugo iskalno aplikacijo.

**Nasvet:** Če je ime aplikacije enako `Default`, lahko za iskanje po vseh zbirkah in zunanjih virih z eno poizvedbo uporabite vzorčno iskalno aplikacijo.

### **timeout**

Podaja čas čakanja, v sekundah, na odziv strežnika za iskanje, preden zahteva za iskanje poteče. To število mora biti celo število (kot je 60 in ne 60.5 ali šestdeset). Če ne podate vrednosti čakalnega časa, bo privzeta vrednost 30 sekund.

### **hostname**

Podaja popolnoma kvalificirano ime gostitelja spletnega strežnika, ki ste ga konfigurirali za primerek WebSphere Application Server. Privzeta vrednost je `localhost`.

Če se želite prepričati, da iskalna aplikacija pravilno deluje, privzeto vrednost spremenite v popolnoma kvalificirano ime gostitelja, konfiguriranega za program WebSphere Application Server. Če je na primer ime gostitelja lokalnega računalnika `MojRacunalknik`, ime gostitelja spletnega strežnika pa je `www.ibm.com`, podajte `www.ibm.com`.

### **protocol**

Podaja protokol za komuniciranje s spletnim strežnikom: `http` ali `https`. Če to pustite prazno, bo privzeta vrednost `http`.

### **port**

Podaja številko vrat spletnega strežnika, ki ste jo konfigurirali za primerek WebSphere Application Server. Privzeta vrednost je `80`, ki je običajna za protokol HTTP. Običajna vrata, ki jih uporablja protokol HTTPS, pa so `443`.

### **trustStore**

Če uporabljate protokol HTTPS, podajte popolnoma kvalificirano pot do datoteke shrambe ključev (datoteka baze podatkov, ki vsebuje javne ključe). Te informacije, imenovane tudi *overjena shramba*, omogočajo, da za overjene komunikacije uporabite protokol SSL (Secure Sockets Layer). Če želite podati pot Windows, poševnico nazaj vnesite kot dvojno poševnico nazaj. Primer: `x:\\Application Server\\webserver.key`

### **trustPassword**

Če uporabljate protokol HTTPS, podajte geslo za podano datoteko shrambe ključev.

### **username**

Iskalna aplikacija bo to vrednost samodejno nastavila na ime uporabnika, ki ga uporabnik poda pri prijavi v iskalno aplikacijo. Ime uporabnika tukaj podajte le v

primeru, če želite nadomestiti privzeto vedenje pri overjanju uporabnikov. To polje bo uporabljeno le v primeru, če ste v programu WebSphere Application Server omogočili globalno zaščito.

**password**

Iskalna aplikacija bo to vrednost samodejno nastavila na geslo, ki ga uporabnik poda pri prijavi v iskalno aplikacijo. Geslo tukaj podajte le v primeru, če ste podali ime uporabnika. To polje bo uporabljeno le v primeru, če ste v programu WebSphere Application Server omogočili globalno zaščito.

**ssoCookieName**

Podaja ime piškotka, ki vsebuje niz žetona enotne prijave (SSO). Privzeta vrednost je LtpaToken.

**proxyHost**

Podaja popolnoma kvalificirano ime gostitelja strežnika proxy, če le-tega potrebujete za dostop do iskalnega strežnika.

**proxyPort**

Podaja vrata za podanega gostitelja strežnika proxy.

**proxyUser**

Podaja ime uporabnika za prijavo na strežnik proxy, če ta zahteva osnovno overjanje.

**proxyPassword**

Podaja geslo za podano ime uporabnika strežnika proxy.

**filter**

Podaja razred, namenjen priklicu dokumentov, navedenih v rezultatih iskanja. Privzeti razred je `com.ibm.es.api.filters.SetDocumentURIFilterFetch`. To vrednost spremenite le v primeru, če imate svoj prilagojeni razred, ki ga želite uporabiti za priklic dokumentov.

**logging.level**

Podaja količino zabeleženih podrobnosti:

**OFF** Sistem ne bo zabeležil nobenega sporočila.

**SEVERE**

Sistem bo zabeležil sporočila, ki označujejo resno napako. To je privzeta vrednost.

**INFO** Zabeležena bodo informativna sporočila.

**FINE** Zabeležena bodo sporočila sledenja z malo podrobnostmi. (Ta možnost ustreza ravni beleženja FINE v razredu Java `java.util.logging.Level`.)

**ALL** Zabeležena bodo vsa sporočila.

## Ikone za tipe virov

Slike, ki predstavljajo tipe virov podatkov, ki jim pripadajo dokumenti v rezultatih iskanja, lahko prilagodite. V datoteki `config.properties` so vnaprej definirane naslednje ikone tipov virov, ki označujejo pajke in zunanje vire, podprte ob namestitvi programa OmniFind Enterprise Edition.

**documentSource.vbr.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom Content Edition. Privzeta ikona je `/images/sourceVBR.gif`.

**documentSource.db2.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom DB2. Privzeta ikona je `/images/sourceDB2.gif`.

**documentSource.cm.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom DB2 Content Manager. Privzeta ikona je /images/sourceCM.gif.

**documentSource.dominodoc.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom Domino Document Manager. Privzeta ikona je /images/sourceDominoDoc.gif.

**documentSource.exchange.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom Exchange Server. Privzeta ikona je /images/sourceExchange.gif.

**documentSource.database.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom baze podatkov JDBC. Privzeta ikona je /images/sourceJDBC.gif.

**documentSource.nntp.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom NNTP. Privzeta ikona je /images/sourceNNTP.gif.

**documentSource.notes.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom Notes. Privzeta ikona je /images/sourceNotes.gif.

**documentSource.quickplace.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom QuickPlace. Privzeta ikona je /images/sourceWorkplace.gif.

**documentSource.seedlist.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom seznama semen. The default icon is /images/sourceSeedlist.gif.

**documentSource.unixfs.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom datotečnega sistema UNIX. Privzeta ikona je /images/sourceUnixFS.gif.

**documentSource.web.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s spletnim pajkom. Privzeta ikona je /images/sourceWeb.gif.

**documentSource.wcm.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom Web Content Management. Privzeta ikona je /images/sourceWorkplace.gif.

**documentSource.wps.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom WebSphere Portal. Privzeta ikona je /images/sourceWPS.gif.

**documentSource.winfs.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da je bil dokument pajkan s pajkom datotečnega sistema Windows. Privzeta ikona je /images/sourceWindowsFS.gif.





### **documentSource.ldap.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da dokument pripada zunanjemu viru, ki je bil izdelan za strežnik LDAP. Privzeta ikona je /images/sourceLDAP.gif.



### **documentSource.jdbc.icon**

Podaja pot in ime slikovne datoteke, ki označuje, da dokument pripada zunanjemu viru, ki je bil izdelan za tabelo baze podatkov JDBC (Java Database Connectivity). Privzeta ikona je /images/sourceJDBC.gif.

## **Ikone odjemalskega prikazovalnika**

Dokumente rezultatov iskanja lahko prikažete v spletnem brskalniku. Dokumente, ki ste jih pajkali s pajki Notes ali pajki Domino Document Manager, konfiguriranimi za protokol DIIOP, lahko prikažete tudi v odjemalski aplikaciji prikazovalnika Lotus Notes.

Če želite dokumente prikazati v odjemalski aplikaciji prikazovalnika, zagotovite, da je naslednja lastnost nastavljena na true:

```
clientViewer.show=true
```

Slike, ki predstavljajo odjemalsko aplikacijo prikazovalnika, lahko prilagodite. V naslednjem primeru ikona Lotus Notes označuje, da dokument lahko prikažete z aplikacijo prikazovalnika:

```
client.notes.icon=/images/notes.gif
client.dominodoc.icon=/images/notes.gif
```

Ikona in povezava na odjemalsko aplikacijo prikazovalnika bosta v rezultatih iskanja prikazani na naslednji način:



### **Odjemalski prikazovalnik**

## **Polja dokumenta**

Pri tipih virov podatkov, ki vsebujejo polja, lahko krmilite, katera polja bodo prikazana v dokumentih rezultatov iskanja.

### **fields.URI prefix=*s\_presledkom\_ločen\_seznam\_imen\_polj***

Pred znak dvopičja (:) v predponi URI morate vnesti poševnico nazaj (\). Če želite seznam imen polj nadaljevati v naslednji vrstici, prejšnjo vrstico zaključite s poševnico nazaj (\). Primer:

```
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\://=
```

## Ikone polj

Pri tipih virov podatkov in dokumentih, ki vsebujejo polja, lahko prilagodite slike, ki predstavljajo polja. Vsa polja nad povzetkom dokumenta vsebujejo unikatno sliko. V datoteki `config.properties` so vnaprej definirane naslednje ikone polj.



### **field.icon.databasetitle**

Podaja pot in ime slikovne datoteke, ki označuje, da polje vsebuje naziv dokumenta. Privzeta ikona je `/images/notesdb.gif`.



### **field.icon.databasename**

Podaja pot in ime slikovne datoteke, ki označuje, da polje vsebuje ime baze podatkov, ki ji dokument pripada. Privzeta ikona je `/images/db2.gif`.



### **field.icon.tablename**

Podaja pot in ime slikovne datoteke, ki označuje, da polje vsebuje ime tabele, ki ji dokument pripada. Privzeta ikona je `/images/table.gif`.



### **field.icon.directory**

Podaja pot in ime slikovne datoteke, ki označuje, da polje vsebuje ime imenika, ki mu dokument pripada. Privzeta ikona je `/images/closedFolder.gif`.



### **field.icon.filename**

Podaja pot in ime slikovne datoteke, ki označuje, da polje vsebuje ime datoteke dokumenta. Privzeta ikona je `/images/document.gif`.

### **field.icon.documentID**

Podaja pot in ime slikovne datoteke, ki označuje, da polje vsebuje identifikator dokumenta. Za spletne dokumente je morda bolje, da uporabite prazno sliko in podate sliko za URL, slike pa ne prikažete uporabniku. Privzeta ikona je `/images/dot.gif`.



### **field.icon.group**

Podaja pot in ime slikovne datoteke, ki označuje, da polje vsebuje identifikator dokumenta. Za spletne dokumente je morda bolje, da uporabite prazno sliko in podate sliko za URL, slike pa ne prikažete uporabniku. Privzeta ikona je `/images/document.gif`.



### **field.icon.from**

Podaja pot in ime slikovne datoteke, ki označuje, da polje označuje osebo, ki je poslala dokument. Privzeta ikona je `/images/author.gif`.



### **field.icon.creator**

Podaja pot in ime slikovne datoteke, ki označuje, da polje označuje izdelovalca dokumenta. Privzeta ikona je `/images/author.gif`.



### **field.icon.author**

Podaja pot in ime slikovne datoteke, ki označuje, da polje označuje avtorja dokumenta. Privzeta ikona je `/images/author.gif`.



### **field.icon.revisionuser**

Podaja pot in ime slikovne datoteke, ki označuje, da polje označuje osebo, ki je pregledala dokument. Privzeta ikona je `/images/author.gif`.

**field.icon.owner**

Podaja pot in ime slikovne datoteke, ki označuje, da polje označuje lastnika dokumenta. Privzeta ikona je `/images/author.gif`.

**field.icon.modifier**

Podaja pot in ime slikovne datoteke, ki označuje, da polje označuje osebo, ki je spremenila dokument. Privzeta ikona je `/images/author.gif`.

## Privzete ikone za polja

Tukaj lahko podate sliko, ki jo bo sistem uporabil v primeru, če za polja, prikazana v rezultatih iskanja, ne konfigurirate nobene ikone. V datoteki `config.properties` je vnaprej definirana naslednja privzeta ikona polja.

**field.defaultIcon**

Podaja pot in ime slikovne datoteke, ki bo privzeta ikona za polja v rezultatih iskanja. Privzeta ikona je `/images/database.gif`.

## Datumska polja

Tukaj lahko podate, katera polja so datumska polja. Imena polj, ki jih podate tukaj, bodo v rezultatih iskanja oblikovana kot datumski podatki. Format datumov se ujema s področnimi nastavitvami v spletnem brskalniku.

**date.fields=s\_presledkom\_ločen\_seznam\_imen\_polj**

Če želite seznam imen polj nadaljevati v naslednji vrstici, prejšnjo vrstico zaključite s poševnico nazaj (`\`).

Zgled:

```
date.fields=modifieddate createddate
```

## Nazivi dokumentov

Za dokumente lahko podate nadomestne nazive in s tem besedilo naziva zamenjate s pomenljivejšimi podatki (kar pomeni, da lahko nazive *počistite*). Namesto prikaza nazivov dokumentov z neinformativno oznako Prosojnica 1, lahko na primer podate, da bo ta oznaka v rezultatih iskanja zadržana. Dobljeni dokument lahko označite tudi z bolj smiselnim poljem, na primer datotečnim imenom.

Prav tako lahko za dokumente podate nadomestne nazive tako, da odstranite besede brez pomena v nazivu dokumenta (kar pomeni, da lahko *okrnite* nazive). Če se na primer večje število dokumentov začne z Microsoft Word -, lahko berljivost rezultatov iskanja povečate tako, da zadržite ponavljajoče se začetno besedilo.

**titles.clean=z\_vejico\_ločen\_seznam\_naslovov****titles.truncatePrefix=z\_vejico\_ločen\_seznam\_predpon**

Z vejico ločeni sezname lahko vsebujejo presledke in druge znake, ne pa tudi vejice.

Če želite seznam nadaljevati v naslednji vrstici, prejšnjo vrstico zaključite s poševnico nazaj (`\`).

Primer:

```
titles.clean=Slide 1, Layout 1, untitled, \
Untitled Document, PowerPoint Presentation, \
(ta stran nima naziva)
```

```
titles.truncatePrefix=Microsoft Word -, Microsoft Powerpoint -
```

## Privzete vrednosti za uporabniške nastavitve

Za stran Nastavitve v iskalni aplikaciji lahko podate privzete vrednosti. Če uporabnik spremeni nastavitve, bodo nove vrednosti veljale samo v trenutni seji uporabnika. V datoteki `config.properties` so vnaprej definirane naslednje nastavitve.

### **preferences.resultsRange=10**

Podaja, da lahko vsaka stran v rezultatih iskanja vsebuje 10 dokumentov.

### **preferences.siteCollapsing=Yes**

Podaja, da bodo URI-ji iz istega vira v rezultatih iskanja skrčeni. Krčenje spletnih mest je na voljo le, če so rezultati razvrščeni po pomembnosti. Pri spletnih virih in virih podatkov NNTP bodo URI-ji, ki ustrezajo URI-ju osnovnega spletnega mesta (kot je `www.ibm.com`) skrčeni samodejno. Pri drugih tipih virov podatkov in spletnih mestih z globljimi ravnmi poti (kot je `www.ibm.com/hr`), morate pravila za krčenje spletnih mest konfigurirati na skrbniški konzoli sistema iskanja v podjetju.

### **preferences.spellCorrections=Yes**

Podaja, da bodo v primeru, ko uporabnik predloži poizvedbo, ki vsebuje možno napačno črkovano besedo, prikazani predlogi za popravke črkovanja. Pomnite, da bo sistem pred izračunavanjem predlogov za črkovanje vedno odstranil izkjučitvene besede.

### **preferences.extendedHighlighting=No**

Podaja, da izrazi poizvedbe v dodatnih poljih (kot je naziv dokumenta), prav tako pa tudi v polju s povzetkom dokumenta ne bodo označeni.

## Privzete zbirke in zunanji viri

Tukaj lahko podate, katere zbirke in zunanji viri bodo na straneh Nastavitve in Zahtevnejše iskanje vnaprej izbrani. Uporabniki lahko privzeti niz uredijo tako, da bodo iskali po manjšem številu zbirk in zunanjih virov, kot so na voljo po privzetku. Če tukaj omejite niz zbirk in zunanjih virov, lahko uporabniki pri spreminjanju svojih nastavitvev in možnosti zahtevnejšega iskanja izberejo poljubno zbirko ali zunanji vir, ki je na voljo iskalni aplikaciji.

### **preferences.defaultCollections=\***

**preferences.defaultCollections=space\_separated\_list\_of\_collection\_IDs** (s presledki ločen seznam ID-jev zbirk)

Podajte zvezdico (\*), če želite omogočiti iskanje po vseh zbirkah in zunanjih virih. (Zbirke in zunanje vire morate na skrbniški konzoli sistema iskanja v podjetju povezati z iskalno aplikacijo.) To je privzeta nastavitvev v datoteki `config.properties`.

Če želite omejiti, kaj bodo uporabniki lahko iskali, če ne spremenijo svojih nastavitvev ali možnosti zahtevnejšega iskanja, podajte ID-je zbirk in zunanjih virov, po katerih lahko uporabniki po privzetku iščejo.

Primer:

```
preferences.defaultCollections=*\npreferences.defaultCollections=coll_id1 coll_id2
```

## Dodatne informacije o rezultatih iskanja

Tukaj lahko omejite količino informacij, navedenih v rezultatih iskanja, obenem pa krmilite, ali lahko uporabniki filtrirajo rezultate iskanja. V datoteki `config.properties` so privzete naslednje nastavitve.

### **refreshButton.show=false**

Nadzira, ali je na strani osnovnega iskanja prikazan gumb **Osveži**. Gumb **Osveži** bo

vedno na voljo pri zahtevnejših iskanjih. Če to možnost nastavite na true, bodo uporabniki vedno lahko osvežili seznam zbirk in zunanjih virov, ki so na voljo za iskanje.

Če uporabljate Prilagojevalnik iskalnih aplikacij, ne potrebujete gumba **Osveži**.

Če ne uporabljate Prilagojevalnika iskalnih aplikacij, boste gumb **Osveži** morda želeli prikazati za preskus sprememb v konfiguracijski datoteki. Ko shranite spremembe, lahko kliknete gumb **Osveži** in tako vidite, kako spremembe vplivajo na iskalno aplikacijo. Če nimate gumba **Osveži**, morate aplikacijo ESSearchServer na aplikacijskem strežniku WebSphere Application Server znova zagnati in s tem uveljaviti spremembe.

Če za iskanje ni na voljo nobene zbirke ali zunanjega vira (če ste na primer podali napačno ime gostitelja, iskalni strežniki niso zagnani ali pa aplikacija ESSearchServer na aplikacijskem strežniku WebSphere Application Server ni bila zagnana), bo gumb **Osveži** prikazan samodejno in vam bo tako v pomoč pri odpravljanju težav.

#### **builtQueryString.show=false**

Nadzira prikaz skladnje celotne razširjene poizvedbe v področju, prikazanem pred seznamom dokumentov rezultatov iskanja. To možnost nastavite na true, če želite prikazati dejansko obdelano poizvedbo.

#### **extraQueryData.show=false**

Nadzira prikaz dodatnih informacij o poizvedbi. To možnost nastavite na true, če želite prikazati informacije o omejitvah ACL, imenih zbirk in zunanjih virov za iskanje in jeziku poizvedbe.

#### **refineResults.show=true**

Nadzira, ali lahko uporabniki rezultate iskanja zožijo tako, da podajo dodatne izraze poizvedbe. Če to možnost nastavite na true, bo v spodnjem delu strani z rezultati iskanja prikazano okence poizvedbe z oznako **Išči po rezultatih**.

#### **sorting.show=true**

Nadzira, ali bo prikazana možnost za razvrščanje rezultatov iskanja. To možnost nastavite na false, če želite zadržati možnosti **Razvrsti po** in **Vrstni red razvrščanja** za razvrščanje rezultatov iskanja.

#### **sourceTypeFilter.show=true**

Nadzira, ali bo v rezultatih iskanja prikazana možnost za filtriranje rezultatov po tipih virov. To možnost nastavite na false, če ne želite, da bi uporabniki rezultate filtrirali po tipu vira.

Če želite uporabnikom preprečiti, da bi rezultate filtrirali po tipu dokumenta, izbrišite izbrane ali vse vnose za tipe dokumentov v konfiguracijski datoteki (`documentType.oznaka=tipi_dokumentov`).

#### **filter.showOnTwoLines=true**

Nadzira, ali bodo možnosti za filtriranje rezultatov po tipu vira in filtriranje rezultatov po tipu datoteke v rezultatih iskanja prikazane v eni ali dveh vrsticah. Uporabniki lahko pri pregledovanju rezultatov iskanja izberejo tip vira in nato še tip datoteke ter tako prikažejo samo dokumente, ki se ujemajo z izbranimi filtri.

Če želite povečati količino prostora, ki je na voljo za prikaz rezultatov iskanja, to lastnost nastavite na false. Če želite izboljšati berljivost filtrov, posebej v primeru, ko razpoložljivi filtri presegajo eno vrstico, lahko to lastnost nastavite na true, tako da bo vsak filter prikazan v svoji vrstici.

#### **clientViewer.show=true**

Nadzira, ali bo za prikaz dokumentov rezultatov iskanja uporabljena odjemalska

aplikacija prikazovalnika Lotus Notes. To možnost nastavite na false, če za prikaz dokumentov Domino ne želite uporabiti aplikacije prikazovalnika.

**showDetails.show=true**

Nadzira prikaz povezav Prikaži podrobnosti in Skrij podrobnosti v rezultatih iskanja. To možnost nastavite na false, če ne želite, da bi uporabniki prikazali dodatne podrobnosti o dokumentih rezultatov iskanja.

**showDetailsImage.show=true**

Nadzira prikaz podrobnosti o dokumentih rezultatov iskanja v oknu. To možnost nastavite na false, če ne želite, da bi uporabniki s postavitvijo kazalke na URI dokumenta prikazali dodatne podrobnosti o dokumentih rezultatov iskanja.

**numberSearchResultsReturned.show=true**

Nadzira, ali bo prikazano skupno število rezultatov iskanja. To možnost nastavite na false, če ne želite, da bi uporabniki videli, koliko dokumentov je sistem vrnil v rezultatih iskanja.

**showMessage.error=true**

Nadzira prikaz sporočil o napakah. To možnost nastavite na false, če na vrhu iskalne aplikacije ne želite prikazati sporočil o napakah.

**showMessage.warning=true**

Nadzira prikaz opozorilnih sporočil. To možnost nastavite na false, če na vrhu iskalne aplikacije ne želite prikazati opozorilnih sporočil.

**showMessage.info=true**

Nadzira prikaz informativnih sporočil. To možnost nastavite na false, če na vrhu iskalne aplikacije ne želite prikazati informativnih sporočil.

**showMessage.success=true**

Nadzira prikaz sporočil o uspehu. To možnost nastavite na false, če na vrhu strani iskalne aplikacije ne želite prikazati sporočil, ki označujejo uspešen zaključek dejanja.

## Prilagojena trak in logotip

Tukaj lahko prilagodite slike, ki se prikažejo v področju traku na vrhu iskalne aplikacije. Privzete slike za OmniFind Enterprise Edition lahko na primer nadomestite s slikami, ki odražajo blagovno znamko vašega podjetja. Če traku ne želite prikazati, lahko eno ali obe vrstici izločite iz kode. Lastnost banner.icon označuje grafiko, ki bo prikazana na levi strani področja traku. Lastnost banner2.icon označuje grafiko, ki bo prikazana na desni strani področja traku.

```
banner.icon=/images/WS_II_0FEdition.gif
banner2.icon=/images/WS_II_mosaic.gif
```

## Prilagojena slika ozadja

Tukaj lahko prilagodite slike, ki so v iskalni aplikaciji prikazane kot ozadje strani. Privzete slike sistema iskanja v podjetju lahko na primer nadomestite s slikami, ki odražajo blagovno znamko vašega podjetja. Če na strani ne želite prikazati ozadja, lahko eno ali več izmed naslednjih vrstic izločite iz kode.

```
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
```

## Povezave

Lastnosti v področju Povezave v datoteki `config.properties` omogočajo, da imena strani iskalne aplikacije na posamezni strani prikažete kot povezave in ne na orodjarni in na straneh z zavihki. Prikaz povezav je uporaben v primeru, če želite iskalno aplikacijo zagnati kot portalski programček in želite zmanjšati količino prostora, namenjenega prikazu iskalne aplikacije na strani portala.

Če se po iskalni aplikaciji raje pomikate z izbiro možnosti na orodjarni in prek strani z zavihki, te vrstice izločite iz kode.

## Zavihki iskanja

Lastnosti v področju Zavihki iskanja v datoteki `config.properties` podajajo imena JSP-jev (Java Server Pages), namenjenih za strani z zavihki na pogledu Iskanja iskalne aplikacije (Osnovno iskanje, Zahtevnejše iskanje in Drevo kategorij). Teh strani ne urejajte, če nimate izkušenj s programiranjem v jeziku Java in z JSP-ji.

Zgledi prilagoditve tega področja vključujejo naslednje:

- Preusmeritev iskalne aplikacije na prilagojene JSP-je, ki nudijo drug videz strani z zavihki.
- Komentiranje vnosov Drevesa kategorij. Če na primer za zbirke ne želite konfigurirati kategorij, vam v iskalni aplikaciji ni treba prikazati strani Drevo kategorij.
- Kopiranje vnosov za strani z zavihki v področje Orodjarne v datoteki `config.properties` in komentiranje teh vrstic. Prikažete lahko na primer samo orodjarno, strani z zavihki pa sploh ne.

## Orodjarne

Lastnosti v področju Orodjarne v datoteki `config.properties` podajajo imena JSP-jev (Java Server Pages), namenjenih za orodjarno v iskalni aplikaciji. Teh strani ne urejajte, če nimate izkušenj s programiranjem v jeziku Java in z JSP-ji.

Zgledi prilagoditve tega področja vključujejo naslednje:

- Preusmeritev iskalne aplikacije na prilagojene JSP-je, ki nudijo drug videz orodjarne.
- Komentiranje vnosov orodjarne za postavke, ki jih ne želite prikazati. V orodjarno morda ne boste na primer želeli vključiti povezave na stran O programu.
- Premik funkcije za prikaz strani Zahtevnejše iskanje iz področja zavihka v datoteki `config.properties` tako, da bo ta možnost na voljo samo v orodjarni.

## Pomenljive oznake za tipe dokumentov

Berljivost filtra za tipe dokumentov lahko izboljšate tako, da dejanska imena tipov dokumentov preslikate v krajše izraze, ki več pomenijo. Tipe dokumentov, ki so na voljo iskalnim aplikacijam, definira razred `AvailableDocumentTypes` API-ja za iskanje in indeksiranje (SI-API). Razpoložljivi tipi dokumentov so zaradi pripravnosti prav tako navedeni na koncu datoteke `config.properties`.

### **documentType.oznaka=s\_presledki\_ločen\_seznam\_tipov\_dokumentov**

Podaja ime, ki bo v rezultatih iskanja prikazano v vrstici s filtrom za tipe dokumentov, in seznam dejanskih tipov dokumentov, ki bodo prikazani, ko uporabnik izbere filter.

Podate lahko na primer oznako `html` in nato vanjo preslikate pripone datotek in tipe MIME različnih spletnih dokumentov. Ko uporabnik želi filtrirati rezultate iskanja in klikne `html`, bodo prikazani samo dokumenti s podanimi priponami in tipi MIME.

V datoteki `config.properties` so vnaprej definirane naslednje preslikave tipov dokumentov:

```
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
```

Če se vrednost, ki jo podate za oznako tipa dokumenta, ujema z imenom lastnosti v datoteki `application.properties`, se prikaže vrednost za lastnost v datoteki `application.properties` in ne vrednost, ki jo podate na tem mestu. Če na primer kot oznako za filter tipa datotek podate `documentType.unixfs`, se vrednost za lastnost `unixfs` v datoteki `application.properties` (**datotečni sistem UNIX**) prikaže kot ime filtra tipa datotek, ki ga lahko kliknete.

## Prilagojeni filtri

Tukaj lahko podate prilagojene poizvedbe, s katerimi filtrirate prikaz dokumentov rezultatov iskanja.

### **filterCustom.oznaka=izrazi\_poizvedbe**

Podaja ime, ki bo v rezultatih iskanja prikazano v vrstici s prilagojenim filtrom, in poizvedbo, ki v primeru, ko uporabnik izbere filter, zoži rezultate iskanja. (Uporabniki lahko pri pregledovanju rezultatov iskanja izberejo prilagojeni filter in tako prikažejo samo dokumente rezultatov iskanja, ki ustrezajo vnaprej definirani poizvedbi.)

Rezultati iskanja so v naslednjem zgledu filtrirani tako, da prikazujejo samo dokumente, ki pripadajo bazi podatkov oddelka za ravnanje s človeškimi viri (hr):  
`filterCustom.samo_baza_podatkov_HR=databasename::hr`

Ko uporabnik klikne filter **samo\_baza\_podatkov\_HR** in s tem filtrira rezultate iskanja, bo sistem obdelal poizvedbo `databasename::hr`. Pri prikazu rezultatov iskanja bodo navedeni samo dokumenti iz baze podatkov hr.

Če se vrednost, ki jo podate za oznako prilagojenega filtra, ujema z imenom lastnosti v datoteki `application.properties`, se prikaže vrednost za lastnost v datoteki `application.properties` in ne vrednost, ki jo podate na tem mestu. Če na primer kot oznako filtra po meri podate `filterCustom.hostData=`, se vrednost lastnosti `hostData` v datoteki `application.properties` (**Nastavitve strežnika**) prikažejo kot ime prilagojenega filtra, ki ga lahko kliknete.

Nekateri prilagojeni filtri so izločeni iz kode in so v datoteki `config.properties` na voljo samo kot zgledi.

## Odkrivanje dvojnikov

Če dodate dokumente v indeks za iskanje v podjetju, se z analizo odstranijo dvojniki, tako da uporabniki v rezultatih iskanja ne vidijo ponovljenih dokumentov. Če želite še dodatno filtrirati rezultate iskanja, lahko podate možnost, s katero zadržite dokumente, ki so skorajda enaki drug drugemu, in preprečite njihov prikaz v rezultatih iskanja.

### **preferences.nearDuplicateDetection=No**

Podaja, da skoraj identični dokumenti med obdelavo poizvedb ne bodo filtrirani.

Če podate **Da**, se dokumenti s podobnimi naslovi in povzetki ne prikažejo v rezultatih iskanja. Prikaže se sporočilo, ki uporabnike obvesti, da so bili nekateri



dokumenti izpuščeni, ker so podobni drugim dokumentom v dobljenem nizu. Uporabniki lahko kliknejo povezavo, s katero onemogočijo zadrževanje dokumentov, in v rezultatih prikažejo vse dokumente.

Če želite zadržati skoraj enake dokumente, mora objekt Poizvedba API-ja za iskanje in indeksiranje (SI-API) za iskalno aplikacijo podajati metodo `setProperty` z nizom `NearDuplicateDetection`, nastavljenim na `Yes` (na primer `query.setProperty("NearDuplicateDetection", "Yes");`).

## Analiza najboljših rezultatov (stolpični grafikon za metapodatkovna polja)

Podate lahko možnosti za prikaz stolpičnih grafikonov, ki ponazarjajo analizo najboljših rezultatov. Po privzetku je analiziranih najboljših 500 rezultatov. Vsak grafikon ustreza enemu polju metapodatkov, vsak stolpec v grafikonu pa ustreza vrednosti polja. Dolžina stolpca nakazuje pogostost pojavitve vrednosti polja. Daljši kot je stolpec, več je pojavitev vrednosti tega polja.

Pri konfiguriranju prostora za pajkanje lahko podate možnosti za polja metapodatkov. Če želite konfigurirati grafikon najboljših rezultatov za polje metapodatkov, morate podati, da polje omogoča iskanje po poljih in da ga je mogoče vrniti v rezultatih iskanja.

Lastnosti, ki jih konfigurirate za stolpični grafikon, so v naslednjem formatu, pri čemer je *number* številka, ki unikatno označuje stolpični grafikon v rezultatih iskanja, *option* je možnost stolpičnega grafikona, *value* pa je vrednost možnosti:

`topResultsChartsnumber.option=value`

**`topResultsChartsnumber.titleKey=application_key`**

Podaja naslov grafikona, pri čemer je *number* število, ki unikatno označuje grafikon v iskalni aplikaciji, *application\_key* pa je oznaka za naslov grafikona. Pri tej oznaki lahko gre za ključ v datoteki `application.properties` iskalne aplikacije ali vrednost, ki jo podate na tem mestu. V naslednjem primeru je oznaka za naslov grafikona podana z vrednostjo ključa `topResults.mostRecentDocuments` v datoteki `application.properties`:

`topResultsCharts3.titleKey=topResults.mostRecentDocuments`

Naslov grafikona je lahko tudi točna vrednost, ki jo podate na tem mestu:

`topResultsCharts3.titleKey=Organizations`

**`topResultsChartsnumber.enable=true`**

Podaja, ali se bo stolpični grafikon prikazal uporabniku, ki si ogleduje rezultate iskanja. Če podate `false`, se grafikon ne prikaže.

**`topResultsChartsnumber.fieldName=field_name`**

Podaja ime metapodatkovnega polja, katerega vrednosti bodo analizirane v tem grafikonu. Na primer `databasetitle`. Za vsak grafikon, ki ga dodate, morate podati drugačno ime polja (enako polje ne sme biti uporabljeno v več grafikonih).

**`topResultsChartsnumber.maxValues.collapsed=number`**

Podaja število skrčenih postavk za prikaz v tem grafikonu. Podajte na primer 5, da v polju pokažete pet najpogostejših vrednosti.

**`topResultsChartsnumber.maxValues.expanded=number`**

Podaja število postavk, ki bodo prikazane v povsem razširjenem grafikonu. Podajte na primer 10, da v razširjenem grafikonu prikažete največ 10 različnih vrednosti.

**`topResultsChartsnumber.fieldValueSeparator=character`**

Podaja znak, s katerim so ločene vrednosti v polju, ki ga želite analizirati. Če polje na primer vsebuje več vrednosti, med seboj ločenih s podpičjem, (na primer `posrednik;prodajalec;zastopnik`), lahko s to lastnostjo podate podpičje (`;`) kot ločilo

med vrednostmi polja, tako da je mogoče vsako vrednost dodati v stolpični grafikon kot posebno postavko. Če ne podate te možnosti, se vrednost celotnega polja doda v grafikon kot ena sama postavka.

#### **topResultsChartsnumber.canUserChangeFieldName=true**

Podaja, ali lahko uporabnik med ogledovanjem rezultatov iskanja izbere drugo polje in si ogleda najboljše rezultate za to polje. Če podate `false`, uporabniki med ogledovanjem rezultatov iskanja ne morejo izbrati drugega polja za analizo.

Če podate `true`, se skupaj z rezultati iskanja za trenutno polje prikaže tudi seznam vseh polj, ki jih je sistem našel kot prvih 500 rezultatov. Če uporabnik izbere polje s tega seznama, se oznaka grafikona spremeni v **Grafikon dinamičnega polja**, rezultati stolpičnega grafikona za izbrano polje pa so prikazani, dokler uporabnik ne izbere drugega polja ali zapre brskalnika. Ko uporabnik naslednjič zažene iskalno aplikacijo, se prikažejo rezultati za izvirno polje.

#### **topResultsChartsnumber.width=number**

Podaja širino prikaza stolpičnega grafikona v pikslih. Na primer: 300.

#### **topResultsChartsnumber.barheight=number**

Podaja višino posameznega stolpca v stolpičnem grafikonu v pikslih. Na primer: 10.

#### **topResultsChartsnumber.color=#color\_code**

Podaja šestnajstiško kodo osnovne barve za stolpec v stolpičnem grafikonu. Privzeta vrednost je modra (#0309C0).

#### **topResultsChartsnumber.color.gradient=#color\_code**

Podaja barvo gradienta za stolpec v stolpičnem grafikonu. Privzeta vrednost je turkizna (#00FFFF). S premikanjem števila rezultatov za posamezno vrednost polja od nizke do visoke pogostosti se spreminja tudi bava stolpca od vrednosti, podane za možnost `color`, do vrednosti, podane za možnost `color.gradient`. Na primer:



#### **topResultsChartsnumber.sortKey=frequency**

Podaja način razvrstitve postavk v stolpičnem grafikonu. Podprte vrednosti:

**none** Postavke v stolpičnem grafikonu niso razvrščene.

**label** Postavke v stolpičnem grafikonu so razvrščene po vrednosti polja. Če konfigurirate lastnosti po meri z lastnostjo `topResultsCharts.number listOfLabels.prefixKey=field_name`, se postavke v stolpičnem grafikonu razvrstijo tudi po imenu oznake.

#### **frequency**

Postavke v stolpičnem grafikonu so razvrščene po številu rezultatov, vrnjenih na vrednost polja.

#### **topResultsChartsnumber.sortOrder=descending**

Za postavke, ki so razvrščene po oznaki ali pogostosti, podaja vrstni red razvrščanja. Podprte vrednosti:

#### **ascending**

Postavke, razvrščene po oznaki, so navedene v abecednem vrstnem redu od a do ž. Pri postavkah, ki so razvrščene po pogostosti, so vrednosti polj, ki se pojavljajo pogosteje, na seznamu prikazane nižje od vrednosti, ki se pojavljajo redkeje.

#### **descending**

Postavke, razvrščene po oznaki, so navedene v nasprotnem abecednem vrstnem redu od ž do a. Pri postavkah, ki so razvrščene po pogostosti, so vrednosti polj, ki se pojavljajo pogosteje, na seznamu prikazane višje od vrednosti, ki se pojavljajo redkeje.

### **topResultsChartsnumber.listOfLabels.prefixKey=field\_name**

Neobvezno. Omogoča podajanje informacij, ki bodo vedno prikazane za določene vrednosti polj, pri čemer *field\_name* označuje polje, katerega vrednost je analizirana za ta grafikon. Oznake, ki jih podate za to lastnost, so vedno prikazane v stolpičnem grafikonu, četudi se vrednost polja ne pojavi v rezultatih iskanja.

Za polje lahko konfigurirate poljubno število oznak. Za vsako oznako lahko podate dve lastnosti v naslednji obliki:

*field\_namenumber.value=value*

*field\_namenumber.displayValue=display\_value*

pri čemer je:

*field\_name*

Ime polja, za katerega konfigurirate oznake.

*number*

Številka, ki unikatno označuje oznako.

*value* Vrednost, ki jo želite prikazati v rezultatih analize.

*display\_value*

Je oznaka, ki bo prikazana v stolpičnem grafikonu. Pri tej oznaki lahko gre za ključ v datoteki `application.properties` iskalne aplikacije ali vrednost, ki jo podate na tem mestu.

Primer:

```
topResultsCharts1.listOfLabels.prefixKey=databasetitle
```

```
databasetitle1.value=JK Enterprises Articles & Papers
```

```
databasetitle1.displayValue=Articles & papers
```

```
databasetitle2.value=JK Enterprises Blank Forms
```

```
databasetitle2.displayValue=Blank forms
```

```
databasetitle3.value=JK Enterprises Bulletins & Guidelines
```

```
databasetitle3.displayValue=Bulletins & Guidelines
```

## **Analiza najboljših rezultatov (HTML po meri)**

Podate lahko možnosti za prikaz grafikona najboljših rezultatov, in sicer tako, da razširite API `com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` z razredom Java po meri. Če za prikazovanje rezultatov uporabite takšen pristop, lahko s HTML oblikujete vrnitev vseh rezultatov iskanja, ne samo poizvedb, s katerimi preiskujete polja metapodatkov. Rezultati so lahko vrnjeni za dokumente v zbirkah za iskanje v podjetju ali iz iskanj po zunanjih repozitorijih in spletnih mestih.

Za iskalno aplikacijo je mogoče konfigurirati poljubno število grafikonov. Z naslednjo lastnostjo dodelite naslov grafikonu:

### **topResultsChartsnumber.titleKey=application\_key**

Podaja naslov grafikona, pri čemer je *number* število, ki unikatno označuje grafikon v iskalni aplikaciji, *application\_key* pa je oznaka za naslov grafikona. Pri tej oznaki lahko gre za ključ v datoteki `application.properties` iskalne aplikacije ali vrednost, ki jo podate na tem mestu. V naslednjem primeru je oznaka za naslov grafikona podana z vrednostjo ključa `topResults.mostRecentDocuments` v datoteki `application.properties`:

```
topResultsCharts3.titleKey=topResults.mostRecentDocuments
```

Vsaka postavka v grafikonu ustreza vrednosti rezultata posameznega iskanja in je sestavljena iz niza lastnosti naslednje oblike, pri čemer je *number* številka, ki unikatno označuje grafikon, *option* je možnost grafikona, *value* pa je vrednost možnosti:

`topResultsChartsnumber.option=value`

**`topResultsChartsnumber.enable=true`**

Podaja, ali bo ta grafikon prikazan uporabniku, ki si ogleduje rezultate iskanja. Če podate `false`, se grafikon ne prikaže.

**`topResultsChartsnumber.maxValues.collapsed=number`**

Podaja število skrčenih postavk za prikaz v tem grafikonu. Podajte na primer 5, da v polju pokažete pet najpogostejših rezultatov, ki izpolnjujejo kriterij.

**`topResultsChartsnumber.width=number`**

Podaja širino prikaza grafikona v pikslih. Na primer: 400.

**`topResultsChartsnumber.dynamicChartClass=custom_class`**

Podaja ime razreda Java po meri, ki razširi API `com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` in definira način prikaza izhodnih podatkov v grafikonu. Primer:

```
topResultsCharts.3.dynamicChartClass=com.ibm.es.searchui.charts.servlet.
DynamicMostRecentDocuments
```

Naslednje lastnosti, ki so vključene v privzeto datoteko `config.properties` za vzorčno iskalno aplikacijo, uporablja vzorčni razred Java `DogearSearchResults`. Informacije o parametrih poizvedb, ki jih lahko vključite v iskalno aplikacijo po meri, najdete v dokumentaciji za API Dogear.

Nastavitev teh lastnosti s prilagojevalnikom iskalnih aplikacij ni mogoča:

**`topResultsChartsnumber.xsl.fileName=style_file`**

Podaja ime in pot do slogovnega lista XSL, ki bo uporabljen za oblikovanje prikaza najboljših rezultatov v grafikonu. Na primer: `/styles/dogear.xsl`.

**`topResultsChartsnumber.url=url`**

Podaja URL, ki bo preiskan.

**`topResultsChartsnumber.url.parameters=ps=number`**

Podaja velikost strani. Vrednost `ps=3` na primer omeji velikost strani na 3.

To je preprosto niz parametrov, kot je definiran v API-ju Dogear REST.

## **Analiza najboljših rezultatov (največje število rezultatov)**

**`topResult.resultSize=number`**

Privzeta in največja vrednost za število rezultatov, ki jih vrne analiza najboljših rezultatov, je 500. To vrednost lahko zmanjšate, povečati pa je ne morete. Če na primer želite podati manjšo številko ali če pride do težav zaradi poteka čakalnega časa za zahteve med analizo najboljših rezultatov. Nastavitev te lastnosti s prilagojevalnikom iskalnih aplikacij ni mogoča.

### **S tem povezani pojmi**



Nastavljanje lastnosti poizvedbe



Javanski razredi za prikazovanje najboljših rezultatov

## **Urejanje lastnosti vzorčne iskalne aplikacije**

Vzorčna iskalna aplikacija za iskanje v podjetju lahko išče v vseh aktivnih zbirkah in zunanjih virih v vašem sistemu. Uredite lahko konfiguracijsko datoteko, da podate možnosti za okolje

vašega spletnega strežnika, kot privzeto aplikacijo lahko uporabite drugo iskalno aplikacijo ali krmilite, katere možnosti so prikazane ob zagonu aplikacije.

## O tej nalogi

Namestitveni program razmesti vzorčno iskalno aplikacijo za iskanje v podjetju na aplikacijski strežnik IBM WebSphere na iskalnih strežnikih za iskanje v podjetju. Če želite konfigurirati to iskalno aplikacijo, uredite konfiguracijsko datoteko `config.properties`, ki je razmeščena z aplikacijo.

Spremembe začnejo veljati šele, ko zaustavite in znova zaženete aplikacijo `ESSearchServer` na aplikacijskem strežniku `WebSphere Application Server`.

## Postopek

Urejanje vzorčne iskalne aplikacije

1. Na iskalni strežnik se prijavite kot skrbnik iskanja v podjetju.

2. S standardnim urejevalnikom besedil uredite datoteko `config.properties`.

Datoteka `config.properties` je nameščena na naslednjem nahajališču, pri čemer je `ES_INSTALL_ROOT` namestitveni imenik `OmniFind Enterprise Edition` na iskalnem strežniku:

```
ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/
ESSearchApplication.war/WEB-INF/config.properties
```

3. Uredite lastnosti, da podate informacije o okolju vašega spletnega strežnika in nastavitvah iskanja, nato pa shranite in zaprite datoteko. Znak lojtra (#) v datoteki nakazuje vrstico z opombami.

4. Zaustavite in znova zaženite aplikacijo `ESSearchServer`.

### AIX,

### Linux

### ali Solaris

```
./stopServer.sh ESSearchServer
```

```
./startServer.sh ESSearchServer
```

### Windows


```
stopServer ESSearchServer
```

```
startServer ESSearchServer
```

Ti skripti se nahajajo v imeniku `WAS_INSTALL_ROOT/AppServer/bin`:

- Za aplikacijski strežnik `WebSphere Application Server` različice 5 je privzeta pot za namestitev `/usr/WebSphere` na sistemih `AIX`, `/opt/WebSphere` na sistemih `Linux` ali `Solaris` in `C:\Program Files\WebSphere` na sistemih `Windows`.
- Za aplikacijski strežnik `WebSphere Application Server` različice 6 je privzeta pot za namestitev `/usr/IBM/WebSphere` na sistemih `AIX`, `/opt/IBM/WebSphere` na sistemih `Linux` ali `Solaris` in `C:\Program Files\IBM\WebSphere` na sistemih `Windows`.

## S tem povezani pojmi

 [Nastavljanje lastnosti poizvedbe](#)

 [Javanski razredi za prikazovanje najboljših rezultatov](#)

## S tem povezana opravila

“Konfiguriranje iskalnih strežnikov, da sprejmejo samo zaščitene (SSL) zahteve” na strani 226

## Prilaganje iskalnih aplikacij

Prilagojevalnik iskalnih aplikacij (Search Application Customizer) je grafični vmesnik, ki ga lahko uporabljate za prilaganje iskalnih aplikacij za iskanje v podjetju ali iskalnih aplikacij po meri.

### Omejitve

Prilagojevalnik iskalnih aplikacij je na voljo kot samostojna aplikacija. Prilagojevalnika iskalnih aplikacij ne morete zagnati s portala WebSphere Portal in z njim prilagoditi iskalnih aplikacij, ki se izvajajo kot portalski programčki. Če želite prilagoditi iskalne aplikacije, ki se izvajajo kot portalski programčki, morate urediti lastnosti in konfigurirati primerek portalskega programčka z vmesnikom za skrbništvo nad portalom WebSphere Portal.

### O tej nalogi

Prilagojevalnik iskalnih aplikacij vam omogoča ponazarjanje sprememb, ki jih želite napraviti, in spreminjanje iskalne aplikacije, ne da bi morali urediti njeno konfiguracijsko datoteko. Tako lahko na primer spremenite slike pasice in ozadja, postavitev iskalnega vmesnika in podate možnosti za delo z rezultati iskanja.

Če v Prilagojevalniku iskalnih aplikacij aplikacij napravite kakršnekoli izbore, prilagojevalnik prikaže učinek teh sprememb. Ko shranite spremembe, posodobite konfiguracijsko datoteko za iskalno aplikacijo.

Spremembe začnejo veljati šele, ko zaustavite in znova zaženete aplikacijo ESSearchServer na aplikacijskem strežniku WebSphere Application Server.

### Nasvet:

Podrobnejše informacije o načinih uporabe Prilagojevalnika iskalnih aplikacij poiščite v publikaciji IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

### Postopek

Prilaganje iskalne aplikacije:

1. Če želite prilagoditi vzorčno iskalno aplikacijo za iskanje v podjetju, v spletni brskalnik vnesite URL Prilagojevalnika iskalnih aplikacij. Primer:  
`http://SearchServer.com/ESSearchApplication/palette.do`  
*SearchServer.com* je gostiteljsko ime iskalnega strežnika.  
Če vaš spletni strežnik ni konfiguriran, da uporablja vrata 80, morate podati tudi pravilno številko vrat. Primer:  
`http://SearchServer.com:9080/ESSearchApplication/palette.do`

**Nasvet:** Če ste skrbnik iskanja v podjetju, lahko Prilagojevalnik iskalnih aplikacij odprete tudi tako, da na skrbniški konzoli iskanja v podjetju izberete možnost **Prilagojevalnik iskanja**.

2. Če želite prilagoditi iskalno aplikacijo po meri, vnesite URL za Prilagojevalnik iskalnih aplikacij in pripnite ime konfiguracijske datoteke za vašo iskalno aplikacijo. Primer:  
`http://SearchServer.com/ESSearchApplication/palette.do?configFile=/WEB-INF/myConfig.properties`  
Če datoteka, ki ste jo podali, ne obstaja, se prikažejo vrednosti v datoteki `config.properties` za vzorčno iskalno aplikacijo.

**Nasvet:** Konfiguracijsko datoteko, ki jo želite uporabiti z iskalno aplikacijo lahko podate tudi tako, da potem, ko ste zagnali Prilagojevalnik iskalnih aplikacij, kliknete **Naloži** in podate ime za datoteko.

3. Če je na aplikacijskem strežniku WebSphere Application Server omogočena globalna zaščita, se prijavite z veljavnim ID-jem uporabnika in geslom.
4. Izberite možnosti, ki jih želite prilagoditi: kot so informacije o iskalnem strežniku, vrste sporočil, ki bodo prikazane, možnosti za poizvedbe in rezultate iskanja ter imena slik, ki jih želite uporabiti za prepoznavanje različnih vrst dokumentov v rezultatih iskanja.  
Za pomoč pri podajanju možnosti kliknite **Pomoč za prilagojevalnik**. Če si želite ogledati učinek določenih sprememb, na primer kako so predstavljeni rezultati iskanja, vnesite poizvedbo in kliknite **Iskanje**.
5. Ko ste zadovoljni s svojimi izbirami, kliknite **Shrani**, da posodobite konfiguracijsko datoteko. Če kliknete **Ponastavi**, se možnosti, prikazane v Prilagojevalniku iskalnih aplikacij, obnovijo na vrednosti v zadnji shranjeni različici konfiguracijske datoteke.
6. Na indeksni strežnik se prijavite kot skrbnik iskanja v podjetju in zaustavite ter znova zaženite aplikacijo ESSearchServer.

**AIX,**

**Linux**

**ali Solaris**

```
./stopServer.sh ESSearchServer
```

```
./startServer.sh ESSearchServer
```

**Windows**

```
stopServer ESSearchServer
```

```
startServer ESSearchServer
```

Ti skripti se nahajajo v imeniku WAS\_INSTALL\_ROOT/AppServer/bin:

- Za aplikacijski strežnik WebSphere Application Server različice 5 je privzeta pot za namestitvev /usr/WebSphere na sistemih AIX, /opt/WebSphere na sistemih Linux ali Solaris in C:\Program Files\WebSphere na sistemih Windows.
- Za aplikacijski strežnik WebSphere Application Server različice 6 je privzeta pot za namestitvev /usr/IBM/WebSphere na sistemih AIX, /opt/IBM/WebSphere na sistemih Linux ali Solaris in C:\Program Files\IBM\WebSphere na sistemih Windows.

**S tem povezana opravila**

“Konfiguriranje iskalnih strežnikov, da sprejmejo samo zaščitene (SSL) zahteve” na strani 226

## Kloniranje vzorčne iskalne aplikacije

Če želite vzorčno iskalno aplikacijo uporabiti kot model za izdelavo vaših lastnih iskalnih aplikacij, lahko kopirate datoteko config.properties ali uporabite Prilagojevalnik iskalnih aplikacij.

### O tej nalogi

Izdelavo vaše iskalne aplikacije po meri olajšate tako, da klonirate konfiguracijske možnosti, ki jih podate za vzorčno iskalno aplikacijo, in nato prilagodite možnosti, ki jih želite spremeniti.

Potem ko klonirate vzorčno iskalno aplikacijo, podate ime vaše konfiguracijske datoteke za začetek nove iskalne aplikacije. Podate tudi ime vaše konfiguracijske datoteke, ki se uporablja za prilagoditev nove iskalne aplikacije s Prilagojevalnikom iskalnih aplikacij.

S kloniranjem vzorčne iskalne aplikacije lahko hitro izdelate iskalne aplikacije za poseben namen ali naslovnike. Tako lahko na primer izdelate eno iskalno aplikacijo za zaposlene na oddelku za ravnanje s človeškimi viri in drugo za prodajalce.

Spremembe začnejo veljati šele, ko zaustavite in znova zaženete aplikacijo ESSearchServer na aplikacijskem strežniku WebSphere Application Server.

## Postopek

Kloniranje vzorčne iskalne aplikacije

1. Urejanje konfiguracijske datoteke:
  - a. Kopirajte datoteko `config.properties` za vzorčno iskalno aplikacijo in jo preimenujte. Datoteka `config.properties` je nameščena na naslednjem nahajališču, pri čemer je `ES_INSTALL_ROOT` namestitveni imenik OmniFind Enterprise Edition na iskalnem strežniku:

```
ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/
ESSearchApplication.war/WEB-INF/config.properties
```

Datoteka, ki jo izdelate, se mora nahajati v podimeniku `WEB-INF`.
  - b. Uredite lastnosti, ki jih želite uporabiti v vaši iskalni aplikaciji po meri in shranite datoteko. Spremeniti boste morali vsaj lastnost `applicationName`, da podate ime vaše iskalne aplikacije.
2. Kloniranje vzorčne iskalne aplikacije s Prilagojevalnikom iskalne aplikacije:
  - a. Zaženite Prilagojevalnik iskalnih aplikacij tako, da dodate ime konfiguracijske datoteke, ki jo želite izdelati. V naslednjem zgledu se izdela datoteka z imenom `myNewFile.properties`:

```
http://ESServer.com/ESSearchApplication/palette.do?configFile=/WEB-INF/
myNewFile.properties
```

Ker ta datoteka še ne obstaja, so uporabljene vrednosti, podane v konfiguracijski datoteki `config.properties` vzorčne iskalne aplikacije.

**Nasvet:** Konfiguracijsko datoteko za iskalno aplikacijo lahko izdelate tudi tako, da potem, ko ste zagnali Prilagojevalnik iskalnih aplikacij, kliknete **Naloži** in podate ime za datoteko. Datoteka se izdela, ko kliknete **Shrani**, da shranite vaše prilagoditvene možnosti.
  - b. Če je na aplikacijskem strežniku WebSphere Application Server omogočena globalna zaščita, se prijavite z veljavnim ID-jem uporabnika in geslom.
  - c. Podajte ime za vašo iskalno aplikacijo in možnosti, ki jih želite uporabiti za iskanje, ter kliknite **Shrani**. Spremembe, ki jih podate, se shranijo v vašo novo konfiguracijsko datoteko v podimeniku `WEB-INF`.

Za pomoč pri podajanju možnosti kliknite **Pomoč za prilagojevalnik**. Če si želite ogledati učinek določenih sprememb, na primer kako so predstavljeni rezultati iskanja, vnesite poizvedbo in kliknite **Iskanje**.
3. Na indeksni strežnik se prijavite kot skrbnik iskanja v podjetju in zaustavite ter znova zaženete aplikacijo ESSearchServer.

**AIX,**

**Linux**

**ali Solaris**

```
./stopServer.sh ESSearchServer
```

```
./startServer.sh ESSearchServer
```



## Windows

```
stopServer ESSearchServer
startServer ESSearchServer
```

Ti skripti se nahajajo v imeniku WAS\_INSTALL\_ROOT/AppServer/bin:

- Za aplikacijski strežnik WebSphere Application Server različice 5 je privzeta pot za namestitvev /usr/WebSphere na sistemih AIX, /opt/WebSphere na sistemih Linux ali Solaris in C:\Program Files\WebSphere na sistemih Windows.
- Za aplikacijski strežnik WebSphere Application Server različice 6 je privzeta pot za namestitvev /usr/IBM/WebSphere na sistemih AIX, /opt/IBM/WebSphere na sistemih Linux ali Solaris in C:\Program Files\IBM\WebSphere na sistemih Windows.

## Analiziranje najboljših rezultatov

Uporabnikom lahko pomagate pri izboljšavi niza rezultatov iskanja tako, da konfigurirate možnosti za analizo najboljših rezultatov.

Glavna naloga analize najboljših rezultatov je razvrščanje trenutnega niza rezultatov iskanja glede na pogostost pojavitve analiziranih vrednosti. Uporabniki lahko filtrirajo rezultate iskanja tako, da izberejo eno od analiziranih vrednosti, ki samodejno doda izbrano vrednost kot novo ključno besedo v kriterije iskanja. Pomen analize najboljših rezultatov je v tem, da uporabnikom omogoča natančno naravnavanje rezultatov iskanja, ne da bi za to morali uporabiti zapleteno ali napredno skladnjo za iskanje.

V aplikaciji za iskanje v podjetju lahko s stolpičnimi grafikoni grafično ponazorite, katere vrednosti polj metapodatkov se v rezultatih iskanja pojavljajo najpogosteje. Izdelate lahko tudi razred Java po meri, s katerim prikažete najboljše rezultate, vključno z rezultati iz iskalnih virov izven podjetja, v formatu HTML.

### Omejitve

Prepričajte se, da so v računalniku, v katerem se izvaja WebSphere Application Server in iskalna aplikacija, nameščene ustrezne pisave za vaš jezik. S tem korakom preverite, ali je ob izdelavi stolpičnega grafikona pisava nastavljena na nastavitvev, ki podpira znake iz oznake grafikona. Ta nastavitev je še posebej pomembna za azijske jezike. Če namestite pisave po namestitvi izdelka WebSphere Application Server, morate znova zagnati WebSphere Application Server, da spremembe stopijo v veljavo.

Če zaženete iskalno aplikacijo kot samostojno aplikacijo, lahko s prilagojevalnikom iskalnih aplikacij konfigurirate grafike najboljših rezultatov ali pa uredite konfiguracijsko datoteko iskalne aplikacije. Če zaženete iskalno aplikacijo kot portalski programček v portalu WebSphere Portal, morate urediti lastnosti in konfigurirati primerek portalskega programčka za skrbniški vmesnik WebSphere Portal. Konfiguriranje možnosti za analizo najboljših rezultatov s prilagojevalnikom iskalnih aplikacij ni mogoče.

### O tej nalogi

Najboljše rezultate lahko predstavite grafično, in sicer tako, da v stolpičnem grafikonu prikažete najpogostejše vrednosti v poljih metapodatkov. Prav tako lahko z razredom Java razširite iskalno aplikacijo in prikažete najboljše rezultate v drugem formatu, na primer tako, da z uporabo HTML predstavite najboljše rezultate v obliki nerazvrščenega seznama.

### Stolpični grafikoni za metapodatkovna polja

Podate lahko možnosti za analizo metapodatkovnih polj in pokažete rezultate analize v stolpičnih grafikoni. Grafikoni se prikažejo ob rezultatih uporabnikove poizvedbe.

Vsak grafikon prikazuje podatke iz posameznega metapodatkovnega polja (na primer velikost dokumenta, avtor, datum in podobno), vsak stolpec v grafikonu pa ustreza vrednosti določenega polja.

Dolžina stolpca kaže število dokumentov, ki vsebujejo določeno vrednost polja, v primerjavi z drugimi dokumenti, ki v tem polju vsebujejo drugačne vrednosti. Daljši kot je stolpec, več je pojavitev vrednosti tega polja. Uporabniki lahko natančno naravnajo rezultate s klikom na stolpec v stolpičnem grafikonu. Vrednost polja, ki ga predstavlja izbrani stolpec, se doda kot dodaten poizvedbeni izraz, dodatni kriteriji iskanja pa skrčijo nove rezultate iskanja.

Pri konfiguriranju prostora za pajkanje lahko izberete možnost, s katero podate možnosti iskanja, ki veljajo za polja z metapodatki. Podate lahko na primer, ali je mogoče metapodatkovno polje preiskovati kot prosto besedilo, iskati po imenu polja, pokazati v rezultatih iskanja, iskati kot parametrične podatke in podobno. Če želite v rezultatih iskanja prikazati grafike, ki ponazarjajo metapodatkovna polja, morate za pajka konfigurirati možnosti metapodatkovnih polj. Podati morate vsaj, da polje omogoča iskanje po poljih in da ga je mogoče prikazati v rezultatih iskanja.

Če pri konfiguriranju možnosti za metapodatkovno polje izberete možnost **Popolno ujemanje**, se v primeru, ko uporabnik v stolpičnem grafikonu izbere določen stolpec, zažene poizvedba za popolno ujemanje. Iskanje s popolnim ujemanjem podaja, da bodo rezultati vrnjeni samo, če se določbe poizvedbe ujemajo s celotno vrednostjo polja. Če vsebuje polje manj vsebine ali dodatno vsebino, ujemanja ni.

Če pri konfiguriranju možnosti metapodatkovnega polja ne izberete možnosti **Complete match**, se v primeru, ko uporabnik v stolpičnem grafikonu izbere določen stolpec, zažene poizvedba po poljih. V tem primeru se mora v polju pojaviti dodaten poizvedbeni izraz, za katerega pa ni nujno, da se ujema s celotno vrednostjo polja.

### Rezultati, oblikovani s HTML

API `com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` lahko razširite z razredom Java po meri. Če za prikazovanje rezultatov uporabite takšen pristop, lahko s HTML oblikujete prikaz vseh rezultatov iskanja, ne samo poizvedb, s katerimi preiskujete polja metapodatkov. Rezultati so lahko vrnjeni za dokumente v zbirkah za iskanje v podjetju ali iz iskanj po zunanjih repozitorijih in spletnih mestih.

Vsak grafikon ustreza posameznemu razredu Java, razred pa podaja način predstavitve najboljših rezultatov v grafikonu.

`com.ibm.es.searchui.charts.servlet.DynamicMostRecentDocuments`, ki je vzorčni razred, v katerem je vzorčna koda za iskanje v podjetju, predstavi najboljše rezultate v obliki nerazvrščenega seznama. Dokumenti so razporejeni po datumu, prikažejo pa se samo naslovi dokumentov in datumi.

Drug vzorčni razred, ki je na voljo skupaj z iskanjem v podjetju, `com.ibm.es.searchui.charts.servlet.DogearSearchResults`, prikazuje, kako lahko uporabnikom ponudite seznam zaznamkov iz pripomočka Lotus Connections Dogear, ki so povezani s poizvedbami uporabnikov.

Ko uporabniki kliknejo rezultat iskanja v obliki prilagojenih izhodnih podatkov, se dokument odpre v novem oknu brskalnika.

### Postopek

Ta postopek prikazuje, kako s prilagojevalnikom iskalnih aplikacij podate, da želite analizirati metapodatkovna polja. Najpogostejše vrednosti iz metapodatkovnih polj so grafično prikazane v stolpičnem grafikonu.

1. Odprite prilagojevalnik iskalnih aplikacij. Če še ni prikazana, naložite konfiguracijsko datoteko iskalne aplikacije, ki jo želite prilagoditi, in kliknite **Uveljaviti**.

2. Pomaknite se navzdol do **Grafikona najboljših rezultatov** in kliknite **Dodaj grafikon**.
3. Ko se prikaže seznam možnosti grafikona, počistite polje **Grafikon po meri** (to možnost izberite samo, če ste za analizo in vrnitev najboljših rezultatov izdelali razred Java po meri) in izberite potrditveno polje **Omogoči grafikon**, s čimer zagotovite, da uporabniki pri ogledovanju rezultatov iskanja vidijo stolpični grafikon.
4. V polju **Naslov grafikona** podajte opisno oznako grafikona. Vrednost, ki jo podate na tem mestu, zamenja besedilo rezerviranega mesta **Nova vrstica grafikona**.
5. Odločite se, ali boste izbrali potrditveno polje **Omogoči izbiro dinamičnih polj**. Če omogočite to možnost, lahko uporabniki pri pregledovanju rezultatov iskanja izberejo različna polja in si ogledajo analizo najboljših rezultatov za to polje.  
Morda boste želeli s to možnostjo uporabnikom omogočiti natančno naravnavanje niza rezultatov iskanja. Uporabnik lahko na primer preišče bazo podatkov in poišče informacije o ženskih uslužbenkah. Po vnosu poizvedbe za iskanje po polju, ki podaja spol uslužbenca (na primer `sex:F`), se vrednosti najboljših rezultatov prikazujejo v obliki stolpcev v stolpičnem grafikonu. Nato uporabnik s seznama polj, ki obstajajo v začetnem nizu rezultatov, izbere `job`. Stolpični grafikoni sedaj prikazujejo rezultate za najboljše delovne funkcije, ki jih opravljajo ženske. Na koncu uporabnik s seznama razpoložljivih polj izbere možnost `designer`. V rezultatih se prikazujejo informacije o ženskih uslužbenkah, ki opravljajo delovno funkcijo oblikovalca.
6. V polje **Ime metapodatkovnega polja** vnesite ime metapodatkovnega polja, katerega vrednosti želite analizirati v grafikonu.
7. Če polje vsebuje več vrednosti, med seboj ločenih z ločevalnim znakom, morate ta znak podati v polju **Ločilo med vrednostmi polja**. Če polje na primer vsebuje dve vrednosti, kot sta ime in priimek stranke, med seboj ločeni s podpičjem, lahko na tem mestu podate podpičje ter posebej dodate vsako vrednost v stolpični grafikon. Če ne identificirate ločevalnega znaka, sistem celotno vrednost polja analizira kot eno postavko.
8. Podajte možnosti za prikaz stolpičnega grafikona, na primer velikost in barvo stolpcev in število stolpcev, ki bo prikazano ob skrčitvi ali razširitvi grafikona.
9. Podajte možnosti za razvrščanje rezultatov analize. Rezultate lahko na primer razvrstite po vrednosti polj ali pogostosti pojavitve vrednosti.
10. Če želite zagotoviti prikaz informacij za določena polja z vrednostmi, četudi se vrednosti ne pojavijo v rezultatih iskanja, kliknite možnost **Dodaj vrstico** v področju **Oznake po meri**. Podajte vrednost polja, ki jo želite prikazati v stolpičnem grafikonu, četudi se vrednosti ne pojavi v rezultatih iskanja, in podajte oznako za ta stolpec v stolpičnem grafikonu.  
Če želite vedno preveriti, ali se v rezultatih iskanja pojavi konkurenčno podjetje, na primer JK Enterprises, četudi število pojavitev ne izpolnjuje pogojev za prikaz v najboljših rezultatih, podajte JK Enterprises kot izvorno vrednost, ki jo želite analizirati, nato pa z opisno oznako označite ta stolpec v stolpičnem grafikonu.
11. Kliknite možnost **Uveljavi**, prikazano ob naslovu grafikona, da uveljavite možnosti, podane za ta grafikon.
12. Če so rezultati iskanja že prikazani v področju iskalne aplikacije, se za trenutno iskanje izvede analiza najboljših rezultatov, sistem pa prikaže novi grafikon. V nasprotnem primeru vnesite poizvedbo, s katero boste preizkusili prikaz stolpičnega grafikona.
13. Če ste zadovoljni z grafikonom, kliknite **Shrani** ter posodobite konfiguracijsko datoteko iskalne aplikacije.

#### S tem povezani pojmi

 Nastavljanje lastnosti poizvedbe

 Javanski razredi za prikazovanje najboljših rezultatov

---

## Dostopanje do iskalnih aplikacij

Do iskalne aplikacije dostopate tako, da v spletnem brskalniku podate URL.

### Preden začnete

Iskalno aplikacijo morate konfigurirati za okolje vašega spletnega strežnika.

### O tej nalogi

Vzorčna iskalna aplikacija je nameščena na iskalnih strežnikih za iskanje v podjetju. S to pripravljeno aplikacijo lahko preskusite zbirke in zunanje vire in jih šele nato omogočite za uporabnike. Lahko pa jo uporabite kot model za izdelavo lastnih iskalnih aplikacij.

### Postopek

Zagon iskalne aplikacije:

1. V spletni brskalnik vnesite URL za iskalno aplikacijo. Primer:  
`http://SearchServer.com/ESSearchApplication/`  
*SearchServer.com* je gostiteljsko ime iskalnega strežnika.  
Če vaš spletni strežnik ni konfiguriran, da uporablja vrata 80, morate podati tudi pravilno številko vrat. Primer:  
`http://SearchServer.com:9080/ESSearchApplication/`
2. Če želite zagnati iskalno aplikacijo po meri, vnesite URL za vzorčno iskalno aplikacijo in pripnite ime konfiguracijske datoteke za vašo iskalno aplikacijo. Primer:  
`http://SearchServer.com/ESSearchApplication/search.do?configFile=/WEB-INF/myConfig.properties`  
Če datoteka, ki ste jo podali, ne obstaja, se prikaže vzorčna iskalna aplikacija za iskanje v podjetju.
3. Če je na aplikacijskem strežniku WebSphere Application Server omogočena globalna zaščita, se v aplikacijo prijavite z veljavnim ID-jem uporabnika in geslom.  
Če je omogočena zaščita katere izmed zbirk, ki so na voljo za iskalno aplikacijo in ta zaščiten zbirka vključuje pajke, ki so konfigurirani za preverjanje uporabniških poverilnic med obdelavo poizvedb, lahko konfigurirate uporabniški profil. Na strani Moj profil podajte poverilnice za dostop do zaščitenih domen. Nato lahko iščete po teh domenah, ne da bi se morali vanje prijaviti.  
Če pajek podpira zaščito z enotno prijavo (SSO), lahko iščete po zaščitenih domenah, ne da bi morali izdelati uporabniški profil.
4. Na strani Iskanje predložite poizvedbo. Iskanje bo potekalo v vseh zbirkah in zunanjih virih, ki so izbrani za iskanje na strani Nastavitve.

---

## Konfiguriranje iskalnih strežnikov, da sprejmejo samo zaščitene (SSL) zahteve

Na iskalnih strežnikih lahko onemogočite vmesnik HTTP in konfigurirate strežnike tako, da bodo sprejeli rezultate iskanja samo prek SSL in zaščitenega vmesnika HTTPS.

### O tej nalogi

Če želite iskalne strežnike konfigurirati tako, da bodo za obdelavo iskalnih zahtev uporabljali samo protokol SSL (plast zaščitenih vtičnic), onemogočite vmesnik HTTP. Prepričajte se tudi, da je na obeh iskalnih strežnikih in na vseh odjemalskih računalnikih, kot je na primer

strežnik WebSphere Portal, na katerem je nameščen portalski programček za iskanje v podjetju, nameščena enaka datoteka shrambe ključev. Datoteka shrambe ključev, imenovana tudi *varna shramba ključev*, vsebuje javne ključe, ki omogočijo uporabo SSL-a v overjenih komunikacijah.

## Postopek

Če želite konfigurirati iskalne strežnike tako, da bodo sprejemali samo zaščitene zahteve, naredite naslednje:

1. Prijavite se kot skrbnik iskanja v podjetju. Če uporabljate namestitev na več strežnikov, na indeksnem strežniku naredite naslednje:
  - a. Zaustavite sistem iskanja v podjetju:  
`esadmin system stopall`
  - b. Uredite datoteko `ES_NODE_ROOT/master_config/nodes.ini`.
  - c. Spremenite vrednost `node_ID.searchserverport` z vrat HTTP (običajno 80) v vrata HTTPS (običajno 443) in shranite datoteko. Če uporabljate namestitev na več strežnikov, posodobite obe vrednosti `node_ID.searchserverport` (po eno za vsak iskalni strežnik).
2. S pomočjo naslednjih korakov posodobite iskalni strežnik. Če uporabljate namestitev na več strežnikov, na obeh iskalnih strežnikih naredite naslednje:
  - a. Če ste izvedli namestitev na več strežnikov, se v iskalni strežnik prijavite kot skrbnik iskanja v podjetju.
  - b. Uredite datoteko `ES_NODE_ROOT/nodeinfo/es.cfg`.
  - c. Posodobite lastnost `TrustStore` in podajte popolnoma kvalificirano pot do datoteke shrambe ključev SSL.
  - d. Posodobite lastnost `HTTPProtocol` ter podajte HTTPS, nato pa shranite datoteko.
  - e. Vnesite naslednji ukaz, pri čemer je `trustStore_password` geslo za datoteko shrambe ključev. S tem ukazom šifirate vrednost gesla in posodobite vrednost `TrustStorePassword` v datoteki `es.cfg`.

### AIX, Linux ali Solaris

```
eschangetrustpw.sh trustStore_password
```

### Windows

```
eschangetrustpw trustStore_password
```

- f. Prepričajte se, da lastnosti `trustStore` in `trustPassword` v datoteki `config.properties` iskalne aplikacije podajata pravilno popolnoma kvalificirano pot in geslo za datoteko shrambe ključev. Te informacije lahko preverite ali spremenite tako, da uredite datoteko `config.properties` ali da uporabite Prilagojevalnik iskalnih aplikacij.
    - g. Če za iskanje v podjetju uporabljate portalski programček za iskanje, poskrbite, da parametra portalskega programčka `trustStore` in `trustPassword` podajata pravilno popolnoma kvalificirano pot in geslo za datoteko shrambe ključev. Te informacije lahko preverite ali spremenite s pomočjo možnosti za upravljanje portalskega programčka v vmesniku za izvajanje skrbništva nad portalom WebSphere Portal.
3. Ponovno zaženite sistem iskanja v podjetju:  
`esadmin system startall`

### S tem povezana opravila

“Urejanje lastnosti vzorčne iskalne aplikacije” na strani 218

“Prilagajanje iskalnih aplikacij” na strani 220

“Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 5.1” na strani 313

## Konfiguriranje iskalnih strežnikov za sprejemanje zahtev prek strežnika proxy

Iskalni strežnik lahko konfigurirate tako, da bo sprejemal zahteve prek strežnikov.

### Postopek

Če želite omogočiti predlaganje zahtev iskalnim strežnikom prek strežnika proxy:

1. Prijavite se kot skrbnik sistema iskanja v podjetju in ga zaustavite. Pri namestitvi na več strežnikih se prijavite na indeksni strežnik.  
`esadmin system stopall`
2. Uredite datoteko `ES_NODE_ROOT/master_config/nodes.ini`.
  - a. Spremenite lastnost `ProxyServer`, da podate popolnoma kvalificirano ime gostitelja za strežnik proxy.
  - b. Spremenite lastnost `ProxyServerPort`, da podate vrata za strežnik proxy.
  - c. Izbirno: Če strežnik proxy zahteva, da so vse zahteve overjene, posodobite lastnost `ProxyServerUserName`, da podate veljavno uporabniško ime za strežnik proxy.
  - d. Izbirno: Če strežnik proxy zahteva, da so vse zahteve overjene, vnesite naslednji ukaz, kjer je `proxyServer_password` geslo za podano uporabniško ime strežnika proxy. Ta ukaz šifrira vrednost gesla in posodobi vrednost `ProxyServerUserPassword` v datoteki `es.cfg`.

#### AIX, Linux ali Solaris

```
exchangeproxypw.sh proxyServer_password
```

#### Windows

```
exchangeproxypw proxyServer_password
```

3. Poskrbite, da lastnosti `proxyHost` in `proxyPort` v konfiguracijski datoteki (datoteka `.properties`) za iskalno aplikacijo podajata pravilno popolnoma kvalificirano ime gostitelja in številko vrat za strežnik proxy.  
Če strežnik proxy zahteva overjanje, poskrbite, da lastnosti `proxyUser` in `proxyPassword` podajata veljavno uporabniško ime in geslo za strežnik proxy. Te informacije lahko preverite ali spremenite z urejanjem konfiguracijske datoteke ali s pomočjo Prilagojevalnika iskalnih aplikacij.
4. Če za iskanje v podjetju uporabljate portalski programček Iskanje, poskrbite, da parametra portalskega programčka `proxyHost` in `proxyPort` podajata pravilno popolnoma kvalificirano ime gostitelja in številko vrat za strežnik proxy.  
Če strežnik proxy zahteva overjanje, poskrbite, da lastnosti `proxyUser` in `proxyPassword` podajata veljavno uporabniško ime in geslo za strežnik proxy. Za preverjanje ali spreminjanje teh informacij uporabite možnosti upravljanja portalskih programčkov v skrbniškem vmesniku WebSphere Portal.
5. Ponovno zaženite sistem iskanja v podjetju:  
`esadmin system startall`

---

## Podpora za zunanje vire

*Zunanji vir* je vir podatkov, po katerem lahko z aplikacijo sistema iskanja v podjetju iščete brez pajkanja, razčlenjevanja ali indeksiranja dokumentov v viru podatkov.

Po naslednjih tipih virov podatkov lahko iščete kot po zunanjih virih:

- Baze podatkov, ki podpirajo protokol Java database connectivity (JDBC). Podprte so samo baze podatkov IBM DB2, Oracle, Microsoft SQL Server 2000 in Microsoft SQL Server 2005. Za vsako tabelo, ki jo omogočite za iskanje, je izdelan ločen zunanji vir.

**Omejitev:** Podpora za baze podatkov SQL Server 2005 je omejena na tabele, ki ne vsebujejo tipa podatkov Variant. Gonilnik JDBC za SQL Server 2005 v sistemih AIX ni podprt.

- Strežniki LDAP (Lightweight Directory Access Protocol). Za vsak strežnik LDAP je izdelan po en zunanji vir.

Po konfiguriranju informacij o zunanjem viru morate le-tega povezati z vsaj eno iskalno aplikacijo. Uporabniki bodo tako lahko iskali po zunanjem viru, sočasno pa bodo lahko izdajali poizvedbe za zbirke, ki ste jih v sistemu iskanja v podjetju izdelali s pajkanjem, razčlenjevanjem in indeksiranjem podatkov.

### S tem povezani pojmi

 Search and index API federators

---

## Dodajanje zunanjih virov v sistem

Pri dodajanju zunanjih virov v sistem iskanja v podjetju morate podati tip vira, ki ga nameravate dodati. Čarovnik vam pomaga podati informacije o viru podatkov in načinu iskanja po le-tem.

### Preden začnete

Če želite v sistem dodati zunanji vir, morate biti član z vlogo skrbnika sistema iskanja v podjetju.

### Omejitve

Če želite po bazi podatkov Oracle iskati kot po zunanjem viru, morate v iskalne strežnike sistema iskanja v podjetju namestiti odjemalski program Oracle.

Gonilnik JDBC za Microsoft SQL Server 2005 ni podprt v sistemih AIX.

### O tej nalogi

Pri dodajanju informacij o zunanjem viru v sistem morate omogočiti uporabnike, ki bodo z aplikacijo sistema iskanja v podjetju izdajali poizvedbe za vir. Omogočite lahko iskanje po strežnikih LDAP (Lightweight Directory Access Protocol) in tabelah baz podatkov Java database connectivity (JDBC).

Pri konfiguriranju strežnika LDAP vam čarovnik pomaga podati informacije, s katerimi lahko sistem vzpostavi povezavo s strežnikom, in možnosti za iskanje po strežniku.

Pri konfiguriranju baze podatkov JDBC vam čarovnik pomaga podati informacije, s katerimi lahko sistem vzpostavi povezavo z bazo podatkov, pomaga izbrati tabele, ki jih lahko omogočite za iskanje, in podati možnosti za iskanje po podatkih v tabelah. Za vsako tabelo, ki jo dodate v sistem, bo ta izdelal ločeni zunanji vir, po katerem je mogoče iskati.

Za informacije o gonilnikih za SQL Server 2000 preglejte temo <http://www.microsoft.com/downloads/details.aspx?familyid=07287B11-0502-461A-B138-2AA54BFDC03A&displaylang=en>. Za informacije o gonilnikih za SQL Server 2005 preglejte temo <http://www.microsoft.com/downloads/details.aspx?familyid=e22bc83b-32ff-4474-a44a-22b6ae2c4e17&displaylang=en>.

## Postopek

Če želite v sistem dodati zunanji vir, storite naslednje:

1. Če želite v sistem iskanja v podjetju vključiti baze podatkov JDBC, izvedite naslednje korake pred dodajanjem zunanjega vira. Ta korak, ki omogoča, da sistem poišče ustrezne gonilnike JDBC, morate izvesti enkrat.
  - a. V strežnik pajkanja se prijavite kot skrbnik sistema iskanja v podjetju.
  - b. Uredite datoteko `ES_INSTALL_ROOT/configurations/interfaces/discovery__interface.ini`, podajte spremenljivki okolja `CLASSPATH` in `LD_LIBPATH` in s tem pot razreda vključite v gonilnike JDBC, pot pa v datoteke knjižnic.
  - c. Uredite datoteko `ES_INSTALL_ROOT/configurations/interfaces/customcommunication__interface.ini`, podajte spremenljivki okolja `CLASSPATH` in `LD_LIBPATH` in s tem pot razreda vključite v gonilnike JDBC, pot pa v datoteke knjižnic.
  - d. Neobvezno: če želite za lokalne ali katalogizirane baze podatkov uporabiti gonilnik Oracle JDBC, v spremenljivko okolja `LD_LIBPATH` dodajte pot do knjižnice Oracle (na primer `LD_LIBPATH=../home/oracle/OraHome1/lib32`) nato pa uredite datoteko `escrset.sh`, podajte pot do knjižnice in izvozite spremenljivko za namestitveni imenik Oracle. Primer:

```
ORACLE_HOME=/home/oracle/OraHome1
export ORACLE_HOME
```
  - e. Znova zaženite sistem iskanja v podjetju, vključno s splošno komunikacijsko plastjo (CCL):

### AIX, Linux ali Solaris

```
esadmin stop
stopccl.sh
startccl.sh
esadmin start
```

### Ukazni poziv Windows

```
esadmin stop
stopccl
startccl
esadmin start
```

### Skrbniška orodja storitev operacijskega sistema Windows

Če želite zagnati CCL v ozadju:

- 1) Vnesite `esadmin stop`.
- 2) Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
- 3) Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaustavi**. Ko se storitev zaustavi, kliknite **Zaženi**.



- 4) Vnesite esadmin start.
2. Kliknite **Zunanji viri** in s tem odprite pogled Zunanji viri.
3. Kliknite možnost **Dodaj zunanji vir**.
4. Izberite tip zunanjega vira, ki ga želite dodati, strežnik LDAP ali bazo podatkov JDBC.
5. Kliknite možnost **Naprej** in začnite s konfiguriranjem zunanjega vira.

Odpre se čarovnik za tip vira, ki ga nameravate izdelati. Sledite pozivom čarovnika in tako konfigurirajte zunanji vir. Če želite boljše spoznati možnosti, ki jih podajate, lahko na katerikoli strani kliknete možnost **Pomoč**.

Naslednja imena gonilnikov JDBC in lokacij vam utegnejo koristiti pri konfiguriranju informacij o povezavi za baze podatkov DB2 in Oracle:

#### **DB2: starejši gonilnik JDBC**

Ime gonilnika: COM.ibm.db2.jdbc.app.DB2Driver

Vzorčno nahajališče: *koren\_namestitve\_db2/java/db2java.zip*

#### **DB2: univerzalni gonilnik JDBC**

Ime gonilnika: com.ibm.db2.jcc.DB2Driver

Vzorčna nahajališča:

*koren\_namestitve\_db2/java/db2jcc.jar*

*koren\_namestitve\_db2/java/db2jcc\_license\_cu.jar*

**Oracle** Ime gonilnika: oracle.jdbc.driver.OracleDriver

Vzorčno nahajališče: *osnova\_oracle/jdbc/lib/ojdbc14.jar*

#### **Microsoft SQL Server 2000**

Ime gonilnika: com.microsoft.jdbc.sqlserver.SQLServerDriver

Vzorčna nahajališča:

*osnova\_mssql\_jdbc/lib/mssqlserver.jar*

*osnova\_mssql\_jdbc/lib/msbase.jar*

*osnova\_mssql\_jdbc/lib/msutil.jar*

#### **Microsoft SQL Server 2005**

Ime gonilnika: com.microsoft.sqlserver.jdbc.SQLServerDriver

Vzorčno nahajališče: *namestitveni\_imenik/sqljdbc\_1.0/podrocne\_nastavitve/sqljdbc.jar*

Primer: *namestitveni\_imenik/sqljdbc\_1.0/enu/sqljdbc.jar*

6. Ko podate možnosti za iskanje po zunanjem viru, kliknite **Dokončaj**.  
Nov zunanji vir bo na pogledu Zunanji viri naveden skupaj z drugimi zunanjimi viri, ki ste jih dodali v sistem.

#### **S tem povezani pojmi**

 Search and index API federators

---

## **Povezovanje iskalnih aplikacij z zunanjimi viri**

Pred iskanjem po zunanjem viru morate le-tega povezati z vsaj eno iskalno aplikacijo.

### **Pređen začnete**

Če želite iskalne aplikacije povezati z zunanjimi viri in nato po njih iskati, morate biti član z vlogo skrbnika sistema iskanja v podjetju.

### **Postopek**

Če želite iskalno aplikacijo povezati z enim ali več zunanjimi viri, storite naslednje:

1. V orodni vrstici skrbniške konzole kliknite **Zaščita**.
2. Na strani Iskanje kliknite **Konfiguriranje iskalnih aplikacij**.

3. Na strani Konfiguriranje iskalnih aplikacij kliknite **Dodajanje iskalne aplikacije**.
4. Vnesite ime iskalne aplikacije.
5. Izberite zunanje vire, po katerih želite iskati z aplikacijo:
  - Kliknite možnost **Vse zbirke in zunanji viri**, če želite, da iskalna aplikacija dostopi do vseh zunanjih virov, ki jih dodate v sistem.
  - Kliknite možnost **Specifične zbirke in zunanji viri**, če želite, da iskalna aplikacija dostopi samo do zunanjih virov, ki jih podate.  
Ko izberete to možnost, se prikaže seznam z imeni zbirk in zunanjih virov. Za vsak zunanji vir, po katerem lahko aplikacija išče, izberite potrditveno polje **Izberi**.
6. Kliknite **Potrdi**.

#### **S tem povezani pojmi**

 [Search and index API federators](#)

---

## Zaščita iskanja v podjetju

Mehanizmi zaščite v iskanju v podjetju vam omogočajo, da zaščitite vire pred nepooblaščenim iskanjem in omejite skrbniške funkcije na specifične uporabnike.

Z iskanjem v podjetju lahko uporabniki iščejo veliko virov podatkov. Da zagotovite, da to počnejo samo uporabniki s pooblastili za dostop do vsebine, in da zagotovite, da samo uporabniki s pooblastili lahko dostopajo do skrbniške konzole, iskanje v podjetju usklajuje in zagotavlja zaščito na več ravneh.

### Nasvet:

Podrobnejše informacije o načinih konfiguriranja zaščite za iskanje v podjetju poiščite v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios. Scenariji opisujejo, kako omogočite globalno zaščito v izdelku WebSphere Application Server z repozitorijem LDAP, namestite komponento za upravljanje identitet in konfigurirate različne pajke ter tako zagotovite uveljavitev zaščite na ravni dokumenta.

### Spletni strežnik

Prva raven zaščite je spletni strežnik. Če na aplikacijskem strežniku WebSphere Application Server omogočite globalno zaščito, lahko uporabnike dodelite skrbniškim vlogam in overite uporabnike, ki izvajajo skrbništvo nad sistemom. Ko se uporabnik prijavi na skrbniško konzolo, so mu na voljo samo funkcije in zbirke, za katere ima skrbniška pooblastila.

Iskalne aplikacije lahko za overjanje uporabnikov, ki iščejo po zbirkah, uporabljajo tudi mehanizme zaščite aplikacijskega strežnika WebSphere Application Server.

### Zaščita na ravni zbirke

Pri izdelavi zbirke lahko omogočite zaščito na ravni zbirke. Te nastavitve ne morete več spremeniti, ko je zbirka izdelana. Če zaščite na ravni zbirke ne omogočite, pozneje ne morete podati nadzora zaščite na ravni dokumenta.

Ko je zaščita na ravni zbirke omogočena:

- Procesi za globalno analizo iskanja v podjetju uveljavijo različna pravila za indeksiranje dvojnikov dokumentov.
- Konfigurirate lahko možnosti za zagotavljanje zaščite na ravni dokumenta, tako, da na primer povežete žetone zaščite z dokumenti, ko so pajkani, zahtevate, da so trenutne poverilnice potrjene med obdelavo poizvedbe, in podate, ali naj se sidrno besedilo v spletnih dokumentih indeksira.
- Zaščito lahko zagotavljate s preslikovanjem iskalnih aplikacij (ne posameznih uporabnikov) v zbirke in zunanje vire, ki jih lahko iščejo. Nato uporabite standardne mehanizme za nadzor dostopa, da dovolite ali zavrnete dostop do iskalnih aplikacij.

Omogočanje zaščite zbirke in kakovost iskanja se medsebojno izključujeta. Omogočanje zaščite zbirke zmanjša informacije, ki so indeksirane za vsak dokument. Stranski učinek je ta, da bo za nekatere poizvedbe manj rezultatov.

### Zaščita na ravni dokumenta

Pri konfiguriranju pajkov za zbirke lahko omogočite zaščito na ravni dokumenta. Podate lahko na primer možnosti za povezovanje žetonov zaščite s podatki, ko pajek zbira podatke. Vaša iskalna aplikacija lahko uporabi te žetone, ki so shranjeni z

dokumenti v indeksu, da zagotovi nadzor dostopa in zagotovi, da lahko samo uporabniki z ustreznimi poverilnicami izvajajo poizvedbe podatkov in si ogledujejo rezultate iskanja.

Za določene tipe virov podatkov lahko konfigurirate možnosti za preverjanje prijavnih priporočil uporabnika s trenutnim nadzorom dostopa med obdelavo poizvedbe. Ta dodatna plast zaščite zagotavlja, da so pooblastila uporabnika preverjena v realnem času z izvirnim virom podatkov. Ta zmožnost lahko zavaruje pred primeri, ko se uporabnikove poverilnice spremenijo po indeksiranju dokumenta in njegovih žetonov zaščite.

Stopnja globalne analize za obdelavo sidrnega besedila ponavadi poveže besedilo, ki se pojavi v enem dokumentu (izvirnem), z drugim dokumentom (ciljnim), v katerem se to besedilo ne pojavi nujno. Ko konfigurirate spletnega pajka, lahko podate, ali želite sidrno besedilo izključiti iz indeksa, če besedilo vsebuje povezavo na dokument, ki ga spletni pajek ne sme pajkati.

Zaščita za vaše zbirke presega mehanizme overjanja in nadzora dostopa, ki jih lahko za zaščito indeksirane vsebine uporablja iskanje v podjetju. Za zaščito pred zlonamernim in nepooblaščenim dostopom do podatkov, ko so v prehodu, obstajajo varovalke. Iskalni strežniki na primer uporabljajo protokole, kot na primer plast zaščitene vtičnice (SSL), zaščitena lupina (SSH) in protokol Secure Hypertext Transfer Protocol (HTTPS), da komunicirajo z indeksnim strežnikom in iskalno aplikacijo.

Dodatna zaščita je zagotovljena s šifriranjem. Geslo za skrbnika iskanja v podjetju, podano ob namestitvi izdelka, je na primer shranjeno v šifriranem formatu. Gesla, ki jih uporabniki podajo v uporabniških profilih, se prav tako shranijo v šifriranem formatu.

Za boljšo zaščito morate zagotoviti, da je strojna oprema strežnika ustrezno izolirana in zaščitena pred nepooblaščenimi vdori. Z namestitvijo požarnega zida lahko zaščitite strežnike iskanja v podjetju pred vdori iz drugega dela vašega omrežja. Prav tako se prepričajte, da na strežnikih iskanja v podjetju ni odprtih vrat. Sistem konfigurirajte tako, da posluša zahteve samo na vratih, ki so izrecno dodeljena dejavnostim in aplikacijam iskanja v podjetju.

---

## Namestitvena zaščita

Namestitveni program za OmniFind Enterprise Edition ustvari okolje za zagotavljanje zaščite, ko uporabniki izvajajo skrbništvo nad zbirkami iskanja v podjetju ali po njih iščejo.

### ID skrbnika za iskanje v podjetju

Med namestitvenim procesom je namestitveni program pozvan, da poda uporabniški ID in geslo za skrbnika iskanja v podjetju. Sistem uporablja podane poverilnice za:

- Overjanje skrbnika iskanja v podjetju, ko se izvajajo naloge izvajanja skrbništva.
- Izdelovanje notranje baze podatkov iskanja v podjetju.
- Zagon vseh sej ali procesov iskanja v podjetju.

Uporabniški ID, podan med namestitvijo, mora biti veljavni uporabniški ID operacijskega sistema s sistemskimi skrbniškimi pooblastili. Ustrezno šifrirane poverilnice shrani namestitveni program v datoteko lastnosti na vsak strežnik iskanja v podjetju.

### Šifriranje

Za zaščito zaupnih podatkov je v uporabi šifriranje, ki v vseh sporočilih, prenesenih prek sistema iskanja v podjetju, kodira del s podatki za overjanje. Ta proces pomeni le malo dodatne obremenitve, ker se šifrirajo samo ID-ji in gesla overjanja. Prav tako se šifrirajo vsa gesla, ki jih shrani sistem (v konfiguracijskih datotekah, bazi podatkov iskanja v podjetju itd.).

## Zaščita aplikacijskega strežnika WebSphere Application Server

Če na iskalnem strežniku še ni nameščen aplikacijski strežnik WebSphere Application Server, potem namestitveni program tiho namesti izdelek z onemogočeno globalno zaščito. Če globalno zaščito na aplikacijskem strežniku WebSphere Application Server pozneje omogočite, potem je za overjanje skrbnika iskanja v podjetju odgovoren aplikacijski strežnik WebSphere Application Server.

Če omogočite globalno zaščito, potem morate v uporabniški register aplikacijskega strežnika WebSphere Application Server dodati ID in geslo skrbnika iskanja v podjetju, ki sta bila podana med namestitvijo, kot je imenik protokola LDAP (Lightweight Directory Access Protocol).

Če globalno zaščito omogočite po namestitvi OmniFind Enterprise Edition, morate posodobiti konfiguracijske vrednosti in zagnati ukaz **eschangewaspw**, da šifirate in shranite poverilnice aplikacijskega strežnika WebSphere Application Server v datoteko lastnosti iskanja v podjetju.

---

## Overjanje v primerjavi z nadzorom dostopa

Za zaščito vsebine pred nepooblaščenimi uporabniki in za nadzor dostopa do skrbniških funkcij, iskanje v podjetju podpira overjanje in pooblaščenje uporabnika (nadzor dostopa).

### Overjanje

Overjanje je vsak proces, s katerim sistem preveri identiteto uporabnika, ki želi dostopati do sistema. Ker nadzor dostopa temelji na identiteti uporabnika, ki zahteva dostop do vira, je overjanje bistveno za učinkovito zaščito.

Overjanje uporabnikov iskanja v podjetju se izvaja prek poverilnic, ki so sestavljene iz najmanj uporabniškega ID-ja in gesla.

Za overjanje uporabnikov, ki dostopajo do skrbniške konzole, iskanje v podjetju poveča podporo overjanja, ki je na voljo z aplikacijskim strežnikom WebSphere Application Server.

### Pooblaščenje (nadzor dostopa)

Pooblaščenje je vsak mehanizem, s katerim sistem dodeli ali prekliče pravico dostopa do nekaterih podatkov ali pravico za izvajanje kakšnega dejanja. Uporabnik se mora pogosto prijaviti v sistem tako, da uporabi neko obliko overjanja. Mehanizmi za nadzor dostopa določajo, katere operacije uporabnik sme opraviti ali katerih ne sme, tako, da primerjajo uporabnikovo identiteto s seznamom nadzora dostopa (ACL). Nadzor dostopa vključuje:

- Dovoljenja za datoteke, kot je pravica do izdelovanja, branja, urejanja in brisanja datoteke.
- Dovoljenja za programe, kot je pravica do izvajanja programa.
- Dovoljenja za podatke, kot je pravica do priklica ali posodobitve informacij v bazi podatkov.

---

## Skrbniške vloge

Iskanje v podjetju uporablja koncept vlog za nadziranje dostopa do raznih funkcij na skrbniški konzoli.

Po namestitvi OmniFind Enterprise Edition (OmniFind Enterprise Edition) namestitveni program konfigurira uporabniški ID in geslo za skrbnika iskanja v podjetju. Ob prvem dostopu do skrbniške konzole se morate prijaviti kot ta uporabnik. Če na aplikacijskem

strežniku WebSphere Application Server ne omogočite globalne zaščite, potem je ta uporabniški ID edini uporabniški ID, ki ga lahko uporabljate za dostop do skrbniške konzole za iskanje v podjetju.

Če omogočite globalno zaščito na aplikacijskem strežniku WebSphere Application Server, lahko kot skrbniške uporabnike iskanja v podjetju vpišete še dodatne uporabnike. Z dodeljevanjem vlog uporabnikom lahko omejite dostop do specifičnih zbirk in krmilite funkcije, ki jih lahko krmili vsak skrbniški uporabnik. Uporabniški ID-ji, ki jih dodelite skrbniškim vlogam v iskanju v podjetju, morajo obstajati v uporabniškem registru aplikacijskega strežnika WebSphere Application Server.

Ko se skrbniški uporabnik prijavi, iskanje v podjetju overi uporabniški ID. V konzoli so na voljo samo zbirke in funkcije, nad katerimi lahko uporabnik izvaja skrbništvo.

Uporabnike lahko vpišete v naslednje skrbniške vloge:

#### **Skrbnik iskanja v podjetju**

Ti uporabniki izdelujejo zbirke in imajo pooblastila za izvajanje skrbništva nad vsemi vidiki sistema iskanja v podjetju. Ko namestite OmniFind Enterprise Edition, podate ID uporabnika in geslo za prvega uporabnika s pooblastilom skrbnika iskanja v podjetju. Po prvi prijavi lahko ta uporabnik dodeli druge uporabnike vlogi skrbnika za iskanje v podjetju.

#### **Skrbnik zbirke**

Ti uporabniki lahko urejajo, nadzorujejo in krmilijo delovanje zbirk, za katere imajo skrbniška pooblastila. Ti uporabniki ne morejo izdelovati zbirk. Skrbniki zbirk lahko nadzorujejo in vodijo dejavnosti na ravni sistema samo, če jim to pooblastilo dodeli skrbnik iskanja v podjetju.

#### **Operater**

Ti uporabniki lahko nadzorujejo in krmilijo delovanje zbirk, za katere imajo skrbniška pooblastila. Tako lahko na primer zaženejo in zaustavijo dejavnosti zbirke, ne morejo pa zbirk izdelovati ali urejati. Operater lahko nadzoruje in vodi dejavnosti na ravni sistema samo, če mu to pooblastilo dodeli skrbnik iskanja v podjetju.

#### **Nadzornik**

Ti uporabniki lahko nadzorujejo zbirke, za katere imajo skrbniška pooblastila. Ti uporabniki ne morejo nadzorovati operacij (kot je zagon in zaustavitev strežnikov), izdelovati zbirk ali jih urejati. Nadzornik lahko opazuje dejavnosti na ravni sistema (ne more pa jih voditi) samo, če mu to pooblastilo dodeli skrbnik iskanja v podjetju.

## **Konfiguriranje skrbniških uporabnikov**

S konfiguriranjem skrbniških vlog lahko omejite dostop do zbirk in krmilite funkcije, ki jih posamezna skrbniška vloga lahko izvaja.

#### **Preden začnete**

Preden uporabnika dodelite skrbniški vlogi, se prepričajte, da je na aplikacijskem strežniku WebSphere Application Server omogočena zaščita. Prav tako preverite, ali v uporabniškem registru WebSphere Application Server obstaja uporabniški ID.

Skrbniške uporabnike lahko konfigurirajo samo člani skrbniške vloge za iskanje v podjetju.

#### **Postopek**

Za dodeljevanje uporabnikov skrbniškim vlogam:

1. Kliknite **Zaščita**, da odprete pogled Zaščita.

2. Na strani Skrbniške vloge kliknite **Dodaj uporabnika**.
3. Vnesite uporabniški ID uporabnika, ki ga želite vpisati, in izberite ustrezno skrbniško vlogo.
4. Če tega uporabnika ne vpisujete kot skrbnika iskanja v podjetju, podajte, ali ta uporabnik lahko dostopa do strani iz orodne vrstice **Sistem**.  
Morda boste želeli nekaterim operaterjem ali skrbnikom zbirke na primer omogočiti, da nadzirajo datoteke dnevnika na ravni sistema.
5. Če tega uporabnika ne vpisujete kot skrbnika iskanja v podjetju, izberite zbirke in zunanje vire, nad katerimi lahko ta uporabnik izvaja skrbništvo.  
Izberete lahko potrditvena polja za posamezne zbirke in zunanje vire, ali pa pustite uporabniku, da izvaja skrbništvo nad vsemi zbirkami in zunanjimi viri.

---

## Zaščita na ravni zbirke

Da zagotovite zaščito na ravni zbirke, konfigurirate možnosti za indeksiranje vsebine in možnosti, da dovolite iskalnim aplikacijam iskanje po specifičnih zbirkah.

Pri izdelovanju zbirke lahko izberete možnost, da omogočite zaščito zbirke. Če izberete to možnost, lahko pozneje konfigurirate nadzor zaščite na ravni dokumenta. Ko je zaščita zbirke omogočena, procesi za globalno analizo iskanja v podjetju prav tako uveljavijo različna pravila za indeksiranje dvojnikov dokumentov.

Ko ste izdelali iskalno aplikacijo, vam njen ID omogoča, da podate, po katerih zbirkah in zunanjih virih lahko iskalna aplikacija išče in kateri uporabniki lahko dostopajo do nje.

## Analiza podvojenih dokumentov in zaščita zbirke

Če omogočite zaščito zbirke, procesi globalne analize ne prepoznajo dvojnikov dokumentov v zbirki.

Med globalno analizo procesi indeksiranja identificirajo dokumente, ki so kopije (ali približne kopije) eden drugega. Nato vse te dokumente povežejo s kanonično predstavitvijo vsebine. S tem, ko dovolite identificiranje dvojnikov dokumentov, lahko zagotovite, da rezultati iskanja ne vsebujejo več dokumentov z isto (ali skoraj isto) vsebino.

Če ob izdelavi zbirke omogočite zaščito zbirke, dvojnikov dokumentov ne bodo identificirani in tako tudi ne bodo povezani s splošno kanonično predstavitvijo. Zato se vsak dokument indeksira ločeno. To zagotavlja, da se ovrednoti nadzor zaščite za vsak dokument in da uporabniki tako iščejo samo dokumente z žetoni zaščite, ki se ujemajo z njihovimi poverilnicami. Vsebina dveh dokumentov je lahko skorajda enaka, vendar dokumenta za vpeljavo zaščite uporabljata različne sezname za nadzor dostopa.

Za dva podvojena dokumenta, dokument\_A in dokument\_B predpostavimo, da ima uporabnik samo pravice za dostop do dokumenta\_B. Če odkrivanje dvojnikov izloči dokument\_B, uporabnik v rezultatih iskanja ne vidi dokumenta, saj so za dokument\_A v veljavi omejitve za dostop.

Onemogočanje analize dvojnikov dokumentov lahko izboljša zaščito dokumentov v zbirki, vendar lahko pride do slabše kakovosti iskanja, če uporabniki v rezultatih iskanja prejmejo več kopij istega dokumenta.

## Identifikatorji iskalne aplikacije

Zmožnost iskanja po različnih zbirkah krmili preslikovanje iskalnih aplikacij v zbirke in zunanje vire, ki jih lahko iščejo. Aplikacija z imenom Default (Privzeta) omogoča vzorcu iskalne aplikacije, da se uporabi za iskanje po vseh zbirkah in zunanjih virih.

Vse iskalne aplikacije morajo posredovati veljavno ime aplikacije (APPID) aplikacijskemu programerskemu vmesniku (API) za iskanje v podjetju. Iskalna aplikacija lahko išče le po zbirkah in zunanjih virih, povezanih s tem APPID-om.

Praden lahko iskalna aplikacija dostopi do zbirke ali zunanjega vira, mora skrbnik iskanja v podjetju povezati iskalno aplikacijo s specifičnimi zbirkami in viri, po katerih lahko išče. Iskalna aplikacija lahko išče po vseh zbirkah in zunanjih virih v sistemu iskanja v podjetju, ali pa samo po zbirkah in zunanjih virih, ki jih podate vi.



Vzorčna iskalna aplikacija (ESSearchApplication) ima datoteko lastnosti, ki podaja aplikacijsko ime, ki se naj uporabi. Privzeto mesto za to datoteko lastnosti je `ES_INSTALL_ROOT\installedApps\ESSearchApplication.ear\ESSearchApplication.war\WEB-INF\config.properties`.

Prvotna vrednost za ime aplikacije je Default (Privzeta). Če to vrednost spremenite, spremenite seznam zbirk in zunanjih virov, ki jih aplikacija ESSearchApplication lahko išče.

Za nadzor nad tem, kateri uporabniki lahko iščejo katere zbirke, morate povezati uporabnike (ali uporabniške skupine) z aplikacijo odjemalca tako, da uporabite standardne funkcije za nadzorovanje dostopa aplikacijskega strežnika WebSphere Application Server, podobno kot bi te funkcije uporabili za omejitev dostopa do URL-ja. Omejite lahko na primer dostop do URL-ja, ki zažene vašo iskalno aplikacijo.

Za več informacij o ID-jih iskalne aplikacije in o tem, kako vključiti nadzor zaščite v vaše prilagojene iskalne aplikacije, glejte API indeksa in iskanja za iskanje v podjetju.

#### S tem povezani pojmi

-  [Search and index API overview](#)
-  [Search and index API security](#)

---

## Zaščita na ravni dokumenta

Če je ob izdelavi zbirke zaščita omogočena, lahko konfigurirate nadzor zaščite na ravni dokumenta. Nadzor dostopa na ravni dokumenta zagotavlja, da rezultati iskanja vsebujejo samo dokumente, ki jih lahko vidi uporabnik s pooblastili, ki je predložil zahtevo za iskanje.

Sistem iskanja v podjetju podpira veliko pristopov za konfiguriranje nadzora zaščite na ravni dokumenta:

- Pred dodajanjem v indeks se dokumenti lahko prefiltrirajo in povežejo z žetoni zaščite.
- Pri nekaterih tipih podatkov lahko rezultate iskanja po-filtrirate, da potrdite uporabnikove poverilnice prijave na osnovi trenutnih podatkov nadzora zaščite. Komponenta za upravljanje identitete iskanja v podjetju lahko šifrira različne poverilnice, ki jih uporabniki potrebujejo za dostop do različnih repozitorijev, in shrani šifrirane poverilnice v profile. Če so iskani viri zaščiteni z izdelkom, ki zagotavlja zaščito enotne prijave (SSO), lahko komponenta za upravljanje identitete krmili dostop do dokumentov, ne da bi od uporabnikov zahtevala izdelavo profilov.
- Pri večini tipov pajka lahko za povezavo žetonov zaščite z dokumenti v indeksu uporabite prilagojen razred (dodatek) Java.
- Pri dokumentih, ki jih pajka spletni pajek, lahko iz indeksa izključite sidrno besedilo v dokumentih, ki vsebujejo povezave do nedovoljenih dokumentov.

#### S tem povezani pojmi

-  [Application security](#)



## Pred- in po-filtriranje rezultatov iskanja

Obstajata dva različna pristopa k filtriranju dokumentov, da se zagotovi, da rezultati iskanja vsebujejo samo dokumente, ki jih lahko vidi uporabnik s pooblastili, ki je predložil zahtevo za iskanje.

- Prvi pristop podvoji izvorne sezname za nadzor dostopa dokumenta (ACL-i) v indeks ob pajkanju in se zanaša na iskalnik, da primerja uporabniške poverilnice z ACL-i dvojnika dokumenta. Predfiltriranje dokumentov in nadzor nad tem, kateri dokumenti se dodajo v indeks, privedeta do najboljših rezultatov. Kljub temu je težavno oblikovati vsa načela zaščite različnih virov v ozadju v indeksu in vpeljati primerjalno logiko na enoten način. Ta pristop prav tako ni tako odziven na spremembe, ki se lahko pojavijo v izvornih ACL-jih.
- Drugi pristop po-filtrira dokumente v nizu rezultatov tako, da upošteva vire v ozadju za trenutne podatke zaščite. Ta pristop omogoča, da so sodelujoči viri v ozadju končni razsodniki dokumentov, vrnjenih uporabniku, in zagotavlja, da niz rezultatov odraža trenutni nadzor dostopa. Ta pristop zmanjša zmogljivost iskanja, saj zahteva, da imajo vsi viri v ozadju povezave. Če vir ni dostopen, je potrebno povezave do dokumentov filtrirati iz niza rezultatov skupaj z dokumenti, za ogled katerih uporabnik nima pooblastil.

**Pomembno:** Pri konfiguraciji z več strežniki se po-filtriranje za nekatere tipe virov izvede na strežniku pajkanja. Če strežnik pajkanja ne deluje zaradi vzdrževalnih del, uporabniki pri poizvedovanju zbirk iskanja v podjetju ne naletijo na rezultate. Prav tako ni rezultatov, če strežniki v ozadju, ki morajo nadzorovati dostop, niso dostopni.

Pri iskanju v podjetju temelji podpora za uveljavljanje nadzora dostopa na kombinaciji teh dveh pristopov. Zasnova zagotavlja najboljšo zmogljivost in ohranja točna načela zaščite izvornih repozitorijev dokumentov. Sistem lahko s tem, ko shrani podatke za nadzor dostopa v indeks, zagotovi začasni (potencialno manjši) niz podatkov, ki ga zatem lahko po-filtrirate, da preverite trenutni nadzor dostopa. V veljavi je predpostavka, da v primeru, ko ima uporabnik dostop do repozitorija, ki je lastnik dokumenta, obstajajo možnosti, da ima uporabnik dostop tudi do dokumenta.

Podatki za nadzor dostopa, ki so shranjeni v indeksu, se razlikujejo glede na tip pajka. Tako lahko na primer pajek Notes shrani nadzor dostopa na ravni baze podatkov in strežnika, pajek QuickPlace pa nadzor dostopa za strežnike, območja in sobe.

Vsi tipi virov podatkov v sistemu iskanja v podjetju podpirajo med pajkanjem zmožnost indeksiranja izvornih seznamov za nadzor dostopa. Nekateri tipi virov podatkov prav tako podpirajo zmožnost po-filtriranja nizov rezultatov in preverjanja uporabnikovih trenutnih poverilnic (ta tip podpore se zagotavlja prek izvornih mehanizmov zaščite ali komponente za upravljanje identitete iskanja v podjetju).

Ta dvojna varnostna zasnova vključuje naslednje naloge:

- Povzemanje izvornih informacij ACL med pajkanjem.
- Shranjevanje informacij ACL o strežniku in bazi podatkov v indeks.
- Izdelovanje uporabnikovega konteksta zaščite ob prijavi uporabnika ali ob inicializaciji seje. Ta naloga mora upoštevati različne identifikatorje, ki jih mora posamezni uporabnik uporabiti, da lahko dostopa do različnih virov v ozadju.
- Obdelava iskanja z uporabnikovim kontekstom zaščite in izdelovanje začasnega niza rezultatov, ki vsebuje samo tiste dokumente, do katerih ima uporabnik dostop na ravni repozitorija.
- Po-filtriranje začasnega niza rezultatov z upoštevanjem virov v ozadju, ki so prispevali dokumente v niz rezultatov za trenutne izvirne informacije ACL-

## Potrjevanje s shranjenimi žetoni zaščite

Če je ob izdelavi zbirke zaščita omogočena, lahko konfigurirate nadzor zaščite na ravni dokumenta tako, da shranite podatke zaščite v indeks.

Po privzetku je vsakemu dokumentu dodeljen javni žeton, ki poskrbi, da je dokument na voljo vsem. Če je zaščita za zbirko omogočena, lahko javni žeton zamenjate z vrednostjo, ki jo poda skrbnik ali z vrednostjo, ki je povzeta iz polja v pajkanem dokumentu. Ko konfigurirate pajka, podate, da želite za omejevanje tega, kateri uporabniki lahko dostopajo do dokumentov, ki jih pajka pajek, uporabljati žetone zaščite.

Ko skrbnik zbirke konfigurira pajka, lahko skrbnik poda možnosti zaščite za posamezne tabele, datotečne sisteme in tako naprej (za različne vire podatkov v območju pajkanja lahko konfigurirate različna pravila zaščite). Skrbnik lahko:

- Poda, da so dokumenti javni (vsi uporabniki lahko iščejo po dokumentih)
- Vsakemu dokumentu dodeli uporabniško definirane žetone zaščite
- Povzame žetone zaščite iz polja v pajkanih podatkih in dodeli povzeti žeton vsakemu dokumentu

Žetoni zaščite (z izjemo privzetega javnega žetona) so v celoti uporabniško definirani. Žeton zaščite lahko predstavlja uporabniški ID, ID skupine, uporabniško vlogo ali katerokoli drugo vrednost, za katero določite, da je veljavna za vir podatkov.

Skrbnik lahko na primer poda, da se za nadzor dostopa do dokumentov, ki jih pajka pajek Notes, uporabi polje `hrDeptName`. Skrbnik lahko prav tako poda, da se, če to polje v dokumentu ne obstaja ali če ne vsebuje podatkov zaščite, za nadzor dostopa do dokumentov uporabi uporabniško definirana žetona `hrgroup1` in `hrgroup2`.

Žetoni zaščite so pajku na voljo prek pajkove konfiguracijske datoteke. Pajek poda za vsak dokument vrednost žetona zaščite kot metapodatke. Komponenta indeksiranja prebere žeton zaščite in ga uveljavi na informacije o objavi za dokument v indeksu. Če skrbnik za izvorni vir podatkov posodobi seznam nadzora dostopa, bo posodobljen nadzor zaščite na voljo ob naslednji gradnji glavnega indeksa ali delta indeksa.

Uveljavite lahko prilagojena poslovna pravila, da s kodiranjem pravil v razredu Java določite vrednost žetonov zaščite. Ko konfigurirate lastnosti pajka, podate ime dodatka, ki ga naj uporablja pajek, ko pajka dokumente. Žetoni zaščite, ki jih doda vaš dodatek, so shranjeni v indeksu in jih lahko uporabite za nadzor dostopa do dokumentov.

## Kako iskalne aplikacije uporabljajo žetone zaščite

Za zagotovitev žetonov zaščite ob iskanju je odgovorna iskalna aplikacija odjemalca. Tako so lahko dokumenti ustrezno filtrirani. Če ni podan noben žeton zaščite, potem se bo med obdelavo iskanja samodejno uveljavil privzeti javni žeton.

Vzorčna iskalna aplikacija za iskanje v podjetju prikazuje, kako lahko vpeljete zaščito na ravni dokumenta. V tem primeru predpostavimo, da je skrbnik dodelil vrednost žetona zaščite skupini dokumentov (v nasprotju z povzemanjem žetona zaščite iz polja v pajkanem dokumentu). Iskalna aplikacija uporablja za določitev dokumentov, do katerih lahko dostopa uporabnik, uporabnikov ID prijave. Iskalna aplikacija ne uporablja dejanskega uporabniškega ID-ja, ampak se zanaša na ID skupine, kateri uporabnik pripada. Z uporabo ID-ja skupine kot žetona zaščite lahko uporabnike dodajate v skupino ali odstranjujete iz nje, ne da bi morali ponovno zgraditi indeks.

Žeton zaščite, ki ga je skrbnik dodelil nizu dokumentov, predstavlja veljavni ID skupine operacijskega sistema. Različnim dokumentom se v območju pajkanja dodelijo različni ID-ji skupine. Primer:

Dokument1-5: Žeton zaščite = Skupina1

Dokument6-10: Žeton zaščite = Skupina2

## Potrjevanje trenutnih poverilnic med obdelavo poizvedbe

Če je zaščita za zbirko omogočena, ko je ta izdelana, vam določeni tipi domen omogočajo, da potrdite uporabnikove trenutne poverilnice, ko uporabnik predloži poizvedbo.

Pred odzivom na poizvedbo se iskalni strežniki povežejo prek vmesnika z izvirnimi repozitoriji, da potrdijo uporabnikova trenutna dovoljenja in nato odstranijo vse dokumente, za katere uporabnik nima pooblastil, da si bi jih ogledoval v iskalnih rezultatih.

Ko konfigurirate naslednje tipe pajkov, lahko izberete možnost, s katero preverite poverilnice uporabnika tako, da primerjate poverilnice s trenutnim nadzorom dostopa, ki ga upravlja izvorni repozitorij. Po pajkanju in indeksiranju dokumentov se uporabi komponenta za upravljanje iskanja v podjetju, ki potrdi uporabnike, ki skušajo iskati po zaščiteneh zbirkah.

- Pajek Content Edition (samo tipi repozitorijev Documentum, FileNet Panagon Content Services, Hummingbird DM, Portal Document Manager in SharePoint)
- Pajek DB2 Content Manager
- Pajek Domino Document Manager
- Pajek Notes
- Pajek QuickPlace
- Pajek datotečnega sistema Windows

Za naslednje tipe pajkov je mogoče trenutne uporabniške poverilnice preveriti takrat, ko uporabniki s portalskim programčkom za iskanje v portalu WebSphere Portal preiskujejo zbirke iskanja v podjetju.

- Pajek Web Content Management
- Pajek WebSphere Portal

### S tem povezani pojmi

“Zagotavljanje zaščite na ravni dokumenta za dokumente Lotus Domino” na strani 256

“Zagotavljanje zaščite na ravni dokumenta za dokumente datotečnega sistema Windows” na strani 259

### S tem povezana opravila

“Konfiguriranje overjenih strežnikov Lotus Domino za preverjanje poverilnic uporabnikov” na strani 257

## Upravljanje identitete iskanja v podjetju

Upravljanje poverilnic za več uporabnikov je pogosta težava podjetja. Sistem iskanja v podjetju reši težavo tako, da priskrbi izbirno komponento za upravljanje identitete.

### Nasvet:

Podrobnejše informacije o načinih nastavitve komponente za upravljanje identitete, ki je primerna za iskanje v podjetju, poiščite v knjigi IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Informacije v podjetju lahko obstajajo v različnih oblikah. Lahko so porazdeljene po podjetju in upravljane z najprimernejšo programsko opremo za nalogo, ki je na voljo. Uporabniki

podjetja lahko uporabijo aplikacijo SQL za dostop do relacijskih baz podatkov ali sistem za upravljanje dokumentov za dostop do dokumentov, ki so pomembni za njihovo delo.

Krmiljenje dostopa do občutljivih informacij v teh repozitorijih običajno zagotavlja upravljajoča programska oprema. Uporabniki se na gostiteljskem sistemu identificirajo z uporabniškim ID-jem in geslom. Potem, ko jih sistem overi, upravljajoča programska oprema krmili, katere dokumente uporabnik lahko vidi in vanje posega, na osnovi definirane pravice dostopa za uporabnika.

Uporabniki imajo pogosto različne uporabniške ID-je in gesla, povezana z vsakim repozitorijem. Podobno, kot se morajo uporabniki identificirati izvornim repozitorijem podjetja, morajo uporabniki tudi zagotoviti poverilnice, preden si ogledajo dokumente v zbirki iskanja v podjetju, ki zahtevajo potrditev trenutnih poverilnic. Uporabniki z več identitetami morajo predložiti ustrezne poverilnice za vsako identiteto.

Če podate, da želite za upravljanje identitete na skrbniški konzoli uporabljati iskanje v podjetju, lahko strežniki iskanja za potrditev uporabnikovih trenutnih poverilnic med obdelavo poizvedb uporabijo naslednje pristope:

- Iskalna aplikacija lahko pozove uporabnika, da registrira poverilnice, ki jih potrebuje za dostop do različnih domen v uporabniškem profilu. Profil, ki je šifriran in shranjen v zaščiteni shrambi podatkov, omogoča uporabniku iskanje zaščitenih domen. Če za domeno, ki zahteva preverjanje veljavnosti trenutnih poverilnic, le-teh ne podate, bodo dokumenti iz te domene izključeni iz rezultatov iskanja.
- Če je dokumente v zbirki pajkal pajek, ki nudi podporo za zaščito z enotno prijavo (SSO), in vi podate, da želite za nadzor dostopa do dokumentov uporabljati zaščito z enotno prijavo, bo sistem za overjanje uporabnikov med trajanjem seje iskanja uporabljal metode zaščite z enotno prijavo. Pri iskanju zaščitenih domen uporabnik ne rabi izdelati profila, ki podaja poverilnice, ali podati uporabniškega ID-ja in gesla.

Ko uporabniki iščejo po zbirkah, ki po predložitvi poizvedbe zahtevajo preverjanje veljavnosti trenutnih poverilnic, bo sistem z uporabo tega profila ali metod enotne prijave zavrnil ali odobril dostop do dokumentov.

## **Pridobivanje uporabnikovih informacij o skupinah**

Za potrditev uporabnikovih poverilnic mora upravljanje identitete pridobiti uporabnikove informacije o skupinah za vsako uporabnikovo identiteto in te informacije dodati nizu konteksta uporabniške zaščite (USC). Te informacije o skupinah se uporabljajo za filtriranje rezultatov v skladu z podatki za nadzor dostopa, ki so shranjeni v indeksu za iskanje v podjetju ali v skladu s podatki za overjanje z enotno prijavo. Komponenta za upravljanje identitete to stori tako, da se s pomočjo žetonov enotne prijave ali z uporabo uporabnikovih poverilnic poveže na sistem v ozadju in zahteva skupine, katerih član je uporabnik.

Ko na skrbniški konzoli konfigurirate možnosti upravljanja identitete, lahko podate, kako pogosto se naj te informacije o skupinah osvežijo. Nove podatke o skupinah lahko povzamete vsakič, ko se uporabnik prijavi v iskalno aplikacijo, lahko pa tudi redno, na primer vsake tri dni.

## **Zaščita brez komponente za upravljanje identitete**

Vsa podjetja ne želijo upravljati več identitet svojih uporabniških skupnosti s komponento za upravljanje identitete iskanja v podjetju. Če na skrbniški konzoli iskanja v podjetju onemogočite komponento za upravljanje identitete, potem je za generiranje niza konteksta uporabniške zaščite odgovorna iskalna aplikacija. Po generiranju se za nastavitev vrednosti omejitev ACL na vsaki poizvedbi uporabi niz USC. Primer:

```
Query q = factory.createQuery("IBM");
q.setACLConstraints("Kontekst uporabniške zaščite v XML");
```

**Nasvet:** Kot pomoč pri pisanju vaše lastne funkcionalnosti za upravljanje identitete je na voljo programski nadzor nad bazo podatkov za upravljanje identitete, ki ga nudi razširitev API-ja za iskanje in indeksiranje (SI-API). Ta API vam omogoča, da generirate USC z objekti Java - niz XML bo tako samodejno zgrajen.

Niz poizvedbe XML mora imeti naslednjo obliko, kjer ... vsebuje v celoti oblikovan niz XML:

```
@SecurityContext::'...'
```

Format niza XML je tak:

```
<identities id="login_UserName">
 <ssoToken>token_value</ssoToken>
 <identity id="security_domain">
 <type>Notes</type>
 <username>domain_UserName</username>
 <password encrypt="no">domain_userPW</password>
 <groups>
 <group id="g1" />
 <group id="g2" />
 </groups>
 <properties>
 <property name="property_name">property_value</property>
 ...
 </properties>
 </identity>
 ...
</identities>
```

### **identities**

Vrednost atributa **id** je uporabniški ID, ki ga posreduje uporabnik ob prijavi v sistem.

### **ssoToken**

Neobvezno: Podaja žeton Enostavno overjanje neodvisnih strank (LTPA), ki se izdelava za uporabnika za čas trajanja seje brskalnika. Ta parameter uporabite samo, če je ciljna domena omogočena za enotno prijavo in je pajek konfiguriran za uporabo zaščite z enotno prijavo.

### **identity**

Vsebuje uporabnikove poverilnice za določen vir podatkov. Vrednost atributa **id** je domena, v kateri so shranjeni informacije o uporabnikovih poverilnicah (če uporabljate Domino, je to ime domene Domino).

### **type**

Identificira tip podatkov in ustreza tipu pajka (Notes, DB2, Exchange Server itd.).

### **username**

Podaja uporabniško ime, ki bo uporabljeno za iskanje v domeni.

### **password**

Podaja geslo za podano uporabniško ime. Atribut **encrypt** mora biti nastavljen na **no** (iskanje v podjetju ne nudi metode šifriranja zunaj komponente za upravljanje identitete).

### **groups**

Podaja imena skupin, ki jim uporabnik pripada. Za vsako ime skupine se uporabi ločen element skupine.

### **properties**

Podaja seznam lastnosti s posebnimi povezavami, kot sta ID in šifrirano geslo skrbnika, uporabljena za izdelavo pajka, ali to, ali je omogočena enotna prijavo za vir.

**property\_name**

Ime lastnosti.

**property\_value**

Vrednost lastnosti.

## Potrditev uporabnika z uporabniškimi profili

Iskalne aplikacije lahko pozovejo uporabnike, da registrira poverilnice, ki jih potrebuje za dostop do različnih domen v uporabniškem profilu.

Za iskanje po domeni, ki zahteva potrditev uporabniških poverilnic ob predložitvi poizvedbe, morajo uporabniki iskalni aplikaciji zagotoviti poverilnice, ki jih uporabljajo za prijavo v domeno. Z upravljanjem identitet iskanja v podjetju lahko uporabniki shranijo poverilnice za neomejeno število domen v uporabniškem profilu. Poverilnice so šifrirane in varno shranjene v sistemu iskanja v podjetju.

Če za domeno, ki zahteva preverjanje veljavnosti trenutnih poverilnic, le-teh ne podate, bodo dokumenti iz te domene izključeni iz rezultatov iskanja.

Uporabniki lahko izdelajo uporabniški profil in registrirajo svoje poverilnice, ko uporabljajo iskalno aplikacijo. V vzorčni iskalni aplikaciji za iskanje v podjetju to zmožnost nudi možnost **Moj profil**. Vaše prilagojene iskalne aplikacije lahko vpeljejo to zmožnost na različne načine.

Zbirke lahko vsebujejo dokumente iz več tipov virov. Zbirka lahko na primer vsebuje dokumente, ki so bili pajkani iz datotečnega sistema Windows in več baz podatkov Lotus Notes. Komponenta za upravljanje identitet razlikuje med različnimi tipi virov in zahteva samo poverilnice, ki so potrebne za dostop do domen, ki zahtevajo potrditev.

Po privzetku je vsaka poverilnica omogočena za iskanje in tako zahteva od uporabnika, da poda uporabniški ID in geslo, ki ustreza zaščitenim domenam. Če je uporabnik pozabil uporabniški ID ali geslo za določeno domeno, lahko domeno onemogočite za iskanje tako, da pustite potrditveno polje prazno. Onemogočanje domene prepreči, da bi bili zaščiteni dokumenti v teh domenah vrnjeni v nizu rezultatov.

Po izdelavi profila lahko uporabnik predloži zahtevo za iskanje. Komponenta za upravljanje identitete ima potrebne informacije za gradnjo uporabnikovega niza konteksta zaščite (USC), ki se bo uporabil pri naslednjih zahtevah za iskanje. Če ne uporabite komponente za upravljanje identitete, mora iskalna aplikacija podati niz USC, ko uporabniki izdajo poizvedbo za domene, ki zahtevajo potrditev trenutnih poverilnic.

Ko uporabnik naslednjč skuša iskati po zbirkah iskanja v podjetju, komponenta za upravljanje identitete ponovi proces preverjanja poverilnic, a tokrat lahko najde uporabnikov profil. Če se nič ne spremeni, se uporabnika samodejno vrne v položaj, kjer lahko predloži zahteve za iskanje, ne da bi moral izdelati profil.

Če komponenta za upravljanje identitete zazna spremembo na katerikoli izmed uporabnikovih poverilnic, se bo ob dostopu do iskalne aplikacije uporabniku samodejno prikazala stran profila. To se na primer zgodi ob spremembah gesel za katerokoli domeno, ki je omogočena za iskanje, ali ko se domena, ki zahteva overjanje, doda v zbirko.

Uporabniki ne rabijo upoštevati priporočila za posodobitev profila, vendar bo to povzročilo, da bodo ti dokumenti izključeni iz rezultatov iskanja.

V vzorčni iskalni aplikaciji, ki je na voljo z iskanjem v podjetju, lahko uporabniki kadarkoli posodobijo profile tako, da v orodni vrstici izberejo **Moj profil**.

## Potrditev uporabnika z zaščito z enotno prijavo

Če je dokumente v zbirki pajkal pajek, ki nudi podporo za zaščito z enotno prijavo (SSO), lahko podate, da želite za nadzor dostopa do dokumentov uporabljati zaščito z enotno prijavo, ko konfigurirate možnosti za upravljanje identitete.

### Omogočanje enotne prijave

Overjanje z enotno prijavo omogoča uporabniku, da je overjen enkrat in da pridobi dostop do mnogih virov, ne da bi moral znova podati poverilnice. V sistemu iskanja v podjetju overjanje z enotno prijavo zmanjša breme upravljanja več uporabniških imen in gesel, ki jih morajo uporabniki podati za dostop do dokumentov v zaščiteneh zbirkah.

IBM aplikacijski strežnik WebSphere Application Server in Lotus Domino podpirata obliko enotne prijave, ki je znana kot Enostavno overjanje neodvisnih strank (LTPA). Ko skuša uporabnik dostopati do enega izmed izdelkov, se mora overiti z uporabniškim imenom in geslom. To uporabniško ime in geslo se preverita za repozitorij LDAP, ki ga oba souporabljata. Po overitvi uporabnika se izdelava piškotek seje, ki vsebuje žeton LTPA. Uporabnik lahko nato dostopa do drugih virov na kateremkoli strežniku, ki ima isto konfiguracijo overjanja, ne da bi moral ponovno podati poverilnice. Ta žeton ostane, vse dokler je seja brskalnika veljavna.

Za omogočanje podpore enotne prijave za uporabo z zbirkami iskanja v podjetju:

- Prepričajte se, da sta na iskalnih strežnikih za iskanje v podjetju omogočena globalna zaščita aplikacijskega strežnika WebSphere Application Server in veljavni register LDAP. Register LDAP je lahko katerikoli veljavni izdelek LDAP, ki ga podpira aplikacijski strežnik WebSphere Application Server.
- Prepričajte se, da je mehanizem overjanja WebSphere konfiguriran za uporabo aktivnega mehanizma overjanja LTPA. Ko konfigurirate LTPA, podajte veljavno, a prilagodljivo ime domene, kot na primer `your.server.com`.
- Prepričajte se, da je bil ključ LTPA izvožen z aplikacijskega strežnika WebSphere Application Server in uvožen v druge izdelke v isti domeni, na katerih želite omogočiti podporo za LTPA.

Potem, ko z brskalnikom preverite, da zgornja konfiguracija zaščite deluje pravilno, lahko uporabite skrbniško konzolo iskanja v podjetju, da konfigurirate pajke, ki podpirajo overjanje z enotno prijavo.

### Enotna prijava in upravljanje identitet

Ko uporabniki iščejo po zbirkah, ki zahtevajo potrditev trenutnih poverilnic, bo sistem z uporabo metod zaščite z enotno prijavo zavrnil ali odobril dostop do dokumentov. Uporabnikom ob iskanju virov, ki podpirajo overjanje z enotno prijavo, ni potrebno podati poverilnic. Komponenta za upravljanje identitete se uporabi, če velja vse od naslednjega:

- Enotna prijava je ustrezno omogočen na aplikacijskem strežniku WebSphere Application Server in ciljnih domenah.
- Zaščita je omogočena v vsaj eni od zbirk, ki jih lahko preišče iskalna aplikacija
- Možnosti za uporabo komponente za upravljanje identitete in zaščite z enotno prijavo so omogočene na skrbniški konzoli iskanja v podjetju.
- Možnost za uporabo zaščite z enotno prijavo in možnosti za zagotavljanje zaščite na ravni dokumenta (kot na primer indeksiranje nadzora dostopa ali potrjevanje trenutnih poverilnic med obdelavo poizvedb) so bile izbrane, ko so bili konfigurirani naslednji tipi pajka:
  - Content Edition (na voljo samo za repozitorije Portal Document Manager-ja)
  - Domino Document Manager (na voljo za pajke, ki uporabljajo protokol DIIOP)
  - Notes (na voljo za pajke, ki uporabljajo protokol DIIOP)

- QuickPlace (na voljo za pajke, ki uporabljajo protokol DIOP)

## Zaščita portalskega programčka za iskanje

Če uporabniki s portalskim programčkom za iskanje preiskujejo iskalne zbirke s portala WebSphere Portal, je zaščita omogočena tudi za dokumente, ki jih pajkajo pajki seznama semen, Web Content Management in WebSphere Portal.

Zaščiteno iskanje je podprto za navedene tipe virov samo, če uporabite portalski programček, ne pa iskalne aplikacije, ki se izvaja izven portala WebSphere Portal. V portalu WebSphere Portal se uporabniške poverilnice pridobijo iz portalskega programčka za iskanje. Potem ko se uporabnik prijavi v WebSphere Portal, vse iskalne zahteve vključujejo uporabnikove podatke o zaščiti (uporabniško ime, članstva v skupinah in podobno). Ker so te informacije vedno na voljo, enotna prijava na osnovi žetonov LTPA ni potrebna.

Če za zaščito spletnih mest in dokumentov na strežniku portala WebSphere Portal uporabite drug izdelek, morate pri konfiguriranju pajka podati možnosti enotne prijave. Če na primer uporabite izdelek, kot je IBM Tivoli Access Manager WebSEAL ali CA SiteMinder SSO Agent for PeopleSoft, morate podati poverilnice, ki pajku omogočijo dostop do dokumentov na strežniku prek enotne prijave. V tem primeru omogočate enotno prijavo za dostop pajka do zaščitene vsebine, ne pa enotne prijave za zaščiteno iskanje.

## Konfiguriranje upravljanja identitet

Da podate, kako naj bodo uporabniške poverilnice potrjene med obdelovanjem poizvedbe, lahko uporabite komponento za upravljanje identitet.

### Praden začnete

Možnosti za upravljanje identitet lahko konfigurirajo samo člani skrbniške vloge za iskanje v podjetju.

### O tej nalogi

Ko uporabniki iščejo po zbirkah, ki med obdelavo poizvedbe zahtevajo preverjanje veljavnosti trenutnih poverilnic, lahko komponenta za upravljanje identitet za iskanje v podjetju uporabi metodi zaščite z uporabniškim profilom ali enotno prijavo (SSO), da zavrne ali odobri dostop do dokumentov.

### Postopek

Za konfiguriranje upravljanja identitet:

1. Kliknite **Zaščita**, da odprete pogled Zaščita.
2. Na strani Iskanje kliknite **Konfiguriranje upravljanja identitet**.
3. Na strani Konfiguriranje upravljanja identitet izberite potrditveno polje za uporabo komponente za upravljanje identitet iskanja v podjetju, ki krmili potrjevanje uporabniških poverilnic med obdelovanjem poizvedbe. Če to potrditveno polje ni izbrano, mora iskalna aplikacija v primeru, ko uporabnik izda poizvedbo za domene, ki pri obdelavi poizvedbe zahtevajo preverjanje veljavnosti trenutnih poverilnic, podati niz uporabniškega konteksta zaščite (USC).
4. Podajte, kako pogosto naj komponenta za upravljanje identitet povzame uporabniške poverilnice iz skupinskih zapisov v uporabniškem registru aplikacijskega strežnika WebSphere Application Server. Podatke o poverilnicah lahko osvežite vsakič, ko uporabnik dostopa do iskalne aplikacije, ali pa po preteku števila dni, ki ste ga podali.
5. Za tipe pajkov, ki podpirajo overjanje z enotno prijavo, podajte, ali naj komponenta za upravljanje identitet pri potrjevanju uporabnikov uporabi žetone zaščite z enotno prijavo



namesto uporabniških profilov. Izberete lahko potrditveno polje za uporabo overjanja z enotno prijavo pri vseh tipih pajka ali pa izberete potrditvena polja za posamezne tipe pajkov.

**Pomembno:** Komponenta za upravljanje identitet uporablja mehanizme zaščite z enotno prijavo samo, če je zaščita z enotno prijavo ustrezno konfigurirana na aplikacijskem strežniku WebSphere Application Server in ciljnih domenah.

## Analiza sidrnega besedila

Če omogočite zaščito zbirke, procesi globalne analize uveljavijo posebna pravila za indeksiranje sidrnega besedila v dokumentih, ki jih pajkajo spletni pajki. Če zaščite zbirke ne omogočite, lahko podate, ali želite indeksirati sidrno besedilo v povezavah na nedovoljene dokumente, ko konfigurirate posamezne spletne pajke.

Sidrno besedilo je informacija znotraj hiperbesedilne povezave, ki opisuje stran, na katero se povezava nanaša. V naslednji povezavi je na primer besedilo Query Syntax sidrno besedilo v povezavi, ki se nanaša na stran syntax.htm:

```
Query Syntax
```

Spletni pajek običajno obiše povezave v dokumentih, da lahko pajka dodatne dokumente in vključi te strani s povezavami v indeks. Med globalno analizo procesi indeksa povežejo sidrno besedilo ne samo z dokumentom, v katerega je vdelano (virni dokument), ampak tudi s ciljnim dokumentom. V zgornjem primeru je sidrno besedilo Query Syntax povezano s ciljno stranjo syntax.htm in z izvorno stranjo, ki vsebuje sidrni konstrukt. Ta povezava omogoča, da se ciljni dokument priključi s poizvedbami, ki podajajo besedilo, ki se pojavi v izvornem dokumentu. Povezava predstavlja varnostno tveganje, če lahko uporabniki gledajo ciljni dokument, ne pa izvirnega.

Če ob izdelavi zbirke omogočite zaščito zbirke, je obdelovanje sidrnega besedila onemogočeno. Sidrno besedilo se več ne indeksira z dokumentom, razen če se dejansko pojavi v dokumentu ali v metapodatkih dokumenta. Ta varnostni nadzor zagotavlja, da uporabniki ne vidijo informacij v dokumentih, do katerih ne smejo dostopati; dokument je vrnjen samo v rezultatih iskanja, če se njegova vsebina ali metapodatki ujemajo s poizvedbo.

S tem, ko omogočite zaščito zbirke, lahko izboljšate zaščito spletnih dokumentov, saj omogočite uporabnikom, da iščejo samo po dokumentih z žetoni zaščite, ki se ujemajo z njihovimi poverilnicami. Če pa rezultati iskanja ne obdelajo sidrnega besedila, ti morebiti ne bodo vključevali vseh dokumentov, ki so potencialno koristni za poizvedbo.

Če zaščite zbirke ne omogočite, lahko podate, ali želite indeksirati sidrno besedilo v povezavah na nedovoljene dokumente, ko konfigurirate napredne lastnosti spletnega pajka.

## Indeksiranje sidrnega besedila v povezavah na nedovoljene dokumente

Če dokument vsebuje povezave na dokumente, ki jih spletni pajek ne sme pajkati, lahko podate, ali želite obdržati sidrno besedilo za te povezave v indeksu, ko konfigurirate spletnega pajka.

### Preden začnete

Možnosti lahko konfigurirate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke za spletnega pajka, ki ga želite konfigurirati.


### O tej nalogi

Napotki v datoteki robots.txt ali v metapodatkih spletnih dokumentov lahko preprečijo, da bi spletni pajek dostopal do dokumentov na spletnem mestu. Če dokument, ki ga spletni pajek sme pajkati, vsebuje povezave na nedovoljene dokumente, lahko podate, kako želite obravnavati sidrno besedilo za te povezave.

Podate lahko, ali želite indeksirati sidrno besedilo na nedovoljene dokumente, ko konfigurirate spletnega pajka. Za največjo zaščito podajte, da ne želite indeksirati sidrnega besedila v povezavah na nedovoljene dokumente. A če ne indeksirate sidrnega besedila, rezultati iskanja morebiti ne bodo vključevali vseh dokumentov, ki so potencialno koristni za poizvedbo.

### Postopek

Za omogočanje ali onemogočanje indeksiranja sidrnega besedila v povezavah na nedovoljene dokumente:

1. Uredite zbirko in na strani Pajkanje poiščite spletnega pajka, ki ga želite konfigurirati ter kliknite  **Lastnosti pajka**.
2. Kliknite **Uredi zahtevnejše lastnosti spletnega pajka**.
3. Za indeksiranje sidrnega besedila v vseh dokumentih, ki jih pajek pajka, izberite potrditveno polje **Indeksiraj sidrno besedilo v povezavah na nedovoljene dokumente**. Uporabniki se bodo lahko poučili o straneh, ki jih spletni pajek ne sme pajkati, tako, da bodo iskali besedilo, ki je v sidrnem besedilu povezav, ki kažejo na te strani.  
Da sidrno besedilo v povezavah na nedovoljene dokumente izločite iz indeksa, počistite to potrditveno polje. Uporabniki se ne bodo mogli poučiti o straneh, ki jih spletni pajek ne sme pajkati. Sidrno besedilo bo izločeno iz indeksa, skupaj s nedovoljenimi dokumenti.
4. Kliknite **Potrdi** in nato na strani Lastnosti spletnega pajka ponovno kliknite **Potrdi**.
5. Če želite uveljaviti spremembe, zaustavite in znova zaženite pajka.

Če želite uveljaviti spremembe za dokumente, ki so bili predhodno indeksirani, jih morate ponovno pajkati, da bodo lahko znova indeksirani. Če je predhodno pajkanje dodalo v indeks informacije o nedovoljenih dokumentih, bodo te informacije izbrisane iz indeksa.

---

## Omogočanje zaščite za iskanje v podjetju

Če nameravate zagotavljati zaščito, ko uporabni, ki izvajajo skrbništvo nad sistemom iskanja v podjetju ali po njem iščejo, morate konfigurirati globalno zaščito na aplikacijskem strežniku WebSphere Application Server. Prav tako morate konfigurirati informacije o zaščiti v konfiguracijskih datotekah iskanja v podjetju, na skrbniški konzoli in iskalnih aplikacijah.

### Nasvet:

Podrobnejše informacije o načinih omogočitve globalne zaščite v WebSphere Application Server z repozitorijem LDAP, vključno s primeri konfiguriranja različnih pajkov z namenom uveljavitve zaščite na ravni dokumentov, poiščite v publikaciji IBM Redbook, IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

### Postopek

Da omogočite zaščito za sistem iskanja v podjetju:

1. Odločite se, kateri tip uporabniškega registra želite uporabiti za overjanje uporabnikov. Veliko skrbnikov aplikacijskega strežnika WebSphere Application Server se na primer odloči uporabiti uporabniški register protokola LDAP (Lightweight Directory Access Protocol).

2. V skrbniški konzoli iskanja v podjetju izberite **Zaščita** in dodelite vsaj enega od uporabnikov v uporabniškem registru aplikacijskega strežnika WebSphere Application Server skrbniški vlogi **Skrbnik iskanja v podjetju**.

Lahko pa tudi dodate ID skrbnika iskanja v podjetju, ki je podan ob namestitvi OmniFind Enterprise Edition is v uporabniški register aplikacijskega strežnika WebSphere Application Server.

**Pomembno:** Ko je globalna zaščita omogočena, lahko samo uporabniški ID-ji, ki so v uporabniškem registru in katerim je dodeljena skrbniška vloga iskanja v podjetju, dostopajo do skrbniške konzole in izvajajo skrbništvo nad sistemom iskanja v podjetju.

3. Za omogočanje globalne zaščite in konfiguriranje uporabniškega registra sledite postopkom v “Konfiguriranje globalne zaščite in uporabniškega registra LDAP na aplikacijskem strežniku WebSphere Application Server”.
4. Če globalno zaščito omogočite po namestitvi izdelka OmniFind Enterprise Edition, morate v sistem iskanja v podjetju vnesti ID uporabnika in geslo za WebSphere Application Server. Te informacije lahko vnesete, če pred tem podate ukaz **eschangewaspw**:
  - Če ste OmniFind Enterprise Edition namestili na en sam strežnik, sledite postopku v “Omogočanje zaščite za sistem iskanja v podjetju za en sam strežnik” na strani 251.
  - Če ste OmniFind Enterprise Edition namestili na več kot en strežnik, sledite postopku v “Omogočanje zaščite za sistem iskanja v podjetju za več strežnikov” na strani 252.
5. Izpolnite ustrezne naloge za tipe dokumentov, ki jih nameravate pajkati in po njih iskati. Za podrobnosti glejte “Zahteve za nastavitve pajka za podporo zaščite” na strani 253.

## Konfiguriranje globalne zaščite in uporabniškega registra LDAP na aplikacijskem strežniku WebSphere Application Server

Da omogočite zaščito v OmniFind Enterprise Edition, morate najprej omogočiti globalno zaščito na aplikacijskem strežniku WebSphere Application Server.

### O tej nalogi

Kot del omogočanja globalne zaščite morate konfigurirati uporabniški register, da overite uporabniške ID-je. Ta naloga opisuje, kako konfigurirati uporabniški register protokola LDAP (Lightweight Directory Access Protocol) na aplikacijskem strežniku WebSphere Application Server ob istem času, ko omogočite globalno zaščito. Čeprav aplikacijski strežnik WebSphere Application Server podpira še druge tipe uporabniških registrov, za overjanje uporabnikov iskanja v podjetju ne morete uporabljati lokalnega registra operacijskega sistema. Za uporabo lokalnega registra operacijskega sistema morajo na iskalnih strežnikih za iskanje v podjetju obstajati uporabniški računi operacijskega sistema za vsakega uporabnika v vašem podjetju.

Ta naloga temelji na različici 6 aplikacijskega strežnika WebSphere Application Server. Če uporabite starejšo različico aplikacijskega strežnika WebSphere Application Server, so lahko privzete poti in nekatere oznake uporabniškega vmesnika drugačne. Ta naloga prav tako uporablja strežnik IBM Tivoli Directory za register LDAP. Če uporabljate drugačen tip registra ali prilagojen register, morate podati ustrezne informacije o vašem registru.

Ta naloga povzema korake, potrebne za konfiguriranje globalne zaščite za uporabo s sistemom iskanja v podjetju. Za podrobna navodila glejte informacijski center *različice 6.0.x aplikacijskega strežnika WebSphere Application Server* na naslednjem URL-ju:  
<http://publib.boulder.ibm.com/infocenter/wasinfo/v6r0/index.jsp>

### Postopek

Da omogočite globalno zaščito na aplikacijskem strežniku WebSphere Application Server:

1. Na iskalnem strežniku za iskanje v podjetju dostopite do naslednjega URL-ja, da odprete skrbniško konzolo aplikacijskega strežnika WebSphere Application Server, kjer je *localhost* localhost ali ime strežnika, kot na primer omnifind.search.xyz.com.  
`http://localhost:9060/ibm/console`
2. Kliknite **Zaščita** in nato **Globalna zaščita**.
3. Nastavite WebSphere za uporabo registra LDAP:
  - a. Pod **Uporabniški registri** kliknite **LDAP**.
  - b. Podajte uporabniški ID in geslo strežnika, ki se uporabljata za zagon aplikacijskega strežnika.
  - c. Za tip registra izberite strežnik IBM Tivoli Directory.
  - d. Podajte gostiteljsko ime strežnika LDAP, ki je naslov IP ali gostiteljsko ime storitve imena domene (DNS). Privzeta številka vrat je 389.
  - e. Podajte osnovno razločevalno ime (DN), ki je začetna točka za iskanje po registru, kot na primer `ou=sales,o=ibm,c=us`.
  - f. Ker nekateri strežniki LDAP ne podpirajo anonimnega povezovanja v času iskanja po registru, podajte DN za aplikacijski strežnik, kot na primer `cn=searchuser,o=ibm,c=us` in nato geslo za aplikacijski strežnik. Aplikacijski strežnik uporablja ta DN in geslo za povezovanje z registrom.
  - g. Za uporabo komuniciranja s plastjo zaščitenih vtičnic (SSL) med WebSphere-om in LDAP-om izberite potrditveno polje **SSL omogočen**.
  - h. Kliknite **Uveljavi** in nato **Potrdi**.
4. Pod **Splošne lastnosti** izberite potrditveni polji **Omogoči globalno zaščito** in **Zagotovi zaščito Java 2**.
5. Za mehanizem aktivnega overjanja izberite Preprosti mehanizem overjanja WebSphere (SWAM).
6. Za aktivni uporabniški register izberite uporabniški register protokola LDAP (Lightweight Directory Access Protocol).
7. Kliknite **Potrdi**.
8. Kliknite povezavo **Shrani** na vrhu strani. Ko vas računalnik pozove, da shranite spremembe, kliknite gumb **Shrani**.
9. V orodni vrstici kliknite **Odjava**.
10. Zaustavite in znova zaženite aplikacijo ESSearchServer.

**AIX,**

**Linux**

**ali Solaris**

```
./stopServer.sh ESSearchServer
```

```
./startServer.sh ESSearchServer
```

**Windows**

```
stopServer ESSearchServer
```

```
startServer ESSearchServer
```

Ti skripti se nahajajo v imeniku `WAS_INSTALL_ROOT/AppServer/bin`:

- Za aplikacijski strežnik WebSphere Application Server različice 5 je privzeta pot za namestitvev `/usr/WebSphere` na sistemih AIX, `/opt/WebSphere` na sistemih Linux ali Solaris in `C:\Program Files\WebSphere` na sistemih Windows.

- Za aplikacijski strežnik WebSphere Application Server različice 6 je privzeta pot za namestitvev /usr/IBM/WebSphere na sistemih AIX, /opt/IBM/WebSphere na sistemih Linux ali Solaris in C:\Program Files\IBM\WebSphere na sistemih Windows.
11. Znova zaženite skrbniško konzolo aplikacijskega strežnika WebSphere Application Server.
  12. Ker se sedaj strežnik zažene v zaščitenem načinu, vnesite uporabniški ID in geslo strežnika, ki ste ju podali, ko ste konfigurirali uporabniški register LDAP za prijavo v konzolo (glejte korak 3b na strani 250).

## Omogočanje zaščite za sistem iskanja v podjetju za en sam strežnik

Če globalno zaščito za strežnik WebSphere Application Server omogočite po namestitvi izdelka OmniFind Enterprise Edition, morate z ukazom **eschangewaspw** posodobiti konfiguracijsko datoteko iskanja v podjetju **es.cfg** in vnesti geslo uporabnika strežnika WebSphere Application Server.

### Preden začnete

Prepričajte se, da datoteka **config.properties** za aplikacijo **ESSearchApplication** poda veljavno uporabniško ime in geslo aplikacijskega strežnika WebSphere Application Server. Privzeta lokacija te datoteke je **ES\_INSTALL\_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war/WEB-INF**.

### O tej nalogi

Ukaz **eschangewaspw** geslo šifrira, preden ga shrani v datoteko **es.cfg**.

### Postopek

Da omogočite obstoječi sistem iskanja v podjetju za en sam strežnik, da bo uporabljal globalno zaščito:

1. Na strežniku iskanja v podjetju se prijavite kot skrbnik iskanja v podjetju in zaustavite sistem iskanja v podjetju:  
`esadmin system stopall`
2. Prepričajte se, da vnos **WASUser** v datoteki **ES\_NODE\_ROOT/nodeinfo/es.cfg** poda veljavno uporabniško ime aplikacijskega strežnika WebSphere Application Server.
3. Zaženite naslednji skript, kjer je **WAS\_password** geslo za uporabniško ime aplikacijskega strežnika WebSphere Application Server, ki je podano v datoteki **ES\_NODE\_ROOT/nodeinfo/es.cfg** (glejte korak 2).

#### AIX, Linux ali Solaris

```
eschangewaspw.sh WAS_password
```

#### Windows

```
eschangewaspw WAS_password
```

4. V okolju Windows izberite **Nadzorna plošča** → **Orodja za upravljanje** → **Storitve** in v storitvi WebSphere Application Server in **ESSearchServer** dodajte enaka uporabniško ime in geslo za strežnik WebSphere Application Server.
5. Ponovno zaženite sistem iskanja v podjetju:  
`esadmin system startall`

## Omogočanje zaščite za sistem iskanja v podjetju za več strežnikov

Če globalno zaščito za strežnik WebSphere Application Server omogočite po namestitvi izdelka OmniFind Enterprise Edition, morate z ukazom **eschangewaspw** posodobiti konfiguracijsko datoteko iskanja v podjetju **es.cfg** in vnesti geslo uporabnika strežnika WebSphere Application Server.

### Preden začnete

Prepričajte se, da datoteka **config.properties** za aplikacijo **ESearchApplication** poda veljavno uporabniško ime in geslo aplikacijskega strežnika WebSphere Application Server. Ta datoteka se po privzetku nahaja v imeniku **ES\_INSTALL\_ROOT/installedApps/ESearchApplication.ear/ESearchApplication.war/WEB-INF** na iskalnih strežnikih.

### O tej nalogi

Ukaz **eschangewaspw** geslo šifrira, preden ga shrani v datoteko **es.cfg**.

### Postopek

Da omogočite obstoječi sistem iskanja v podjetju za več strežnikov, da bo uporabljal globalno zaščito:

1. Na indeksnem strežniku iskanja v podjetju naredite naslednje:
  - a. Prijavite se kot skrbnik sistema za iskanje v podjetju in le-tega zaustavite:  
`esadmin system stopall`
  - b. Prepričajte se, da vnos **WASUser** v datoteki **ES\_NODE\_ROOT/nodeinfo/es.cfg** poda veljavno uporabniško ime aplikacijskega strežnika WebSphere Application Server.
  - c. Zaženite naslednji skript, kjer je **WAS\_password** geslo za uporabnika aplikacijskega strežnika WebSphere Application Server, ki je podano v datoteki **ES\_NODE\_ROOT/nodeinfo/es.cfg** (glejte korak 1b).

#### AIX, Linux ali Solaris

```
eschangewaspw.sh WAS_password
```

#### Windows

```
eschangewaspw WAS_password
```

2. Opravite naslednje korake na drugem iskalnem strežniku (za konfiguracijo z dvema strežnikoma) ali na strežniku pajkanja in obeh iskalnih strežnikih (za konfiguracijo s štirimi strežniki):
  - a. Prijavite se kot skrbnik iskanja v podjetju.
  - b. Zaženite naslednji skript, kjer je **WAS\_password** geslo za uporabnika aplikacijskega strežnika WebSphere Application Server, ki je podano v datoteki **ES\_NODE\_ROOT/nodeinfo/es.cfg** (glejte korak 1b).

#### AIX, Linux ali Solaris

```
eschangewaspw.sh WAS_password
```

#### Ukazni poziv Windows

```
eschangewaspw WAS_password
```

3. V okolju Windows izberite **Nadzorna plošča** → **Orodja za upravljanje** → **Storitve** in v storitvi **WebSphere Application Server** in **ESearchServer** dodajte enaka uporabniško ime in geslo za strežnik WebSphere Application Server.
4. Na indeksnem strežniku iskanja v podjetju ponovno zaženite sistem iskanja v podjetju:

## Zahteve za nastavitve pajka za podporo zaščite

Za pridobitev informacij, ki omogočajo, da zaščita na ravni dokumenta prične veljati, morajo pajki imeti dovoljenje za dostop do izvornih podatkov o zaščiti. Pri nekaterih tipih podatkov je za konfiguriranje zaščitene okolja potrebno izvesti dodatne korake.

Tabela 7. Zahteve za nastavitve pajka za podporo zaščite

---

### Pajki Content Edition

---

Preden izdelate pajka za dostop do repozitorijev v neposrednem načinu, konfigurirajte sistem WebSphere Information Integrator Content Edition, da se bo izvajal v neposrednem načinu, in konfigurirajte spojnik za strežnik pajkanja.

Preden izdelate pajka za dostop do repozitorijev v načinu strežnika, zaženite skript (`escrvbr.sh` na AIX, Linux ali Solaris, ali `escrvbr.vbs` na Windows), da konfigurirate strežnik pajkanja.

Ko ste pajka konfigurirali, podajte uporabniški ID in geslo, ki omogoča pajku, da dostopa do vsakega repozitorija, ki bo pajkan. Za vsak repozitorij v območju pajkanja lahko po potrebi podate različni uporabniški ID in geslo.

S tem povezane teme:

- “Način neposrednega dostopa do repozitorijev Content Edition” na strani 41
- “Način strežniškega dostopa do repozitorijev WebSphere II Content Edition” na strani 42
- “Konfiguriranje strežnika pajkanja v sistemu UNIX za WebSphere II Content Edition” na strani 42
- “Konfiguriranje strežnika pajkanja v sistemu Windows za WebSphere II Content Edition” na strani 43

---

### Pajki DB2

---

Preden izdelate pajka, zaženite skripto (`escrdb2.sh` na AIX, Linux ali Solaris, ali `escrdb2.vbs` na Windows), da konfigurirate strežnik pajkanja.

Ko ste konfigurirali pajka za pajkanje oddaljenih, nekatalogiziranih baz podatkov, podajte uporabniški ID in geslo, ki omogoča pajkanje vsake baze podatkov na ciljnim strežniku baze podatkov. Za vsako bazo podatkov v območju pajkanja lahko po potrebi podate različni uporabniški ID in geslo.

S tem povezane teme:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2” na strani 46
- “Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2” na strani 47

---

### Pajki DB2 Content Manager

---

Preden izdelate pajka, zaženite skripto (`escrcm.sh` na AIX, Linux ali Solaris, ali `escrcm.vbs` na Windows), da konfigurirate strežnik pajkanja.

Ko ste pajka konfigurirali, podajte uporabniški ID in geslo, ki omogoča pajku, da dostopa do vsakega strežnika, ki bo pajkan. Za vsak strežnik v območju pajkanja lahko po potrebi podate različni uporabniški ID in geslo.

S tem povezane teme:

- “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajke programa DB2 Content Manager” na strani 53
- “Konfiguriranje strežnika pajkanja v sistemu Windows za pajke programa DB2 Content Manager” na strani 54

---

### Pajki Domino Document Manager, Notes in QuickPlace

---

Tabela 7. Zahteve za nastavitve pajka za podporo zaščite (nadaljevanje)

---

<p>Za pajkanje strežnikov Lotus Domino, ki uporabljajo protokol Klic oddaljenemu postopku Notes (NRPC):</p> <ul style="list-style-type: none"><li>• Na sistemu AIX se prepričajte, da je modul I/O Completion Port nameščen in na voljo na strežniku pajkanja.</li><li>• Preden izdelate pajka, zaženite skripto (<code>escrnote.sh</code> na AIX, Linux ali Solaris, ali <code>escrnote.vbs</code> na Windows), da konfigurirate strežnik pajkanja.</li><li>• Na strežniku pajkanja za iskanje v podjetju mora biti nameščen strežnik Domino, ki mora biti član domene Domino, ki bo pajkana.</li><li>• Za potrditev trenutnih poverilnic uporabnika, ko uporabnik predloži zahtevo za iskanje, mora biti strežnik Domino, ki bo pajkan, konfiguriran kot strežnik Lotus Domino Trusted Server.</li><li>• Pri konfiguriranju pajka podajte pot za datoteko uporabniškega ID-ja Lotus Notes, ki je pooblaščen za dostop do strežnika, kot je <code>c:\Program Files\lotus\notes\data\name.id</code> ali <code>/local/notesdata/name.id</code>, in geslo za to datoteko ID-ja.</li></ul>	<p>S tem povezane teme:</p> <ul style="list-style-type: none"><li>• “Konfiguriranje V/I zaključnih vrat v AIX za pajkanje virov Lotus Domino” na strani 73</li><li>• “Konfiguriranje strežnika pajkanja v sistemu UNIX za pajkanje virov Lotus Domino” na strani 69</li><li>• “Konfiguriranje strežnika pajkanja v sistemu Windows za pajkanje virov Lotus Domino” na strani 71</li><li>• “Konfiguriranje overjenih strežnikov Lotus Domino za preverjanje poverilnic uporabnikov” na strani 257</li><li>• “Konfiguriranje strežnikov, ki uporabljajo protokol DIIOP” na strani 72</li><li>• “Konfiguriranje strežnika QuickPlace za uporabo zaščite Lokalni uporabnik” na strani 258</li><li>• “Konfiguriranje pomoči za imenike na strežniku QuickPlace” na strani 259</li></ul>
<p>Za pajkanje strežnikov Lotus Domino, ki uporabljajo protokol Domino Internet Inter-ORB (DIIOP):</p> <ul style="list-style-type: none"><li>• Na sistemu AIX se prepričajte, da je modul I/O Completion Port nameščen in na voljo na strežniku pajkanja.</li><li>• Strežnik pajkanja konfigurirajte tako, da lahko uporablja protokol.</li><li>• Pri konfiguriranju pajka podajte popolnoma kvalificiran uporabniški ID Lotus Notes, ki je pooblaščen za dostop do strežnika, kot je <b>Uporabniško ime/Kraj/Moje podjetje</b>, in geslo za ta uporabniški ID.</li></ul>	
<p>Za pajkanje strežnikov QuickPlace morate konfigurirati strežnik QuickPlace tako, da bo podpiral zaščito Lokalni uporabnik ali Imeniška podpora, glede na tip zaščite, ki ga želite uporabiti.</p>	

---

### Pajki Exchange Server

---

<p>Pri konfiguriranju pajka podajte uporabniški ID, ki ima pooblastila za dostop do javnih map na strežniku Exchange, ki bo pajkan, in geslo za ta uporabniški ID.</p> <p>Da bo pajek pri pajkanju podatkov uporabljal upravljanje ključa strežnika Exchange in protokol Plast zaščitene vtičnice (Secure Sockets Layer - SSL), prav tako podajte popolnoma kvalificirano pot do datoteke skladišča ključev in geslo, ki omogoča pajku, da dostopa do te datoteke. Datoteka skladišča ključev mora biti na strežniku pajkanja za iskanje v podjetju.</p>	<p>S tem povezana tema:</p> <ul style="list-style-type: none"><li>• “Preverjanje dostopa do zaščitene dokumentov strežnika Exchange Server” na strani 256</li></ul>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------

---

### Pajki baze podatkov JDBC

Pri konfiguriranju pajka lahko podate uporabniški ID in geslo, ki omogoča pajkanje tabel v ciljni bazi podatkov. Za vsako bazo podatkov v območju pajkanja lahko po potrebi podate različni uporabniški ID in geslo.

### Pajki NNTP

---



Tabela 7. Zahteve za nastavitve pajka za podporo zaščite (nadaljevanje)

---

Strežniki NNTP, ki bodo pajkani, morajo dovoljevati strežniku pajkanja, da bere podatke.

---

### **Pajki datotečnega sistema UNIX**

---

Podimeniki AIX, Linux in Solaris, ki bodo pajkani, morajo dovoljevati strežniku pajkanja, da bere podatke.

---

### **Spletni pajki**

---

Spletni pajek upošteva protokol za izključitev robotov. Če spletni strežnik v strežniški imenik najvišje ravni vključi datoteko robots.txt, bo pajek datoteko analiziral in nato spletna mesta v tem strežniku pajkanjal le, če mu je to dovoljeno. Za informacije o tem protokolu glejte <http://www.robotstxt.org/wc/exclusion.html>.

S tem povezane teme:

- “Spletna mesta, zaščitena z osnovnim overjanjem HTTP” na strani 90
- “Spletna mesta, zaščitena z overjanjem na osnovi obrazcev” na strani 91

Ko konfigurirate spletnega pajka:

- Morate podati posredniško ime uporabnika za pajka. Pravila v datotekah robots.txt strežnikov, ki bodo pajkani, lahko to ime podajo tako, da dovoli ali zavrne dostop.
- Neobvezno: Če spletni strežnik uporablja osnovno overjanje HTTP za omejevanje dostopa do spletnih mest, lahko podate poverilnice overjanja, ki omogočajo spletnemu pajku, da dostopa do strani, zaščiteneh z geslom.
- Neobvezno: Če spletni strežnik uporablja obrazce HTML za omejevanje dostopa do spletnih mest, lahko podate poverilnice overjanja, ki omogočajo spletnemu pajku, da dostopa do strani, zaščiteneh z geslom.

---

### **Pajki seznama semen, Web Content Management in WebSphere Portal**

---

Pred izdelavo pajka, morate zagnati namestitveni skript, da iskanje v podjetju integrirate s strežnikom portala WebSphere Portal. Za različice portala WebSphere Portal je na voljo več skriptov.

S tem povezana tema:

- “Namestitveni skripti za integracijo sistema iskanja v podjetju v program WebSphere Portal” na strani 312

Pri konfiguriranju pajka podajte popolnoma kvalificirano razločevalno ime, ki omogoča pajku, da priključiti strani s strežnika, ki bo pajkan, kot je uid=admin,cn=RegularEmployees,ou=Software Group,o=IBM,c=US, in podajte geslo za to ime. Razločevalno ime se mora ujemati z razločevalnim imenom, konfiguriranim v portalu WebSphere Portal.

Prepričajte se, da so dovoljenja za uporabniško razločevalno ime, ki ga podate, definirana v komponenti portala WebSphere Portal Nadzor dostopa do portala (PAC). Pajek s pomočjo PAC pridobi podatke o krmiljenju dostopa za dokumente, v katerih pajka.

---

### **Pajki datotečnega sistema Windows**

---

Tabela 7. Zahteve za nastavitve pajka za podporo zaščite (nadaljevanje)

---

Podimeniki, ki bodo pajkani, morajo dovoljevati strežniku pajkanja, da bere podatke. Pri konfiguriranju pajka za pajkanje oddaljenih datotečnih sistemov podajte uporabniški ID, ki omogoča pajku, da dostopa do oddaljenih podatkov, ter podajte geslo za ta uporabniški ID.	S tem povezane teme: <ul style="list-style-type: none"><li>• “Zagotavljanje zaščite na ravni dokumenta za dokumente datotečnega sistema Windows” na strani 259</li><li>• “Zaščiteno iskanje po overjenih domenah Windows” na strani 261</li></ul>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Za potrditev trenutnih poverilnic uporabnika, ko uporabnik predloži zahtevo za iskanje, se prepričajte, da so računi domene pravilno konfigurirani. Zahteve za nastavitve računov domene za datoteke, ki so bile pajkane na lokalnem računalniku, se razlikujejo od zahtev za datoteke, ki so bile pajkane na oddaljenem strežniku Windows.

---

## Preverjanje dostopa do zaščiteneh dokumentov strežnika Exchange Server

Za uporabo pajka Exchange Server za pajkanje dokumentov, ki so zaščiteni s požarnim zidom, morate preveriti, ali strežnik pajkanja lahko dostopa do strežnika javne mape Microsoft Exchange Server.

### O tej nalogi

Če strežnik pajkanja ne more dostopati do zaščitene strežnika Exchange Server, boste prejeli od strežnika sporočilo HTTP koda 501 (Ni implementirano). Prav tako se vam lahko prikažejo sporočila, ki navajajo, da je bil prejet nepričakovan odziv HTTP.

### Postopek

Da se prepričate, da strežnik pajkanja lahko dostopa do dokumentov iz zaščitenega zida:

1. Na strežniku pajkanja poženite spletni brskalnik.
2. Premaknite se na URL za strežnik javnih map Exchange Server, v katerem želite izvesti pajkanje. Primer: <http://exchange.yourCompany.com/public/>
3. Preverite, da lahko odprete stran strežnika Exchange Server.

Če ne morete dostopati do strežnika Exchange Server, se posvetujte s skrbnikom za vašo organizacijo.

## Zagotavljanje zaščite na ravni dokumenta za dokumente Lotus Domino

Če uporabljate strežnik Domino, na katerem želite izvesti pajkanje, protokol klica oddaljeni proceduri Notes (NRPC), morate konfigurirati strežnik pajkanja, da se lahko zagotavlja nadzor dostopa na ravni dokumenta.

Za zagotavljanje zaščite za dokumente na strežniku Domino, ki uporablja protokol NRPC, morate na strežnik pajkanja namestiti strežnik Domino. Ta strežnik Domino mora biti član vaše domene Domino. Za namestitev in konfiguriranje strežnika Domino sledite navodilom v dokumentaciji Lotus Domino.

Prav tako morate izpolniti naslednje naloge, da lahko iskalni strežniki preverijo, ali ima uporabnik, ki išče po zaščiteni zbirki, pooblastila za prikaz dokumentov, ki ustrezajo iskalnim kriterijem:

- “Konfiguriranje overjenih strežnikov Lotus Domino za preverjanje poverilnic uporabnikov” na strani 257.

- “Konfiguriranje globalne zaščite in uporabniškega registra LDAP na aplikacijskem strežniku WebSphere Application Server” na strani 249.

#### S tem povezani pojmi

“Potrjevanje trenutnih poverilnic med obdelavo poizvedbe” na strani 241

“Pajki Notes” na strani 66

## Konfiguriranje overjenih strežnikov Lotus Domino za preverjanje poverilnic uporabnikov

Za zagotavljanje zaščite za dokumente, ki jih je pajkal pajek Notes, ki uporablja protokol klica oddaljeni proceduri Notes (NRPC), je potrebno strežnike Domino, ki bodo pajkani, konfigurirati, da bodo overjeni strežniki Lotus Domino.

### Praden začnete

Ta postopek je potreben, če želite zagotoviti zaščito na ravni dokumenta, ko iščete oddaljene baze podatkov. Ta postopek ni potreben za iskanje baz podatkov, ki so lokalne za strežnik pajkanja.

Za konfiguriranje overjenih strežnikov mora biti na pajku nameščen strežnik Domino. Ta strežnik Domino mora biti član vaše domene Domino.

### O tej nalogi

Ko konfigurirate možnosti zaščite na ravni dokumentov za pajka Notes, podate, ali želite zagotoviti nadzor dostopa tako, da potrdite trenutne uporabnikove poverilnice, ko uporabnik predloži poizvedbo. Za zagotavljanje tega tipa zaščite morajo biti strežniki Domino, ki bodo pajkani, overjeni strežniki Lotus Domino.

Ko uporabniki iščejo domeno, ki zahteva potrditev njihovih trenutnih poverilnic, overjeni strežnik omogoči ID-ju strežnika Domino, da preklopi kontekst na trenutni uporabniški ID. Baza podatkov Domino se odpre, kot da bi jo odprl trenutni uporabnik, in zagotovljene so vse informacije seznama za nadzor dostopa baze podatkov.

Tovrstna zmožnost preklapljanja med konteksti je običajno na voljo samo za baze podatkov, ki so shranjene v imeniku `data` na lokalnem strežniku Domino. Z različico 6.5.1 Lotusa Domino je ta zmožnost na voljo prek overjenega strežnika. Za konfiguriranje overjenega strežnika skrbnik Domina poda, kateri strežniki Domino lahko izvajajo občutljive operacije, kot je delovanje kot drug uporabnik v primeru dostopa do baze podatkov iz oddaljenega računalnika.

### Postopek

Za konfiguriranje overjenega strežnika izvedite naslednje korake na vseh strežnikih Domino, ki jih pajka pajek Notes:

1. Na strežniku Domino za odpiranje odjemalca Skrbnik Lotusa Domino uporabite datoteko z ID-jem skrbnika za domeno Domino.
2. Kliknite **Datoteka** in nato izberite **Odpri strežnik**.
3. Vnesite ime strežnika Domino, za katerega želite omogočiti zmožnosti overjenega strežnika.
4. Izberite zavihek **Konfiguracija**.
5. Razširite objekt **Strežnik**, izberite dokument **Trenutni strežnik** in kliknite **Urejanje strežnika**.
6. Izberite zavihek **Zaščita**, pomaknite se na dno strani, poiščite vnos **Overjeni strežniki** in kliknite puščico navzdol.

7. Podajte eno od naslednjih možnosti:

#### **LocalDomainServers**

Izberite to možnost, če naj bodo vsi strežniki v domeni Domino overjeni strežniki.

#### *server\_name*

Podajte ime strežnika Domino, ki ga želite pajkati in po njem iskati kot overjenem strežniku.

Če je strežnik Domino, ki bo pajkan, v drugi domeni Domino, potem morate podati ime strežnika ali izbrati skupino **OtherDomainServers**. Prav tako morate slediti postopkom Domino za navzkrižno potrjevanje datoteke z ID-jem strežnika Domino za iskanje v podjetju z drugo domeno Domino. Za informacije o teh postopkih glejte dokumentacijo strežnika Domino.

8. Kliknite **Shrani in zapri**, da shranite vaše spremembe.
9. Zaustavite in ponovno zaženite oddaljene strežnike Domino, ki ste jim omogočili, da delujejo kot overjeni strežniki.

#### **S tem povezani pojmi**

“Potrjevanje trenutnih poverilnic med obdelavo poizvedbe” na strani 241

“Pajki Notes” na strani 66

## **Konfiguriranje strežnika QuickPlace za uporabo zaščite Lokalni uporabnik**

Če nameravate konfigurirati pajka QuickPlace, da bo uporabljal možnost Lokalni uporabnik za izvajanje zaščite, morate konfigurirati imenik Domino na strežniku Lotus QuickPlace, preden izdelate pajka.

### **O tej nalogi**

Ko konfigurirate pajka QuickPlace, izberete način zaščite za pajka, ki ga bo uporabljal za izvajanje zaščite na ravni dokumenta. Če izberete način Lokalni uporabnik, se morate prepričati, da so vsi lokalni uporabniški ID-ji in lokalne skupine registrirani v imeniku Domino (hierarhija imenika Domino mora ustrezati hierarhiji QuickPlace).

Prav tako se morate prepričati, da sta uporabniški ID in geslo, ki ju podate za pajka, registrirana v imeniku Domino in imata dovoljenje za branje baze podatkov, ki bo pajkana.

Za uporabo QuickPlace-a je potrebno samo uporabniško ime. Za pajkanje virov QuickPlace pa je potreben v celoti razširjen uporabniški ID. Razširjeni uporabniški ID je v naslednjem formatu:

`username/placename/QP/domainname`

Naslednjipostopek uporabite, da ugotovite v celoti razširjeno različico uporabniškega ID-ja, da se prepričate, da ta uporabniški ID ima pooblastila za branje baze podatkov QuickPlace, in da dodate uporabniški ID v imenik Domino. Imenik Domino mora vsebovati uporabniški ID, ki se bo uporabil za pajkanje baz podatkov QuickPlace in vseh lokalnih uporabnikov in lokalnih skupin (hierarhija imenika Domino mora ustrezati hierarhiji QuickPlace).

### **Postopek**

Za konfiguriranje strežnika QuickPlace za uporabo zaščite Lokalni uporabnik:

1. Potrdite dovoljenja za uporabniški ID:
  - a. Na strežniku QuickPlace odprite dokument Strežnik.

- b. Odprite stran Datoteke in nato odprite seznam za nadzor dostopa (ACL) za bazo podatkov, ki jo želite pajkati.
  - c. Potrdite, da ID Lokalnega uporabnika, za uporabo katerega bo pajek konfiguriran, obstaja v ACL-u in da ima ta uporabniški ID dovoljenje za branje baze podatkov.  
V celoti razširjeno obliko tega uporabniškega ID-ja morate podati v koraku 2.
2. Uporabnika dodajte v imenik Domino:
- a. Na strežniku QuickPlace odprite dokument Strežnik.
  - b. V postavki drevesa ljudi v sklopu strani Ljudje in skupine dodajte v celoti razširjen uporabniški ID, ki ste ga potrdili v koraku 1.
  - c. V polju **Internetno geslo** podajte geslo za ta uporabniški ID.

## Konfiguriranje pomoči za imenike na strežniku QuickPlace

Če nameravate konfigurirati pajka QuickPlace, da bo uporabljal imenik LDAP za uveljavljanje zaščite, morate na strežniku Lotus QuickPlace izdelati bazo podatkov pomoči za imenike (Directory Assistance), preden konfigurirate pajka.

### Omejitve

Na strežniku QuickPlace, ki ga želite pajkati, se morajo izvajati naloge DIIOP in HTTP.

### Postopek

Za konfiguriranje pomoči za imenike LDAP na strežniku QuickPlace:

1. Izdelajte bazo podatkov pomoči za imenike:
  - a. Na strežniku QuickPlace odprite dokument Strežnik.
  - b. S pomočjo predloge **Pomoč za imenike(6)** izdelajte bazo podatkov. Ta predloga se nahaja na strežniku.
  - c. Za izdelavo dokumenta v bazi podatkov kliknite **Dodaj pomoč za imenike**.
  - d. Odprite zavihek Osnovno in v polju **TipDomene** izberite **LDAP**.
  - e. Odprite zavihek Kontekst poimenovanja in se prepričajte, da je potrditveno polje **Overjen za poverilnice** izbrano.
  - f. Odprite zavihek LDAP in podajte informacije o strežniku LDAP.
  - g. Shranite in zaprite dokument Strežnik.
2. Strežnik QuickPlace konfigurirajte, da bo uporabljal bazo podatkov pomoči za imenike:
  - a. Na strežniku QuickPlace odprite dokument Strežnik.
  - b. Odprite zavihek Osnovno in v polju **Ime baze podatkov pomoči za imenike** podajte ime baze podatkov, ki ste jo izdelali v koraku 1.
  - c. Shranite in zaprite dokument Strežnik.  
Strežnik QuickPlace lahko sedaj uporablja strežnik LDAP kot sekundarni imenik Domino.

## Zagotavljanje zaščite na ravni dokumenta za dokumente datotečnega sistema Windows

Če želite omogočiti preverjanje veljavnosti trenutnih poverilnic v primeru, ko uporabnik išče dokumente, pajkane s pajkom datotečnega sistema Windows, morate informacije o računih domene konfigurirati na strežniku pajkanja in strežniku Microsoft Windows.

Ko konfigurirate pajka datotečnega sistema Windows, podate, ali želite pajkati podimenike na lokalnem računalniku ali podimenike na oddaljenem računalniku. Če ste omogočili zaščito zbirke, lahko prav tako podate možnosti za nadzorovanje dostopa do dokumentov v pajkanih podimenikih.

Če želite zagotoviti nadzor dostopa tako, da potrdite trenutne uporabnikove poverilnice, ko uporabnik predloži poizvedbo, se morate prepričati, da so računi domene pravilno konfigurirani. Zahteve za nastavitev računov domene za datoteke, ki so bile pajkane na lokalnem računalniku, se razlikujejo od zahtev za datoteke, ki so bile pajkane na oddaljenem strežniku Windows.

**Pomembno:** Uporabniških poverilnic ni mogoče potrditi med obdelovanjem poizvedbe, če držita oba naslednja pogoja:

- Strežnik Windows, ki bo pajkan, ni član domene.
- Imenik, ki bo pajkan, je oddaljeni imenik, kot je \\servername\hostname.

## Potrditev z lokalnimi podatki za nadzor dostopa

Za potrditev trenutnih uporabniških poverilnic sistem uporablja tako informacije o lokalnem računu uporabnika kot tudi informacije o računu domene (če računalnik pripada domeni Windows). Za potrditev poverilnic med obdelovanjem poizvedbe morata biti v informacijah o zaščiti za dokumente, ki bodo iskani, navedeni obe imeni.

### Lokalni računi

Pri lokalnem računu je uporabniško ime v naslednjem formatu:

*COMPUTER\_NAME\USERNAME*

Za prijavo uporabniki podajo samo uporabniško ime, a ustrezno podana dodelitev pravic za uporabnika Windows uporablja polno ime. Če je na primer uporabniško ime lokalnega računa abcuser, je lahko polno ime računa WINSERVER1\abcuser.

Ko uporabniki uporabljajo iskalno aplikacijo in konfigurirajo profil za iskanje po zaščiteneh dokumentih na lokalnem sistemu, morajo podati uporabniško ime, ki ga uporabljajo za prijavo v Windows (na primer abcuser).

### Računi domene

Pri računu domene je uporabniško ime v naslednjem formatu:

*DOMAIN\_NAME\USERNAME*

Za prijavo uporabniki podajo te informacije v naslednjem formatu:

*USERNAME@DOMAIN\_NAME*

Če na primer konfigurirate dodelitve uporabniških pravic za datoteko in izberete domeno WIN1\abcuser, se račun prikaže kot abcuser@win1.company.com.

Ko uporabniki uporabljajo iskalno aplikacijo in konfigurirajo profil, ki jim omogoča iskanje po dokumentih v zaščiteni domeni, morajo podati uporabniško ime, ki ga uporabljajo za prijavo v Windows (na primer abcuser@win1.company.com).

Za zagotavljanje trenutne potrditve poverilnic na lokalnih računalnikih morajo imeti uporabniški računi, ki jih uporablja strežnik pajkanja, naslednje pravice uporabnika Windows. Za dodelitev uporabniških pravic uporabite skrbniška orodja Windows: **Skrbniška orodja** → **Lokalna načela zaščite** → **Lokalna načela** → **Dodelitev lokalnih uporabniških pravic**.

- Uporabniški ID, pod katerim se izvaja strežnik pajkanja, mora imeti pravico **Delovanje kot del operacijskega sistema**. Te pravice se konfigurirajo za skrbniškega uporabnika za iskanje v podjetju na strežniku pajkanja ob namestitvi izdelka OmniFind Enterprise Edition.

- Uporabniki morajo imeti uporabniško pravico **Lokalna prijava**.

## Potrditev z oddaljenimi lokalnimi podatki za nadzor dostopa do domene

Za operacijski sistem Windows je vsak imenik, ki se prične z `\\servername`, oddaljeni imenik. Primer:

```
\\software\utilities\IBM
```

Za dostop do oddaljenega imenika uporabniki podajo svoja uporabniška imena v naslednjem formatu:

```
USERNAME@DOMAIN NAME
```

Ko uporabniki uporabljajo iskalno aplikacijo on konfigurirajo profil, ki jim omogoča iskanje po zaščiteneh dokumentih na oddaljenem sistemu, morajo podati uporabniško ime, ki ga uporabljajo za dostop do oddaljenega sistema Windows (na primer `abcuser@win1.company.com`).

Za zagotavljanje trenutne potrditve poverilnic na oddaljenih računalnikih morajo imeti uporabniški računi naslednje pravice uporabnika Windows. Za dodelitev uporabniških pravic uporabite skrbniška orodja Windows: **Skrbniška orodja** → **Načela zaščite domene**.

- Strežnik pajkanja in strežnik Windows, na katerih bo izvedeno iskanje, morata biti člana iste domene.
- Uporabniški ID, pod katerim se izvaja strežnik pajkanja, mora imeti pravico **Delovanje kot del operacijskega sistema**. Te pravice se konfigurirajo za skrbniškega uporabnika za iskanje v podjetju na strežniku pajkanja ob namestitvi izdelka OmniFind Enterprise Edition.
- Uporabniki morajo imeti uporabniško pravico **Prijava kot paketna obdelava**.

### S tem povezani pojmi

“Potrjevanje trenutnih poverilnic med obdelavo poizvedbe” na strani 241

“Pajki datotečnega sistema Windows” na strani 102

“Zaščiteno iskanje po overjenih domenah Windows”

## Zaščiteno iskanje po overjenih domenah Windows

Da bi uveljavil zaščito na ravni dokumentov za oddaljene datotečne sisteme Windows, sistem iskanja v podjetju podpira preverjanje seznamov za nadzor dostopa (ACL) v overjenih domenah.

## Konfiguriranje pajka

Če želite konfigurirati pajka datotečnega sistema Windows za podporo overjenih domen, morate podati možnosti v novi konfiguracijski datoteki. Konzola za skrbništvo nad iskanjem v podjetju ne nudi podpore za konfiguriranje te možnosti.

1. Za podporo zaščiti na ravni dokumentov v overjenih domenah Windows uredite naslednjo datoteko:

```
ES_NODE_ROOT/master_config/session_ID/winfscrawler_ext.xml
```

**Nasvet:** Če želite ugotoviti ID seje pajka datotečnega sistema Windows, ki ga želite konfigurirati, lahko nadzirate pajka v konzoli za skrbništvo nad iskanjem v podjetju ali pa izdate ukaz **esadmin report collections**.

2. Podajte ime domene Windows in ime NETBIOS imenika Active Directory. Primer:

```

<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. Zaustavite in znova zaženite pajka, da spremembe stopijo v veljavo.

### Omejitve

- Dokumenti ne smejo vsebovati ACL-jev iz več domen Windows. Skupine in uporabniki domen morajo pripadati po eni domeni Windows na zbirko.
- Če želite zagotoviti podporo za overjanje dostopa do oddaljenih datotečnih sistemov, se morajo strežniki Windows izvajati v isti domeni Windows ali v overjenih domenah Windows.
- Pajek datotečnega sistema Windows prebere ime NETBIOS imenika Active Directory, povezanega s strežnikom Windows, ki bo pajkan, in s pomočjo imena NETBIOS filtrira ACL datoteke. Imenik Active Directory, ki ga združi strežnik pajkanja, overi drugi imenik Active Directory, ki definira uporabniške in skupinske račune.
- Podani uporabniški račun, prek katerega bo pajek dostopal do oddaljenega strežnika Windows, mora biti v domeni Windows, za katero želite uveljaviti in preveriti nadzor dostopa.
- Operacijski sistem Windows dovoljuje le en račun za povezavo omrežnih map na enem datotečnem strežniku. Drugi računi se ne morejo istočasno povezati z istim datotečnim strežnikom. Zaradi tega ne morete konfigurirati različnih računov za različne pajke in pajkati istega strežnika Windows, tudi če se pajki nahajajo v različnih zbirkah.

#### S tem povezani pojmi

“Zagotavljanje zaščite na ravni dokumenta za dokumente datotečnega sistema Windows” na strani 259

---

## Onemogočanje zaščite za iskanje v podjetju

Zaščito za aplikacijo za iskanje v podjetju lahko onemogočite na strežniku aplikacije WebSphere Application Server. Če ste pred tem že konfigurirali nadzor zaščite na ravni dokumenta, lahko podate, da se nadzor ne upošteva. Nastavitve zaščite prav tako vplivajo na to, kako so skrčeni rezultati prikazani med rezultati iskanja.

## Onemogočanje zaščite za aplikacijo podjetja na aplikacijskem strežniku WebSphere Application Server

Za nadzor nad tem, katere dejavnosti iskanja v podjetju zahtevajo overjanje uporabnika, lahko onemogočite globalno zaščito za posamezne aplikacije podjetja na aplikacijskem strežniku WebSphere Application Server.

### O tej nalogi

Namestitveni program OmniFind Enterprise Edition razmesti v WebSphere Application Server tri aplikacije podjetja:

- Aplikacija ESAdmin vsebuje vmesnik za skrbniško konzolo iskanja v podjetju.
- Aplikacija ESSearchApplication vsebuje vmesnik za vzorčno iskalno aplikacijo.
- Aplikacija ESSearchServer zagotavlja vso oddaljeno komuniciranje za izvedbo SIAPI iskanja v podjetju in omogoča vmesniku SIAPI, da komunicira z iskalnimi strežniki.



Po privzetku vse tri aplikacije podjetja podpirajo globalno zaščito WebSphere Application Server. Ko te aplikacije zaznajo, je globalna zaščita omogočena, začnejo overjati vse zahteve, ki jih prejmejo.

Nekatere organizacije morebiti želijo omogočiti ali onemogočiti zaščito za določeno aplikacijo podjetja. Morda boste želeli overiti vse uporabnike, ki dostopajo skrbniške konzole iskanja v podjetju, ne pa tudi uporabnikov, ki uporabljajo vmesnik SI-API ali vzorčno iskalno aplikacijo.

## Postopek

Za onemogočanje zaščite za aplikacijo podjetja:

1. Na iskalnem strežniku zaženite skrbniško konzolo WebSphere Application Server. Skrbniško konzolo lahko odprete na naslednje načine:
  - Uporabite meni Windows **Zaženi**, da izberete program.
  - Za različico 5 aplikacijskega strežnika WebSphere Application Server odprite spletni brskalnik in pojdite na `http://hostname:port/admin`, kjer je *hostname* ime gostitelja iskalnega strežnika in *port* številka vrat za skrbniško konzolo WebSphere Application Server. Vrata skrbniške konzole so običajno 9090.
  - Za različico 6 aplikacijskega strežnika WebSphere Application Server odprite spletni brskalnik in pojdite na `http://hostname:port/ibm/console`, kjer je *hostname* ime gostitelja iskalnega strežnika in *port* številka vrat za skrbniško konzolo WebSphere Application Server. Vrata skrbniške konzole so običajno 9060.
2. Ko ste pozvani, da podate uporabniški ID in geslo, vnesite ID in geslo skrbnika, ki sta bila podana, ko ste omogočili globalno zaščito v WebSphere Application Server.
3. Ko ste se prijavi na skrbniško konzolo, kliknite **Aplikacije** in nato **Aplikacije podjetja**.
4. Izberite potrditveno polje zraven imena aplikacije podjetja, za katero želite onemogočiti zaščito.
5. Pomaknite se navzdol in kliknite povezavo **Preslikaj vloge zaščite na uporabnike/skupine**.
6. Poiščite vlogo **VsiOverjeni** in izberite potrditveno polje pod stolpcem **Vsi?**
7. Kliknite **Potrdi**.
8. Kliknite povezavo **Shrani**, da shranite vaše spremembe.
9. Če uporabljate razmestitev omrežja WebSphere, izberite potrditveno polje **Sinhroniziranje sprememb z Nodes**.
10. Kliknite **Shrani**.
11. Zaustavite in znova zaženite aplikacijo ESSearchServer.

**AIX,**

**Linux**

**ali Solaris**

```
./stopServer.sh ESSearchServer
```

```
./startServer.sh ESSearchServer
```

**Windows**

```
stopServer ESSearchServer
```

```
startServer ESSearchServer
```

Ti skripti se nahajajo v imeniku `WAS_INSTALL_ROOT/AppServer/bin`:

- Za aplikacijski strežnik WebSphere Application Server različice 5 je privzeta pot za namestitvev `/usr/WebSphere` na sistemih AIX, `/opt/WebSphere` na sistemih Linux ali Solaris in `C:\Program Files\WebSphere` na sistemih Windows.

- Za aplikacijski strežnik WebSphere Application Server različice 6 je privzeta pot za namestitvev /usr/IBM/WebSphere na sistemih AIX, /opt/IBM/WebSphere na sistemih Linux ali Solaris in C:\Program Files\IBM\WebSphere na sistemih Windows.

## Onemogočanje zaščite na ravni dokumenta

Uporabnikom lahko omogočite, da iščejo po zbirki ne glede na to, ali je z dokumenti v indeksu povezan kakršenkoli nadzor dostopa. Za pajke, ki podpirajo trenutno potrditev poverilnic, lahko prav tako omogočite uporabnikom, da iščejo po zbirki brez potrjevanja trenutnega nadzora dostopa med obdelavo poizvedbe.

### Preden začnete

Zaščito na ravni dokumenta za vse dokumente v zbirki lahko omogočite ali onemogočite samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

### Omejitve

Možnosti zaščite na ravni dokumentov lahko podate samo, če ste za zbirko pri njeni izdelavi omogočili zaščito.

### O tej nalogi

Pajke lahko konfigurirate tako, da povežejo žetone zaščite z dokumenti, ko so ti pajkani. Vaša iskalna aplikacija lahko uporabi te žetone, ki so shranjeni v indeksu, za zagotavljanje nadzora dostopa, ko uporabniki iščejo po zbirki. Pri nekaterih pajkih lahko prav tako podate, da želite potrditi trenutni nadzor dostopa, ki je povezan z dokumenti v njihovih izvirnih repozitorijih, ko uporabniki predložijo poizvedbo.

Da odstranite te omejitve zaščite lahko podate, naj iskalni strežniki ne upoštevajo nobenega žetona zaščite, ki ga posreduje poizvedba. Prav tako lahko omogočite uporabnikom, da izvajajo poizvedbe na dokumentih, ne da bi se njihove poverilnice primerjale s trenutnim nadzorom dostopa.

Morda boste želeli začasno onemogočiti zaščito na ravni dokumenta, če preizkušate novo zbirko ali če morate odpraviti težavo z iskalno aplikacijo.

### Postopek

Za onemogočanje nadzora zaščite na ravni dokumenta:

1. Uredite zbirko, izberite stran Splošno in kliknite **Omogočanje in onemogočanje zaščite na ravni dokumenta**.
2. Na strani Zaščita na ravni dokumenta za vse dokumente izberite potrditveno polje **Ne upoštevaj nadzora dostopa na ravni dokumenta v indeksu**, če ne želite, da bi se žetoni zaščite, ki so jih pajki povezali z dokumenti, uporabljali, ko uporabniki izvedejo poizvedbo na zbirki.

Pajki bodo v dokumente dodajali žetone zaščite, vendar pa iskalni strežniki teh ne bodo upoštevali in bodo dovolili uporabnikom, da iščejo predhodno zaščitene dokumente.

3. Izberite potrditveno polje **Ne potrdi trenutnih poverilnic med obdelavo poizvedbe**, če ne želite potrditi trenutnega nadzora dostopa, ki je povezan z dokumenti v njihovih izvirnih repozitorijih, ko uporabniki predložijo poizvedbe. To potrditveno polje je na voljo le za dokumente, pajkane s pajki, ki podpirajo to zmožnost.

Če izberete to potrditveno polje, bodo ostale možnosti zaščite na ravni dokumenta še vedno v veljavi. Če ste na primer podali možnosti za shranjevanje nadzora dostopa v

indeks, ko ste konfigurirali pajka, bo ta nadzor zaščite še naprej veljal, dokler ne izberete potrditvenega polja **Prezri nadzor dostopa na ravni dokumenta v indeksu**.

## Onemogočanje zaščite za skrčene rezultate iskanja

Če je zaščita zbirke onemogočena, rezultatov iskanja iz istega spletnega mesta ni mogoče skrčiti v rezultatih iskanja, razen če podate, da med obdelavo poizvedbe ne želite potrditi uporabniških poverilnic.

### Preden začnete

Potrditev trenutnih poverilnic za vse dokumente v zbirki lahko omogočite ali onemogočite samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

### Omejitve

Možnosti zaščite na ravni dokumentov lahko podate samo, če ste za zbirko pri njeni izdelavi omogočili zaščito.

### O tej nalogi

Ob potrditvi trenutnih poverilnic se vsak vir dokumenta pregleda in morebiti usmeri v potrjevanje. Potrditev trenutnih poverilnic morate onemogočiti, če želite, da se dokumenti, ki imajo isto predpono URI ali ki pripadajo predhodno konfigurirani skrčeni skupini URI, skrčijo v iskalnih rezultatih.

### Postopek

Za onemogočanje potrditve trenutnih poverilnic, da se dokumenti lahko skrčijo v iskalnih rezultatih:

1. Uredite zbirko, izberite stran Splošno in kliknite **Omogočanje in onemogočanje zaščite na ravni dokumenta**.
2. Na strani Zaščita na ravni dokumenta za vse dokumente izberite potrditveno polje **Pri obdelavi poizvedb ne potrdi trenutnih poverilnic**.
3. Nadzirajte zbirko, izberite stran Iskanje in ustavite ter ponovno zaženite procese iskalnega strežnika.

Ko uporabniki izvedejo poizvedbo v zbirki, se dokumenti, ki imajo isto predpono URI ali ki pripadajo spletnim mestom, ki so konfigurirani za krčenje, skrčijo v iskalnih rezultatih. V vzorčni iskalni aplikaciji si lahko uporabniki ogledajo skrčene rezultate tako, da kliknejo povezavo **Več rezultatov z istega vira**.



---

## Zagon in zaustavitev sistema iskanja v podjetju

Ko izdelate zbirko, morate zagnati strežnike za pajkanje, razčlenjevanje in indeksiranje podatkov (iskalni strežniki se zaženejo samodejno). Po spreminjanju zbirke morate zaustaviti in znova zagnati strežnike.

Večina strežnikov za iskanje v podjetju se lahko izvaja neprenehoma ali v skladu z razporedi, ki jih podate. Tako lahko na primer podate razporede za gradnjo glavnih in delta indeksov. Potem ko zaženete sistem iskanja v podjetju, je običajno potrebno zaustaviti in znova zagnati postopke strežnikov samo, če spremenite konfiguracijske nastavitve (kot je posodabljanje kategorij ali povečanje velikosti iskalnega predpomnilnika).

Če spremenite vsebino zbirke ali spremenite pravila za to, kako pajki zbirajo podatke iz virov v vašem podjetju, morate zaustaviti in znova zagnati iskalne strežnike, da bodo spremembe začele veljati. Če ne spremenite pravil za pajkanje, se spletni pajek izvaja neprenehoma, drugi pajki pa se izvajajo v skladu z razporedom, ki ste ga podali.

Da bi izboljšali razpoložljivost iskalnih strežnikov, kadar indeksni strežnik in skrbniška konzola nista dosegljiva, lahko podate ukaze, ki zaženejo iskalne strežnike za zbirko v samostojnem načinu.

---

## Zagon sistema iskanja v podjetju

Če želite uporabnikom omogočiti iskanje po zbirki, morate zagnati systemske procese, nato pa zagnati še strežnike, ki pajkajo, razčlenjujejo in indeksirajo zbirko (iskalni strežniki se samodejno zaženejo).

### **Preden začnete**

Konfigurirajte vire podatkov, ki jih želite pajkati, in podajte možnosti za način razčlenjevanja, indeksiranja in iskanja podatkov. Če želite uporabnikom na primer omogočiti pregledovanje podrobnosti o kategorijah v rezultatih iskanja, pred zagonom razčlenjevalnika konfigurirajte kategorije.

Če želite zagnati strežnike za iskanje v podjetju, morate imeti vlogo skrbnika za iskanje v podjetju, skrbnika zbirke ali operatorja s pooblastili za izvajanje skrbništva nad to zbirko.

Strežnike za iskanje v podjetju za posamezno zbirko morate zagnati v pravilnem vrstnem redu. Šele ko zaženete pajka in prepajkate podatke, lahko indeksirate prepajkane podatke.

### **Omejitve**

Če želite zagnati sistem iskanja v podjetju, morate uporabiti uporabniški račun, ki ga je mogoče overiti s pomočjo lokalnega overjanja. Če poskusite zagnati sistem z računom datotečnega sistema Andrew File System (AFS), sistem javi napake.

### **Postopek**

Če želite zagnati sistem iskanja v podjetju, naredite naslednje:





1. Če uporabite iskanje v podjetju v dvo- ali štiristrežniški konfiguraciji, se prijavite kot skrbnik iskanja v podjetju in na vsakem strežniku zaženite CCL (skupna komunikacijska plast).



**AIX, Linux ali Solaris**  
startccl.sh -bg

**Ukazni poziv Windows**  
startccl

### **Skrbniška orodja storitev operacijskega sistema Windows**

Če želite zagnati CCL v ozadju:

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
  - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
2. Zaženite komponente sistema iskanja v podjetju:
- a. V strežnik za iskanje v podjetju se prijavite kot skrbnik iskanja v podjetju.
  - b. Zaženite vse komponente sistema: **esadmin system startall**  
Zgornji ukaz zažene naslednje procese in aplikacije:
    - Spletni strežnik (v večstrežniški konfiguraciji se spletni strežnik zažene na obeh iskalnih strežnikih)
    - Aplikaciji ESSearchServer in ESAdmin na aplikacijskem strežniku WebSphere (v večstrežniški konfiguraciji se aplikaciji zaženeta na obeh iskalnih strežnikih)
    - Seja ESAdmin na indeksnem strežniku
    - CCL v računalniku, kjer ste zagnali ukaz, če se CCL še ne izvaja
    - Omrežni strežnik baz podatkov za iskanje v podjetju
    - Informacijski center za iskanje v podjetju (v večstrežniški konfiguraciji se informacijski center zažene na obeh iskalnih strežnikih)
3. Zaženite skrbniško konzolo za iskanje v podjetju in se prijavite kot skrbnik iskanja v podjetju. Če uporabljate skrbniške vloge, se lahko prijavite kot skrbnik zbirke ali operater, ki ima pooblastila za dostop do zbirke, ki jo želite zagnati.
4. Na pogledu Zbirke poiščite zbirko, nad katero želite izvajati skrbništvo, in kliknite  **Nadzor**.
5. Na strani Pajkanje za vsakega pajka, ki ga želite zagnati, kliknite  **Zaženi**.
- Ko zaženete spletnega pajka, začne ta nemudoma pajkati podatke. Tovrstni tipi pajkov se izvajajo brez prekinitev in pajkajo ter vnovič pajkajo spletne dokumente.
  - Če zaženete katerega od drugih tipov pajkov, zaženete sejo pajka. Pajek začne pajkati na predviden datum in ob predvidenem času. Če niste načrtovali časa zagona pajka ali če želite pajka zagnati prej, nadzirajte pajka in za vsak vir podatkov, ki ga želite pajkati, kliknite ikono za zagon.  
Potem ko se pajek zažene, se lahko izvaja brez prekinitev. Če ste načrtovali zagon pajka, se bo pajek znova zagnal na predvidene dni ob predvidenih časih.
6. Po pajkanju podatkov odprite stran Razčlenjevanje in kliknite  **Zaženi**, da zaženete razčlenjevalnik.  
Razčlenjevalnik se lahko izvaja brez prekinitev. Ponavadi vam ni treba zaustaviti razčlenjevalnika, razen če opravite spremembe v načinu razčlenjevanja podatkov (na primer posodobitev kategorij ali preslikav polj XML).
7. Neobvezno: Če želite vsiliti zagon procesov indeksiranja, namesto da bi čakali na zagon indeksiranja ob predvidenem dnevu in času, odprite stran Indeksiranje in v **Glavnem področju** kliknite  **Zaženi**.  
Procesi indeksiranja se lahko izvajajo brez prekinitev. Indeks bo zgrajen na predviden dan in ob predvidenih časih.

**Nasvet:** Iskalni strežniki se samodejno zaženejo, izvajajo pa se lahko neprekinjeno. Praviloma vam ni treba zaustaviti iskalnih strežnikov, razen če opravite spremembe v iskalnem predpomnilniku ali nastavitvah za povzetke dokumentov. Če želite zagnati iskalne strežnike, odprite stran Iskanje, kliknite najprej  **Zaustavi**, nato pa še  **Zaženi**.

Da bi izboljšali razpoložljivost iskalnih strežnikov, kadar skrbniška konzola ni dosegljiva, lahko podate ukaze, ki zaženejo iskalne strežnike za zbirko v samostojnem načinu. Če se indeksni strežnik ne izvaja, skrbniška konzola ni na voljo.

#### **S tem povezani pojmi**

“Skrbniške vloge” na strani 235

#### **S tem povezana opravila**

“Prijava na skrbniško konzolo” na strani 16

“Krmiljenje, katere komponente bodo zagnane ali zaustavljene” na strani 270

“Izvajanje skrbništva nad iskalnimi strežniki v samostojnem načinu” na strani 271

---

## **Zaustavitev sistema iskanja v podjetju**

Če spremenite konfiguracijo strežnika za iskanje v podjetju ali če morate odpraviti nastalo težavo, bo morda treba zaustaviti strežnik in ga znova zagnati.

### **Preden začnete**


Če želite zaustaviti strežnike za iskanje v podjetju, morate imeti vlogo skrbnika za iskanje v podjetju, skrbnika zbirke ali operatorja s pooblastili za izvajanje skrbništva nad to zbirko.

### **O tej nalogi**

Strežnike za iskanje v podjetju je mogoče zaustaviti vsakega posebej. Če na primer zaustavite in znova zaženete pajka, da bi uveljavili spremembe, ki ste jih opravili v konfiguriranju, vam ni treba zaustaviti in znova zagnati tudi iskalnih strežnikov.

### **Postopek**

Če želite zaustaviti strežnike za iskanje v podjetju, naredite naslednje:

1. Na pogledu Zbirke poiščite zbirko, nad katero želite izvajati skrbništvo, in kliknite  **Nadzor**.
2. Na strani Pajkanje poiščite pajka, nad katerim želite izvajati skrbništvo, in ga zaustavite ali začasno ustavite.

Če spremenite prostor za pajkanje ali lastnosti pajka, zaustavite pajka in ga znova zaženite, da uveljavite opravljene spremembe. V kolikor ste spremenili prostor za pajkanje in bi radi uveljavili spremembe za dokumente, ki so že indeksirani, morate dokumente tudi znova prepajkati.

**Nasvet:** Morda bo sistem prikazal sporočilo, da je zahtevana operacija potekla, četudi se proces v ozadju še zmeraj izvaja. Da bi ugotovili, ali se je naloga že končala, na skrbniški konzoli kliknite možnost **Osveži** (ne kliknite **Osveži** v spletnem brskalniku). Proces je končan, če statusna ikona pajka nakazuje, da je zaustavljen.

3. Na strani Razčlenjevanje kliknite  **Zaustavi**, da zaustavite pajka.

Potem ko spremenite pravila za razčlenjevanje podatkov, zaustavite razčlenjevalnik in ga znova zaženite, da uveljavite opravljene spremembe. Spremembe stopijo v veljavo samo za na novo pajkane dokumente. Če želite spremembe uveljaviti tudi za dokumente, ki so že v indeksu, morate zagnati popolno pajkanje in znova prepajkati vse dokumente, s čimer omogočite njihovo vnovično razčlenjevanje in indeksiranje.

4. Na strani Indeksiranje kliknite  **Zaustavi**, da zaustavite indeks, ki je pravkar v postopku gradnje.  
Gradnjo indeksa lahko zaustavite tudi med nadziranjem čakalne vrste indeksov. To storite tako, da na orodjarni izberete možnost **Sistem**, odprete stran Indeksiranje, nato pa kliknete  **Zaustavi** za indeks, katerega gradnjo želite zaustaviti.
5. Na strani Iskanje kliknite  **Zaustavi**, da zaustavite iskalne strežnike. Ponavadi je treba zaustaviti iskalne strežnike in jih nato znova zagnati samo v primeru, ko spremenite iskalni predpomnilnik ali nastavitve za povzetek dokumentov.
6. Če želite zaustaviti sistem iskanja v podjetju namesto posameznih strežnikov, naredite naslednje:
  - a. V strežnik za iskanje v podjetju se prijavite kot skrbnik iskanja v podjetju.
  - b. Zaustavite vse komponente sistema: **esadmin system stopall**  
Ukaz zaustavi naslednje procese in aplikacije:
    - Spletni strežnik (v večstrežniški konfiguraciji se spletni strežnik zaustavi na obeh iskalnih strežnikih)
    - Aplikaciji ESSearchServer in ESAdmin na aplikacijskem strežniku WebSphere Application Server (v večstrežniški konfiguraciji se aplikacije zaustavijo na obeh iskalnih strežnikih)
    - Seja ESAdmin na indeksnem strežniku
    - CCL (skupna komunikacijska plast) za iskanje v podjetju v računalniku, v katerem ste izdali ukaz
    - Omrežni strežnik baz podatkov za iskanje v podjetju
    - Informacijski center za iskanje v podjetju (v večstrežniški konfiguraciji se informacijski center zaustavi na obeh iskalnih strežnikih)

#### S tem povezana opravila

“Prijava na skrbniško konzolo” na strani 16

“Krmiljenje, katere komponente bodo zagnane ali zaustavljene”

“Izvajanje skrbništva nad iskalnimi strežniki v samostojnem načinu” na strani 271

---

## Krmiljenje, katere komponente bodo zagnane ali zaustavljene

Z ukazoma **esadmin system startall** in **esadmin system stopall** lahko krmilite, katere komponente bodo zagnane ali zaustavljene.

### O tej nalogi

Datoteka `ES_INSTALL_ROOT/default_config/AutoRunComponents.properties` vsebuje seznam komponent za iskanje v podjetju, ki jih je mogoče zagnati ali zaustaviti z ukazoma **esadmin system startall** in **esadmin system stopall**. Po privzetku se vse navedene komponente zaženejo in zaustavijo s tema ukazoma.

Če želite preprečiti zagon ali zaustavitev določenih komponent, lahko uredite datoteko lastnosti.

### Postopek

Če želite podati, katere komponente bodo zagnane ali zaustavljene takrat, ko zaženete ali zaustavite sistem iskanja v podjetju, naredite naslednje:

1. V strežnik, v katerem nameravate izdati ukaz **esadmin system startall** ali **esadmin system stopall**, se prijavite kot skrbnik iskanja v podjetju.



2. Uredite datoteko ES\_INSTALL\_ROOT/default\_config/AutoRunComponents.properties.
3. Če želite preprečiti zagon komponente, dodajte polje za komponento z imenom `Component.startable.component_ID=false`, pri čemer je `component_ID` komponenta, ki je ne želite zagnati.
4. Če želite preprečiti zaustavitev komponente, dodajte polje za komponento z imenom `Component.stopable.component_ID=false`, pri čemer je `component_ID` komponenta, ki je ne želite zaustaviti.
5. Shranite in zapustite datoteko.  
Ko naslednjič izdate ukaz **esadmin system startall** ali **esadmin system stopall**, spremenjena komponenta ne bo zagnana oziroma zaustavljena, kar je v skladu s spremembami, ki ste jih opravili v datoteki lastnosti.

**Zgled:** V tem primeru bo strežnik HTTP zagnan z ukazom **esadmin system startall** (privzeta nastavitve), ne bo pa zaustavljen z ukazom **esadmin system stopall** (dejanje določa poudarjena vrstica v primeru):

```
#####
Details of component 3.
#####
Component.name.3=IBM HTTP server
Component.impl.class.3=com.ibm.es.control.util.component.impl.HTTPControlImpl
Component.nodes.3=search
By default all components are startable
Component.stopable.3=false
```

#### S tem povezana opravila

“Zagon sistema iskanja v podjetju” na strani 267

“Zaustavitev sistema iskanja v podjetju” na strani 269

---

## Izvajanje skrbništva nad iskalnimi strežniki v samostojnem načinu

Če želite zagotoviti visoko razpoložljivost iskalnih strežnikov, lahko iskalne strežnike zaženete za posamezne zbirke, četudi se indeksni strežnik ne izvaja.

### Omejitve

Če želite zaustaviti ali zagnati iskalne strežnike, morate biti skrbnik iskanja v podjetju.

Zagon in zaustavitev iskalnih strežnikov v samostojnem načinu nista mogoča iz skrbniške konzole iskanja v podjetju. Če se indeksni strežnik ne izvaja, dostop do skrbniške konzole ni mogoč.

Preden boste lahko zagnali in zaustavili strežnike v samostojnem načinu, se morate prepričati, da so bili pajek, razčlenjevalnik ter indeksni in iskalni strežniki zagnani vsaj enkrat. S tem zagotovite, da so vse potrebne datoteke na iskalnih strežnikih sinhronizirane.

Če je za zbirko omogočena zaščita na ravni dokumentov, se prepričajte, da je zagnan tudi strežnik pajkanja. S tem zagotovite, da bo mogoče uveljaviti krmilne elemente za zaščito na ravni dokumentov. Če se strežnik pajkanja ne izvaja, sistem v rezultatih iskanja vrne samo dokumente, ki ne zahtevajo overjanja.

### O tej nalogi

Če indeksni strežnik ni dosegljiv, lahko uporabnikom zagotovite nadaljnje iskanje v sistemu tako, da iskalne strežnike zaženete v samostojnem načinu. Ukaze za zagon in zaustavitev iskalnih strežnikov lahko izdate iz poljubnega strežnika iskanja v podjetju, ko gre za

namestitev z več strežniki. Ukazi poskusijo zagnati ali zaustaviti iskalne strežnike za podano zbirko na vseh razpoložljivih iskalnih strežnikih.

## Postopek

Če želite zagnati ali zaustaviti iskalne strežnike v samostojnem načinu, naredite naslednje:

1. Če želite zagnati iskalne strežnike za zbirko, ko se indeksni strežnik ne izvaja:

- a. V strežnik pajkanja in iskalne strežnike se prijavite kot skrbnik iskanja v podjetju, nato pa zaženite storitev CCL (skupna komunikacijska plast):

### AIX, Linux ali Solaris

```
startccl.sh -bg
```

### Ukazni poziv Windows

```
startccl
```

### Skrbniška orodja storitev operacijskega sistema Windows

Če želite zagnati CCL v ozadju:

- 1) Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
- 2) Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.

- b. Na iskalnih strežnikih zaženite skript **startServer**, ki se nahaja v imeniku **WAS\_INSTALL\_ROOT/AppServer/bin**, s čimer zaženete aplikacijo **ESSearchServer** na strežniku **WebSphere Application Server**:

### AIX, Linux ali Solaris

```
./startServer.sh ESSearchServer
```

### Windows

```
startServer ESSearchServer
```

- c. Vnesite naslednji ukaz, v katerem *collection\_id* označuje zbirko, ki je lastnik iskalnih strežnikov, ki jih želite zagnati:

```
esadmin startSearch -cid collection_id
```

2. Če želite zaustaviti iskalne strežnike za zbirko, ko se indeksni strežnik ne izvaja:

- a. V strežnik za iskanje v podjetju se prijavite kot skrbnik iskanja v podjetju.
- b. Vnesite naslednji ukaz, v katerem *collection\_id* označuje zbirko, ki je lastnik iskalnih strežnikov, ki jih želite zaustaviti:

```
esadmin stopSearch -cid collection_id
```

### S tem povezana opravila

“Zagon sistema iskanja v podjetju” na strani 267

“Zaustavitev sistema iskanja v podjetju” na strani 269

---

## Nadziranje dejavnosti iskanja v podjetju

Med nadziranjem dejavnosti sistema in zbirke lahko prikažete status različnih procesov, oprezate za morebitnimi težavami ali prilagajate konfiguracijske nastavitve, da izboljšate učinkovitost.

S pomočjo skrbniške konzole za iskanje v podjetju lahko nadzirate sistem in po potrebi prilagodite delovanje. Za vse glavne dejavnosti (pajkanje, razčlenjevanje, indeksiranje in iskanje) je mogoče prikazati podrobne statistične podatke. Statistični podatki zajemajo informacije o povprečnem odzivnem času in napredku, na primer koliko dokumentov je sistem prepajkal ali indeksiral v času trajanja določene seje.

S klikom ikon lahko zaženete ali zaustavite večino dejavnosti. Operacije, ki so na voljo, vam omogočajo, da zaustavite dejavnost, spremenite njeno konfiguracijo, odpravite nastalo težavo ali znova zaženete obdelavo, ko ste pripravljeni na nadaljevanje.

### S tem povezana opravila

“Zagon sistema iskanja v podjetju” na strani 267

“Zaustavitev sistema iskanja v podjetju” na strani 269

---

## Ocenjevanje števila dokumentov v zbirki

Ko izdelate ali uredite zbirko za iskanje v podjetju, ocenite, koliko dokumentov bo zbirka predvidoma vsebovala. Sistem na podlagi tega števila oceni pomnilniške in diskovne vire, ki jih potrebuje za zbirko, vendar ne da bi omejeval velikost zbirke.

### Preden začnete

Če želite spremeniti ocenjeno velikost zbirke, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

### O tej nalogi

Ko zbirka doseže ocenjeno velikost, sistem ne preneha z dodajanjem dokumentov v indeks. Če konfigurirate opozorila za zbirko in izberete možnost, po kateri vas sistem obvesti, da je število dokumentov v indeksu preseglo omejitvev, privzeta omejitev ustreza vrednosti, ki jo podate kot oceno za število dokumentov v zbirki. Sistem spremlja podano oceno in odstotek, ki podaja mejo za opozorilo, in pošlje elektronsko pošto, ko se bo število dokumentov, konfiguriranih za zbirko, približalo največji vrednosti.

### Postopek

Če želite podati oceno za potencialno velikost zbirke, naredite naslednje:

1. Uredite zbirko, izberite stran Splošno in kliknite **Konfiguriraj splošne možnosti**.
2. V polje **Ocenjeno število dokumentov** vnesite število, s katerim podate, kako velika bo predvidoma vaša zbirka. Privzeta vrednost je 1 000 000 dokumentov.

---

## Nadziranje zbirke


Prikažete lahko splošne informacije o statusu posamezne komponente v zbirki ali izberete možnosti za prikaz podrobnih informacij o posameznih komponentah in URI-jih.


### Praden začnete


Zbirke lahko nadzirajo vsi skrbniški uporabniki iskanja v podjetju. Če želite zagnati ali zaustaviti komponente oziroma omogočiti ali onemogočiti razporede, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik ali operator zbirke.

### Postopek

Če želite nadzirati zbirko, naredite naslednje:

1. Na pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**. Prikažejo se informacije o trenutnem statusu posamezne komponente zbirke.

**Nasvet:** Če urejate zbirko in ste že na strani Splošno, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.

2. Če želite prikazati podrobnejše informacije o URI-ju, kliknite  **Podrobnosti URI-ja**. Morda boste želeli na primer preveriti, ali je določen URI v indeksu oziroma ali je bil indeks, v katerem je URI, prekopiran na iskalne strežnike.
3. Če želite nadzirati posamezno komponento in prikazati podrobnejše statistične podatke o njeni dejavnosti, kliknite ikono **Status**.

---

## Prikaz podrobnosti o URI-jih

Prikažete lahko podrobne informacije o URI-ju. Prikažete lahko trenutne in zgodovinske informacije o tem, kako je dokument, ki ga predstavlja ta URI, pajkan, indeksiran in iskan.

### Praden začnete

Praden predložite zahtevo za prikaz poročila URI ali pošljete poročilo na elektronski naslov, se prepričajte, da je komponenta, od katere želite prejeti informacije, dejavna. Če na primer želite videti podrobnosti o načinu pajkanja in indeksiranja dokumenta ter iskanja po njem, se prepričajte, da se spletni pajek, indeksni strežnik in iskalni strežniki izvajajo. Za sledenje izbrisanemu dokumentu morate konfigurirati možnosti beleženja za sledenje dokumentu.

### O tej nalogi

Zbiranje informacij o URI-ju je dolgotrajen postopek. Izberete lahko možnost za prikaz zahtevanih informacij, nato pa počakate, da se prikažejo. Učinkovitejša pot je, da pošljete poročilo na elektronski naslov, ki ga podate.


Praden prejmete poročilo, morate zagotoviti, da ste za sistem iskanja v podjetju konfigurirali informacije o poštnem strežniku. Te informacije lahko podate, ko konfigurirate možnosti za elektronsko pošto na strani Dnevnik v prikazu Sistem.


Strežnik indeksov in iskalni strežniki lahko nudijo informacije o vseh URI-jih (kot na primer ali je URI v indeksu, ali je bil prekopiran v iskalne strežnike). Za ogled informacij o pajkanju dokumenta, morate podati URI za dokument, ki ga je pajkal spletni pajek.

### Postopek

Za ogled podrobnosti o URI-ju:

1. Na pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**.

**Nasvet:** Če urejate zbirko in ste že na strani Splošno, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.

2. Kliknite  **Podrobnosti o URI-ju**.
3. Na strani Podrobnosti o URI-ju vnesite URI, za katerega želite informacije.
4. Izberite potrditvena polja za tip informacij, ki jih želite prikazati:

**Podrobnosti o pajku (na voljo samo za spletne pajke)**

To potrditveno polje izberite, da prikazete informacije o tem, kako je spletni pajek pajkal dokument in o trenutnem statusu dokumenta v območju pajkanja.

**Podrobnosti o indeksu**

To potrditveno polje izberite, da vidite, ali je bil dokument indeksiran in prekopiran na iskalne strežnike.

**Podrobnosti o iskanju**

To potrditveno polje izberite, da prikazete informacije o tem, kako je mogoče iskati po URI-ju in ali je dokument na voljo za iskanje.

**Dokumenti, ki jih je zbrisal razčlenjevalnik**

To potrditveno polje izberite, če vas zanima, ali je bil dokument izbrisan iz sistema iskanja v podjetju med razčlenjevanjem in če je bil, kakšen je vzrok brisanja dokumenta.

**Dokumenti, izbrisani iz indeksa**

To potrditveno polje izberite, če vas zanima, ali je bil dokument izbrisan iz sistema iskanja v podjetju med indeksiranjem ali analizo in če je bil, kakšen je vzrok brisanja dokumenta.

5. Za prikaz poročila kliknite **Prikaz poročila** in počakajte.
6. Kliknite **Pošlji poročilo**, da pošljete poročilo na elektronski naslov za poznejši prikaz.
  - a. Na strani Pošlji podrobno poročilo URI vnesite elektronski naslov za prejemanje poročila v polju **Elektronski naslov za obvestila**.
  - b. Kliknite **Pošlji poročilo**.

**S tem povezani pojmi**

“Sledenje dokumentov” na strani 288

**S tem povezana opravila**

“Prikazovanje poročil o odstranjenih dokumentih” na strani 289

**S tem povezane povezave**

“Formati URI v indeksu iskanja v podjetju” na strani 107

---

## Nadziranje pajkov


Ogledate si lahko splošne informacije o statusu vsakega pajka v zbirki ali izberete možnosti za prikaz podrobnih informacij o dejavnostih pajka.

**Preden začnete**


Če vam vaša skrbniška vloga onemogoča nadziranje zbirk, si lahko ogledate statistične podatke pajka, ne morete pa spremeniti njegovega vedenja (kot je zagon ali ustavitev pajka).


**Postopek**

Za nadziranje pajka:

1. V pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**.

2. Odprite stran Pajek.

**Nasvet:** Če urejate zbirko in ste že na strani Pajek, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje pajkov.

3. Če se pajek izvaja ali je začasno ustavljen in želite videti podrobne informacije o stanju pajka, kliknite  **Podrobnosti**. Tipi statističnih podatkov, ki jih vidite, se razlikujejo glede na tip pajka.

Če vam vaša skrbniška vloga omogoča izvajanje skrbništva nad procesi za zbirko, lahko zaženete, zaustavite in začasno ustavite pajka, medtem ko si ogledujete podrobnosti o dejavnosti pajka. Če je za pajka možno ustvariti razpored, lahko prav tako omogočite in onemogočite razpored pajkanja.

4. Če je pajek zaustavljen ali začasno ustavljen in želite začeti sejo pajka, kliknite  **Zaženi** ali **Nadaljuj**.

#### **Za spletne pajke:**



Če je bil pajek ustavljen, začne ponovno pajkati in preišče celotno območje pajkanja. Če ste pajka začasno zaustavili, se pajkanje nadaljuje na začetku cilja, kjer ste pajkanje prekinili.

Če želite, da pajek nemudoma začne popolno pajkanje, kliknite ikono **Podrobnosti** nato **Znova začni popolno pajkanje**. Pajek začne pajkati po celotnem območju pajkanja, vključno s stranmi, ki se od zadnjega pajkanja niso spremenile. Vnovično pajkanje vseh dokumentov boste izvedli tako, da na primer spremenite pravila za razčlenjevanje dokumentov in ta pravila želite uveljaviti za dokumente, ki so bili predhodno indeksirani.

#### **Pri vseh ostalih tipih pajkov:**

Če je bil pajek ustavljen, bo ta začel s pajkanjem ob razporejenem datumu in času. Ko pajek vir podatkov prvič pajka, bo izvedel popolno pajkanje. Ko se razporejeno pajkanje ponovi, bo pajek izvedel pajkanje vseh popravkov v viru podatkov (dodatkov dokumentom, izbrisanih dokumentov in sprememb) ali samo dodatkov in sprememb dokumentov. Tip pajkanje lahko konfigurirate v razporedu pajka.

Če pajka niste razporedili, oziroma ne želite, da bi se ta zagnal prej, kliknite ikono **Podrobnosti**. Nato v področju s podrobnostmi o prostoru za pajkanje kliknite ikono za tip pajkanja, ki ga želite zagnati: popolno pajkanje, vsi popravki ali samo novi in spremenjeni dokumenti. Za vsak vir podatkov, ki ga želite pajkati (kot je na primer strežnik, baza podatkov ali podmapa), morate klikniti ustrezno ikono za zagon.

5. Če se pajek izvaja in ga želite ustaviti, kliknite  **Ustavi** ali  **Začasno ustavi**. Pajek preneha s pajkanjem podatkov, dokler ga znova ne zaženete.

Če pajka ponovno zaženete, bo s pajkanjem nadaljeval na začetku cilja, kjer ste pajkanje prekinili. Pajek DB2 bo na primer s pajkanjem nadaljeval v prvi vrstici tabele, ki jo je pajkal, ko ste ga začasno zaustavili.

---

## **Prikaz podrobnosti o dejavnostih spletnega pajka**


S prikazom podrobnosti o dejavnosti spletnega pajka lahko ocenite vsestransko zmogljivost in prilagodite lastnosti spletnega pajka in definicije območja pajkanja, kot je potrebno.


### **Preden začnete**


Vsi skrbniški uporabniki iskanja v podjetju lahko nadzorujejo dejavnosti pajka. Da lahko zaženete ali ustavite pajka, morate biti član s skrbniško vlogo v iskanju v podjetju, skrbnik zbirke ali operater za zbirko.

## Postopek

Za ogled podrobnosti o dejavnosti spletnega pajka:

1. V pogledu Zbirke poiščite zbirko s spletnim pajkom, ki ga želite nadzirati, in kliknite  **Nadzor**.
2. Odprite stran Pajek.

**Nasvet:** Če urejate zbirko in ste že na strani Pajek, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje pajkov.

3. Če se spletni pajek, ki ga želite nadzirati, izvaja ali je začasno ustavljen, kliknite  **Podrobnosti**.
4. Na strani podrobnosti za spletnega pajka si oglejte ali izberite naslednje možnosti za ogled statističnih podatkov o trenutni in pretekli dejavnosti pajka.
  - Za prikaz števila niti, ki aktivno pajkajo spletna mesta in števila niti, ki niso dejavna, kliknite **Podrobnosti niti**.
  - Za informacije o spletnih mestih, ki jih pajek aktivno pajka, kliknite **Aktivna spletna mesta**.
  - Kliknite **Nedavno pajkani URL-ji**. Te informacije povejo, kaj je pajek nedavno pajkal. Če se postavke na seznamu ne spremenijo, ko osvežite stran, potem pajkanje ne poteka.
  - Za ogled poročil o pretekli dejavnosti pajka kliknite **Zgodovina pajka**.
  - V območju **status URL-ja** vnesite URL, o katerem želite informacije.
    - a. Za statusne informacije o URL-ju kliknite **Podrobnosti URL-ja**. Podrobnosti URL-ja lahko zahtevate samo za URL-je, ki so že bili pajkani.
    - b. Za podajanje informacij, ki jih želite vključiti v poročilo o spletnem mestu, kateremu pripada URL, kliknite **Podrobnosti spletnega mesta**. Zahtevate lahko podrobnosti spletnega mesta za predhodno pajkano spletno mesto ali za spletno mesto, ki še ni bilo pajkano.

To možnost na primer uporabite, da vidite, ali je URL v območju pajkanja, ali je že bil pajkan ali pravkar odkrit, kdaj mora biti ponovno pajkan in informacije o zadnjem poskusu pajkanja spletnega mesta. Prav tako lahko zaprosite za ogled vsebine datoteke robots.txt za spletno mesto, kar vam lahko pomaga ugotoviti, zakaj se na spletnem mestu izvaja pajkanje.

## Podrobnosti o nitih spletnega pajka

Spletnega pajka lahko nadzirate, da vidite število niti, ki aktivno pajkajo spletna mesta in število niti, ki niso dejavne.

Ko si ogledujete podrobnosti o spletnem pajku med nadziranjem zbirke, lahko vidite status niti pajka. Stanja, ki jih boste najverjetneje videli, so:

### Čakam

Pomeni, da nit nima URL-ja, ki bi ga lahko pajkala. Do tega pride takrat, ko nit konča pajkanje in pajek ne more najti več URL-jev, da bi pajkal dovolj hitro. Če je na primer lastnost pajka, ki krmili, kako dolgo mora pajek čakati, preden lahko priključite drugo stran iz istega spletnega mesta, previsoka, lahko ta prepreči dovolj hitro dobavo URL-jev.

### Pridobivam

Pomeni, da nit prenaša stran s spletnega mesta.

### Dokončano

Pomeni, da nit pošilja strani, ki jih je pajkala, preostalemu pajku, ni pa še pripravljena na pajkanje drugega URL-ja.

## Odloženo

Pomeni, da je pajek začasno ustavljen

V najboljšem primeru vse niti nenehno pridobivajo strani. Če so niti pogosto v dokončanem stanju, potem je mogoče, da ima baza podatkov težave s prepustnostjo.

Če so niti pogosto v stanju čakanja, ponovno preglejte podane vrednosti za polje **Največje število aktivnih gostiteljev** v lastnostih pajka. Če je vrednost nizka, potem morebiti v območju pajkanja ni dovolj spletnih mest, da bi niti ostale dejavne, ali pa ni dovolj primernih URL-jev za pajkanje. Do nizke dejavnosti lahko pride med drugim zaradi nepravilnosti v iskanju DNS in iskanju robot.

## Aktivna mesta spletnega pajka

Nadzirate lahko spletnega pajka, da vidite informacije o spletnih mestih, ki jih pajek aktivno pajka.

Ko si ogledujete podrobnosti o spletnem pajku med nadziranjem zbirke, lahko vidite statistične podatke o aktivnih spletnih mestih. Statistični podatki kažejo:

- Koliko URL-jev je pajek prenesel iz svoje notranje baze podatkov v spomin za pajkanje v tem času
- Koliko URL-jev je pajek skušal pajkati do tega trenutka
- Koliko časa ostane, preden bo spletno mesto deaktivirano in odstranjeno iz spomina za to ponovitev pajka
- Kako dolgo je bilo spletno mesto v spominu

Te informacije se spreminjajo iz trenutka v trenutek, ko pajek napreduje skozi pravila pajkanja, ki so zanj konfigurirana. V najboljšem primeru je število aktiviranih URL-jev blizu vrednosti, ki je konfigurirana za polje **Največje število aktivnih gostiteljev** v lastnostih spomina pajka.

Če je število aktiviranih URL-jev blizu nič, potem pajek ne najde primernih URL-jev. Do takšne nizke dejavnosti lahko pride zaradi neuspešnih iskanj DNS, težav s povezljivostjo omrežja, napakami v bazah podatkov in težavami z definicijami območja pajkanja. Primer:

- Če je bilo v spominu dlje časa veliko spletnih mest, a le malo URL-jev pajkanih, gre za težave s povezljivostjo omrežja.
- Če na seznamu ni dovolj spletnih mest, gre za težave z definicijami območja pajkanja ali težave z iskanjem DNS.
- Če so spletna mesta pajkana ob razumni hitrosti, a v spominu ostane veliko nepajkanih URL-jev, uredite lastnosti spomina pajka in prilagodite vrednost časovne omejitve v polju **Čas, ko lahko vsak gostitelj ostane dejaven**, tako da spletna mesta ostanejo v spominu dlje časa.

## Hitrost pajkanja spletnega pajka

Nadzirate lahko spletnega pajka, da vidite informacije o tem, kako hitro pajek prenaša strani iz spletnih mest.

Ko si ogledujete podrobnosti o spletnem pajku med nadziranjem zbirke, lahko vidite statistične podatke o tem, kako hitro pajek pajka podatke (hitrost pajkanja). Prav tako lahko vidite statistične podatke o tem, koliko URL-jev je pajek prepajkal od takrat, ko se je pričela trenutna seja.

Hitrost pajkanja je število pajkanih strani na sekundo. To število je odvisno od mnogih lastnosti, ki jih lahko konfigurirate za spletnega pajka:



- Število niti pajka
- Število dejavnih spletnih mest
- Čas, ko mora pajek čakati, preden lahko priključite drugo stran iz istega spletnega strežnika

Če ima pajek eno aktivno spletno mesto na nit pajka in mora pajek počakati dve sekundi, preden lahko priključite drugo stran iz istega spletnega strežnika, potem pajek ne more pajkati hitreje kot eno stran na nit v dveh sekundah. Če na primer pajek uporablja privzeto število niti (200), potem lahko pajek pajka 100 strani na sekundo za 200 niti.

Če obstaja dvakrat toliko spletnih mest, kot je niti pajka, in mora pajek počakati dve sekundi, preden lahko priključite drugo stran iz istega spletnega strežnika, potem lahko pajek doseže eno stran na nit v sekundi. V tem primeru bi omejevalni dejavniki postali omrežne hitrosti prenosa in prepustnost baze podatkov. Ko se hitrost pajkanja poravnava z številom niti pajka, dejavnimi spletnimi mesti in pajkovim časom čakanja, je to znak visoke zmogljivosti pajka.

Drug dejavnik, ki ga morate upoštevati pri nadziranju zmogljivosti spletnega pajka, je število URL-jev, ki jih je pajek prepajkal od pričetka trenutne seje. Povprečje dolgotrajne prepustnosti izračunate tako, da to število delite s celotnim časom izvajanja pajka. Če se to število ne povečuje, je pajek končal ali pa ne more nadaljevati. Napredek pajka lahko na primer onemogočijo napake pri povezljivosti omrežja, napake v bazi podatkov in nepravilnosti pri iskanju DNS.

## Izdelava poročil za spletnega pajka

S prikazom poročil o pretekli dejavnosti spletnega pajka lahko ocenite vsestransko zmogljivost in prilagodite lastnosti spletnega pajka in definicije območja pajkanja, kot je potrebno.

### Predn začnete

Če vam vaša skrbniška vloga onemogoča nadziranje zbirk, si lahko ogledate statistične podatke pajka in izdelate poročila o dejavnosti pajka, ne morete pa spremeniti njegovega vedenja (kot je zagon ali ustavitev pajka).


### O tej nalogi


Informacije o dejavnosti spletnega pajka vam lahko dajo različni tipi poročil. Pri določenih tipih poročil dobite informacije takoj, ko se zberejo iz pajkove notranje baze podatkov. Izdelava poročila spletnega mesta in poročilo o statusnih kodah HTTP traja nekaj časa. Če izdelate ta tipa poročil, lahko za prejetje poročila podate elektronski naslov, namesto da čakate na rezultate, da bodo vrnjeni na skrbniški konzoli za iskanje v podjetju.


Za informacije o tem, kako si razlagati statistične podatke v poročilu, kliknite **Pomoč**, medtem ko nadzirate spletnega pajka in izdelujete poročila.

### Postopek

Za izdelavo poročil spletnega pajka:

1. V pogledu Zbirke poiščite zbirko s spletnim pajkom, ki ga želite nadzirati, in kliknite  **Nadzor**.
2. Odprite stran Pajek.

**Nasvet:** Če urejate zbirko in ste že na strani Pajek, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje pajkov.

3. Če se spletni pajek, za katerega želite izdelati poročila, izvaja ali je začasno ustavljen, kliknite  **Podrobnosti**.
4. Na strani podrobnosti za spletnega pajka izberite možnost za tip poročila, ki ga želite izdelati:
  - V območju **Povzetek statusa pajka** kliknite **Zgodovina pajka**, da izdelate poročila o pajku in vseh spletnih mestih, ki jih odkrije ali pajka.
  - V območju **status URL-ja** podajte URL določenega spletnega mesta, za katerega želite izdelati poročilo, in nato kliknite **Podrobnosti spletnega mesta**.
5. Za zgodovino pajka in poročila spletnega mesta lahko izberete potrditveno polje za vsak statistični podatek, ki ga želite videti v poročilu, nato pa kliknite **Prikaz poročila**.  
Pri teh tipih statističnih podatkov pajek vrne poročilo na skrbniško konzolo takoj, ko pridobi informacije iz svoje notranje baze podatkov.
6. Če izdelujete poročilo o zgodovini pajka, lahko podate možnosti za izdelavo poročila spletnega mesta, nato pa kliknite **Zaženi poročilo**.  
To poročilo se izdela s statističnimi podatki, katere ste vključili, in shrani v datoteko, ki jo podate (ime datoteke mora biti absolutno). Podate lahko, da želite prejemati elektronsko sporočilo, ko je poročilo izdelano.
7. Če izdelujete poročilo o zgodovini pajka, lahko podate možnosti za izdelavo poročila o statusni kodici HTTP, nato pa kliknite **Zaženi poročilo**.  
To poročilo vsebuje informacije o številu porazdeljenih statusnih kod HTTP na spletno mesto. Poročilo se shrani v datoteko, ki jo podate (ime datoteke mora biti absolutno). Podate lahko, da želite prejemati elektronsko sporočilo, ko je poročilo izdelano.  
S pomočjo tega poročila lahko preverite, katere strani vračajo veliko statusnih kod 4xx (ki sporočajo, da sistem ni našel strani), statusnih kod 5xx (ki sporočajo težave s strežnikom), statusnih kod 6xx (ki nakazujejo težave s povezljivostjo) in tako naprej.  
To poročilo je najbolj koristno takrat, ko je pajek že dejaven nekaj časa (kot na primer pajek, ki je dejaven nekaj tednov). Pomaga vam lahko pri ugotavljanju izginulih spletnih mest, pravkar prispelih spletnih mest, mest z velikim številom URL-jev (ki lahko kažejo na odvečno pajkanje baze podatkov Lotus Notes) in mest s povratnim datotečnim sistemom, ki ga streže strežnik HTTP. Če spletna mesta z velikim številom statusnih kod HTTP ne prispevajo k indeksu, lahko zmogljivost pajka izboljšate tako, da odstranite spletna mesta iz območja pajkanja.

## Statusne kode HTTP, vrnjene spletnemu pajku

Pri nadziranju spletnega pajka si lahko ogledujete informacije o statusnih kodah HTTP, ki jih pajek prejme od strani, ki jih skuša pajkati.

### Povzetek tabele

Pri nadziranju zgodovine spletnega pajka ali statusa določenega URL-ja lahko vidite informacije o statusnih kodah HTTP, ki so bile vrnjene pajku. S pomočjo teh informacij lahko upravljate območje pajkanja in optimizirate zmogljivost pajka. Če na primer pajek prejme veliko število statusnih kod HTTP za URL in statusne kode navajajo, da strani na tem mestu ni mogoče pajkati, lahko povečate zmogljivost tako, da odstranite ta URL iz območja pajkanja.

Naslednja tabela navaja statusne kode HTTP in način, kako si jih spletni pajek razlaga. Vrednosti od 100 do 505 so standardne statusne kode HTTP (za več informacij glejte <http://www.w3.org/Protocols/rfc2616/rfc2616.html>). Preostale statusne kode so last iskanja v podjetju in spletnega pajka.

Tabela 8. Statusne kode HTTP iz spletnega pajka

Koda	Opis	Koda	Opis	Koda	Opis	Koda	Opis
NULL	Nepajkan	400	Neveljavna zahteva	500	Notranja napaka strežnika	693	Napaka pri izbiranju (URLFetcher)
100	Nadaljaj	401	Nepooblaščen	501	Ni implementiran	694	Napaka pri pisanju (URLFetcher)
101	Preklapljanje protokolov	402	Zahtevano plačilo	502	Neveljaven prehod	695	Nepopolno oglavje bloka (URLFetcher)
200	Uspešno	403	Prepovedano	503	Storitev ni na voljo	699	Nepričakovana napaka (URLFetcher)
201	Izdelano	404	Ni bil najden	504	Čakalni čas prehoda	700	Napaka pri razčlenjevanju (ni konca oglavja)
202	Sprejeto	405	Metoda ni dovoljena	505	Različica HTTP ni podprta	710	Napaka pri razčlenjevanju (oglavje)
203	Nepooblašcene informacije	406	Nesprejemljivo	611	Napaka pri branju	720	Napaka pri razčlenjevanju (ni kode HTTP)
204	Ni vsebine	407	Zahtevano overjanje proxy-ja	612	Napaka pri povezovanju	730	Napaka pri razčlenjevanju (telo)
205	Ponovno nastavite vsebino	408	Časovna omejitev zahteve	613	Čakalni čas branja	740 ali 4044	Izvzet z datoteko robots.txt
206	Delna vsebina	409	Neskladje	614	Usklajevanje SSL-a ni uspelo	741	Roboti začasno niso na voljo
300	Več izbir	410	Pošel	615	Druga napaka pri branju	760	Izvzet z definicijo območja pajkanja
301	Trajno premaknjen	411	Zahtevana dolžina	616	Nepравilnost FBA	761	Ne dovoli lokalno območje pajkanja; dovoli globalno
302	Najden	412	Predpogoj ni uspel	617	Napaka pri kodiranju	770	Neveljaven protokol ali nestandardna sistemska vrata
303	Glejte ostalo	413	Prevelik objekt zahteve	618	Preusmeritev brez URL-ja preusmeritve	780	Izvzet z izločitvami tipa datoteke
304	Nespremenjen	414	URI zahteve je predolg	680	Napaka pri iskanju DNS	786	Neveljaven URL
305	Uporabite proxy	415	Nepodprt tip medija	690	Nepравilno oblikovan URL	2004	Ni oznake META indeksa
306	(Neuporabljen)	416	Zahtevanemu obsegu ni mogoče ugoditi	691	Izgubljena povezava (URLFetcher)	3020	Mehka preusmeritev

Tabela 8. Statusne kode HTTP iz spletnega pajka (nadaljevanje)

Koda	Opis	Koda	Opis	Koda	Opis	Koda	Opis
307	Začasna preusmeritev	417	Pričakovanje ni uspelo	692	Čakalni čas pisanja (URLFetcher)		

## Opombe k tabeli

### Statusne kode 4xx

Kodo 400 (neveljavna zahteva) boste redko videli. V skladu z standardom statusnih kod HTTP kode 4xx nakazujejo, da odjemalec (pajek) ni uspel. Težava pa se ponavadi nahaja na strežniku ali v URL-ju, ki ga je pajek prejel kot povezavo. Nekateri spletni strežniki na primer ne dopuščajo URL-jev, ki skušajo prikrmariti navzgor od korena spletnega mesta (kot je <http://xyz.ibm.com/../../sales>). Drugi spletni strežniki s tem pomikanjem navzgor nimajo težav in ne upoštevajo operatorja nadrejenega imenika (..) ko je pajek že pri korenu.

Nekateri strežniki obravnavajo zahtevo za koren spletnega mesta kot napako, nekatere zastarele povezave pa lahko zahtevajo operacije, ki se jih več ne prepozna ali niso več v uporabi. Ko se od aplikacijskega strežnika zahteva stran, ki je več ne streže, ta vrne izjemo, ki povzroči, da spletni strežnik vrne statusno kodo HTTP 400, saj zahteva ni več veljavna.

- 615** Pomeni, da je strežnik pajkanja, ki prenaša podatke s spletnih mest, naletel na nepričakovano izjemo. Mnogo takšnih statusnih kod lahko kaže na težave s pajkom.

### Statusne kode 61x

Z izjemo kode 615, kažejo statusne kode 61x na težave, ki lahko nastanejo pri pajkanju, kot je recimo časovna omejitev. Pri naslednjih statusnih kodah bo morebiti potrebno izvesti popravke:

#### 611, 612 in 613

Počasna spletna mesta in nizka zmogljivost omrežja so lahko vzrok za te težave.

- 611** Pomeni, da se je napaka pojavila, ko je pajek priklical dokument.
- 612** Pomeni, da se je napaka pojavila, ko se je pajek poskušal povezati s spletnim strežnikom.
- 613** Pomeni, da se je časovna omejitev pojavila, ko je pajek iskal dokument.

- 614** Pomeni, da pajek ne more pajkati zaščiteneh (HTTPS) spletnih mest. Če menite, da bi morala biti ta spletna mesta dostopna, preverite, ali so certifikati na strežniku pajkanja in ciljnem spletnem strežniku pravilno nastavljeni. Če je spletno mesto na primer potrjeno s priznanimi pooblastili za certifikate, lahko v varno shrambo, ki jo uporablja pajek, dodate nova pooblastila za certifikate.

Prav tako si oglejte, kako so lastnoročno podpisani certifikati konfigurirani na spletnih mestih, ki jih poskušate pajkati. Pajek je konfiguriran tako, da sprejme lastnoročno podpisane certifikate. Nekatera spletna mesta izdelajo lastnoročno podpisan certifikat za korenski URL (kot je <http://sales.ibm.com/>), ki ga nato skušajo uporabiti v poddomenah (kot je <http://internal.sales.ibm.com/>). Pajek ne more sprejeti certifikatov, ki se uporabljajo na ta način. Lastnoročno podpisane certifikate sprejema samo, če se ime domene predmeta ([sales.ibm.com](http://sales.ibm.com/)) in podpisnik certifikata ujemata z imenom domene strani, ki je zahtevana.

**616** Pomeni, da se obrazec za prijavo za overjanje na osnovi obrazca še vedno pojavi pri prenosu po ponovnem overjanju.

Če informacije, ki so na voljo v konfiguracijski datoteki FBA (obrazec za prijavo ter podatki overjanja, kot so uporabniško ime, geslo itd.), ne uspejo overiti pajka, se statusna koda 616 dodeli vsem stranem, ki temeljijo na overjanju na osnovi obrazca. Naloga skrbnika je, da izve, zakaj konfiguracija FBA ne deluje.

**617** Pomeni, da ni mogoče izdelati niza iz bajtne vsebine dokumenta, ker je kodirni niz (charset) neveljaven ali ker dokument vsebuje neveljavne bajte.

**618** Pomeni, da URL preusmerjanja ni veljaven, ko pajek prejme naslednje statusne kode HTTP. Možno je, da mesto oglavja odgovora HTTP neveljavno.

301 Trajno premaknjen  
302 Najden

**680** Pomeni, da pajek ni mogel pridobiti naslovov IP za gostitelje v območju pajka, morebiti zaradi težav z dostopom do omrežja. Ta tip težave pomeni, da pajek ne more pajkati celotnih spletnih mest in ne samo, da ni mogel pajkati nekaterih URL-jev. Mnogo takšnih statusnih kod znatno zmanjša prepustnost.

#### Statusne kode 69x

Statusne kode od 690 do 699 se nikoli ne zapišejo v pajkovo trajno bazo podatkov. Te kode predstavljajo rezultate, ki ne ponazarjajo resničnega rezultata prenosa iz oddaljenega gostitelja, ampak trenutni položaj znotraj pajka, kot je recimo komponenta, ki se zapre, medtem ko druga čaka na rezultat ali pošilja rezultat. Te statusne kode se pojavijo v nekaterih dnevnikih, a ne v trajnih zapisih, in jih tako ne smete uporabljati kot vrednosti, nastavljene z izbiro.

#### Statusne kode 7xx

Kode 7xx povečini nastanejo zaradi pravil v območju pajkanja:

##### 710 - 730

Pomenijo, da pajek zaradi težav ni mogel do konca izvesti prenosa ali da je pajek na spletnem mestu naletel na neveljavne podatke HTML. Če velikokrat srečate te statusne kode, se posvetujte s predstavnikom za podporo iskanja v podjetju.

##### 740 ali 4044

Pomenita, da vsebine datoteke ni mogoče indeksirati, ker je dokument izključen zaradi omejitev v datoteki spletnega mesta robots.txt.

**740** Pomeni, da je možno sidrne povezave, ki kažejo na izključen dokument, vključiti v indeks.

**4044** Pomeni, da so sidrne povezave v dokumentih, ki kažejo na izključen dokument, prav tako izključene iz indeksa.

**741** Pomeni, da spletno mesto ima datoteko robots.txt, ki omogoča pajkanje, vendar prenos ni uspel. Če zaporedoma ne more pajkati URL-ja, se ta odstrani iz območja pajkanja. Če se vam ta tip statusne kode pogosto prikazuje, preverite, ali ciljno mesto začasno ali trajno ni na voljo. Če ciljno spletno mesto ni več na voljo, ga odstranite iz območja pajkanja.

Preostale statusne kode 7xx se večinoma pojavijo, ko spremenite območje pajkanja, pajek pa se izvaja že kar nekaj časa. Te statusne kode običajno ne pomenijo težav, ki bi jih bilo potrebno odpraviti.

**3020** Pomeni, da dokument s statusno kodo 200 vsebuje oglavje mesta, ki napoti uporabniškega posrednika na drug URL.

---

## Nadziranje razčlenjevalnika

Nadziranje razčlenjevalnika, ko želite videti informacije o dokumentih, ki jih razčlenjevalnik analizira, preden so dodani v indeks iskanja v podjetju. S pomočjo možnosti lahko pregledate statistične podatke in izvajate skrbništvo nad delovanjem razčlenjevalnika.

### Preden začnete

Če vam vaša skrbniška vloga onemogoča nadziranje zbirk, si lahko ogledate status razčlenjevalnika, ne morete ga pa zagnati ali ustaviti.

### O tej nalogi


Pri nadziranju razčlenjevalnika vidite posnetek dejavnosti razčlenjevalnika, ki nudi statistične podatke o dejavnostih razčlenjevalnika v določenem trenutku. Statistični podatki vam povejo število dokumentov, ki so bili pajkani in so v postopku razčlenjevanja ali nanj čakajo, ter število dokumentov, ki so bili pajkani in čakajo na shranitev v indeks.

Ko je razčlenjevalnik dejaven, vam sporočila nudijo dodatne informacije o stanju razčlenjevalnika. Primer:


- Možno je, da razčlenjevalnik aktivno razčlenjuje dokumente.
- Možno je, da je razčlenjevalnik v mirovanju. Razčlenjevalnik miruje, dokler ni na voljo več dokumentov za razčlenjevanje. Če se pojavijo napake, razčlenjevalnik počaka na ponovni zagon. Razčlenjevalnik se sam ponovno zažene, če ni na voljo storitev razčlenjevalnika (samodejni ponovni zagon se na primer pojavi, ko ni mogoče vzpostaviti povezave do storitve razčlenjevalnika ali ko so vsi razčlenjevalnikovi navidezni računalniki Java zasedeni z drugimi zbirkami).
- Razčlenjevalnik lahko začasno ustavite (razčlenjevalnik lahko na primer začasno ustavite, dokler ni končana gradnja indeksov).

### Postopek


Za nadziranje razčlenjevalnika za zbirko:

1. V pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**.
2. Odprite stran Razčleni.


**Nasvet:** Če urejate zbirko in ste že na strani Razčleni, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.

3. Če se razčlenjevalnik izvaja in želite videti podrobne informacije o stanju dejavnosti razčlenjevanja, kliknite  **Podrobnosti**.

Če vam vaša skrbniška vloga omogoča izvajanje skrbništva nad procesi za zbirko, lahko medtem, ko si ogledujete podrobnosti o dejavnostih razčlenjevanja, zaženete in ustavite razčlenjevalnik.

4. Če je razčlenjevalnik zaustavljen in ga želite zagnati, kliknite  **Zaženi**.

Ko prvič izdelate zbirko, razčlenjevalnika zaženete po tem, ko pajek začne s pajkanjem podatkov. S tem boste zagotovili, da ima razčlenjevalnik na voljo podatke za analiziranje in kategoriziranje. Razen, če spremenite pravila razčlenjevanja, lahko pustite, da se razčlenjevalnik neprekinjeno izvaja.

5. Če se razčlenjevalnik izvaja in ga želite zaustaviti, kliknite  **Ustavi**.

Razčlenjevalnik morate zaustaviti in nato znova zagnati, če spremenite pravila za razčlenjevanje. Če na primer spremenite konfiguracijo razčlenjevalnika, ga morate ustaviti in ponovno zagnati, preden bodo začele spremembe veljati.

---

## Nadziranje dejavnosti indeksa za zbirko


Nadziranje indeksa za zbirko, ko želite videti potek gradnje indeksa, omogočanje in onemogočanje razporeda indeksov, zagon in zaustavitev delovanja indeksiranja.

### Preden začnete






Vsi skrbniški uporabniki iskanja v podjetju lahko nadzorujejo dejavnosti indeksov. Da lahko zaženete ali ustavite gradnjo indeksa oziroma omogočite ali onemogočite razpored indeksa, morate biti član s skrbniško vlogo v iskanju v podjetju, skrbnik zbirke ali operater za zbirko.

### Postopek

Za nadziranje čakalne vrste indeksov za zbirko:

1. V pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**.
2. Odprite stran Indeksiranje.

**Nasvet:** Če urejate zbirko in ste že na strani Indeksiranje, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.

3. Če ste izdelali razpored indeksa in le-tega ne želite graditi ob načrtovanem datumu in času, kliknite  **Onemogoči razpored**. Gradnja indeksa se bo izvedla šele, ko razpored omogočite ali sami zaženete proces gradnje indeksa.
4. Če ste izdelali razpored indeksa, vendar je razpored gradnje onemogočen, kliknite  **Onemogoči razpored**. Gradnja indeksa se bo izvedla na datum in čas, ki ste ga podali v razporedu indeksa.
5. Če je indeks zaustavljen in ga želite zagnati, kliknite  **Zaženi**.  
Običajno se indeksiranje izvaja redno. Če indeksiranje med gradnjo ustavite ali onemogočite razpored indeksa, lahko gradnjo indeksa izsilite tako, da kliknete **Zaženi**.
6. Če je gradnja indeksa v poteku in jo želite zaustaviti, kliknite  **Ustavi**.  
Gradnjo delta indeksa na primer ustavite takrat, ko želite vsiliti gradnjo glavnega indeksa potem, ko ste spremenili tip kategorizacije, ki se uporablja v zbirki.
7. Če se med gradnjo indeksa pojavijo napake, kliknite  **Napaka**.  
Prikaže se stran Vsebina datoteke dnevnika, kjer lahko vidite dodatne informacije o napakah pri indeksiranju. Na tej strani lahko izberete posamezna sporočila o napakah in tako vidite podrobnosti o težavi.

---

## Nadziranje čakalne vrste indeksov za iskanje v podjetju

V čakalni vrsti indeksov si lahko ogledate status vseh gradenj indeksov, ustavite indeks, ki je v postopku gradnje, ali zbrisete indeks iz vrste.

### Preden začnete




Da lahko izvajate skrbništvo nad čakalno vrsto indeksov, morate biti član s skrbniško vlogo v iskanju v podjetju.

### O tej nalogi

Naenkrat lahko gradite več indeksov, vendar je lahko naenkrat v čakalni vrsti le en indeks na zbirko. Ko konfigurirate možnosti indeksa za sistem, podate, koliko indeksov si lahko hkrati deli čakalno vrsto in vire indeksiranja.

### Postopek

Za nadziranje čakalne vrste indeksov:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Izberite stran Indeksiranje.  
Prikaže se seznam zbirk z indeksi v čakalni vrsti indeksov. Za vsak indeks lahko vidite tip indeksa v gradnji (delta ali celoten), čas, ko je indeks vstopil v čakalno vrsto ter čas, ko je se pričela gradnja indeksa (če gradnja poteka).
3. Za izvajanje skrbništva nad posameznim indeksom kliknite ikono **Status**.  
Ogledate si lahko na primer, koliko časa še preostane do dokončanja indeksa, ali koliko dokumentov je v indeksu, lahko pa tudi onemogočite razpored indeksov.
4. Indeks v gradnji ustavite tako, da kliknete  **Ustavi**.  
Če ste na primer spremenili pravila kategorij, lahko ustavite gradnjo delta indeksa ter tako vsilite pričetek gradnje glavnega indeksa.  
Indeks po zaustavitvi zaženete tako, da bodisi počakate, da indeks vstopi v čakalno vrsto ob naslednjem načrtovanem času, bodisi kliknete ikono **Status** za nadziranje indeksa in nato  **Zaženi**, da pričnete z gradnjo indeksa.
5. Za odstranitev indeksa iz čakalne vrste indeksov kliknite  **Odstrani**.

---

## Nadziranje iskalnih strežnikov





Lahko si ogledate podrobne informacije o statusu za dejavnost iskalnega strežnika za določeno zbirko ali pa podrobne informacije o statusu za iskalne strežnike po celotnem vašem sistemu iskanja v podjetju.

### Preden začnete

Vsi skrbniški uporabniki iskanja v podjetju lahko nadzorujejo iskalne strežnike za zbirke, za katere imajo skrbniška pooblastila. Da lahko nadzirate vse iskalne strežnike v sistemu iskanja v podjetju, morate biti član s skrbniško vlogo v iskanju v podjetju.

Da lahko zaženete ali ustavite iskalni strežnik, morate biti član s skrbniško vlogo v iskanju v podjetju, skrbnik zbirke ali operater za zbirko.

### Postopek

1. Za nadziranje iskalnih strežnikov za posamezno zbirko:
  - a. V pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**.
  - b. Odprite stran Iskanje.  
  
**Nasvet:** Če urejate zbirko in ste že na strani Iskanje, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.
2. Da lahko nadzirate vse iskalne strežnike v sistemu iskanja v podjetju:
  - a. Kliknite **Sistem**, da odprete pogled Sistem.
  - b. Izberite stran Iskanje.
3. Če je iskalni strežnik zaustavljen in ga želite zagnati, kliknite  **Zaženi**.
4. Če se iskalni strežnik izvaja in ga želite zaustaviti, kliknite  **Ustavi**.  
Če omogočite ali onemogočite iskalni predpomnilnik, spremenite velikost iskalnega predpomnilnika ali spremenite hitre povezave, morate zaustaviti ali ponovno zagnati iskalne strežnike, da bodo spremembe začele veljati.
5. Za ogled povzetka o tem, koliko časa iskalni strežnik porabi za obdelavo iskalnih zahtev, kliknite **Odzivni časgodovina**.  
Poročilo prikazuje povprečni čas v milisekundah, ki ga je iskalni strežnik potreboval za odziv na iskalne poizvedbe na določeni datum.



Povprečni odzivni čas pove, kako dobro deluje sistem, ter ustreza kakovosti storitve. Povečan odzivni čas lahko pomeni, da je sistem pod visoko obremenitvijo. Sistem lahko bremenijo število zbirk v iskanju in velikost zbirke.

6. Za ogled seznama najpogosteje predloženih poizvedb, kliknite **Pogoste poizvedbe**.

Poročilo prikazuje ključne besede v 50 najpogosteje predloženih poizvedbah in kolikokrat so uporabniki predložili določeno poizvedbo.

S pregledovanjem najpogostejših poizvedb lahko ugotovite kandidate za hitre povezave. Z izdelavo hitrih povezav lahko izboljšate kakovost iskanja za več uporabnikov. Zagotovite lahko, da so med rezultati iskanja zmeraj vrnjeni ustrezni dokumenti.

Prav tako lahko ustvarite povezave do virov, ki odgovarjajo na poizvedbe s portala podjetja. Če na primer uporabniki pogosto iščejo informacije o stroških, izdelajte na vaši intranet domači strani povezavo na stran, ki vsebuje informacije o postopkih poročila o stroških.

7. Za ogled seznama nazadnje predloženih poizvedb, kliknite **Zadnje poizvedbe**.

Poročilo prikazuje ključne besede v 50 najnovejših poizvedbah.

S pregledovanjem zadnjih poizvedb lahko ugotovite trenutne trende in nujna stanja v organizaciji. Tako lahko na primer opazite povečano zanimanje za kakšno temo.

Povečano zanimanje lahko pomeni, da je za to temo potrebna hitra povezava ali da morate poskrbeti, da bo tema na voljo uporabnikom še na druge načine (kot je povezava na portalu podjetja).

## Spreminjanje načina izračuna statističnih podatkov poizvedb

Spremenite lahko način, na katerega sistem izračunava število priljubljenih poizvedb in nedavnih poizvedb.

### O tej nalogi

Med nadzorom rezultatov iskanja lahko izberete možnosti za ogled seznama 50 najbolj priljubljenih poizvedb in seznama 50 nazadnje obdelanih poizvedb. V privzeti konfiguraciji iskalnega strežnika se poizvedbe z enakovrednim nizom poizvedbe in različnimi nastavitvami obsega štejejo kot neodvisni rezultati. Zaradi tega se na primer število poizvedb poveča, ko uporabnik klikne možnost za ogled naslednje strani rezultatov.

Način izračunavanja statističnih podatkov o poizvedbah lahko spremenite tako, da uredite datoteko `runtime-generic.properties` za iskalni strežnik. Če nastavite parameter **distinctRecentQueryCheck** na vrednost `true`, sistem kot neodvisne poizvedbe šteje le poizvedbe, ki vrnejo začetno stran z rezultati.

Če želite urediti lastnosti iskalnega strežnika, se morate prijaviti kot skrbnik iskanja v podjetju. Da lahko zaženete ali ustavite iskalni strežnik, morate biti član s skrbniško vlogo v iskanju v podjetju, skrbnik zbirke ali operater za zbirko.

### Postopek

Če želite spremeniti način izračunavanja statističnih podatkov o poizvedbah, naredite naslednje:

1. Prijavite se kot skrbnik iskanja v podjetju. Ta ID uporabnika je bil podan med namestitvijo OmniFind Enterprise Edition. V konfiguraciji z več strežniki se prijavite v iskalni strežnik.
2. Uredite naslednjo datoteko, pri čemer je `collection_ID` ID zbirke, ki jo želite spremeniti, `node_ID` pa označuje iskalni strežnik, na katerem boste opravili spremembo:  
`ES_NODE_ROOT/master_config/collection_ID.runtime.node_ID/runtime-generic.properties`

3. Spremenite vrednost parametra **distinctRecentQueryCheck** v **true**, nato pa shranite in zaprite datoteko.
4. V konfiguraciji z več strežniki ponovite zgornje korake še na drugem iskalnem strežniku.
5. Če želite spremeniti vedenje za statistične podatke o poizvedbah še za druge zbirke, ponovite zgornje korake za vsako zbirko, ki jo želite spremeniti.
6. Če želite, da spremembe stopijo v veljavo, se prijavite v konzolo za skrbništvo nad iskanjem v podjetju in znova zaženite iskalne strežnike:
  - a. Kliknite **Sistem**, da odprete pogled Sistem.
  - b. Izberite stran Iskanje.
  - c. Za vsak iskalni strežnik, ki ste ga spremenili, kliknite  **Zaustavi**.
  - d. Za vsak iskalni strežnik, ki ste ga spremenili, kliknite  **Zaženi**.

---

## Nadziranje poslušalnika podatkov

Nadziranje poslušalnika podatkov za ogled statusa in podrobnosti o dejavnosti aplikacije poslušalnika podatkov za odjemalca.

### Preden začnete

Da lahko nadzirate poslušalnik podatkov, morate biti član s skrbniško vlogo v iskanju v podjetju.

**Pomembno:** Poslušalnik podatkov v naslednjih izdajah ne bo podprt. Za razvoj odjemalskih aplikacij za iskanje v podjetju namesto API-jev poslušalnika podatkov uporabite raje API-je za iskanje in indeksiranje (SI-API). Naslednje informacije so namenjene uporabnikom, ki so predhodno izdelali aplikacije poslušalnika podatkov.

### Postopek

Za nadziranje poslušalnika podatkov:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Na strani Poslušalnik podatkov si oglejte ikone statusa, da preverite, ali je poslušalnik podatkov dejaven ali ustavljen.
3. Če se poslušalnik podatkov izvaja in želite videti podrobne informacije o stanju dejavnosti aplikacije odjemalca, kliknite  **Podrobnosti**.

Ikone statusa na strani Podrobnosti poslušalnika podatkov kažejo, ali se poslušalnik podatkov izvaja ali je ustavljen. Statistični podatki kažejo, koliko zahtev čaka na obdelavo, trenutno stanje vsake niti, ki dela na zahtevah aplikacije odjemalca, in koliko niti je dejavnih za dano stanje niti.

4. Če spremenite številko vrat za poslušalnika podatkov ali če spremenite največje število dokumentov, ki so lahko v začasni shrambi, kliknite  **Ponovni zagon**.

Poslušalnik podatkov se zažene, ko zaženete sistem iskanja v podjetju. Poslušalnika podatkov ni potrebno znova zagnati, razen če spremenite eno ali obe konfiguracijski možnosti.

---

## Sledenje dokumentov

Dokumente je mogoče odstraniti iz sistema v različnih fazah obdelave. Podate lahko možnosti, s katerimi ugotovite, kdaj je bil dokument odstranjen in kaj je bil vzrok za njegovo odstranitev.

Če razčlenjevalnik naleti na težavo, ki preprečuje razčlenjevanje dokumenta, sistem zabeleži sporočilo o odstranjenem dokumentu s kodo vzroka. (Ta tip napake ne povzroči odstranitve starejših različic dokumenta iz indeksa).

Tudi informacije o dokumentih, odstranjenih v fazi indeksiranja, se zabeležijo. URI-je in vzorce URI-jev je mogoče na primer izrecno izbrisati. Dokument je morda pajkal pajek, ki ste ga pozneje zbrisali. Izvorni dokument morda ne obstaja več (z dokumentom je povezana negativna koda HTTP) ali pa je z dokumentom povezana koda HTTP neznana. Dokument je mogoče odstraniti tudi, če za dokument, ki zahteva globalno analizo, manjkajo informacije o rangu.

Če veste, da je bil dokument pajkan, vendar se ta ne pojavi v indeksu, lahko s skrbniško konzolo za iskanje v podjetju sledite toku dokumenta skozi sistem. S pomočjo podrobnih poročil lahko izveste, kje in zakaj je bil dokument odstranjen. Poročilo lahko na primer kaže, da je bil dokument nepričakovano odstranjen med globalno analizo, ali da je skrbnik na primer odstranil URI iz indeksa.

#### **S tem povezana opravila**

“Prikaz podrobnosti o URI-jih” na strani 274

## **Konfiguriranje datotek dnevnika za sledenje dokumentov**

Če želite ugotoviti kdaj, kje in zakaj je bil dokument odstranjen iz sistema, lahko konfigurirate datoteke dnevnika, ki bodo izsledile informacije o odstranjenih dokumentih.

### **Preden začnete**

Če želite konfigurirati možnosti za sledenje odstranjenih dokumentov, morate imeti vlogo skrbnika za iskanje v podjetju ali biti skrbnik zbirke.

### **O tej nalogi**

Da bi sistem preprečil, da bi datoteke dnevnika zavzele preveč diskovnega prostora, krožno zamenjuje datoteke dnevnika in pri vsaki spremembi trenutnega datuma zažene novo datoteko dnevnika. Če ena datoteka dnevnika doseže svojo največjo dovoljeno velikost in se datum ne spremeni, sistem izdelava novo datoteko dnevnika. Ko je doseženo največje dovoljeno število datotek dnevnika, je najstarejša datoteka dnevnika zavržena, da je lahko izdelana nova.

### **Postopek**

Če želite konfigurirati datoteke dnevnika za sledenje dokumentov, naredite naslednje:

1. Uredite zbirko, izberite stran Dnevnik in kliknite **Konfiguriraj sledenje dokumentov**.
2. Prepričajte se, da je na strani Sledenje dokumentov izbrano potrditveno polje za sledenje dokumentov.
3. Podajte število datotek dnevnika, ki bodo v uporabi za beleženje informacij o dokumentih, ki so bili odstranjeni iz sistema. Te datoteke dnevnika souporabljajo vse seje, v katerih je mogoče odstraniti dokumente.

## **Prikazovanje poročil o odstranjenih dokumentih**

Prikazati je mogoče podrobne informacije o dokumentih, ki so bili izbrisani iz sistema iskanja v podjetju. Te informacije so na voljo samo, če je za zbirko omogočeno sledenje dokumentov.

### **Preden začnete**

Preden predložite zahtevo za prikaz sporočila o odstranjenih dokumentih ali pošljete poročilo na elektronski naslov, se prepričajte, da so seje, iz katerih želite sprejeti informacije, dejavne.

Če želite izvedeti več o dokumentih, ki so bili odstranjeni med razčlenjevanjem ali indeksiranjem, se prepričajte, da so seje razčlenjevalnika in indeksa za zbirko zagnane.

Preden prejmete sporočilo, preverite, ali so informacije o poštnem strežniku konfigurirane za iskanje v podjetju. Te informacije lahko podate, ko konfigurirate možnosti za elektronsko pošto na strani Dnevnik v prikazu Sistem.


### O tej nalogi

Zbiranje informacije o odstranjenih dokumentih terja precej časa. Izberete lahko možnost za pregled informacij in počakate, da jih sistem prikaže. Učinkovitejša pot je, da pošljete poročilo na elektronski naslov, ki ga podate.


Če je bil dokument odstranjen, poročilo vsebuje datum in čas odstranitve dokumenta, raven resnosti napake, komponento in sejo, v katerih je prišlo do napake, in sporočilo o napaki.

### Postopek

Če želite prikazati podrobnosti o odstranjenih dokumentih, naredite naslednje:

1. Na pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**.

**Nasvet:** Če urejate zbirko in ste že na strani Splošno, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.

2. Kliknite  **Podrobnosti o URI-ju**.
3. Na strani Podrobnosti o URI-ju vnesite URI, za katerega želite informacije.
4. Izberite potrditvena polja za tip informacij, ki jih želite prikazati:

#### Dokumenti, ki jih je zbrisal razčlenjevalnik

To potrditveno polje izberite, da preverite, ali je bil dokument odstranjen med razčlenjevanjem in, če je bil, vzrok za njegovo odstranitev.

#### Dokumenti, izbrisani iz indeksa

To potrditveno polje izberite, da preverite, ali je bil dokument odstranjen med indeksiranjem ali analiziranjem in, če je bil, vzrok za njegovo odstranitev.

5. Izberite način, na katerega želite prikazati poročilo:
  - Za prikaz poročila kliknite **Prikaz poročila** in počakajte.
  - Kliknite **Pošlji poročilo**, da pošljete poročilo na elektronski naslov za poznejši prikaz. Na strani Pošiljanje podrobnega poročila URI v polje **Elektronski naslov za obveščanje** vnesite elektronski naslov, na katerega želite prejemati poročila, nato pa kliknite **Pošlji poročilo**.

#### S tem povezana opravila

“Prikaz podrobnosti o URI-jih” na strani 274

#### S tem povezane povezave

“Formati URI v indeksu iskanja v podjetju” na strani 107

## Prikazovanje datotek dnevnika o odstranjenih dokumentih


Prikažete lahko zabeležena sporočila o dokumentih, ki so bili odstranjeni iz sistema iskanja v podjetju. Te informacije so na voljo samo, če je za zbirko omogočeno sledenje dokumentov.


### O tej nalogi


Če želite prikazati poročilo o odstranjenem dokumentu, morate poznati URI dokumenta. V datotekah dnevnika za odstranjene dokumente lahko preverite datum in čas odstranitve dokumenta, raven resnosti napake, komponento in sejo, v katerih je prišlo do težave, ter podrobno sporočilo o napaki.

### Postopek

Če želite prikazati datoteke dnevnika za odstranjene dokumente, naredite naslednje:

1. Na pogledu Zbirke poiščite zbirko, ki jo želite nadzirati, in kliknite  **Nadzor**.

**Nasvet:** Če urejate zbirko in ste že na strani Splošno, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.

2. Kliknite  **Datoteke dnevnika za odstranjene dokumente**.
3. Na strani Datoteke dnevnika za odstranjene dokumente izberite datoteko dnevnika, ki jo želite prikazati. Ime posamezne datoteke dnevnika prikazuje, ali je dokument odstranil razčlenjevalnik (pd) ali pa je bil odstranjen med gradnjo indeksa (in), ter vključuje tudi datum izdelave datoteke. Če je na isti dan izdelanih več datotek dnevnika istega tipa, številčna pripona označuje zaporedje izdelave datotek na ta dan. Primer:

```
dropped_doc_in_20060525.log
dropped_doc_pd_20060524.log (vsebuje najnovejše vnose s tega dne)
dropped_doc_pd_20060524.log.1
dropped_doc_pd_20060524.log.2 (vsebuje najstarejše vnose s tega dne)
```

4. Kliknite **Prikaži dnevnik**.

Za vsako sporočilo je na strani Vsebina datoteke dnevnika prikazan datum in čas izdaje sporočila, raven resnosti sporočila, ime seje, ki je izdala sporočilo, ter ID sporočila in besedilo napake.

S klikom gumbov se lahko v datoteki dnevnika pomaknete na prvo stran, zadnjo stran, prejšnjo stran ali naslednjo stran. Prav tako lahko podate številko strani in se pomaknete neposredno na to stran.

5. Če želite prikazati podrobnejše informacije o sporočilu, kliknite možnost  **Podrobnosti**.

Na strani Podrobnosti sporočila dnevnika so prikazani gostiteljsko ime strežnika za iskanje v podjetju, kjer je bilo izdano sporočilo, ime datoteke, ki je izdelala napako, ime funkcije in številka vrstice, kjer je prišlo do napake, ID procesa in ID niti.

Z gumbi se lahko pomaknete na prejšnje ali naslednje sporočilo v datoteki dnevnika.



---

## Datoteke dnevnika in opozorila

Izberete lahko tipe sporočil, ki jih želite zabeležiti za zbirko in za sistem, podate možnosti za izdelovanje in prikaz datotek dnevnika, prejemanje opozoril in elektronske pošte o sporočilih.

Med običajnim delovanjem bodo komponente iskanja v podjetju sporočila dnevnika zapisale v skupno datoteko dnevnika. Ta datoteka dnevnika je v imeniku `ES_NODE_ROOT/logs` na indeksnem strežniku. Za prikaz teh podatkov skupnega dnevnika lahko uporabite skrbniško konzolo.

Če pride do težave, kot je na primer odpoved omrežnih komunikacij, bodo komponente sporočila dnevnika zapisale v imenik `logs` na strežniku, kjer je komponenta nameščena. Za prikaz teh lokalnih datotek dnevnikov uporabite prikazovalnik datotek v tem računalniku, kot je pomožni program `tail` v sistemu UNIX. Za prikaz teh tipov datotek dnevnikov ne morete uporabiti skrbniške konzole.

Ko konfigurirate datoteke dnevnika, lahko izberete tipe sporočil, ki jih želite zabeležiti (kot so sporočila o napaki in opozorilna sporočila), podate, kako pogosto naj se stare datoteke dnevnika zavržejo, da se napravi prostor za nove datoteke dnevnika, podate največjo velikost datotek dnevnika in izberete jezik sporočil. Prav tako lahko podate možnosti za prejemanje elektronske pošte, ko se pojavijo določeni dogodki ali ko se zabeležijo določena sporočila ali tipi sporočil.

Ko nadzirate datoteke dnevnika, lahko izberete, katere datoteke dnevnika želite odpreti. Vsebino datoteke dnevnika lahko filtrirate, tako da vidite samo sporočila specifične ravni resnosti (kot na primer samo sporočila o napakah) ali sporočila, ki jih je izvedla določena seja iskanja v podjetju. Ko si ogledujete datoteko dnevnika, si lahko ogledate tudi podrobnosti o posameznih sporočilih. Morda boste želeli videti ime funkcije, ki je izdala sporočilo, in ostale informacije, ki vam lahko pomagajo, da, če je potrebno, ustrezno ukrepate.

### S tem povezani pojmi

 [Messages for enterprise search](#)

---

## Opozorila

Iskanje v podjetju lahko konfigurirate tako, da zapiše sporočila v datoteko dnevnika, kadarkoli zazna, da so se zgodili določeni dogodki.

Sporočila, ki jih sprožijo dogodki, imenovana opozorila, vas obveščajo o težavah, ki jih morda želite odpraviti, kot na primer vir, ki mu primanjkuje razpoložljivega prostora. Ko konfigurirate opozorila za iskanje v podjetju, podate težave, ki jih naj nadzoruje sistem. Ob vsaki pojavitvi težave sistem samodejno zapiše sporočilo v datoteko dnevnika.

Če želite biti o težavi neposredno obveščeni, lahko podate možnosti za prejemanje elektronskih sporočil, kadarkoli se zabeleži katero od nadzorovanih sporočil.

Konfigurirate lahko opozorila za dogodke na ravni zbirke in za dogodke, ki se pojavijo na ravni sistema. Na ravni zbirke lahko sistem:

- Nadzoruje število dokumentov, ki jih vsak pajek pajka, in izda opozorilno sporočilo, ko bo največje dovoljeno število dokumentov kmalu doseženo.
- Nadzoruje število dokumentov, dodanih v indeks za vašo zbirko, in izda opozorilno sporočilo, ko bo največje dovoljeno število dokumentov kmalu doseženo.

- Vas obvesti, ko čas, ki je potreben za odziv na zahteve iskanja, presega podano mejo.

Na ravni sistema lahko sistem nadzoruje prostor na disku na vsakem strežniku iskanja v podjetju in izda opozorilno sporočilo, ko je raven razpoložljivega prostora nizka.

## Konfiguriranje opozoril na ravni zbirke

S konfiguriranjem opozoril lahko zagotovite, da se sporočila zapišejo v datoteko dnevnika vsakič, ko se zgodijo določeni dogodki na ravni zbirke. Prav tako lahko prejmete elektronsko pošto vsakič, ko se zabeležijo sporočila o teh dogodkih.

### Predn začnete

Opozorila za zbirko lahko konfigurirate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.

### Postopek

Konfiguriranje opozoril na ravni zbirke:

1. Uredite zbirko, izberite stran Dnevnik in kliknite **Konfiguriranje opozoril**.
2. Če želite, da sistem nadzoruje število dokumentov, ki jih posamezni pajek pajka, storite naslednje:
  - a. Izberite potrditveno polje **Ko število dokumentov, pajkanih s katerimkoli pajkom, doseže delež največjega dovoljenega števila**
  - b. V polju **Odstotek** podajte, ko želite, da se sporočilo zabeleži. To število podajte kot odstotek največjega števila dokumentov, ki jih pajek lahko pajka (**Največje število dokumentov za pajkanje** podate, ko konfigurirate lastnosti pajka). Privzeta vrednost je 90 odstotkov.

Ker lahko za različne pajke konfigurirate različne omejitve, bo sistem za vsakega pajka zabeležil ločena sporočila. Če na primer uporabite privzeti prag za obvestila, pajku DB2 omogočite pajkanje 2.000.000 dokumentov, pajku Notes pa omogočite pajkanje 1.000.000 dokumentov; sistem bo eno sporočilo zabeležil, ko pajek DB2 pajka 1.800.000 dokumentov, drugo sporočilo pa, ko pajek Notes pajka 900.000 dokumentov.
3. Če želite, da sistem nadzoruje število dokumentov, ki se dodajajo v indeks, storite naslednje:
  - a. Izberite potrditveno polje **Ko število dokumentov v zbirki doseže odstotek ocenjene velikosti**.
  - b. V polju **Odstotek** podajte, ko želite, da se sporočilo zabeleži. To število podajte kot odstotek ocenjenega števila dokumentov, ki ga lahko zbirka vsebuje. Privzeta vrednost je 85 odstotkov.

Polje **Omejitev** kaže trenutno ocenjeno velikost zbirke. Če želite to vrednost spremeniti, za zbirko odprite stran Splošno, izberite možnost za konfiguriranje splošnih možnosti in v polju **Ocenjeno število dokumentov** podajte novo vrednost.

**Opozorilo:** Ta omejitev in ocenjeno število dokumentov, ki jih konfigurirate za zbirko, sta namenjena samo za nadzor rasti zbirke. Zato ne omejujeta absolutne velikosti indeksa.
4. Če želite, da vas sistem obvesti, ko čas, potreben za odgovor na zahteve za iskanje, presega omejitev, naredite naslednje:
  - a. Izberite potrditveno polje **Ko odzivni čas iskanja doseže omejitev**.
  - b. V polju **Omejitev** vnesite število sekund, ki se vam zdi sprejemljivo za najdaljši odzivni čas iskanja.



Ko je to število preseženo, sistem zabeleži sporočilo o dogodku. Če na primer ohranite privzeto vrednost, bo sistem sporočilo v dnevniku izdelal vsakič, ko povprečni čas iskalnega strežnika za odziv na zahteve za iskanje doseže pet sekund ali več.

Običajni odzivni časi so manj kot pol sekunde. Povprečja, večja kot ena sekunda, lahko kažejo na to, da morate prilagoditi operacijski sistem in tako povečati zmogljivost ali pa je težava v nastavitvah konfiguracije iskalnega strežnika. Povečate lahko na primer količino prostora, ki ga dodelite za iskalni predpomnilnik.

#### 5. Kliknite **Potrdi**.

Če želite prejeti elektronsko pošto, ko sistem zabeleži sporočila o teh dogodkih, odprite stran **Dnevnik** in nato kliknite **Konfiguriranje možnosti elektronske pošte za sporočila**, tako da lahko podate elektronski naslov. ID-ji sporočil za opozorila, ki ste jih omogočili, se samodejno dodajo seznamu ID-jev sporočil, za katere bo poslana elektronska pošta.

Preden lahko prejmete elektronsko pošto, morate prav tako zagotoviti, da so informacije o vašem poštnem strežniku konfigurirane. Za to mora skrbnik iskanja v podjetju izbrati v orodni vrstici **Sistem**, odpreti stran **Dnevnik** in klikniti **Konfiguriranje možnosti elektronske pošte za sporočila**.

#### **S tem povezana opravila**

“Sprejemanje elektronske pošte o zabeleženih sporočilih” na strani 298

## Konfiguriranje opozoril na ravni sistema


S konfiguriranjem opozoril lahko zagotovite, da se sporočila zapišejo v datoteko dnevnika vsakič, ko se zgodijo določeni dogodki na ravni sistema. Prav tako lahko prejmete elektronsko pošto vsakič, ko se zabeležijo sporočila o teh dogodkih.

#### **Preden začnete**

Za konfiguriranje opozoril na ravni sistema morate biti skrbnik iskanja v podjetju.

#### **Postopek**

Konfiguriranje opozoril na ravni sistema:

1. Kliknite **Sistem**, da odprete pogled **Sistem**.
2. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
3. Na strani **Dnevnik** kliknite **Konfiguriranje opozoril**.
4. Če želite, da sistem nadzira velikost razpoložljivega prostora, ki je na voljo na strežnikih iskanja v podjetju, izberite potrditveno polje **Ko količina razpoložljivega prostora datotečnega sistema doseže delež celotnega prostora**.
5. V polju **Odstotek** podajte, kdaj želite, da vas sistem obvesti o tem, da je količina razpoložljivega prostora na strežniku majhna. To število podajte kot odstotek celotnega prostora v datotečnem sistemu. Privzeta vrednost je 80 odstotkov.  
Če je vaš sistem iskanja v podjetju nastavljen na več strežnikih, bo sistem za vsak strežnik izdelal ločeno sporočilo dnevnika. Sporočilo vas na primer obvesti, ko je na strežniku pajkanja na voljo malo prostora; ločena sporočila pa vas obvestijo glede omejitev prostora na indeksnem strežniku in iskalnih strežnikih.
6. Kliknite **Potrdi**.

Če želite prejeti elektronsko pošto, ko sistem zabeleži sporočilo o tem dogodku, odprite stran **Dnevnik** in nato kliknite **Konfiguriranje možnosti elektronske pošte za sporočila**, tako da lahko podate elektronski naslov in informacije o vašem poštnem strežniku.

---

## Konfiguriranje datotek dnevnika

Podate lahko vrste sporočil, ki jih želite beležiti, in možnosti za izdelavo datotek dnevnika.

### Preden začnete

Datoteke dnevnika na ravni zbirke lahko konfigurirate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke. Za konfiguriranje datotek dnevnika na ravni sistema morate biti skrbnik iskanja v podjetju.



### O tej nalogi

Da bi sistem preprečil, da bi datoteke dnevnika zavzele preveč diskovnega prostora, krožno zamenjuje datoteke dnevnika in pri vsaki spremembi trenutnega datuma zažene novo datoteko dnevnika. Če ena datoteka dnevnika doseže svojo največjo dovoljeno velikost in se datum ne spremeni, sistem izdelava novo datoteko dnevnika. Ko je doseženo največje dovoljeno število datotek dnevnika, je najstarejša datoteka dnevnika zavržena, da je lahko izdelana nova.

Za prejemanje elektronske pošte o zabeleženih datotekah najprej podate informacije o načinu pošiljanja elektronske pošte. Nato podate, za katera sporočila želite prejemati elektronsko pošto.

### Postopek

Konfiguriranje datotek dnevnika iskanja v podjetju:

1. Če želite konfigurirati možnosti za izdelovanje in menjavo datotek dnevnika na ravni sistema:
  - a. Kliknite **Sistem**, da odprete pogled Sistem.
  - b. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
  - c. Na strani Dnevnik kliknite **Konfiguriranje možnosti datoteke dnevnika**. Prikaže se stran Možnosti datoteke dnevnika na ravni sistema.
2. Če želite konfigurirati možnosti za izdelovanje in menjavo datotek dnevnika na ravni zbirke:
  - a. V pogledu Zbirke poiščite zbirko, za katero želite podati možnosti, in kliknite  **Uredi**.
  - b. Na strani Dnevnik kliknite **Konfiguriranje možnosti datoteke dnevnika**. Prikaže se stran Možnosti datoteke dnevnika na ravni zbirke.
3. V polju **Tip informacij za beleženje** izberite tipe sporočil, ki jih želite beležiti:

#### Samo sporočila o napakah

Sporočila o napakah kažejo, da je prišlo ne nezaželenega stanja ali nepričakovanega vedenja in da se proces ne more nadaljevati. Izvesti morate dejanje, da težavo odpravite.

#### Sporočila o napakah in opozorilna sporočila

Opozorilna sporočila lahko kažejo na možno navzkrižje ali neskladnost, vendar ne povzročijo zaustavitve procesa. Ta možnost je privzeta.

#### Vsa sporočila

Informativna sporočila podajajo splošne informacije o sistemu in trenutni nalogi in ne zahtevajo nobenega ukrepa.

**Pomembno:** Izbira te možnosti lahko negativno vpliva na zmogljivost sistema. Vsa sporočila beležite samo, če morate odpraviti težave ali če vam tako naroči osebje IBM-ove službe za podporo.

4. V polju **Največja velikost posamezne datoteke dnevnika** vnesite največje število megabajtov za posamezno datoteko dnevnika. Privzeta vrednost je 10 MB.  
Ko datoteka dnevnika doseže to velikost, je izdelana nova datoteka. Izdelovanje datotek poteka do največjega števila, ki ga dovolite. Če so datoteke dnevnika manjše, je njihovo pregledovanje učinkovitejše.
5. V polje **Največje število datotek dnevnika** vnesite največje število datotek dnevnika, ki jih želite izdelati. Privzeta vrednost je 16.  
Če želite, da bodo starejša sporočila dnevnika na voljo za pregledovanje, povečajte to vrednost. Če vas bolj zanimajo novejša sporočila in vzdrževanje daljše zgodovine dejavnosti ni potrebno, to vrednost zmanjšajte.
6. V polju **Privzete področne nastavitve** izberite jezik, ki ga želite uporabiti za beleženje sporočil. Privzeta vrednost je angleščina.
7. Kliknite **Potrdi**.
8. Če želite uveljaviti spremembe, vnesite naslednja ukaza, s katerima zaustavite sistem iskanja v podjetju in ga zaženete znova.

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

## Konfiguriranje informacij strežnika SMTP

Preden lahko prejmete elektronsko pošto o dejavnostih iskanja v podjetju, morate konfigurirati informacije o strežniku Simple Mail Transfer Protocol (SMTP).

### Preden začnete

Informacije o strežniku SMTP lahko konfigurirajo samo člani skrbniške vloge za iskanje v podjetju.


### O tej nalogi

Prejemanje elektronske pošte vam omogočajo mnoge skrbniške funkcije iskanja v podjetju. Preden lahko prejmete elektronsko pošto od teh funkcij, morate podati informacije o vašem strežniku SMTP:

- Če konfigurirate opozorila na ravni zbirke ali na ravni sistema, lahko elektronsko pošto prejmete, kadarkoli se ta sporočila zabeležijo. Prav tako lahko prejmete elektronsko pošto, ko se zabeležijo ostala sporočila, ne samo sporočila, ki jih sprožijo nadzorovani dogodki.
- Če želite videti podrobne informacije o URI-ju v indeksu ali dokumentu, ki je bil izbrisan iz sistema iskanja v podjetju, lahko poročilo prejmete po elektronski pošti.
- Če nadzirate spletnega pajka in podate, da želite izdelati poročila o zgodovini spletnega pajka, lahko po izdelavi poročila prejmete elektronsko pošto.

### Postopek

Konfiguriranje informacij o vašem spletnem strežniku SMTP:

1. Kliknite **Sistem**, da odprete pogled Sistem.
2. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
3. Na strani Dnevnik kliknite **Konfiguriranje možnosti elektronske pošte za sporočila**.

4. Na strani Možnosti elektronske pošte za sistemska sporočila v polju **Poštni strežnik SMTP za uporabo pri pošiljanju elektronske pošte** vnesite popolnoma kvalificirano gostiteljsko ime ali naslov IP strežnika SMTP, ki ga želite uporabiti.  
Sistem ta strežnik uporablja za pošiljanje elektronske pošte na podane naslove.
5. V polju **Pogostost preverjanja elektronske pošte** podajte, kako pogosto naj sistem preveri, ali obstajajo nova ustrezna sporočila, in nato zanje pošlje elektronsko pošto.  
Sistem bo vsa sporočila za določeni naslov elektronske pošte združil v eno sporočilo in ga nato poslal ob podanih intervalih.
6. Kliknite **Potrdi**.

## Sprejemanje elektronske pošte o zabeleženih sporočilih

Podate lahko možnosti za prejemanje elektronskih sporočil, kadarkoli se zabeležijo določena sporočila ali določeni tipi sporočil.

### Preden začnete

Možnosti elektronske pošte za sporočila na ravni sistema lahko podajajo samo člani skrbniške vloge za iskanje v podjetju. Možnosti elektronske pošte za sporočila na ravni zbirke lahko konfigurirate samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če ste skrbnik zbirke.


Preden lahko prejimate elektronsko pošto, morate konfigurirati informacije o strežniku Simple Mail Transfer Protocol (SMTP) tako, da je elektronsko pošto mogoče dostaviti.

### O tej nalogi

Ko konfigurirate opozorila, lahko izberete možnost za beleženje sporočil, ko se zgodijo določeni dogodki. Če te možnosti omogočite, lahko nato konfigurirate možnostmi za prejemanje elektronske pošte samodejno, ko sistem zabeleži ta sporočila. Prav tako lahko podate možnosti za prejemanje elektronske pošte, ko se zabeležijo ostala sporočila, ne samo sporočila, ki jih sprožijo dogodki.


### Postopek

Konfiguriranje možnosti elektronske pošte za sporočila:

1. Če želite prejemati elektronsko pošto o sistemskih sporočilih:
  - a. Kliknite **Sistem**, da odprete pogled Sistem.
  - b. Kliknite  **Uredi**, da preklopite v pogled za urejanje sistema.
  - c. Na strani Dnevnik kliknite **Konfiguriranje možnosti elektronske pošte za sporočila**.
  - d. Na strani Možnosti elektronske pošte za sistemska sporočila izberite potrditveno polje **Pošlji elektronsko pošto o sporočilih na ravni sistema**.
  - e. V polju **Elektronski naslov za prejemanje elektronske pošte** vnesite enega ali več elektronskih naslovov. Običajno informacije o sistemskih sporočilih prejme skrbnik sistema iskanja v podjetju.  
Vsak naslov ločite z vejico. Primer:  
steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com.
  - f. Če želite prejeti elektronsko pošto o vseh sporočilih o napaki, ki so zabeležena, izberite potrditveno polje **Pošlji elektronsko pošto o vseh sporočilih o napaki**.
  - g. Če želite elektronsko pošto prejeti le, ko sistem zabeleži določena sporočila na ravni sistema, vnesite ID-je sporočil za ta sporočila v polje **Pošlji elektronsko pošto o določenih sporočilih**. Vnesite po en ID sporočila na vrstico. Primer:

FFQC4819E  
FFQ00005E

Po privzetku je navedenih mnogo ID-jev sporočil (za opis teh sporočil kliknite **Pomoč**).

- h. Kliknite **Potrdi**.
2. Če želite prejemati elektronsko pošto o sporočilih za zbirke:
- a. Kliknite **Zbirke**, da odprete pogled Zbirke.
  - b. Na seznamu zbirk poiščite zbirko, ki jo želite konfigurirati in kliknite  **Uredi**.
  - c. Na strani Dnevnik kliknite **Konfiguriranje možnosti elektronske pošte za sporočila**.
  - d. Na strani Možnosti elektronske pošte za sporočila zbirke izberite potrditveno polje **Pošlji elektronsko pošto o sporočilih na ravni zbirke**.
  - e. V polje **Elektronski naslov za prejetje elektronske pošte** vnesite enega ali več elektronskih naslovov. Običajno dobi informacije o sporočilih na ravni zbirke skrbnik zbirke.

Vsak naslov ločite z vejico. Primer:

steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com.

- f. Če želite prejeti elektronsko pošto o vseh sporočilih o napaki, ki so zabeležena, izberite potrditveno polje **Pošlji elektronsko pošto o vseh sporočilih o napaki**.
- g. Če želite elektronsko pošto prejeti le, ko sistem zabeleži določena sporočila na ravni zbirke, vnesite ID-je sporočil za ta sporočila v polje **Pošlji elektronsko pošto o določenih sporočilih**. Vnesite po en ID sporočila na vrstico. Primer:

FFQC4819E  
FFQ00005E

Po privzetku je navedenih mnogo ID-jev sporočil (za opis teh sporočil kliknite **Pomoč**).

- h. Kliknite **Potrdi**.

#### S tem povezani pojmi

 Messages for enterprise search

#### S tem povezana opravila

“Konfiguriranje opozoril na ravni zbirke” na strani 294

“Konfiguriranje opozoril na ravni sistema” na strani 295

## Spreminjanje velikosti dnevnika poizvedb

Velikost datotek dnevnika, ki ste jih izdelali za obdelavo poizvedb, lahko povečate ali zmanjšate tako, da uredite konfiguracijsko datoteko. Skrbniška konzola za iskanje v podjetju ne nudi podpore za to nalogo.

### O tej nalogi

Med obdelavo poizvedb se podatki dnevnika zapisujejo v datoteko *collection\_ID\_OmniFindQueryLog\_date.log*, pri čemer *collection\_ID* označuje zbirko, ki jo želite konfigurirati, *date* pa datum izdelave datoteke dnevnika. Povečajte ali zmanjšajte velikost datoteke dnevnika, odvisno od tega, koliko podatkov želite zabeležiti pred izdelavo nove datoteke dnevnika.

### Postopek

Če želite spremeniti velikost datoteke dnevnika obdelave poizvedb, naredite naslednje:

1. Prijavite se kot skrbnik iskanja v podjetju. V konfiguraciji sistema iskanja v podjetju z več strežniki, se prijavite v indeksni strežnik.
2. Odprite datoteko ES\_ROOT\_NODE/master\_config/collection\_ID.runtime.node1/runtime-generic.properties.
3. Poiščite lastnost **MaxFileSize**. Povečajte ali zmanjšajte vrednost, da povečate ali zmanjšate velikost datotek dnevnika, nato pa shranite spremembe.
4. Za sistem iskanja v podjetju z enim strežnikom:
  - a. S konzolo za skrbništvo nad iskanjem v podjetju lahko nadzirate zbirko, ki ste jo spremenili, in zaustavite iskalne strežnike.
  - b. Znova zaženite aplikacijo ESSearchServer na strežniku WebSphere Application Server.
  - c. V skrbniški konzoli znova zaženite iskalne strežnike, ki ste jih zaustavili.
  - d. Odprite iskalno aplikacijo v novem oknu brskalnika.
5. Za sistem iskanja v podjetju z več strežniki:
  - a. Se prijavite na indeksni strežnik kot skrbnik iskanja v podjetju.
  - b. Vnesite naslednje ukaze, da znova zaženete sistem iskanja v podjetju:

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

---


## Prikaz datotek dnevnika


Datoteke dnevnika, ki jih komponente sistema in zbirk zapišejo v skupno datoteko dnevnika, si lahko ogledate. Prav tako lahko podate filtre, s katerimi prikažete sporočila z določeno ravnijo resnosti in sporočila iz določenih sej sistema iskanja v podjetju.

### Preden začnete

Vsi skrbniški uporabniki iskanja v podjetju si lahko ogledujejo datoteke dnevnika za zbirke, za katere imajo skrbniška pooblastila. Datoteke dnevnika na ravni sistema si lahko ogledujete samo, če ste član s skrbniško vlogo v iskanju v podjetju ali če imate dovoljenje za dostop do orodne vrstice **System**.

### Postopek

1. Za prikaz datotek dnevnika za posamezno zbirko:
  - a. Kliknite **Zbirke**, da odprete pogled Zbirke.
  - b. Na seznamu zbirk poiščite zbirko, ki si jo želite ogledati, in kliknite  **Nadzor** ter odprite stran Dnevnik.

**Nasvet:** Če urejate zbirko in ste že na strani Dnevnik, lahko kliknete  **Nadzor** ter spremenite pogled za nadziranje zbirke.

2. Za prikaz datotek dnevnika na ravni sistema:
  - a. Kliknite **System**, da odprete pogled Sistem.
  - b. Izberite stran Dnevnik.
3. V polju **Datoteka dnevnika** izberite datoteko dnevnika, ki si jo želite ogledati. Ime vsake datoteke dnevnika vključuje tip datoteke dnevnika (kot je ime sistemsko ime ali ime zbirke) in datum izdelave datoteke. Če je na nek dan izdelana več kot ena datoteka dnevnika, potem vrstni red izdelave prikazuje številčna pripona. Primer:

*log\_file\_type\_20060526.log* (vsebuje najnovejše vnose na ta dan)  
*log\_file\_type\_20060526.log.1*  
*log\_file\_type\_20060526.log.2* (vsebuje najstarejše vnose na ta dan)  
*log\_file\_type\_20060525.log* (vsebuje najnovejše vnose na ta dan)  
*log\_file\_type\_20060525.log.1*  
*log\_file\_type\_20060525.log.2*  
*log\_file\_type\_20060525.log.3* (vsebuje najstarejše vnose na ta dan)

4. Da prikazete samo sporočila določenih ravni resnosti, izberite ustrezna potrditvena polja v polju **Resnost**.
5. Da prikazete samo sporočila iz določenih sej, izberite ustrezna potrditvena polja v polju **Seja**.
6. Kliknite **Prikaži dnevnik**.

Za vsako sporočilo je na strani Vsebina datoteke dnevnika prikazan datum in čas izdaje sporočila, raven resnosti sporočila, ime seje, ki je izdala sporočilo, ter ID sporočila in besedilo napake.

S klikom gumbov se lahko v datoteki dnevnika pomaknete na prvo stran, zadnjo stran, prejšnjo stran ali naslednjo stran. Prav tako lahko podate številko strani in se pomaknete neposredno na to stran.

7. Če želite prikazati podrobnejše informacije o sporočilu, kliknite možnost  **Podrobnosti**.

Na strani Podrobnosti sporočila dnevnika so prikazani gostiteljsko ime strežnika za iskanje v podjetju, kjer je bilo izdano sporočilo, ime datoteke, ki je izdelala napako, ime funkcije in številka vrstice, kjer je prišlo do napake, ID procesa in ID niti.

Z gumbi se lahko pomaknete na prejšnje ali naslednje sporočilo v datoteki dnevnika.





---

# Varnostno kopiranje in obnavljanje sistema iskanja v podjetju

Skripti izdelovanja varnostnih kopij in obnavljanja omogočajo izdelavo varnostne kopije in obnovitev sistema iskanja v podjetju.

## Za kaj skripti izdelajo varnostno kopijo?

Skripti izdelajo varnostno kopijo naslednjih datotek in jih obnovijo:

- konfiguracijske datoteke iz imenika ES\_NODE\_ROOT/master\_config,
- datoteke baze podatkov za pajke, vključno z vsemi metapodatki pajka, kot so na primer tisti iz zadnjega pajkanja v virih podatkov,
- vse datoteke v imeniku ES\_NODE\_ROOT/data,
- indeksne datoteke za zbirke, ki so konfigurirane z neprivzetimi podatkovnimi imeniki.

## Struktura imenika varnostnih kopij

Skript za izdelovanje varnostnih kopij izdela pod imenikom, ki ga podate pri izvajanju skripta, naslednje podimenike. ID skrbnika za iskanje v podjetju mora imeti dovoljenje za pisanje v imenik, ki ga podate.

### **master\_config**

Vsebuje konfiguracijske datoteke iz imenika ES\_NODE\_ROOT/master\_config.

### **database**

Vsebuje datoteke baze podatkov s strežnika pajkanja.

### **data**

Vsebuje indeksne datoteke z indeksnega strežnika.

## Nasveti za uporabo

- Izdelate lahko varnostne kopije podatkov z enega računalnika in jih obnovite na drugem, vendar upoštevajte naslednje:
  - Datotek, ki so bile varnostno prekopirane iz ene različice OmniFind Enterprise Edition, ne morete obnoviti v sistemu, v katerem se izvaja druga različica OmniFind Enterprise Edition.
  - Podatke morate obnoviti v sistemu, ki vsebuje enako ali večje število strežnikov za iskanje v podjetju. Če na primer izdelate varnostno kopijo sistema iskanja v podjetju, ki se izvaja na enem strežniku, lahko obnovite podatke v sistemu, ki uporablja dva ali štiri strežnike za iskanje v podjetju. Podatkov, ki so bili varnostno prekopirani iz sistema s štirimi strežniki, ne morete obnoviti v sistemu, ki uporablja dva strežnika ali enega.
  - Obnovitev datotek, ki so bile varnostno prekopirane iz operacijskega sistema v sistem, ki uporablja drug operacijski sistem, ni mogoča. Če ste na primer namestili iskanje v podjetju v sistem AIX in ga želite izvajati v sistemu Linux, morate na strežnike Linux namestiti nov sistem iskanja v podjetju.
- Glavni indeks zgradite, preden začnete izdelovati varnostne kopije, tako da bodo varnostno prekopirani najnovejši indeksirani podatki.
- Vse nastavitve za namestitveni imenik (ES\_INSTALL\_ROOT), podatkovni imenik (ES\_NODE\_ROOT) ter za ID in geslo skrbnika iskanja v podjetju morajo biti enake v varnostno prekopiranem sistemu in sistemu, v katerem obnovite podatke.
- Za konfiguracijo z več strežniki izdelajte varnostne kopije sistema in ga obnovite z indeksnega strežnika za iskanje v podjetju. Ker so vsi podatki pajka shranjeni v bazah podatkov na strežniku pajkanja, skripti zaženejo oddaljene ukaze za izdelovanje varnostnih kopij in obnavljanje podatkov pajka.

- Za izdelovanje varnostnih kopij sistemskih datotek iskanja v podjetju v drugem imeniku morate imeti na voljo dovolj diskovnega prostora. Skripti izdelovanja varnostnih kopij in obnavljanja ne preverjajo datotek.
- Med izvajanjem skriptov za varnostno kopijo in obnovitev se zaustavijo vse seje sistema. Če se želite izogniti prikazu nepravilnih ali neskladnih sistemskih informacij, med izvajanjem skriptov ne uporabljajte skrbniške konzole za iskanje v podjetju.
- Če sistem ne uspe zaradi nepopravljive napake, morate znova namestiti OmniFind Enterprise Edition in zagnati skript obnavljanja.

---

## Izdelava varnostne kopije sistema iskanja v podjetju

Varnostno kopijo sistema iskanja v podjetju izdelate s pomočjo skripta `esbackup.sh` za AIX, Linux ali Solaris, ali skripta `esbackup.bat` za Microsoft Windows.

### Omejitve

ID skrbnika za iskanje v podjetju mora imeti dovoljenje za pisanje v imenik, ki ga podate, ko zaženete skript izdelovanja varnostnih kopij.

Med izvajanjem skriptov za varnostno kopijo in obnovitev se zaustavijo vse seje sistema. Če se želite izogniti prikazu nepravilnih ali neskladnih sistemskih informacij, med izvajanjem skriptov ne uporabljajte skrbniške konzole za iskanje v podjetju.

**Opozorilo:** Če pritisnete Ctrl+C, da prekinete skript za izdelovanje varnostnih kopij, sistem preide v neskladno stanje. Za zagon vseh storitvenih sej in vseh izvajajočih se sej, ki so bile med postopkom varnostnega kopiranja zaustavljene, morate vnesti naslednji ukaz:

```
esadmin system startall
```

### Postopek

Postopek za izdelavo varnostne kopije sistema iskanja v podjetju:

1. V indeksni strežnik se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
2. Če se plast za skupno komunikacijo (CCL) za iskanje v podjetju ne izvaja, jo zaženite:

**AIX, Linux ali Solaris**  
`startccl.sh -bg`

**Ukazni poziv Windows**  
`startccl`

#### Skrbniška orodja za storitve Windows

Če želite zagnati CCL v ozadju:

- a. Zaženite storitve Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
  - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
3. Pri konfiguraciji z več strežniki se prepričajte, da se CCL izvaja na vseh strežnikih. Po potrebi ponovite prejšnje korake.
  4. Zaženite varnostno kopiranje, pri čemer je *imenik\_varnostnih\_kopij* imenik za varnostno prekopirane podatke:

**AIX, Linux ali Solaris**  
`esbackup.sh imenik_varnostnih_kopij`

### Ukazni poziv Windows

`esbackup.bat imenik_varnostnih_kopij`

### S tem povezane povezave

“Ukazi za iskanje v podjetju, povratne kode in ID-ji sej” na strani 335

---

## Obnovitev sistema iskanja v podjetju

Ko znova namestite OmniFind Enterprise Edition, lahko uporabite skript `esrestore.sh` za AIX, Linux ali Solaris oziroma skript `esrestore.bat` za Microsoft Windows, da obnovite sistem iskanja v podjetju.

### Omejitve

Med izvajanjem skriptov za varnostno kopijo in obnovitev se zaustavijo vse seje sistema. Če se želite izogniti prikazu nepravilnih ali neskladnih sistemskih informacij, med izvajanjem skriptov ne uporabljajte skrbniške konzole za iskanje v podjetju.

Datoteke, katerih varnostno kopijo ste napravili na eni različici OmniFind Enterprise Edition, ne morete obnoviti na sistemu, ki se izvaja na drugi različici OmniFind Enterprise Edition. Poleg tega mora imeti sistem, na katerega obnavljate podatke, enako ali več strežnikov za iskanje v podjetju, kot jih je imel sistem, na katerem ste izdelali varnostno kopijo.

### Postopek

Obnovitev sistema iskanja v podjetju:

1. V indeksni strežnik se prijavite kot skrbnik iskanja v podjetju. Ta ID uporabnika je bil podan med namestitvijo OmniFind Enterprise Edition.
2. Če se plast za skupno komunikacijo (CCL) za iskanje v podjetju ne izvaja, jo zaženite:

#### AIX, Linux ali Solaris

`startccl.sh -bg`

#### Ukazni poziv Windows

`startccl`

#### Skrbniška orodja storitev operacijskega sistema Windows

Če želite zagnati CCL v ozadju:

- a. Zaženite storitve operacijskega sistema Windows: **Start** → **Programi** → **Skrbniška orodja** → **Storitve**.
  - b. Z desno miškino tipko kliknite **IBM OmniFind Enterprise Edition**, nato pa kliknite **Zaženi**.
3. Pri konfiguraciji z več strežniki se prepričajte, da se CCL izvaja na vseh strežnikih. Po potrebi ponovite prejšnje korake.
  4. Zaustavite krmilnik:  
`esadmin stop`
  5. Obnovite podatke iskanja v podjetju, pri čemer je `backup_directory` imenik, kamor ste izdelali varnostno kopijo datotek:

#### AIX, Linux ali Solaris

`esrestore.sh backup_directory`

#### Ukazni poziv Windows

`esrestore.bat backup_directory`

### S tem povezane povezave

“Ukazi za iskanje v podjetju, povratne kode in ID-ji sej” na strani 335

---

## Izvažanje in uvažanje konfiguracij zbirke

Uvozite in izvozite lahko posamezne konfiguracije zbirke. Uvozijo in izvozijo se samo konfiguracijske datoteke zbirke, ne pa tudi podatki zbirke.

Izvozite lahko zbirko iz enega sistema iskanja v podjetju, nato pa jo uvozite v drugi sistem iskanja v podjetju. Zbirke lahko izvažate in uvažate samo med sistemi, v katerih se izvaja enaka različica iskanja v podjetju. Izvozite lahko na primer zbirko iz sistema različice 8.4 in jo nato uvozite v sistem različice 8.4 ali 8.4.0.150, ne morete pa izvoziti zbirke iz različice sistema 8.3 in je uvoziti v sistem različice 8.4.

Če izvozite zbirko, jo lahko uvozite samo v sistem iskanja v podjetju, ki uporablja enak operacijski sistem. Tako na primer ne morete izvoziti zbirke iz sistema iskanja v podjetju, ki je nameščen na sistemu Linux in je uvoziti v sistem iskanja v podjetju, ki je nameščen na sistemu Windows.

### Izvažanje zbirk

Pri izvozu zbirke izdajte ukaz **esadmin export**, da izvozite vsako zbirko posebej. Sistem ne podpira izvažanja zbirk prek konzole za skrbništvo nad iskanjem v podjetju.

1. Prijavite se kot skrbnik iskanja v podjetju. V konfiguraciji z več strežniki se morate prijaviti v indeksni strežnik. (Če poskusite ukaz zagnati iz drugega strežnika za iskanje v podjetju, ne uspe).
2. Vnesite naslednje ukaze:

```
esadmin export -help
esadmin export -cid collection_id [-fname export_filename] [-verbose]
```

Kjer je:

#### **-help**

Nudi pomoč za ukaz.

#### **-cid** *collection\_id*

Podaja ID zbirke, ki jo boste izvozili.

**Nasvet:** Če želite ugotoviti ID zbirke, ki jo nameravate izvoziti, lahko uporabite konzolo za skrbništvo nad iskanjem v podjetju ali izdate ukaz **esadmin report collections**.

#### **-fname** *export\_filename*

Podaja pot do izvozne datoteke. Če datotečno ime ni absolutno, sistem predpostavi, da gre za imenik ES\_NODE\_ROOT/dump. Če izpustite to možnost, sistem v imeniku ES\_NODE\_ROOT/dump izdela datoteko, ki uporablja naslednja pravila o poimenovanju:

```
_export_yyyyMMdd_HHmssz.zip
```

pri čemer je:

*yyyymmdd*

Podaja trenutno leto, mesec in dan zagona ukaza za izvoz.

*HHmss*

Podaja uro, minute in sekunde zagona ukaza za izvoz.

*z* Časovni odmik od časa GMT ob zagonu ukaza za izvoz. Odmik za pacifiški standardni čas je na primer GMT -0800.

### **-verbose**

Prikazuje informacije, ki lahko nudijo pomoč pri odpravljanju težav z izvozom.

## **Uvažanje zbirk**

Pri uvozu zbirke izdajte ukaz **esadmin import**, da uvozite vsako zbirko posebej. Sistem ne podpira uvažanja zbirk prek konzole za skrbništvo nad iskanjem v podjetju.

1. Prijavite se kot skrbnik iskanja v podjetju. V konfiguraciji z več strežniki se morate prijaviti v indeksni strežnik. (Če poskusite ukaz zagnati iz drugega strežnika za iskanje v podjetju, ne uspe).
2. Vnesite naslednje ukaze:

```
esadmin import -help
esadmin import -fname import_filename
 [-cid new_collection_id]
 [-name new_collection_name]
 [-colDataDir new_collection_data_directory]
 [-force]
 [-verbose]
```

Kjer je:

### **-help**

Nudi pomoč za ukaz.

### **-fname *import\_filename***

Podaja pot do uvozne datoteke. Če datotečno ime ni absolutno, sistem predpostavi, da gre za imenik ES\_NODE\_ROOT/dump.

### **-cid *new\_collection\_id***

Podaja novi ID zbirke, če je treba uvoziti zbirko z drugim ID-jem.

**Nasvet:** Če želite ugotoviti ID zbirke, ki jo nameravate uvoziti, lahko uporabite konzolo za skrbništvo nad iskanjem v podjetju ali izdate ukaz **esadmin report collections**.

### **-name *new\_collection\_name***

Podaja novo ime zbirke, če je treba uvoziti zbirko z drugim imenom.

### **-colDataDir *new\_collection\_data\_directory***

Podaja podatkovni imenik zbirke. Če ga izpustite, sistem dodeli privzeti imenik.

### **-force**

Vsili uvoz zbirke, ki že obstaja v ciljnem sistemu. Sistem s pomočjo ID-ja uvožene zbirke ugotovi, katera zbirka bo zamenjana v ciljnem sistemu.

### **Pomembno:**

- Potem ko uvozite zbirko, ne morete znova uvoziti zbirke z istim ID-jem zbirke, dokler s to možnostjo ne vsilite zbirke, ki jo nameravate uvoziti.
- Zbirka, ki jo uvozite na silo, ne obdrži motorjev za analizo besedila, slovarjev in vtičnikov pajka. Te se preprišejo z informacijami, vsebovanimi v uvoženi zbirki.
- Vsi pajki, povezani z zbirko, se odstranijo in zamenjajo s pajki, ki so podani v uvožni datoteki. Ker se pajki odstranijo, se odstranijo tudi vsi metapodatki pajka in dokumenti, ki še niso bili razčlenjeni. Po uvozu zbirke morate znova pajkati vse dokumente, da znova izdelate podatke za uvožene pajke.

### **-verbose**

Prikazuje informacije, ki lahko nudijo pomoč pri odpravljanju težav z uvozom.

## Nasveti za uporabo

### Motorji za analizo besedila in slovarji

Vsi motorji za analizo besedila in slovarji (kot so slovarji sinonimov, slovarji izključitvenih besed in slovarji poviševalnih besed), ki so povezani z zbirko v sistemu za izvoz, se ne izvozijo skupaj s konfiguracijskimi podatki zbirke. Izvozijo se samo povezave na motorje za analizo besedila in slovarje.

Če ob uvozu v ciljnim sistemu obstaja motor za analizo besedila ali slovar z enakim imenom, se poveže z uvoženo zbirko. V nasprotnem primeru se prikaže opozorilo, ki sporoča, da ni bilo mogoče vzpostaviti povezav. Zaradi tega morate naložiti motorje za analizo besedil in slovarje v ciljni sistem in uporabiti enaka imena, kot ste jih uporabili v uvoženi zbirki.

Če motor za analizo besedil ali slovar s tem imenom ne obstaja v ciljnim sistemu, se povezava prekine. Če se povezava prekine, bo zbirka v ciljnim sistemu pravilno delovala, vendar zbirke ne bodo uporabljale ustreznih motorjev za analizo besedil ali slovarjev.

### Pomnilniški modeli

Pomnilniški model, konfiguriran za izvozni sistem, se ne ohrani. Pomnilniški model, konfiguriran za ciljni sistem, ostane konfiguriran. Če poskusite uvoziti zbirko v sistem, ki je konfiguriran za manjši pomnilniški model, se prikaže opozorilo. Vedenje zbirke morda ne bo takšno, kot bi pričakovali, prav tako pa lahko resno vpliva na uporabo virov v manjših sistemih. Če se želite izogniti težavam, izvozite zbirko v sistem, ki uporablja enak pomnilniški model ali večji pomnilniški model kot izvozni sistem.

**Pajki** Prepričajte se, da so predpogoji za pajka v ciljnim sistemu nameščeni in konfigurirani.

Uvoženi pajki delujejo le, če je mogoč dostop do virov podatkov, po katerih pajkajo pajki. Če zbirka na primer vključuje pajka datotečnega sistema, ki pajka lokalni datotečni sistem, pajek lahko začne s pajkanjem le, če ciljni imenik vsebuje enako imeniško strukturo.

Vtičniki pajkov se ne izvozijo. Pri uvozu zbirke se prikaže opozorilo, nato pa se uvoz nadaljuje. Potem ko uvozite zbirko, posodobite lastnosti pajka in povežite vtičnike pajka. Poskrbite za razmestitev vtičnikov pajka v ciljni sistem, s čimer zagotovite brezhibno delovanje pajka.

### S tem povezane povezave

“Ukazi za iskanje v podjetju, povratne kode in ID-ji sej” na strani 335

---

## Integracija v Lotus Notes različice 8

Razširite lahko zmožnosti iskanja odjemalskih razmestitev IBM Lotus Notes različice 8 tako, da razmestite vtičnik za iskanje v OmniFind Enterprise Edition Lotus Notes. Vtičnik uporabnikom omogoča, da preiskujejo zbirke za iskanje v podjetju iz iskalne vrstice odjemalca Lotus Notes.

Če želite integrirati funkcije za iskanje v podjetju v Lotus Notes, morate izdelati spletno mesto s posodobitvami za vtičnik. Uporabniki lahko nato namestijo vtičnik v lokalne sisteme odjemalca Lotus Notes različice 8.

---

### Izdelava spletnega mesta s posodobitvami za vtičnik za iskanje v podjetju

Če želite integrirati iskanje v podjetju v Lotus Notes različice 8, morate razmestiti spletno mesto s posodobitvami Eclipse za iskanje v podjetju na spletni strežnik v vaši organizaciji. S pomočjo spletnega mesta s posodobitvami lahko uporabniki razmestijo vtičnik za iskanje v podjetju v lokalne namestitve odjemalca Lotus Notes.

#### O tej nalogi

Datoteke, potrebne za izdelavo spletnega mesta s posodobitvami za iskanje v podjetju, se nahajajo v datoteki `com.ibm.es.notes.search.plugin_8.4.0.150.zip`.

#### Postopek

Če želite izdelati mesto s posodobitvami Eclipse za iskanje v podjetju, naredite naslednje:

1. Prekopirajte datoteko `com.ibm.es.notes.search.plugin_8.4.0.150.zip` na spletni strežnik v vaši organizaciji.
2. Odpakirajte vsebino datoteke zip v podimenik imenika v skupni rabi, ki se nahaja na namestitveni poti spletnega strežnika.

Če želite na primer razmestiti mesto s posodobitvami na spletni strežnik, ki ga uporabljate za iskanje v podjetju, odpakirajte vsebino datoteke zip v imenik `C:\Program Files\IBM\HTTP Server\htdocs\en_US\omnifind`.

---

### Nameščanje vtičnika za iskanje v podjetju v odjemalca Lotus Notes različice 8

Če želite uporabiti funkcije za iskanje v podjetju za obdelavo poizvedb, lahko dodate vtičnik za iskanje v podjetju v iskalno vrstico Lotus Notes.

#### O tej nalogi

V zgornjem desnem kotu odjemalca Lotus Notes je prikazana iskalna vrstica. Iskalna vrstica vsebuje več vtičnikov, ki jih nudi osnovna namestitev odjemalca Lotus Notes. Potem ko namestite vtičnik za iskanje v podjetju, lahko iz iskalne vrstice iščete po zbirkah za iskanje v podjetju.

#### Postopek

Če želite namestiti vtičnik za iskanje v podjetju v odjemalca Lotus Notes, naredite naslednje:

1. Odprite datoteko NOTES\_INSTALL\_ROOT\notes.ini, pri čemer je NOTES\_INSTALL\_ROOT običajno c:\lotus\notes.
2. Na konec datoteke dodajte lastnost z imenom OMNIFIND\_ENTERPRISE\_EDITION\_SEARCH\_SERVER\_URL in podajte `http://hostname:port/ESSearchApplication/search.do?q=` kot vrednost, pri čemer je *hostname* ime gostitelja iskalnega strežnika za iskanje v podjetju, *port* pa so vrata spletnega strežnika.
3. Shranite in zaprite datoteko.
4. Odprite odjemalsko aplikacijo Lotus Notes.
5. Izberite možnost **Datoteka** → **Aplikacija** → **Namesti**.
6. Izberite izbirni gumb **Poišči nove funkcije za namestitvev**.
7. Kliknite **Dodaj oddaljeno lokacijo**.
8. Vnesite ime nove lokacije, na primer `enterprise search`. V polje URL vnesite korenski URL spletnega mesta za posodobitev iskanja v podjetju. Če ste na primer odpakirali datoteko `com.ibm.es.notes.search.plugin_8.4.0.150.zip` v podimenik `omnifind` korenskega imenika IBM HTTP Server English, je URL, ki ga podate `http://hostname:port/omnifind/`.
9. Kliknite **Dokončaj**, da shranite novo oddaljeno spletno mesto. .
10. Izberite potrditveno polje, prikazano ob imenu spletnega mesta in kliknite **Naprej**.
11. Preglejte licenčno pogodbo, izberite možnost **Sprejemem pogoje licenčne pogodbe** in kliknite **Naprej**.
12. Kliknite **Dokončaj**, da namestite vtičnik. Če prejmete opozorilo, da nameščate nepodpisan vtičnik, izberite izbirni gumb **Namesti ta vtičnik** in kliknite **Potrdi**.
13. V prikazanem pozivu kliknite **Da**, da znova zaženete odjemalca Lotus Notes.

Kliknete lahko meni, prikazan ob iskalni vrstici Lotus Notes in kot iskalno možnost izberete **OmniFind Enterprise Edition**. Če vnesete poizvedbeni izraz in kliknete gumb **Išči**, se zahteva preusmeri do strežnika za iskanje v podjetju, rezultati pa upodobijo v oknu spletnega brskalnika Lotus Notes.



---

## Integracija s portalom WebSphere Portal

Zmožnosti iskanja programa IBM WebSphere Portal lahko razširite tako, da v izdelku WebSphere Portal za portalske programčke sistema iskanja v podjetju razmestite portalski programček Iskanje.

### Točke integracije

Namestitveni program OmniFind Enterprise Edition nudi namestitvene skripte za integracijo sistema iskanja v podjetju v program WebSphere Portal. Po zagonu teh skriptov lahko sistem iskanja v podjetju integrirate v program WebSphere Portal na več načinov:

#### Portalski programček Iskanje za sistem iskanja v podjetju

Program WebSphere Portal nudi uporabnikom eno točko dostopa za delo z aplikacijami, vsebino, procesi in osebami. Sestav WebSphere Portal omogoča integracijo novih aplikacij, imenovanih portalski programčki, in razmestitev le-teh brez vpliva na druge aplikacije na portalu.

Če razmestite portalski programček Iskanje za sistem iskanja v podjetju v WebSphere Portal, lahko za iskanje po zbirkah sistema iskanja v podjetju in delo z rezultati iskanja uporabite vmesnik programa WebSphere Portal. S pomočjo nastavitve konfiguracije WebSphere Portal lahko zagotovite, da ima portalski programček sistema iskanja v podjetju enak videz in način dela kot drugi portalski programčki v okolju WebSphere Portal.

#### Center za iskanje WebSphere Portal

Center za iskanje WebSphere Portal nudi osrednjo začetno točko za iskanje po vseh virih, ki so v programu WebSphere Portal na voljo za iskanje. Portalska programčka Center za iskanje in Univerzalno iskanje omogočata iskanje po vsebini WebSphere Portal in drugih zbirkah, ki so registrirane v Centru za iskanje.

Če za integracijo sistema iskanja v podjetju v program WebSphere Portal različice 5.1 uporabite namestitvene skripte, bo sistem na stran v vmesniku Center za iskanje dodal stran Iskanje v podjetju. S to stranjo lahko iščete samo po zbirkah sistema iskanja v podjetju ali pa vnesete poizvedbo, ki išče po zbirkah sistema iskanja v podjetju in drugih zbirkah, ki so na voljo v Centru za iskanje.

Če za integracijo sistema iskanja v podjetju v izdelek WebSphere Portal različice 6 uporabite namestitvene skripte, bodo funkcije sistema iskanja v podjetju integrirane kot zvezne storitve, ki jih lahko uporabite za iskanje po zbirkah sistema iskanja v podjetju in drugih zbirkah, ki so na voljo v Centru za iskanje.

#### Iskalna vrstica WebSphere Portal

Zgornji desni vogal vseh tem vmesnika WebSphere Portal vključuje iskalno vrstico. Privzeto vedenje te vrstice je preusmeritev vseh zahtev za iskanje na privzeti iskalnik Centra za iskanje. Če želite za obdelavo poizvedb uporabiti zmogljivejše funkcije sistema iskanja v podjetju, lahko to privzeto vedenje spremenite tako, da bodo vse zahteve za iskanje preusmerjene na portalski programček Iskanje v sistemu iskanja v podjetju.

#### Pajki WebSphere Portal in Web Content Management

Če želite spletna mesta WebSphere Portal in spletna mesta IBM Workplace Web Content Management vključiti v indeks sistema iskanja v podjetju, lahko uporabite skrbniško konzolo sistema iskanja v podjetju in konfigurirate pajke WebSphere Portal in Web Content Management. Nato lahko po indeksirani vsebini iščete s pomočjo portalskega programčka sistema iskanja v podjetju ali iskalne aplikacije.

Pajek WebSphere Portal lahko pajka spletna mesta WebSphere Portal različice 5.1 in WebSphere Portal različice 6. Pajek Web Content Management lahko pajka spletna mesta na strežniku portala WebSphere Portal različice 6.

#### **Dokumenti IBM Lotus Quickr**

Če želite vključiti dokumente IBM Lotus Quickr v indeks za iskanje v podjetju, lahko s pomočjo konzole za skrbništvo nad iskanjem v podjetju konfigurirate pajka seznama semen. Nato lahko s portalskim programčkom za iskanje v podjetju v portalu WebSphere Portal ali samostojno iskalno aplikacijo preiščete indeksirano vsebino.

Pajek seznama semen lahko pajka knjižnice vsebine (dokumentov) Lotus Quickr na strežniku portala WebSphere Portal različice 6.

#### **Koristi integracije**

Sistem iskanja v podjetju izboljša okolje iskanja WebSphere Portal, saj nudi podporo za iskanje po širšem naboru tipov virov podatkov. S portalskim programčkom Iskanje za sistem iskanja v podjetju lahko iščete po spletnih mestih in tudi po drugih tipih virov podatkov, ki jih sistem iskanja v podjetju podpira.

Sistem iskanja v podjetju prav tako omogoča koristi pri stopnjevanosti. Iskalnik portala je uporaben za majhna in srednja podjetja, kjer en sam strežnik zadostuje za podporo iskanja. Če želite podpirati zmogljivosti na ravni podjetja, lahko obremenitev zaradi iskanja v podjetju razporedite na več strežnikov, pri čemer dva strežnika nudita podporo za obdelavo iskanj.

---

## **Nastavitveni skripti za integracijo sistema iskanja v podjetju v program WebSphere Portal**

Če želite sistem iskanja v podjetju integrirati v izdelek IBM WebSphere Portal, lahko zaženete namestitvene skripte, ki so na voljo v namestitvenem programu OmniFind Enterprise Edition.

Datoteko JAR, ki vsebuje namestitvene skripte za vašo različico izdelka WebSphere Portal morate prekopirati s strežnika sistema iskanja v podjetju na strežnik, kjer je nameščen izdelek WebSphere Portal. Namestitveni skripti storijo naslednje:

- razmestijo datoteke EAR, s pomočjo katerih lahko v izdelku WebSphere Portal uporabite sistem iskanja v podjetju in izdelajo pajke za dodajanje vsebin WebSphere Portal in IBM Workplace Web Content Management v zbirke sistema iskanja v podjetju;
- razmestijo datoteke WAR, ki jih potrebuje portalski programček sistema iskanja v podjetju;
- v izdelku WebSphere Portal izdelajo strani in jim dodelijo datoteke portalskega programčka sistema iskanja v podjetju;
- vse potrebne datoteke JAR prekopirajo v namestitvene imenike WebSphere Portal (za datoteke JAR, ki so že v namestitvenih imenikih, bo sistem pred kopiranjem datotek JAR, namenjenih za sistem iskanja v podjetju, izdelal varnostne kopije);
- nudijo točko vdelave za WebSphere Information Integrator Content Edition, ki omogoča iskanje dokumentov upravljalnika dokumentov portala;

Ko zaženete skripte, morate za posodobitev lastnosti portalskega programčka za iskanje in podajanje informacij o iskalnem strežniku sistema iskanja v podjetju uporabiti skrbniški vmesnik programa WebSphere Portal.

## Nasveti za uporabo

- Skripti nastavijo vse točke integracije med sistemom iskanja v podjetju in programom WebSphere Portal. Ne morete na primer izbirno namestiti portalski programček in ne namestiti datotek EAR, ki podpirajo pajke WebSphere Portal in Web Content Management.
- Če ne namestite WebSphere Information Integrator Content Edition in se kasneje odločite, da želite za iskanje dokumentov upravljalnika dokumentov portala uporabiti portalski programček, morate zagnati skript, s katerim sistem iskanja v podjetju odstranite iz programa WebSphere Portal. Nato lahko znova zaženete namestitveni skript in podate namestitveno pot za WebSphere Information Integrator Content Edition.
- Skripti zaustavijo program WebSphere Portal in ga znova zaženejo. Morda boste želeli skripte zagnati po normalnem delovnem času in preveriti, ali nerazpoložljivost storitev portala vpliva na skupino uporabnikov.
- Če med izvajanjem namestitvenega skripta pride do napak, ga znova zaženite. Naloge, ki ste jih med prvim poskusom uspešno dokončali, bodo morda javile napake, vendar pa se bo namestitveni proces nadaljeval in dokončal preostale naloge.
- Ko po izvajanju namestitvenega skripta prvič dostopite do strani portalskega programčka Iskanje v podjetju, se bo stran počasi prikazovala, ker mora sistem prevesti datoteke JSP (Java Server Pages) za portalski programček.

---

## Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 5.1

Če želite v sistem iskanja v podjetju integrirati izdelek WebSphere Portal različice 5.1.0 ali novejše, uporabite skript `wp5_install`.

### O tej nalogi

Datoteke, potrebne za integracijo sistema iskanja v podjetju v izdelek WebSphere Portal so na voljo v datoteki `es.wp5.install.jar`. Ko to datoteko razpakirate, bo sistem povzel naslednje datoteke:

- `ESSearchPortlet.war`
- `ESSearchAdapterPortlet.war`
- `ESSearchAdapter.ear`
- `ESPACServer.ear`
- `esapi.jar`
- `siapi.jar`
- `es.security.jar`
- skriptne datoteke, paketne datoteke, datoteke XML in JACL, ki so potrebne za namestitev

### Postopek

Če želite sistem iskanja v podjetju integrirati v sistem WebSphere Portal različice 5.1, storite naslednje:

1. Datoteko `es.wp5.install.jar` prekopirajte s strežnika sistema iskanja v podjetju na strežnik WebSphere Portal, nato pa jo s pomočjo ukaza Java **JAR** (ali ukaza **TAR**) razpakirajte.
2. Neobvezno: če želite omogočiti podporo za integracijo v upravljalnik dokumentov WebSphere Portal (PDM), storite nekaj izmed naslednjega:
  - Zaženite namestitveni program WebSphere Information Integrator Content Edition, izberite možnost za nameščanje samo spojnika in na strežnik WebSphere Portal namestite samo spojnik PDM.

- V strežniku WebSphere Portal izdelajte imeniško strukturo WebSphere Information Integrator Content Edition in naslednje datoteke iz obstoječe namestitve WebSphere Information Integrator Content Edition prekopirajte na strežnik WebSphere Portal, kjer CE\_ROOT podaja korenski namestitveni imenik 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. V ukazni vrstici zaženite ukaz **wp5\_install.bat** (v sistemu Windows) ali ukaz **wp5\_install.sh** (v sistemih AIX, Linux ali Solaris). Naslednji zgled prikazuje parametre, ki so zaradi lažje berljivosti navedeni v ločenih vrsticah in jih morate podati v ukazu:

```
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**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Application Server.

#### **WASUser**

Ime skrbniškega uporabnika za WebSphere Application Server; potrebno je le, če ste v izdelku WebSphere Application Server omogočili globalno zaščito.

#### **WASPassword**

Geslo za skrbniškega uporabnika izdelka WebSphere Application Server, če ste ga podali.

#### **WPSDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Portal.

#### **WPSUser**

Ime skrbniškega uporabnika izdelka WebSphere Portal.

#### **WSPassword**

Geslo za podanega skrbniškega uporabnika izdelka WebSphere Portal.

#### **WPSHost**

Ime gostitelja in številka vrat strežnika WebSphere Portal.

#### **IICEDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Information Integrator Content Edition; potrebna je le, če ste prej na strežnik WebSphere Portal namestili spojnik upravljalnika dokumentov portala.

4. Ko zaženete skript (WebSphere Portal se bo zaustavil in znova zagnal), portalski programček Iskanje v podjetju posodobite tako, da bo označeval iskalni strežnik:
  - a. V izdelek WebSphere Portal se prijavite z ID-jem in geslom skrbnika portala.
  - b. V zgornjem desnem vogalu kliknite možnost **Skrbništvo**.
  - c. Na usmerjevalnem področju na levi kliknite možnost **Upravljanje portalskih programčkov** nato pa kliknite **Portalski programčki**.
  - d. Možnost **Išči po** spremenite v **Naslov vsebuje**.
  - e. V polje **Najdi** vnesite iskanje v podjetju in kliknite gumb **Najdi**.
  - f. Ko se na desni prikažejo nove ikone, kliknite ikono ključa in tako portalski programček za iskanje konfigurirajte za sistem iskanja v podjetju.

g. Na seznamu parametrov portalskega programčka spremenite naslednje parametre:

**hostname**

Podajte popolnoma kvalificirano ime gostitelja iskalnega strežnika sistema iskanja v podjetju.

**port**

Podajte številko vrat, ki jo program WebSphere Application Server uporablja v iskalnem strežniku za iskanje v podjetju. Privzeta vrednost je 80 (privzeta vrednost za komunikacije SSL je 443).

**username**

Če ste v programu WebSphere Application Server v iskalnem strežniku omogočili globalno zaščito, podajte ime uporabnika, ki je veljavno v registru uporabnikov programa WebSphere Application Server.

**password**

Če ste podali ime uporabnika WebSphere Application Server, podajte ustrezno geslo.

**protocol**

Podajte protokol, namenjen za komunikacije med programom WebSphere Portal in iskalnim strežnikom. Privzetek je HTTP. Če uporabljate SSL, podajte HTTPS.

**trustStore**

Če uporabljate SSL, podajte popolnoma kvalificirano pot (z imenom datoteke) do shrambe potrdil SSL.

**trustPassword**

Če uporabljate SSL, podajte geslo za podano datoteko trustStore.

**ssoCookieName**

Podajte ime piškotka, ki vsebuje niz žetona enotne prijave (SSO). Privzeta vrednost je `LtpaToken`.

**proxyHost**

Če je za dostop do iskalnega strežnika sistema iskanja v podjetju zahtevan strežnik proxy, podajte popolnoma kvalificirano ime gostitelja strežnika proxy.

**proxyPort**

Če ste podali strežnik proxy, podajte številko vrat strežnika proxy.

**proxyUser**

Če strežnik proxy zahteva osnovno overjanje, podajte ime uporabnika, namenjeno za prijavo vanj.

**proxyPassword**

Če ste za strežnik proxy podali ime uporabnika, podajte ustrezno geslo.

h. Kliknite **Potrdi** in s tem shranite spremembe.

## Konfiguriranje iskalne vrstice izdelka WebSphere Portal različice 5.1 za uporabo sistema iskanja v podjetju

Izdelek WebSphere Portal različice 5.1.0 ali novejši lahko konfigurirate tako, da bo v primeru, ko uporabniki v iskalni vrstici predložijo poizvedbo, namesto privzetega iskalnika programa WebSphere Portal uporabil sistem iskanja v podjetju.

### Preden začnete

Pred preusmeritvijo zahtev za iskanje na sistem iskanja v podjetju morate zagnati namestitveni skript **wp5\_install** in tako sistem iskanja v podjetju vgraditi v program



6. Primerek aplikacijskega strežnika WebSphere Portal zaustavite in ga znova zaženite.

## Odstranjevanje sistema iskanja v podjetju iz izdelka WebSphere Portal različice 5.1

Če želite sistem iskanja v podjetju odstraniti iz izdelka WebSphere Portal različice 5.1.0 ali novejše, uporabite skript `wp5_uninstall`.

### O tej nalogi

Ko sistem iskanja v podjetju odstranite iz izdelka WebSphere Portal, parametri portalskega programčka, ki ste jih za portalski programček Iskanje v podjetju podali kot del namestitvenega procesa, ne bodo shranjeni.

Ko zaženete skript, bo ta zaustavil strežnik WebSphere Portal. Ko odstranite programsko opremo sistema iskanja v podjetju, bo skript znova zagnal strežnik WebSphere Portal.

### Postopek

Če želite sistem iskanja v podjetju odstraniti iz izdelka WebSphere Portal različice 5.1, storite naslednje:

V ukazni vrstici zaženite ukaz **`wp5_uninstall.bat`** (v sistemu Windows) ali ukaz **`wp5_uninstall.sh`** (v sistemih AIX, Linux ali Solaris). Naslednji zglede prikazuje parametre, ki so zaradi lažje berljivosti navedeni v ločenih vrsticah in jih morate podati v ukazu:

```
wp5_uninstall.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"
```

#### WASDir

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Application Server.

#### WASUser

Ime skrbniškega uporabnika za WebSphere Application Server; potrebno je le, če ste v izdelku WebSphere Application Server omogočili globalno zaščito.

#### WASPassword

Geslo za skrbniškega uporabnika izdelka WebSphere Application Server, če ste ga podali.

#### WPSDir

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Portal.

#### WPSUser

Ime skrbniškega uporabnika izdelka WebSphere Portal.

#### WSPassword

Geslo za podanega skrbniškega uporabnika izdelka WebSphere Portal.

#### WPSHost

Ime gostitelja in številka vrat strežnika WebSphere Portal.

---

## Nastavitev sistema iskanja v podjetju v programu WebSphere Portal različice 6

Če želite v sistem iskanja v podjetju integrirati izdelek WebSphere Portal različice 6, uporabite skript `wp6_install`.

### O tej nalogi

Datoteke, potrebne za integracijo sistema iskanja v podjetju v izdelek WebSphere Portal so na voljo v datoteki `es.wp6.install.jar`. Ko to datoteko razpakirate, bo sistem povzel naslednje datoteke:

- `ESSearchPortlet.war`
- `ESPACServer.ear`
- `esapi.jar`
- `es.search.provider.jar`
- `es.security.jar`
- ikone za tipe virov iskalne aplikacije, uporabljene na strani z rezultati ponudnika iskanja
- skriptne datoteke, paketne datoteke, datoteke XML in JACL, ki so potrebne za namestitvev

### Postopek

Če želite sistem iskanja v podjetju integrirati v sistem WebSphere Portal različice 6, storite naslednje:

1. Datoteko `es.wp6.install.jar` prekopirajte s strežnika sistema iskanja v podjetju na strežnik WebSphere Portal, nato pa jo s pomočjo ukaza Java **JAR** (ali ukaza **TAR**) razpakirajte.
2. Neobvezno: če želite omogočiti podporo za integracijo v upravljalnik dokumentov WebSphere Portal (PDM), storite nekaj izmed naslednjega:
  - Zaženite namestitveni program WebSphere Information Integrator Content Edition, izberite možnost za nameščanje samo spojnika in na strežnik WebSphere Portal namestite samo spojnik PDM.
  - V strežniku WebSphere Portal izdelajte imeniško strukturo WebSphere Information Integrator Content Edition in naslednje datoteke iz obstoječe namestitve WebSphere Information Integrator Content Edition prekopirajte na strežnik WebSphere Portal, kjer `CE_ROOT` podaja korenski namestitveni imenik WebSphere Information Integrator Content Edition:
3. V ukazni vrstici zaženite ukaz **wp6\_install.bat** (v sistemu Windows) ali ukaz **wp6\_install.sh** (v sistemih AIX, Linux ali Solaris). Naslednji zgled prikazuje možnosti, ki so zaradi lažje berljivosti navedene v ločenih vrsticah. Možnosti morate podati v ukazu:

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

```
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

Popolnoma kvalificirana pot do imenika profilov WebSphere Portal. Privzeta pot v sistemih AIX je `/usr/IBM/WebSphere/AppServer/profiles/wp_profile`, v sistemih



Linux in Solaris /opt/IBM/WebSphere/AppServer/profiles/wp\_profile, v sistemih Windows pa C:\Program Files\IBM\WebSphere\profiles\wp\_profile.

#### **WASDir**

Popolnoma kvalificirana pot do korenkega imenika strežnika WebSphere Application Server. Obvezna samo v sistemih AIX, Linux in Solaris. Privzeta pot do korenkega imenika v sistemih AIX je /usr/IBM/WebSphere/AppServer, v sistemih Linux in Solaris /opt/IBM/WebSphere/AppServer, v sistemih Windows pa C:\Program Files\IBM\WebSphere\AppServer.

#### **WASUser**

Ime skrbniškega uporabnika za WebSphere Application Server; potrebno je le, če ste v izdelku WebSphere Application Server omogočili globalno zaščito.

#### **WASPassword**

Geslo za skrbniškega uporabnika izdelka WebSphere Application Server, če ste ga podali.

#### **WPSDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Portal.

#### **WPSUser**

Ime skrbniškega uporabnika izdelka WebSphere Portal.

#### **WPSPassword**

Geslo za podanega skrbniškega uporabnika izdelka WebSphere Portal.

#### **WPSHost**

Ime gostitelja in številka vrat strežnika WebSphere Portal.

#### **IICEDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Information Integrator Content Edition; potrebna je le, če ste prej na strežnik WebSphere Portal namestili spojnik upravljalnika dokumentov portala.

4. Potem ko zaženete skript ter zaustavite in znova zaženete WebSphere Portal, posodobite portalski programček za iskanje v podjetju tako, da bo označeval iskalni strežnik:
  - a. V izdelek WebSphere Portal se prijavite z ID-jem in geslom skrbnika portala.
  - b. V spodnjem levem vogalu kliknite možnost **Skrbništvo**.
  - c. Na usmerjevalnem področju na levi kliknite možnost **Upravljanje portalskih programčkov** nato pa kliknite **Portalski programčki**.
  - d. Možnost **Išči po** spremenite v **Naslov vsebuje**.
  - e. V polje **Najdi** vnesite iskanje v podjetju in kliknite gumb **Najdi**.
  - f. Ko se na desni prikažejo nove ikone, kliknite ikono ključa in tako portalski programček za iskanje konfigurirajte za sistem iskanja v podjetju.
  - g. Na seznamu parametrov portalskega programčka spremenite naslednje parametre:

##### **hostname**

Podajte popolnoma kvalificirano ime gostitelja iskalnega strežnika sistema iskanja v podjetju.

##### **port**

Podajte številko vrat, ki jo program WebSphere Application Server uporablja v iskalnem strežniku za iskanje v podjetju. Privzeta vrednost je 80 (privzeta vrednost za komunikacije SSL je 443).

##### **username**

Če ste v programu WebSphere Application Server v iskalnem strežniku omogočili globalno zaščito, podajte ime uporabnika, ki je veljavno v registru uporabnikov programa WebSphere Application Server.

**password**

če ste podali ime uporabnika WebSphere Application Server, podajte ustrezno geslo.

**protocol**

Podajte protokol, namenjen za komunikacije med programom WebSphere Portal in iskalnim strežnikom. Privzetek je HTTP. Če uporabljate SSL, podajte HTTPS.

**trustStore**

Če uporabljate SSL, podajte popolnoma kvalificirano pot (z imenom datoteke) do shrambe potrdil SSL.

**trustPassword**

Če uporabljate SSL, podajte geslo za podano datoteko trustStore.

**ssoCookieName**

Podajte ime piškotka, ki vsebuje niz žetona enotne prijave (SSO). Privzeta vrednost je `LtpaToken`.

**proxyHost**

Če je za dostop do iskalnega strežnika sistema iskanja v podjetju zahtevan strežnik proxy, podajte popolnoma kvalificirano ime gostitelja strežnika proxy.

**proxyPort**

Če ste podali strežnik proxy, podajte številko vrat strežnika proxy.

**proxyUser**

Če strežnik proxy zahteva osnovno overjanje, podajte ime uporabnika, namenjeno za prijavo vanj.

**proxyPassword**

Če ste za strežnik proxy podali ime uporabnika, podajte ustrezno geslo.

h. Kliknite **Potrdi** in s tem shranite spremembe.

## Konfiguriranje Centra za iskanje izdelka WebSphere Portal različice 6 za sistem iskanja v podjetju

Izdelek WebSphere Portal različice 6 lahko konfigurirate tako, da bo v primeru, ko uporabniki v Centru za iskanje WebSphere Portal predložijo poizvedbo, iskal po zbirkah sistema iskanja v podjetju.

### Omejitve

Če so zbirke sistema iskanja v podjetju, po katerih želite iskati, zaščitene, morajo uporabniki zagnati portalski programček Iskanje sistema iskanja v podjetju in konfigurirati profil uporabnika. Profil je šifriran in bo shranjen v zaščiteno shrambo sistema iskanja v podjetju. Profil mora obstajati, šele nato lahko uporabniki v Centru za iskanje WebSphere Portal predložijo poizvedbe za iskanje po zaščitениh zbirkah.

### O tej nalogi

Center za iskanje v izdelku WebSphere Portal različice 6 podpira možnosti zveznega iskanja v več zbirkah. Zbirke lahko vsebujejo različne tipe vsebin, kot so knjižnice dokumentov portala in vsebina portala (strani in portalski programčki). Po zagonu namestitvenih skriptov za integracijo sistema iskanja v podjetju v program WebSphere Portal lahko Center za iskanje konfigurirate tako, da bo iskal tudi po zbirkah sistema iskanja v podjetju.

### Postopek

Če želite Center za iskanje konfigurirati tako, da bo iskal po zbirkah sistema iskanja v podjetju, storite naslednje:

1. V izdelek WebSphere Portal se prijavite z ID-jem in geslom skrbnika portala.
2. V spodnjem levem vogalu kliknite možnost **Skrbništvo**.
3. V usmerjevalnem področju na levi strani kliknite **Delo z iskanjem**, nato pa kliknite **Upravljanje iskanja**.
4. Kliknite **Storitve iskanja** in nato še **Nova storitev iskanja**.
5. V polju **Izvedba storitev iskanja** izberite storitev iskanja v podjetju, nato pa v besedilno okence **Ime storitve** vnesite zeleno ime storitve.
6. Na seznamu parametrov spremenite naslednje parametre:

**hostname**

Podajte popolnoma kvalificirano ime gostitelja iskalnega strežnika sistema iskanja v podjetju.

**port**

Podajte številko vrat, ki jo program WebSphere Application Server uporablja v iskalnem strežniku za iskanje v podjetju. Privzeta vrednost je 80 (privzeta vrednost za komunikacije SSL je 443).

**username**

Če ste v programu WebSphere Application Server v iskalnem strežniku omogočili globalno zaščito, podajte ime uporabnika, ki je veljavno v registru uporabnikov programa WebSphere Application Server.

**password**

Če ste podali ime uporabnika WebSphere Application Server, podajte ustrezno geslo.

**protocol**

Podajte protokol, namenjen za komunikacije med programom WebSphere Portal in iskalnim strežnikom. Privzete je HTTP. Če uporabljate SSL, podajte HTTPS.

**trustStore**

Če uporabljate SSL, podajte popolnoma kvalificirano pot (z imenom datoteke) do shrambe potrdil SSL.

**trustPassword**

Če uporabljate SSL, podajte geslo za podano datoteko trustStore.

**ssoCookieName**

Podajte ime piškotka, ki vsebuje niz žetona enotne prijave (SSO). Privzeta vrednost je LtpaToken.

**proxyHost**

Če je za dostop do iskalnega strežnika sistema iskanja v podjetju zahtevan strežnik proxy, podajte popolnoma kvalificirano ime gostitelja strežnika proxy.

**proxyPort**

Če ste podali strežnik proxy, podajte številko vrat strežnika proxy.

**proxyUser**

Če strežnik proxy zahteva osnovno overjanje, podajte ime uporabnika, namenjeno za prijavo vanj.

**proxyPassword**

Če ste za strežnik proxy podali ime uporabnika, podajte ustrezno geslo.

7. Kliknite **Potrdi** in s tem shranite spremembe.

## Konfiguriranje iskalne vrstice izdelka WebSphere Portal različice 6 za uporabo sistema iskanja v podjetju

Izdelek WebSphere Portal različice 6 lahko konfigurirate tako, da bo v primeru, ko uporabniki v iskalni vrstici predložijo poizvedbo, namesto privzetega iskalnika programa WebSphere Portal uporabil sistem iskanja v podjetju.

### Preden začnete

Pred preusmeritvijo zahtev za iskanje na sistem iskanja v podjetju morate zagnati namestitveni skript **wp6\_install** in tako sistem iskanja v podjetju vgraditi v program WebSphere Portal. Prav tako pa morate posodobiti parametre portalskega programčka Iskanje v podjetju tako, da bodo označevali ime gostitelja, vrata in druge informacije o iskalnem strežniku sistema iskanja v podjetju.

### O tej nalogi

Zgornji desni vogal vseh tem vmesnika WebSphere Portal vključuje iskalno vrstico. Privzeto vedenje te vrstice je preusmeritev vseh zahtev za iskanje na portalski programček Center za iskanje. Če želite za obdelavo poizvedb uporabiti zmogljivejše funkcije sistema iskanja v podjetju, lahko to privzeto vedenje spremenite tako, da bodo vse zahteve za iskanje preusmerjene na portalski programček Iskanje v sistemu iskanja v podjetju.

Če preusmerite iskalno vrstico, sprememba vpliva na strani, ki uporabljajo enako temo WebSphere Portal kot portalski programček za iskanje v podjetju, zato morajo te strani klicati datoteko `banner_searchControl.jspf`. Strani, ki uporabljajo drugo temo ali ki ne kličejo datoteke `banner_searchControl.jspf`, še naprej uporabljajo privzeti portalski programček Center za iskanje.

Potem ko opravite to nalogo, boste lahko uporabili Center za iskanje le, če razveljavite spremembe (obnovite lahko na primer prvotno datoteko `banner_searchControl.jspf`).

### Postopek

Če želite v primeru, ko uporabnik predloži poizvedbo v iskalni vrstici programa WebSphere Portal, uporabiti portalski programček sistema iskanja v podjetju, storite naslednje:

1. Zaustavite primerek aplikacijskega strežnika WebSphere Portal.
2. Na strežniku WebSphere Portal se pomaknite v imenik `WPS_PROFILE_ROOT/installedApps/ime_vozlišča/wps.ear/wps.war/themes/html/ime_trenutne teme`, kjer je `ime_vozlišča` ime vozlišča strežnika WebSphere Portal, `ime_trenutne teme` pa je ime trenutno uveljavljene teme za strežnik WebSphere Portal. Ime privzete teme za strežnik WebSphere Portal je IBM.
3. Izdelajte varnostno kopijo datoteke `banner_searchControl.jspf` tako, da jo prekopirate in preimenujete (na primer v `banner_searchControl.jspf.BACKUP`).
4. Uredite datoteko `banner_searchControl.jspf` in zamenjajte vsebino z naslednjim besedilom. V atributu `action=` elementa obrazca zamenjajte `localhost:10038` z imenom gostitelja in številko vrat svojega strežnika 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>
```

```

 <portal-fmt:text key="search.theme.control.label" bundle="nls.engine"/></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. Odprite datoteko **banner.jspf** in jo shranite. Ta korak, ki posodobi spremenjeni datum datoteke in s tem zagotovi, da sistem datoteko znova prevede, ni obvezen, če namesto privzete teme programa WebSphere Portal uporabljate svojo temo.
6. Odprite datoteko **Default.jsp** in jo shranite.
7. Znova zaženite primerek aplikacijskega strežnika WebSphere Portal.

## Nastavitev portalskega programčka za iskanje v podjetju za Lotus Quickr

Portalski programček za iskanje v podjetju lahko v portalu WebSphere Portal različice 6 nastavite tako, da bo preiskoval vire Lotus Quickr.

### Preden začnete

Zaženite ukaz **wp6\_install.bat** (v sistemih Windows) ali ukaz **wp6\_install.sh** (v sistemih AIX, Linux ali Solaris) in opravite postopke za namestitev iskanja v podjetju v portal WebSphere Portal različice 6.

### Postopek

Če želite nastaviti portalske programčke za iskanje v podjetju v portalu WebSphere Portal različice 6 tako, da bodo preiskovali vire Lotus Quickr, naredite naslednje:

1. Posodobite parametre portalskega programčka v konfiguraciji WebSphere Portal:
  - a. Prijavite se v WebSphere Portal z ID-jem in geslom skrbnika Lotus Quickr.
  - b. Kliknite **Skrbništvo nad spletnim mestom**, nato pa kliknite **Zahtevnejše skrbništvo**.
  - c. Na usmerjevalnem področju na levi kliknite možnost **Upravljanje portalskih programčkov** nato pa kliknite **Portalski programčki**.
  - d. Možnost **Išči po** spremenite v **Naslov vsebuje**.
  - e. V polje **Najdi** vnesite iskanje v podjetju in kliknite gumb **Najdi**.
  - f. Ko se na desni prikažejo nove ikone, kliknite ikono ključa in tako portalski programček za iskanje konfigurirajte za sistem iskanja v podjetju.
  - g. Na seznamu parametrov portalskega programčka spremenite naslednje parametre:

#### hostname

Podajte popolnoma kvalificirano ime gostitelja iskalnega strežnika sistema iskanja v podjetju.

#### port

Podajte številko vrat, ki jo program WebSphere Application Server uporablja v iskalnem strežniku za iskanje v podjetju. Privzeta vrednost je 80 (privzeta vrednost za komunikacije SSL je 443).

#### username

Če ste v programu WebSphere Application Server v iskalnem strežniku omogočili globalno zaščito, podajte ime uporabnika, ki je veljavno v registru uporabnikov programa WebSphere Application Server.

**password**

če ste podali ime uporabnika WebSphere Application Server, podajte ustrezno geslo.

**protocol**

Podajte protokol, namenjen za komunikacije med programom WebSphere Portal in iskalnim strežnikom. Privzetek je HTTP. Če uporabljate SSL, podajte HTTPS.

**trustStore**

Če uporabljate SSL, podajte popolnoma kvalificirano pot (z imenom datoteke) do shrambe potrdil SSL.

**trustPassword**

Če uporabljate SSL, podajte geslo za podano datoteko trustStore.

**ssoCookieName**

Podajte ime piškotka, ki vsebuje niz žetona enotne prijave (SSO). Privzeta vrednost je `LtpaToken`.

**proxyHost**

Če je za dostop do iskalnega strežnika sistema iskanja v podjetju zahtevan strežnik proxy, podajte popolnoma kvalificirano ime gostitelja strežnika proxy.

**proxyPort**

Če ste podali strežnik proxy, podajte številko vrat strežnika proxy.

**proxyUser**

Če strežnik proxy zahteva osnovno overjanje, podajte ime uporabnika, namenjeno za prijavo vanj.

**proxyPassword**

Če ste za strežnik proxy podali ime uporabnika, podajte ustrezno geslo.

- h. Kliknite **Potrdi** in s tem shranite spremembe.
2. Za dostop do portalskega programčka po namestitvi naredite naslednje:
  - a. Prijavite se na strežnik Lotus Quickr.
  - b. V oknu brskalnika spremenite URL v naslednje: `http://host_name:port/lotus/myquickr/ESSearchPortlet`

## Odstranjevanje sistema iskanja v podjetju iz izdelka WebSphere Portal različice 6

Če želite sistem iskanja v podjetju odstraniti iz izdelka WebSphere Portal različice 6, uporabite skript `wp6_uninstall`.

### O tej nalogi

Ko sistem iskanja v podjetju odstranite iz izdelka WebSphere Portal, parametri portalskega programčka, ki ste jih za portalski programček Iskanje v podjetju podali kot del namestitvenega procesa, ne bodo shranjeni.

Ko zaženete skript, bo ta zaustavil strežnik WebSphere Portal. Ko odstranite programsko opremo sistema iskanja v podjetju, bo skript znova zagnal strežnik WebSphere Portal.

### Postopek

Če želite sistem iskanja v podjetju odstraniti iz izdelka WebSphere Portal različice 6, storite naslednje:

V ukazni vrstici zaženite ukaz **wp6\_uninstall.bat** (v sistemu Windows) ali ukaz **wp6\_uninstall.sh** (v sistemih AIX, Linux ali Solaris). Naslednji zglede prikazuje parametre, ki so zaradi lažje berljivosti navedeni v ločenih vrsticah in jih morate podati v ukazu:

```
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**

Popolnoma kvalificirana pot do imenika profilov WebSphere Portal.

#### **WASDir**

Popolnoma kvalificirana pot do korenkega imenika strežnika WebSphere Application Server. Obvezna samo v sistemih AIX, Linux in Solaris.

#### **WASUser**

Ime skrbniškega uporabnika za WebSphere Application Server; potrebno je le, če ste v izdelku WebSphere Application Server omogočili globalno zaščito.

#### **WASPassword**

Geslo za skrbniškega uporabnika izdelka WebSphere Application Server, če ste ga podali.

#### **WPSDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Portal.

#### **WPSUser**

Ime skrbniškega uporabnika izdelka WebSphere Portal.

#### **WSPassword**

Geslo za podanega skrbniškega uporabnika izdelka WebSphere Portal.

#### **WPSHost**

Ime gostitelja in številka vrat strežnika WebSphere Portal.

---

## **Integracija iskanja v podjetju z gručnimi sistemi WebSphere Portal**

Portalski programček za iskanje v podjetju lahko nastavite za izvajanje v gručnem sistemu WebSphere Portal različice 6.

### **Nasveti za uporabo**

- Preden zaženete namestitvene skripte za iskanje v podjetju, se prepričajte, da se izvaja upravljalnik za omrežno razmestitev izdelka WebSphere Application Server in da se izvajajo vsa vozlišča v gruči.
- Skripti nastavijo vse točke integracije med sistemom iskanja v podjetju in programom WebSphere Portal. Ne morete na primer izbirno namestiti portalskega programčka, ne da bi namestiti tudi datoteke EAR, ki podpirajo pajke WebSphere Portal in Web Content Management.
- Skripti zaustavijo in znova zaženejo vse primerke strežnika WebSphere Portal v gruči. Priporočamo, da skripte zaženete po obratovalnem času, s čimer poskrbite, da nedosegljivost storitev portala ne vpliva na uporabniško skupnost.
- Če pride med izvajanjem namestitvenih skriptov do napak, znova zaženite namestitveni skript. Naloge, ki ste jih med prvim poskusom uspešno dokončali, bodo morda javile napake, vendar pa se bo namestitveni proces nadaljeval in dokončal preostale naloge.
- Ko po izvajanju namestitvenega skripta prvič dostopite do strani portalskega programčka Iskanje v podjetju, se bo stran počasi prikazovala, ker mora sistem prevesti datoteke JSP (Java Server Pages) za portalski programček.

## Nastavitev iskanja v podjetju v gručnem sistemu WebSphere Portal

Če želite integrirati sistem iskanja v podjetju v gručni sistem WebSphere Portal različice 6, uporabite skript `wp6_cluster_install`.

### O tej nalogi

Datoteke, potrebne za integracijo sistema iskanja v podjetju v izdelek WebSphere Portal so na voljo v datoteki `es.wp6.install.jar`. Ko to datoteko razpakirate, bo sistem povzel naslednje datoteke:

- `ESSearchPortlet.war`
- `ESPACServer.ear`
- `esapi.jar`
- `es.search.provider.jar`
- `es.security.jar`
- skriptne datoteke, paketne datoteke, datoteke XML in JACL, ki so potrebne za namestitev

### Postopek

Če želite integrirati iskanje v portalu z gručnim sistemom WebSphere Portal različice 6, naredite naslednje:

1. Prekopirajte datoteko `es.wp6.install.jar` s strežnika za iskanje v podjetju v vsako vozlišče gruče, v katerem je nameščen WebSphere Portal, nato pa z ukazom Java **JAR** (ali ukazom **TAR**) odpakirajte datoteko.
2. Neobvezno: Če želite nuditi podporo integraciji s pripomočkom WebSphere Portal Document Manager (PDM), opravite enega od naslednjih korakov za vsako vozlišče v gruči:
  - Zaženite namestitveni program WebSphere Information Integrator Content Edition, izberite možnost za nameščanje samo spojnika in na strežnik WebSphere Portal namestite samo spojnik PDM.
  - V strežniku WebSphere Portal izdelajte imeniško strukturo WebSphere Information Integrator Content Edition in naslednje datoteke iz obstoječe namestitve WebSphere Information Integrator Content Edition prekopirajte na strežnik WebSphere Portal, kjer `CE_ROOT` podaja korenski namestitveni imenik WebSphere Information Integrator Content Edition:
3. V ukazni vrstici zaženite ukaz `wp6_cluster_copyFiles.bat` (v Windows) ali ukaz `wp6_cluster_copyFiles.sh` (v sistemih AIX, Linux ali Solaris) za vsako vozlišče v gruči, kjer je nameščen WebSphere Portal. Naslednji primeri prikazujejo možnosti v ločenih vrsticah, saj so tako preglednejši. Ko sami podajate možnosti, jih morate podati v okviru enega ukaza:

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

```
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. V ukazni vrstici zaženite ukaz **wp6\_cluster\_install.bat** (v Windows) ali ukaz **wp6\_cluster\_install.sh** (v AIX, Linux ali Solaris). Naslednji zglede prikazuje možnosti, ki so zaradi lažje berljivosti navedene v ločenih vrsticah. Možnosti morate podati v ukazu:

```
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**

Ime gruče, v kateri je nameščen WebSphere Portal.

#### **WSPProfileDir**

Popolnoma kvalificirana pot do imenika profilov WebSphere Portal. Privzeta pot je /usr/IBM/WebSphere/AppServer/profiles/wp\_profile v sistemih AIX, /opt/IBM/WebSphere/AppServer/profiles/wp\_profile v sistemih Linux in Solaris ter C:\\Program Files\\IBM\\WebSphere\\profiles\\wp\_profile v sistemih Windows.

#### **WASDir**

Popolnoma kvalificirana pot do korenkega imenika strežnika WebSphere Application Server. Obvezna samo v sistemih AIX, Linux in Solaris. Privzeta pot do korenkega imenika je /usr/IBM/WebSphere/AppServer v sistemih AIX in /opt/IBM/WebSphere/AppServer v sistemih Linux ali Solaris.

#### **WASUser**

Ime skrbniškega uporabnika za WebSphere Application Server; potrebno je le, če ste v izdelku WebSphere Application Server omogočili globalno zaščito.

#### **WASPassword**

Geslo za skrbniškega uporabnika izdelka WebSphere Application Server, če ste ga podali.

#### **WPSDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Portal.

#### **WPSUser**

Ime skrbniškega uporabnika izdelka WebSphere Portal.

#### **WSPPassword**

Geslo za podanega skrbniškega uporabnika izdelka WebSphere Portal.

#### **WPSHost**

Ime gostitelja in številka vrat strežnika WebSphere Portal.

#### **webServerName**

Ime definicije spletnega strežnika, ki ji pripada WebSphere Portal.

#### **webServerNodeName**

Ime vozlišča strežnika WebSphere Application Server, ki ji pripada definicija spletnega strežnika.

#### **IICEDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Information Integrator Content Edition; potrebna je le, če ste prej na strežnik WebSphere Portal namestili spojnik upravljalnika dokumentov portala.

5. Potem ko se skript konča, odprite spletni brskalnik in se s strežnika za omrežno razmestitev prijavite v skrbniško konzolo WebSphere. Naslov je ponavadi `http://hostname:9060/ibm/console`.
6. Razširite razdelek **Strežniki** in izberite možnost **Spletni strežniki**.
7. Izberite okence **Izbira**, prikazano ob spletnem strežniku, in kliknite gumb **Izdelaj vtičnik**.
8. Izberite okence **Izbira**, prikazano ob spletnem strežniku, in kliknite gumb **Razširi vtičnik**.
9. Prijavite se v skrbniško konzolo.
10. Posodobite portalski programček za iskanje v podjetju ter označite iskalni strežnik:
  - a. V izdelek WebSphere Portal se prijavite z ID-jem in geslom skrbnika portala.
  - b. V spodnjem levem vogalu kliknite možnost **Skrbništvo**.
  - c. V navigacijskem območju na levi kliknite možnost **Upravljanje portalskih programčkov** nato pa kliknite **Portalski programčki**.
  - d. Možnost **Išči po** spremenite v **Naslov vsebuje**.
  - e. V polje **Najdi** vnesite iskanje v podjetju in kliknite gumb **Najdi**.
  - f. Ko se na desni prikažejo nove ikone, kliknite ikono ključa in tako portalski programček za iskanje konfigurirajte za sistem iskanja v podjetju.
  - g. Na seznamu parametrov portalskega programčka spremenite naslednje parametre:

**hostname**

Podajte popolnoma kvalificirano ime gostitelja iskalnega strežnika sistema iskanja v podjetju.

**port**

Podajte številko vrat, ki jo program WebSphere Application Server uporablja v iskalnem strežniku za iskanje v podjetju. Privzeta vrednost je 80 (privzeta vrednost za komunikacije SSL je 443).

**username**

Če ste v programu WebSphere Application Server v iskalnem strežniku omogočili globalno zaščito, podajte ime uporabnika, ki je veljavno v registru uporabnikov programa WebSphere Application Server.

**password**

če ste podali ime uporabnika WebSphere Application Server, podajte ustrezno geslo.

**protocol**

Podajte protokol, namenjen za komunikacije med programom WebSphere Portal in iskalnim strežnikom. Privzetek je HTTP. Če uporabljate SSL, podajte HTTPS.

**trustStore**

Če uporabljate SSL, podajte popolnoma kvalificirano pot (z imenom datoteke) do shrambe potrdil SSL.

**trustPassword**

Če uporabljate SSL, podajte geslo za podano datoteko trustStore.

**ssoCookieName**

Podajte ime piškotka, ki vsebuje niz žetona enotne prijave (SSO). Privzeta vrednost je `LtpaToken`.

**proxyHost**

Če je za dostop do iskalnega strežnika sistema iskanja v podjetju zahtevan strežnik proxy, podajte popolnoma kvalificirano ime gostitelja strežnika proxy.

**proxyPort**

Če ste podali strežnik proxy, podajte številko vrat strežnika proxy.

**proxyUser**

Če strežnik proxy zahteva osnovno overjanje, podajte ime uporabnika, namenjeno za prijavo vanj.

**proxyPassword**

Če ste za strežnik proxy podali ime uporabnika, podajte ustrezno geslo.

h. Kliknite **Potrdi** in s tem shranite spremembe.

## Odstranjevanje iskanja v podjetju iz gručnega sistema WebSphere Portal

Če želite odstraniti iskanje v podjetju iz gručnega sistema WebSphere Portal različice 6, uporabite skript `wp6_cluster_uninstall`.

**O tej nalogi**

Ko sistem iskanja v podjetju odstranite iz izdelka WebSphere Portal, parametri portalskega programčka, ki ste jih za portalski programček Iskanje v podjetju podali kot del namestitvenega procesa, ne bodo shranjeni.

Ko zaženete skript, bo ta zaustavil strežnik WebSphere Portal. Ko odstranite programsko opremo sistema iskanja v podjetju, bo skript znova zagnal strežnik WebSphere Portal.

**Postopek**

Če želite odstraniti iskanje v podjetju iz gručnega sistema WebSphere Portal različice 6, naredite naslednje:

1. V ukaznem pozivu zaženite ukaz **wp6\_cluster\_uninstall.bat** (v Windows) ali ukaz **wp6\_cluster\_uninstall.sh** (v AIX, Linux ali Solaris) na enem od vozlišč v gruči. Naslednji zgled prikazuje parametre, ki so zaradi lažje berljivosti navedeni v ločenih vrsticah in jih morate podati v ukazu:

```
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**

Ime gruče, v kateri je nameščen WebSphere Portal.

**WSPProfileDir**

Popolnoma kvalificirana pot do imenika profilov WebSphere Portal. Privzeta pot je `/usr/IBM/WebSphere/AppServer/profiles/wp_profile` v sistemih AIX, `/opt/IBM/WebSphere/AppServer/profiles/wp_profile` v sistemih Linux in Solaris ter `C:\\Program Files\\IBM\\WebSphere\\profiles\\wp_profile` v sistemih Windows.

**WASDir**

Popolnoma kvalificirana pot do korenkega imenika strežnika WebSphere Application Server. Obvezna samo v sistemih AIX, Linux in Solaris. Privzeta pot do

korenskega imenika je /usr/IBM/WebSphere/AppServer v sistemih AIX in /opt/IBM/WebSphere/AppServer v sistemih Linux ali Solaris.

**WASUser**

Ime skrbniškega uporabnika za WebSphere Application Server; potrebno je le, če ste v izdelku WebSphere Application Server omogočili globalno zaščito.

**WASPassword**

Geslo za skrbniškega uporabnika izdelka WebSphere Application Server, če ste ga podali.

**WPSDir**

Popolnoma kvalificirana pot do namestitvenega imenika WebSphere Portal.

**WPSUser**

Ime skrbniškega uporabnika izdelka WebSphere Portal.

**WSPassword**

Geslo za podanega skrbniškega uporabnika izdelka WebSphere Portal.

**WPSHost**

Ime gostitelja in številka vrat strežnika WebSphere Portal.

**webServerName**

Ime definicije spletnega strežnika, ki ji pripada WebSphere Portal.

**webServerNodeName**

Ime vozlišča strežnika WebSphere Application Server, ki ji pripada definicija spletnega strežnika.

2. Potem ko se skript konča, odprite spletni brskalnik in se s strežnika za omrežno razmestitev prijavite v skrbniško konzolo WebSphere. Naslov je ponavadi <http://hostname:9060/ibm/console>.
3. Razširite razdelek **Strežniki** in izberite možnost **Spletni strežniki**.
4. Izberite okence **Izbira**, prikazano ob spletnem strežniku, in kliknite gumb **Izdelaj vtičnik**.
5. Izberite okence **Izbira**, prikazano ob spletnem strežniku, in kliknite gumb **Razširi vtičnik**.
6. Prijavite se v skrbniško konzolo.

---

## Selitev iz izdelka WebSphere Portal v iskanje v podjetju

Iskanje v podjetju nudi čarovnika za selitev, s katerim lahko preselite zbirke in na pravilih temelječe taksonomije iz izdelka IBM WebSphere Portal v iskanje v podjetju.

V iskanju v podjetju se taksonomija imenuje *drevo kategorij*. Potem ko preselite taksonomijo, s skrbniško konzolo za iskanje v podjetju uredite drevo in pravila kategorij. Po selitvi zbirke pa s skrbniško konzolo izvajate skrbništvo nad zbirko.

Če želite preseliti taksonomije in zbirke, zaženite čarovnika za selitev na indeksnem strežniku iskanja v podjetju.

---

## Selitev zbirke s portala WebSphere Portal

Če želite preseliti zbirke in na pravilih temelječe taksonomije s portala WebSphere Portal v iskanje v podjetju, pripravite zbirke v portalu WebSphere Portal, nato pa jih s čarovnikom za selitev preselite.

### Preden začnete

Če nameravate preseliti taksonomije in zbirke, preselite datoteke taksonomij, preden z istim postopkom preselite zbirke. S takšnim pristopom zagotovite, da bodo preseljena pravila za kategorizacijo delovala tudi s preseljenimi zbirkami.

### Postopek

Če želite preseliti zbirko (in neobvezno tudi taksonomijo) s portala WebSphere Portal v iskanje v podjetju, naredite naslednje:

1. V iskalniku portala WebSphere Portal zaustavite vse procese pajka v zbirkah, ki jih želite preseliti, in odobrite ali zavrnite vse dokumente v teku. (Iskanje v podjetju ne podpira koncepta dokumentov v teku.)
2. S portalskimi programčki iskalnika portala izvozite nastavitve vsake zbirke, ki jo želite preseliti, v datoteke XML.
3. Če je indeksni strežnik za iskanje v podjetju nameščen na ločenem strežniku, prekopirajte izvožene datoteke XML na indeksni strežnik.
4. V indeksni strežnik za iskanje v podjetju se prijavite kot skrbnik iskanja v podjetju. ID uporabnika je bil podan pri namestitvi izdelka OmniFind Enterprise Edition.
5. Pomaknite se do namestitvenega imenika iskanja v podjetju:

UNIX: `cd $ES_INSTALL_ROOT/bin`  
Windows: `cd %ES_INSTALL_ROOT%\bin`

6. Če želite preselite zbirki z omogočeno zaščito, z naslednjim ukazom zaženite čarovnika za selitev, nato pa kliknite **Naprej**.

UNIX: `./eswpsmigrate.sh`  
Windows: `eswpsmigrate.bat`

7. Če želite za preseljene zbirke onemogočiti zaščito na ravni zbirke, vnesite naslednji ukaz, da zaženete čarovnik za selitev, nato pa kliknite **Naprej**.

UNIX: `./eswpsmigrate.sh disable.security`  
Windows: `eswpsmigrate.bat disable.security`

8. Izberite možnost **Preseli nastavitve iskanja iz iskalnika portala v portalu WebSphere Portal**, nato pa kliknite **Naprej**.
9. Pomaknite se do imenika, v katerem so izvožene konfiguracijske datoteke iskalnika portala, izberite datoteke, ki jih želite preseliti, nato pa kliknite **Naprej**. Sistem analizira in preveri izbrane konfiguracijske datoteke.
10. Za vsako zbirko vnesite naslednje informacije, nato pa kliknite **Naprej**, da začnete seliti zbirke v iskanje v podjetju:
  - Ime zbirke, ki ga želite uporabiti v iskanju v podjetju.
  - Kriterij, po katerem sistem ugotovi pomembnost dokumentov v zbirki. Faktor statičnega rangiranja lahko ne obstaja, lahko pa temelji na datumih dokumentov ali na številu povezav do spletnih dokumentih z drugih spletnih dokumentov.
  - Tip kategorizacije, ki jo želite uporabiti za to zbirko. Če podate, da faktor ne obstaja, se informacije o taksonomiji ne preselijo v iskanje v podjetju. Če izberete na pravih temelječe kategorije, se taksonomija preseli v iskanje v podjetju skupaj z zbirko.Če pride med selitvijo do napak, si oglejte datoteko MigrationWizard.log, ki se nahaja v imeniku, v katerem je nameščen čarovnik za selitev.  
Sedaj lahko s skrbniško konzolo za iskanje v podjetju konfigurirate dodatne nastavitve za preseljene zbirke.

**Zahteve:** Ko konfigurirate lastnosti spletnega pajka za preseljeno zbirko, morate podati elektronski naslov, na katerega želite prejemati obvestila o pajku, ter ime uporabnikovega posrednika (če potrebujete pomoč, kliknite **Pomoč** med konfiguriranjem lastnosti spletnega pajka).

11. Iz skrbniške konzole za iskanje v podjetju zaženite procese pajkanja, razčlenjevanja in indeksiranja za preseljeno zbirko.
12. Potem ko določite, da je preseljena zbirka omogočena za iskanje v podjetju, izbrišite izvorno zbirko v iskalniku portala.
13. Neobvezno: Če želite uporabnikom omogočiti iskanje po preseljenih zbirkah portala v portalu WebSphere Portal, kot skrbnik portala WebSphere Portal naredite naslednje.
  - a. Skupaj z namestitvijo portala WebSphere Portal razmestite portalski programček za iskanje v podjetju.

V gruči strežnikov WebSphere Portal je treba to storiti na strežniku, na katerem je nameščen upravljavnik razmestitve WebSphere Application Server. Upravljavnik razmestitve distribuira portalski programček za iskanje na enega od drugih strežnikov v gruči strežnikov WebSphere Portal.
  - b. Dodajte portalski programček za iskanje v podjetju na ustrezne strani portala.

V portalu WebSphere Portal se nadzor dostopa do portalskega programčka za iskanje oblikuje glede na dostopnost do določenih strani in portalskih programčkov. Četudi preselite nastavitve zbirke, mora skrbnik strežnika WebSphere Portal ročno prestaviti portalski programček.

---

## Nastavitve preseljene zbirke

Ko preselite zbirke s portala IBM WebSphere Portal, čarovnik za selitev izdela privzete nastavitve za zbirke in pajke.

Če zbirke iskalnika portala in zbirke za iskanje v podjetju vsebujejo isto nastavitvev, čarovnik pri selitvi zbirke v iskanje v podjetju uporabi nastavitvev iskalnika portala. V primeru nastavitvev, ki obstajajo samo v iskanju v portalu, čarovnik uporabi nastavitve, ki jih podate pri selitvi zbirke ali privzete nastavitve za zbirke v iskanju v portalu.

## Nastavitve, ki obstajajo v iskalniku portala in iskanju v podjetju

Čarovnik za selitev za vsako preseljeno zbirko preseli naslednje nastavitve:

- Spletna mesta iskalnika portala znotraj zbirke iskalnika portala
- Jezik zbirke
- Taksonomijo (ali drevo kategorij) in pravila za na pravilih temelječe kategorije, kadar zbirka za iskanje v podjetju uporablja na pravilih temelječo kategorizacijo

Vsako spletno mesto iskalnika portala v zbirki se združi v spletnega pajka za iskanje v podjetju. Čarovnik za selitev preseli naslednje nastavitve pajka:

- Začetni URL-ji
- Število vzporednih procesov pajkanja
- Globina pajkanja
- Pretek čakalnega časa (v sekundah) za priklic dokumenta
- Privzet znakovni nabor
- Pravila za pajkanje spletnih mest (vključi ali izključi)

## Nastavitve, ki obstajajo samo v iskanju v podjetju

Ko preselite zbirko, podate informacije o njej. Čarovnik za selitev preseli nastavitve in s pomočjo privzetih nastavitvev za zbirke v iskanju v podjetju konfigurira vsako zbirko, ki jo preselite.

Konfiguracije zbirke in spletnega pajka lahko spremenite z uporabo skrbniške konzole za iskanje v podjetju. Vrednosti, prikazane v oklepajih (), so privzete nastavitve preseljenih podatkov.

- Ime zbirke
- Strategija statičnega rangiranja dokumentov
- Tip kategorizacije, ki je v uporabi (na pravilih temelječ ali nikakršen)
- Ali bo uporabljen iskalni predpomnilnik in koliko odgovorov na poizvedbe lahko vsebuje (da, 5000)
- Ali se bo izvajal nadzor odzivnega časa iskanj in ali bo v primeru prekoračitve podane omejitve izdano opozorilo (da, 5 sekund)
- Ali bo v uporabi nadzor dostopa (ne)
- Razpored za gradnjo delta indeksov
- Razpored za gradnjo glavnega indeksa
- Raven podrobnosti v dnevniku (vsa sporočila)

Čarovnik za selitev za vsakega spletnega pajka izdela tudi naslednje nastavitve:

- Ime pajka
- Opis pajka
- Največja dolžina strani
- Nastavitve zaščite dokumenta
- Tipi MIME dokumenta, ki jih je treba pajkati, če jih je mogoče uveljaviti za izvorni tip podatkov

Preden zaženete na novo preseljenega spletnega pajka, preglejte vse lastnosti pajka in nastavitve prostora za pajkanje ter se prepričajte, da so podane vse zahtevane vrednosti (obvezna polja so označena z rdečo zvezdico). Še posebej natančno preverite, ali ste podali elektronski naslov za sprejemanje obvestil o pajku in ime posrednika uporabnika pajka. Če

potrebujete pomoč pri konfiguriranju lastnosti spletnega pajka, kliknite **Pomoč**.

---

## **Datoteka dnevnika čarovnika za selitev**

Čarovnik za selitev zapiše vsa sporočila v datoteko WpsMigratorLog.log, ki se nahaja v imeniku, v katerem je nameščen tudi čarovnik za selitev.

Za vsako preseljeno zbirko datoteka dnevnika WpsMigratorLog.log vsebuje vrednosti za vse nastavitve, ki so bile prebrane iz iskalnika v portalu WebSphere, in podaja, kje in kako so bile te nastavitve uvožene v zbirke za iskanje v podjetju.



---

## Ukazi za iskanje v podjetju, povratne kode in ID-ji sej

Ukaze lahko uporabljate za ugotavljanje težav, statusa različnih delov sistema, za začetek in zaustavitev sej ali za zagon in zaustavitev sistema.

V namestitvi z več strežniki lahko ukaze zaženete s kateregakoli strežnika v vašem sistemu. Kljub temu morate ukaze zagnati s strežnika indeksov. Strežnik indeksov oziroma strežnik krmilnika lahko dostopa do informacij vseh drugih strežnikov v sistemu.

Večina ukazov ima naslednje formate:

```
esadmin ime_ukaza argumenti
esadmin session_ID dejanje -možnost
```

Za več informacij o vseh ukazih vnesite `esadmin help`. Za več informacij o specifičnem ukazu vnesite `esadmin dejanje help`.

### Ukazi za iskanje v podjetju esadmin

Naslednje ukaze vnesite v eno samo vrstico.

Tabela 9. Ukazi za iskanje v podjetju **esadmin**

Ukaz	Opis
<code>esadmin system startall</code>	<p>Zažene komponente za iskanje v podjetju na vseh strežnikih za iskanje v podjetju, vključno s spletnim strežnikom, aplikacijo ESSearchServer in informacijskim centrom na iskalnih strežnikih; seje pajka na strežnikih pajka; in seje indeksov na strežnikih indeksov. Zažene skupno komunikacijsko plast (CCL) samo na lokalnem strežniku. Za recikliranje CCL-a morate ročno zaustaviti in ponovno zagnati CCL na vsakem oddaljenem strežniku za iskanje v podjetju.</p> <p><b>Vzorec ukaza:</b> <code>esadmin system startall</code></p>
<code>esadmin system stopall</code>	<p>Zaustavi komponente za iskanje v podjetju na vseh strežnikih za iskanje v podjetju, vključno z informacijskim centrom, aplikacijo ESSearchServer in spletnim strežnikom na iskalnih strežnikih; seje pajka na strežnikih pajka; in seje indeksov na strežnikih indeksov. Zaustavi CCL samo na lokalnem strežniku. Za recikliranje CCL-a morate ročno zaustaviti in ponovno zagnati CCL na vsakem oddaljenem strežniku za iskanje v podjetju.</p> <p><b>Vzorec ukaza:</b> <code>esadmin system stopall</code></p>
<code>esadmin system checkall</code>	<p>Preveri status vseh komponent za iskanje v podjetju na vseh strežnikih za iskanje v podjetju.</p> <p><b>Vzorec ukaza:</b> <code>esadmin system checkall</code></p>

Tabela 9. Ukazi za iskanje v podjetju **esadmin** (nadaljevanje)

Ukaz	Opis
esadmin crawler_session_id start	<p>Prične sejo pajka. Ta ukaz ne prične dejavnosti pajkanja.</p> <p><b>Vzorec ukaza:</b> esadmin col1.WEB1.esadmin start</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5310I WEBCrawler1 (sid: col1.WEB1.esadmin) is not running. FFQC5314I Result: 0</p>
esadmin crawler_session_id startCrawl	<p>Prične pajkanje.</p> <p><b>Vzorec ukaza:</b> esadmin col3.DB21.esadmin startCrawl</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) is already running. PID: 23650 FFQC5314I Result: 0</p>
esadmin crawler_session_id pause	<p>Začasno ustavi pajkanje.</p> <p><b>Vzorec ukaza:</b> esadmin col3.DB21.esadmin pause</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) is already running. PID: 23650 FFQC5314I Result: 0</p>
esadmin crawler_session_id resume	<p>Nadaljuje pajkanje.</p> <p><b>Vzorec ukaza:</b> esadmin col3.DB21.esadmin resume</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) is already running. PID: 23650 FFQC5314I Result: 0</p>
esadmin crawler_session_id stopCrawl	<p>Zaustavi pajkanje.</p> <p><b>Vzorec ukaza:</b> esadmin col3.DB21.esadmin stopCrawl</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) is already running. PID: 23650 FFQC5314I Result: 0</p>
esadmin crawler_session_id stop	<p>Zaustavi sejo pajka.</p> <p><b>Vzorec ukaza:</b> esadmin col3.DB21.esadmin stop</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) is already running. PID: 23650 FFQC5314I Result: 0</p>

Tabela 9. Ukazi za iskanje v podjetju **esadmin** (nadaljevanje)

Ukaz	Opis
esadmin crawler_session_id getCrawlerStatus	<p>Pridobi status pajka. Vrnjene informacije so odvisne od tega, ali je pajek spletni pajek ali pajek za vse ostale vire podatkov.</p> <p><b>Zgled spletnega pajka:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlerStatus</pre> <p><b>Možne povratne kode in sporočila za spletnega pajka:</b></p> <pre>FFQC5303I WebCrawler1 (sid: col1.WEB1.esadmin) is already running. PID: 23650</pre> <p><b>Zgled nespletnega pajka:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlerStatus</pre> <p><b>Možne povratne kode in sporočila za nespletnega pajka:</b></p> <pre>FFQC5303I db2crawler (sid: db2col.DB2_96945) is already running. PID: 5936</pre> <p>Za več informacij o povratnih statusnih sporočilih glejte "Podrobne informacije za statusne ukaze" na strani 339.</p>
<pre>esadmin dscrawler_session_id getCrawlSpaceStatus  esadmin web_crawler_session_id getCrawlStatus -selections value</pre>	<p>Pridobi splošni status območja pajkanja za vsakega pajka, ki ni spletni pajek.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatus</pre> <p><b>Vzorci sporočil in povratnih kod:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) is already running. PID: 23650</pre> <p>Pridobi splošni status območja pajkanja za spletnega pajka.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlStatus</pre> <p>Za več informacij o povratnih statusnih sporočilih glejte "Podrobne informacije za statusne ukaze" na strani 339.</p>
<pre>esadmin dscrawler_session_id getCrawlSpaceStatusDetail -ts target_server_id  esadmin webcrawler_session_id getCrawlDetailsPerSite -url string -selections num -threshold num</pre>	<p>Pridobi podrobni status območja pajkanja za vsakega pajka, ki ni spletni pajek. Če ne podate možnosti ciljnega strežnika, bodo vrnjeni podatki za vse ciljne strežnike. Če na primer pajek DB2 pajka bazi podatkov FOUNTAIN in SAMPLE, vi pa ne podate možnosti ciljnega strežnika, bo vrnjen status vseh tabel v bazah podatkov FOUNTAIN in SAMPLE.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatusDetail -ts FOUNTAIN</pre> <p><b>Vzorci sporočil in povratnih kod:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) is already running. PID: 23650</pre> <p>Pridobi podrobni status območja pajkanja za spletnega pajka.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlDetailsPerSite</pre> <p>Za več informacij o povratnih statusnih sporočilih glejte "Podrobne informacije za statusne ukaze" na strani 339.</p>

Tabela 9. Ukazi za iskanje v podjetju **esadmin** (nadaljevanje)

Ukaz	Opis
<pre>esadmin monitor getCollectionParserMonitorStatus -cid collection_ID</pre>	<p>Pridobi status razčlenjevalnika.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin monitor getCollectionParserMonitorStatus -cid coll</pre> <p><b>Vzorci sporočil in povratnih kod:</b></p> <pre>FFQC5303I Monitor (node1) (sid: monitor) is already running. PID: 12543</pre> <p>Za več informacij o povratnih statusnih sporočilih glejte “Podrobne informacije za statusne ukaze” na strani 339.</p>
<pre>esadmin startMain -cid collection_id</pre>	<p>Prične gradnjo glavnega indeksa.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin startMain -cid coll</pre> <p><b>Vzorci sporočil in povratnih kod:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) is already running. PID: 25917 FFQC5314I Result: 1117671147056</pre>
<pre>esadmin startDelta -cid collection_id</pre>	<p>Prične gradnjo delta indeksa.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin startDelta -cid coll</pre> <p><b>Vzorci sporočil in povratnih kod:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) is already running. PID: 4548 FFQC5314I Result: 1117670603408</pre>
<pre>esadmin controller startIndexBuild -cid collection_id -buildType type -detectChanges</pre>	<p>Zaženite gradnjo glavnega ali delta indeksa in podajte, da naj se gradnja nadaljuje samo, če sistem ne odkrije sprememb, ki bi jih bilo treba uveljaviti za indeks.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin controller startIndexBuild -cid col_1 -buildType main -detectChanges</pre>
<pre>esadmin monitor getCollectionIndexMonitorStatus -cid collection_id -buildType [main   delta] -numrecords lastNrecords</pre>	<p>Pridobi status gradnje glavnega indeksa ali delta indeksa. Možnost <code>numrecords</code> prikazuje zapise statusa gradnje indeksa <i>N</i>. Če <code>numrecords</code> izpustite, bo vrnjen status za zadnjih 20 građenj indeksov.</p> <p><b>Vzorec ukaza:</b></p> <pre>esadmin monitor getCollectionIndexMonitorStatus -cid coll -buildType main -numrecords 4</pre> <p><b>Vzorci sporočil in povratnih kod:</b></p> <pre>FFQC5303I Monitor (node1) (sid: monitor) is already running. PID: 12649</pre> <p>Za več informacij o povratnih statusnih sporočilih glejte “Podrobne informacije za statusne ukaze” na strani 339.</p>

Tabela 9. Ukazi za iskanje v podjetju **esadmin** (nadaljevanje)

Ukaz	Opis
esadmin startSearch -cid <i>collection_id</i>	<p>Prične procese iskalnih strežnikov.</p> <p><b>Vzorec ukaza:</b> esadmin startSearch -cid coll</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I Controller (node1) (sid: controller) is already running. PID: 25917 FFQC5314I Result: 0</p>
esadmin stopSearch -cid <i>collection_id</i>	<p>Zaustavi procese iskalnega strežnika.</p> <p><b>Vzorec ukaza:</b> esadmin stopSearch -cid coll</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I Controller (node1) (sid: controller) is already running. PID: 15292 FFQC5314I Result: 0</p>
esadmin monitor getCollectionSearchMonitorStatus -cid <i>collection_id</i> esadmin searchmanager_session_id getStatus -cid <i>collection_id</i>	<p>Pridobi status iskalnega strežnika.</p> <p><b>Vzorec ukaza:</b> esadmin monitor getCollectionSearchMonitorStatus -cid coll</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I Monitor (node1) (sid: monitor) is already running. PID: 12649</p> <p>Vrne podrobne informacije o statusu indeksa iskanja za zbirko na danem iskalnem strežniku. Na vsak iskalni strežnik je ena seja upravljalnika iskanja. Vsaka seja upravljalnika iskanja je odgovorna za nadziranje in delovanje iskalnih indeksov na specifičnem iskalnem strežniku.</p> <p><b>Vzorec ukaza:</b> esadmin searchmanager.node1 getStatus -cid coll</p> <p><b>Vzorci sporočil in povratnih kod:</b> FFQC5303I Search Manager (node1) (sid: searchmanager.node1) is already running. PID: 15711 FFQC5314I Result: PID=18390 CacheHits=3 QueryRate=1 Port=44008 SessionId=coll.runtime.node1 CacheHitRate=0.333 ResponseTime=70 Status=1 SessionName=coll.runtime.node1.1</p> <p>Za več informacij o povratnih statusnih sporočilih glejte "Podrobne informacije za statusne ukaze".</p>

## Podrobne informacije za statusne ukaze

Nekateri ukazi lahko vrnejo obširne informacije. Ta razdelek opisuje informacije, ki zadevajo status pajka in status območja pajkanja. Tabela iz razdelka "Ukazi za iskanje v podjetju

esadmin” na strani 335 podaja informacije, ki jih lahko vrne posamezni ukaz esadmin. Ta razdelek opisuje informacije, ki jih vrnejo naslednji ukazi:

- Status spletnega pajka
- Status nespletnega pajka
- Status območja pajkanja za spletnega pajka
- Status območja pajkanja za nespletne pajke
- Podrobni status območja pajkanja za spletnega pajka
- Podrobni status območja pajkanja za nespletne pajke
- Status razčlenjevalnika
- Status gradnje indeksa
- Status iskalnega strežnika
- Podrobni status iskalnega strežnika

**Status spletnega pajka:** Ko zaženete ukaz za pridobitev statusa spletnega pajka, ukaz vrne informacije v formatu dokumenta XML. Ukaz za status spletnega pajka lahko vrne naslednje informacije:

```
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>
```

Naslednja tabela opisuje vsak element XML in njegove možne atribute, ki jih vrne ukaz za status spletnega pajka:

Tabela 10. Informacije o statusu spletnega pajka

Element	Atributi	Opis
CrawlerStatus	<ul style="list-style-type: none"> <li>• CrawlerThreadStateDist</li> <li>• ActiveBucketList</li> <li>• CrawlRate</li> <li>• RecentlyCrawledURLList</li> <li>• NumURLsThisSession</li> </ul>	Status pajka.

Tabela 10. Informacije o statusu spletnega pajka (nadaljevanje)

Element	Atributi	Opis
CrawlerRunLevel Value	<ul style="list-style-type: none"> <li>String (English) "Not started": Seja pajka obstaja, a še ni prejela sporočila za pričetek procesiranja dokumentov.</li> <li>"Started": Pajek se je zagnal.</li> <li>"Running": Pajek je končal inicializacijo in zagon ter aktivno pajka.</li> <li>"Paused": Pajek je moral začasno ustaviti aktivno pajkanje, ne sme pa ga prekiniti.</li> <li>"Stopping": Pajek je prejel znak za ustavitev in se bo ustavil.</li> <li>"Error": Pajek je v nepopravljivem stanju in ga je potrebno ustaviti in ponovno zagnati, da se pajkanje nadaljuje.</li> </ul>	Informacije o dejavnosti pajka.
CrawlerThreadState State	String (English)	Dejavnost niti pajka. To polje prikazuje, kaj dela nit (ali niti).
ActiveBucket	<ul style="list-style-type: none"> <li>URL: String (URL spec) Protokol, gostitelj in vrata, katerih URL-ji so pajkani.</li> <li>NumActURLs: Celó število (pozitivno) Število URL-jev v vsebniku, ko je ta postal razpoložljiv za pajkanje (aktiviran).</li> <li>NumProcURLs: Celó število (nenegativno) Število URL-jev iz vsebnika, ki so bili zaenkrat procesirani, ali pajkani ali zavrženi.</li> <li>TimeRem: Celó število Število preostalih sekund dočasne prekinitve vsebnika.</li> <li>Duration: Celó število (nonnegative) Število preteklih sekund od aktivacije vsebnika.</li> </ul>	Trenutna dejavnost podanega spletnega mesta.
CrawlRate	Vrednost: Celó število (nenegativno)  Število pajkanih strani na sekundo (vsi vsebniki skupaj).	Merjenje prepustnosti pajka.
RecentlyCrawledURL	URL: String (URL spec)  Niz, ki podaja protokol, gostitelja, vrata in pajkano datoteko.	Nedavno pajkana stran.
NumURLsThisSession	Vrednost: Celó število (nenegativno)	Število pajkanih URL-jev od takrat, ko je ta primerek pajka (procesa) pričel pajkati.

**Non-Web crawler status:** Ko zaženete ukaz za pridobitev statusa pajka za nespletnega pajka, ukaz vrne informacije v formatu dokumenta XML. Ukaz **getCrawlerStatus** za pajke, ki ne pajkajo po spletu, lahko vrne naslednje informacije:

```
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>

```

Naslednje tabele opisujejo elemente in attribute XML za vsakega pajka za iskanje v podjetju, razen za spletnega pajka. Te informacije vrne ukaz za status pajka.

Tabela 11. Statusne informacije za pajke NNTP, DB2, baza podatkov JDBC in Notes

Ime elementa in atributa	Pajek NNTP	Pajki DB2 in baze podatkov JDBC	Pajek Notes
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka
NumberOfServers	Število strežnikov NNTP v območju pajkanja.	Število baz podatkov v območju pajkanja.	Število baz podatkov v območju pajkanja.
NumberOfCompletedServers	Število pajkanih strežnikov NNTP.	Število pajkanih baz podatkov.	Število pajkanih baz podatkov.
NumberOfTargets	Število novičarskih skupin v območju pajkanja.	Število baz podatkov v območju pajkanja.	Število pogledov in map v območju pajkanja.
NumberOfCompletedTargets	Število pajkanih novičarskih skupin.	Število pajkanih tabel.	Število pajkanih pogledov in map.
NumberOfCompletedRecords	Število pajkanih člankov.	Število pajkanih zapisov.	Število pajkanih dokumentov.
RunningThreads	Število niti pajka.	Število niti pajka.	Število niti pajka.

Tabela 12. Statusne informacije za pajke Exchange Server, DB2 Content Manager in Content Edition

Ime elementa in atributa	Pajek strežnika Exchange Server	Pajek DB2 Content Manager	Pajek Content Edition
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka
NumberOfServers	Število strežnikov Exchange Server v območju pajkanja.	Število strežnikov Content Manager v območju pajkanja.	Število repozitorijev v območju pajkanja.
NumberOfCompletedServers	Število pajkanih strežnikov Exchange Server.	Število pajkanih strežnikov Content Manager.	Število pajkanih repozitorijev.
NumberOfTargets	Število podmap v območju pajkanja.	Število tipov postavk v območju pajkanja.	Število razredov v območju pajkanja.
NumberOfCompletedTargets	Število pajkanih podmap.	Število pajkanih tipov postavk.	Število pajkanih razredov postavk.
NumberOfCompletedRecords	Število pajkanih dokumentov.	Število pajkanih dokumentov.	Število pajkanih dokumentov.
RunningThreads	Število niti pajka.	Število niti pajka.	Število niti pajka.

Tabela 13. Statusne informacije za pajke QuickPlace, Domino Document Manager, datotečnega sistema UNIX in datotečnega sistema Windows

Ime elementa in atributa	Pajek QuickPlace	Pajek Domino Document Manager	Pajki datotečnih sistemov UNIX in Windows
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)



Tabela 13. Statusne informacije za pajke QuickPlace, Domino Document Manager, datotečnega sistema UNIX in datotečnega sistema Windows (nadaljevanje)

Ime elementa in atributa	Pajek QuickPlace	Pajek Domino Document Manager	Pajki datotečnih sistemov UNIX in Windows
StatusMessage	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka
NumberOfServers	Število mest v območju pajkanja.	Število knjižnic v območju pajkanja.	Nespremenljiva vrednost 1.
NumberOfCompletedServers	Število pajkanih mest.	Število pajkanih knjižnic.	0 ali 1, če so vsi podimeniki pajkani.
NumberOfTargets	Število baz podatkov mesta in sobe v območju pajkanja.	Število kabinetov v območju pajkanja.	Število podimenikov v območju pajkanja.
NumberOfCompletedTargets	Število pajkanih baz podatkov mesta in sobe.	Število pajkanih kabinetov.	Število pajkanih podimenikov.
NumberOfCompletedRecords	Število pajkanih dokumentov.	Število pajkanih dokumentov.	Število pajkanih datotek.
RunningThreads	Število niti pajka.	Število niti pajka.	Število niti pajka.

Tabela 14. Statusne informacije za pajke WebSphere Portal in Web Content Management

Ime elementa in atributa	Pajek WebSphere Portal	Pajek Web Content Management
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka	Status: 0 - Mirovanje, 1 - Izvajanje, 2 - Premor, -1 - Napaka
NumberOfServers	Število strežnikov v prostoru za pajkanje.	Število spletnih mest v prostoru za pajkanje.
NumberOfCompletedServers	Število pajkanih strežnikov.	Število pajkanih strani.
NumberOfTargets	Število strežnikov v prostoru za pajkanje.	Število spletnih mest v prostoru za pajkanje.
NumberOfCompletedTargets	Število pajkanih strežnikov.	Število pajkanih strani.
NumberOfCompletedRecords	Število pajkanih dokumentov.	Število pajkanih dokumentov.
RunningThreads	Število niti pajka.	Število niti pajka.

**Status območja pajkanja za spletnega pajka:** Ko zaženete ukaz za pridobitev statusa območja pajkanja za spletnega pajka, ukaz vrne informacije v formatu dokumenta XML. Ukaz za status območja spletnega pajkanja lahko vrne naslednje informacije:

Tabela 15. Vrednosti maske izbire za ukaz za status območja pajkanja spletnega pajka

Bit maske	Izbere
1	Število strani v neobdelani podatkovni shrambi.
2	Število odkritih spletnih mest.
4	Število spletnih mest z DNS.
8	Število spletnih mest brez DNS.
16	Število odkritih URL-jev.
32	Število unikatnih shranjenih strani.
64	Število pajkanih URL-jev.
128	Število nepajkanih URL-jev.

Tabela 15. Vrednosti maske izbire za ukaz za status območja pajkanja spletnega pajka (nadaljevanje)

Bit maske	Izbere
256	Število zapadlih URL-jev.
512	Porazdeljevanje statusnih kod HTTP.

Vse vrednosti predstavljajo kumulativne skupne vsote za vse seje, ki uporabljajo trenutno notranjo bazo podatkov:

```
<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>
```

Povratni podatki vsebujejo katerekoli od naslednjih elementov, lahko tudi vse ali nobenega:

Tabela 16. Statusne informacije o prostoru za pajkanje za spletnega pajka

Element	Atribut	Opis
CrawlerStatus	<ul style="list-style-type: none"> <li>NumPagesInRDS</li> <li>NumSitesDiscovered</li> <li>NumSitesWithDNS</li> <li>NumSitesWithoutDNS</li> <li>NumURLsDiscovered</li> <li>NumUniquePagesSaved</li> <li>NumURLsCrawled</li> <li>NumURLsUncrawled</li> <li>NumURLsOverdue</li> <li>HTTPCodeDist</li> </ul>	Informacije o kumulativnem stanju pajkanja (vse seje), ki jih je možno hitro pridobiti.
NumPagesInRDS	<p>Vrednost: Nenegativno celo število</p> <p>Koliko strani je trenutno v vmesnem območju (samo tega pajka) neobdelane podatkovne shrambe (RDS).</p>	Kako polna postaja neobdelana podatkovna shramba (RDS) (samo prispevki tega pajka).
NumSitesDiscovered	<p>Vrednost: Nenegativno celo število</p> <p>Koliko gostiteljev je bilo odkritih s pajkanjem (ali prek semen).</p>	Merilo pajkove pokritosti domene, ki bo pajkana (štetje gostiteljev).
NumSitesWithDNS	<p>Vrednost: Nenegativno celo število</p> <p>Število gostiteljev s povezanimi naslovi IP (ki jih razreši pajek v ozadju).</p>	Merilo za to, kako učinkovito lahko pajek pridobiva naslove IP za gostitelje, odkrite s pomočjo imen DNS v URL-jih.

Tabela 16. Statusne informacije o prostoru za pajkanje za spletnega pajka (nadaljevanje)

Element	Atribut	Opis
NumSitesWithoutDNS	Vrednost: Nenegativno celo število Število gostiteljev brez povezanih naslovov IP (ki jih razreši pajek v ozadju).	Merilo za to, kako učinkovito lahko pajek pridobiva naslove IP za gostitelje, odkrite s pomočjo imen DNS v URL-jih.
NumURLsDiscovered	Vrednost: Nenegativno celo število Koliko unikatnih URL-jev je obiskal pajek.	Merilo pajkove pokritosti domene, ki bo pajkana (štetje URL-jev).
NumUniquePagesSaved	Vrednost: Nenegativno celo število Koliko unikatnih strani je bilo zapisanih v RDS za nadaljnjo obdelavo z drugimi komponentami za iskanje v podjetju.	Doprinos tega pajka k velikosti indeksa.
NumURLsCrawled	Vrednost: Nenegativno celo število Koliko unikatnih URL-jev je pajkal pajek.	Merilo za pajkovo zmožnost celovito obdelati podatke. To število se razlikuje od števila strani, zapisanih v RDS, saj niso vse pajkane strani zapisane v RDS.
NumURLsOverdue	Vrednost: Nenegativno celo število Koliko unikatnih URL-jev je primernih za ponovno pajkanje.	Merilo za pajkovo zmožnost preiti spletno območje.

**Status območja pajkanja za nespletne pajke:** Ko zaženete ukaz za pridobitev statusa območja pajkanja za nespletnega pajka, ukaz vrne informacije v formatu dokumenta XML. Ukaz `getCrawlSpaceStatus` lahko za pajke, ki ne pajkajo po spletu, vrne naslednje informacije:

```
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>
```

Naslednje tabele opisujejo elemente in attribute XML za vsakega pajka za iskanje v podjetju, razen za spletnega pajka. Te informacije vrne ukaz za status območja pajkanja. Ko je raven združitve 0, je `Server@Name` za pajke Notes server name + database name. Ko je raven združitve 1, je `Server@Name` server name + directory name.

Tabela 17. Statusne informacije o prostoru za pajkanje za pajke NNTP, DB2, baze podatkov JDB in Notes

Ime elementa in atributa	Pajek NNTP	Pajki DB2 in baze podatkov	
		JDBC	Pajek Notes
Server@Name	Ime strežnika News	Ime baze podatkov	Ime baze podatkov ali imenika

Tabela 17. Statusne informacije o prostoru za pajkanje za pajke NNTP, DB2, baze podatkov JDB in Notes (nadaljevanje)

Ime elementa in atributa	Pajek NNTP	Pajki DB2 in baze podatkov JDBC	Pajek Notes
Server/Status	Status: (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• 5: Načrtovan</li> <li>• -1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• 5: Načrtovan</li> <li>• -1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• 5: Načrtovan</li> <li>• -1: Napaka</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• 5: Načrtovan</li> <li>• -1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• 5: Načrtovan</li> <li>• -1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• 5: Načrtovan</li> <li>• -1: Napaka</li> </ul>
Server/NumberOfTargets	Število novičarskih skupin v območju pajkanja.	Število baz podatkov v območju pajkanja.	Število pogledov in map ali imenikov v prostoru za pajkanje.
Server/NumberOfCompletedTargets	Število pajkanih novičarskih skupin.	Število pajkanih tabel.	Število pajkanih pogledov in map ali imenikov.
Server/NumberOfErrors	Število napak.	Število napak.	Število napak
Server/StartTime	Čas zagona, če je na voljo.	Čas zagona, če je na voljo.	Čas zagona, če je na voljo.
Server/EndTime	Čas konca, če je na voljo.	Čas konca, če je na voljo.	Čas konca, če je na voljo.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>• 1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>• 2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>• 1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>• 2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>• 1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>• 2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>
Server/ScheduleTime	Čas načrtovanja, če je na voljo.	Čas načrtovanja, če je na voljo.	Čas načrtovanja, če je na voljo.
Server/TotalTime	Skupni čas, če je na voljo.	Skupni čas, če je na voljo.	Skupni čas, če je na voljo.

Tabela 17. Statusne informacije o prostoru za pajkanje za pajke NNTP, DB2, baze podatkov JDB in Notes (nadaljevanje)

Ime elementa in atributa	Pajek NNTP	Pajki DB2 in baze podatkov JDBC	Pajek Notes
Server/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.	0, 1: <ul style="list-style-type: none"> <li>0: Pajek Notes pajka dokumente v navadnem načinu. (Ostali pajki, z izjemo pajka Notes, vedno vrnejo 0.)</li> <li>1: Pajek Notes pajka dokumente v imeniškem načinu.</li> </ul>

Tabela 18. Statusne informacije o prostoru za pajkanje za pajke Exchange Server, DB2 Content Manager in Content Edition

Ime elementa in atributa	Pajek strežnika Exchange Server	Pajek DB2 Content Manager	Pajek Content Edition
Server@Name	Ime strežnika Exchange Server.	Strežniki DB2 Content Manager.	Ime repozitorija.
Server/Status	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>
Server/NumberOfTargets	Število podmap v območju pajkanja.	Število tipov postavk v območju pajkanja.	Število razredov postavk v območju pajkanja.
Server/NumberOfCompletedTargets	Število pajkanih podmap.	Število pajkanih tipov postavk.	Število pajkanih razredov postavk.
Server/NumberOfErrors	Število napak.	Število napak.	Število napak.
Server/StartTime	Čas zagona, če je na voljo.	Čas zagona, če je na voljo.	Čas zagona, če je na voljo.
Server/EndTime	Čas konca, če je na voljo.	Čas konca, če je na voljo.	Čas konca, če je na voljo.

Tabela 18. Statusne informacije o prostoru za pajkanje za pajke Exchange Server, DB2 Content Manager in Content Edition (nadaljevanje)

Ime elementa in atributa	Pajek strežnika Exchange Server	Pajek DB2 Content Manager	Pajek Content Edition
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>
Server/ScheduleTime	Čas načrtovanja, če je na voljo.	Čas načrtovanja, če je na voljo.	Čas načrtovanja, če je na voljo.
Server/TotalTime	Skupni čas, če je na voljo.	Skupni čas, če je na voljo.	Skupni čas, če je na voljo.
Server/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.

Tabela 19. Statusne informacije o prostoru za pajkanje za pajke QuickPlace, Domino Document Manager, datotečnega sistema UNIX in datotečnega sistema Windows

Ime elementa in atributa	Pajek QuickPlace	Pajek Domino Document Manager	Pajki datotečnih sistemov UNIX in Windows
Server@Name	Imenik območja	Baza podatkov knjižnice	Nespremenljiva vrednost lokalnega gostitelja.
Server/Status	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>
Server/NumberOfTargets	Število baz podatkov mesta in sobe v območju pajkanja.	Število kabinetov v območju pajkanja.	Število podimenikov v območju pajkanja.
Server/NumberOfCompletedTargets	Število pajkanih baz podatkov mesta in sobe.	Število pajkanih kabinetov.	Število podimenikov v območju pajkanja.
Server/NumberOfErrors	Število napak.	Število napak.	Število napak.

Tabela 19. Statusne informacije o prostoru za pajkanje za pajke QuickPlace, Domino Document Manager, datotečnega sistema UNIX in datotečnega sistema Windows (nadaljevanje)

Ime elementa in atributa	Pajek QuickPlace	Pajek Domino Document Manager	Pajki datotečnih sistemov UNIX in Windows
Server/StartTime	Čas zagona, če je na voljo.	Čas zagona, če je na voljo.	Čas zagona, če je na voljo.
Server/EndTime	Čas konca, če je na voljo.	Čas konca, če je na voljo.	Čas konca, če je na voljo.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>
Server/ScheduleTime	Čas načrtovanja, če je na voljo.	Čas načrtovanja, če je na voljo.	Čas načrtovanja, če je na voljo.
Server/TotalTime	Skupni čas, če je na voljo.	Skupni čas, če je na voljo.	Skupni čas, če je na voljo.
Server/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.

Tabela 20. Statusne informacije o prostoru za pajkanje za pajka WebSphere Portal in Web Content Management

Ime elementa in atributa	Pajek WebSphere Portal	Pajek Web Content Management
Server@Name	Strežnik WebSphere Portal	Semenski URL za iskanje Web Content Management
Server/Status	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>5: Načrtovan</li> <li>-1: Napaka</li> </ul>
Server/NumberOfTargets	Število strežnikov v prostoru za pajkanje.	Število spletnih mest v prostoru za pajkanje.
Server/NumberOfCompletedTargets	Število pajkanih strežnikov.	Število pajkanih strani.
Server/NumberOfErrors	Število napak.	Število napak.
Server/StartTime	Čas zagona, če je na voljo.	Čas zagona, če je na voljo.

Tabela 20. Statusne informacije o prostoru za pajkanje za pajka WebSphere Portal in Web Content Management (nadaljevanje)

Ime elementa in atributa	Pajek WebSphere Portal	Pajek Web Content Management
Server/EndTime	Čas konca, če je na voljo.	Čas konca, če je na voljo.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>• 1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>• 2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pajek ni konfiguriran za načrtovanje glede na konfiguracijske datoteke pajka.</li> <li>• 1: Pajek je konfiguriran za načrtovanje, a je bilo načrtovanje onemogočeno za sejo</li> <li>• 2: Pajek je konfiguriran za načrtovanje, načrtovanje za sejo je omogočeno</li> </ul>
Server/ScheduleTime	Čas načrtovanja, če je na voljo.	Čas načrtovanja, če je na voljo.
Server/TotalTime	Skupni čas, če je na voljo.	Skupni čas, če je na voljo.
Server/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.

**Podrobni status območja pajkanja za spletnega pajka:** Ko zaženete ukaz za pridobitev podrobnega statusa območja pajkanja za spletnega pajka, ukaz vrne informacije v formatu dokumenta XML. Ukaz za podrobni status območja pajkanja lahko vrne naslednje informacije:

Tabela 21. Vrednosti maske izbire za ukaz za podrobni status območja pajkanja spletnega pajka

Bit maske	Izbere
1	Število strani v neobdelani podatkovni shrambi.
2	Število odkritih spletnih mest.
4	Število spletnih mest z DNS.
8	Število spletnih mest brez DNS.
16	Število odkritih URL-jev.
32	Število unikatnih shranjenih strani.
64	Število pajkanih URL-jev.
128	Število nepajkanih URL-jev.
256	Število zapadlih URL-jev.
512	Porazdeljevanje statusnih kod HTTP.

Vzorec vrnjenih informacij:

```
<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>

```

Naslednja tabela opisuje vsako polje, ki je vrnjeno za podrobni status območja pajkanja spletnega pajka:

Tabela 22. Podrobne statusne informacije o prostoru za pajkanje za spletnega pajka

Element	Atributi	Opis
CrawlDetailsPerSite	<ul style="list-style-type: none"> <li>LastActivationTime:</li> <li>LastActivationDuration:</li> <li>IPAddressList:</li> <li>RobotsContent:</li> <li>HTTPCodeDist:</li> </ul>	Informacije o podrobnem stanju enega spletnega mesta, ki jih je možno hitro pridobiti.
Site	URL	URL korenške strani spletnega mesta.
NumURLsDiscovered	Vrednost	Število URL-jev, ki so bili odkriti s spletnega mesta.
NumURLsOverdue	Vrednost	Število URL-jev, ki so primerni za ponovno pajkanje s spletnega mesta.
NumURLsCrawled	Vrednost	Število URL-jev, ki so bili pajkani za spletno mesto.
NumURLsUncrawled	Vrednost	Število URL-jev, ki še niso pajkani za spletno mesto.
NumURLsOverdueBy	<p>Prag, Vrednost: Celo število (pozitivno ali negativno)</p> <p>Vrednost predstavlja število URL-jev, ki ustrezajo kriterijem za vnovično pajkanje. Prag podaja, kako dolgo so URL-ji čakalni na vnovično pajkanje. Prag se meri kot odmik v sekundah od trenutnega časa. Če je prag negativen, to pomeni, da je vnovično pajkanje URL-jev že zamujeno. Če je prag pozitiven, to pomeni, da bo izvedeno vnovično pajkanje URL-jev.</p>	Število URL-jev, ki so postali primerni za vnovično pajkanje nekaj sekund nazaj, ali postajajo primerni za vnovično pajkanje v naslednjih nekaj sekundah.
NumURLsActivated	Vrednost	Število URL-jev, ki so prišli v spomin med zadnjim pregledovanjem tega spletnega mesta in so na voljo nitim pajka.
LastActivationTime	Vrednost	Število sekund od časa, ko so URL-ji tega spletnega mesta zadnjič vstopili v spomin.
LastActivationDuration	Vrednost	Število sekund, ko so bili URL-ji te strani nazadnje v spominu in na voljo nitim pajka.
IPAddressList	IPAddress	Vsi znani naslovi IP za gostitelja strežnika tega spletnega mesta.
IPAddress	Vrednost	Naslov IPv4, zapisan v pikah, za gostitelja strežnika tega spletnega mesta.

Tabela 22. Podrobne statusne informacije o prostoru za pajkanje za spletnega pajka (nadaljevanje)

Element	Atributi	Opis
RobotsContent	Besedilo	Besedilo iz datoteke robotov, če obstaja.
HTTPCodeDist	HTTPCode	Porazdeljevanje kod HTTP iz poskusov prenosa tega spletnega mesta.
HTTPCode	Code: Celo število Statusna koda HTTP ali druga notranja koda.	Kolikokrat se je določena statusna koda HTTP pojavila med pajkanjem tega spletnega mesta.

**Podrobni status območja pajkanja za nespletne pajke:** Ko zaženete ukaz za pridobitev podrobnega statusa območja pajkanja za nespletne pajke, ukaz vrne informacije v formatu dokumenta XML. Ukaz **getCrawlSpaceStatusDetail** lahko za pajke, ki ne pajkajo po spletu, vrne naslednje informacije:

```
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. Podrobne statusne informacije o prostoru za pajkanje za pajke NNTP, DB2, baze podatkov JDBC in Notes

Ime elementa in atributa	Pajek NNTP	Pajki DB2 in baze podatkov JDBC	Pajek Notes
Target@Name	Ime skupine News	Ime tabele	Ime pogleda ali mape
Target@CrawlType	Ni na voljo.	0,1 (DB2); 0 (baza podatkov JDBC) <ul style="list-style-type: none"> <li>0: Aktivno pajkanje (Navadno)</li> <li>1: Pasivno pajkanje (DB2 Event Publishing)</li> </ul>	0
Target/Status	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>

Tabela 23. Podrobne statusne informacije o prostoru za pajkanje za pajke NNTP, DB2, baze podatkov JDBC in Notes (nadaljevanje)

Ime elementa in atributa	Pajek NNTP	Pajki DB2 in baze podatkov JDBC	Pajek Notes
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>
Target/NumberOfRecords	Številka zadnjega članka na strežniku.	Število pajkanih zapisov.	Število pajkanih dokumentov.
Target/NumberOfCompletedRecords	Število pajkanih člankov.	Število pajkanih zapisov.	Število pajkanih dokumentov.
Target/NumberOfInsertedRecords	Število na novo objavljenih člankov.	Število vstavljenih zapisov.	Število vstavljenih zapisov.
Target/NumberOfUpdatedRecords	Ni na voljo.	Število posodobljenih zapisov.	Število posodobljenih zapisov.
Target/NumberOfDeletedRecords	Ni na voljo.	Število izbranih zapisov.	Število izbranih zapisov.
Target/StartTime	Datum in čas, ko se je pajek zadnjič zagnal.	Datum in čas, ko se je pajek zadnjič zagnal.	Datum in čas, ko se je pajek zadnjič zagnal.
Target/EndTime	Datum in čas, ko je bilo pajkanje končano.	Datum in čas, ko je bilo pajkanje končano.	Datum in čas, ko je bilo pajkanje končano.
Target/TotalTime	Čas, ki ga je pajek porabil za pajkanje.	Čas, ki ga je pajek porabil za pajkanje.	Čas, ki ga je pajek porabil za pajkanje.
Target/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.	0, 1: <ul style="list-style-type: none"> <li>0: Pajek pajka dokumente v običajnem načinu.</li> <li>Pajek pajka dokumente v imeniškem načinu.</li> </ul>
Target/LastUpdatedTime	Ni na voljo.	Zadnji posodobljen čas: <ul style="list-style-type: none"> <li>0: Aktivno pajkanje (Navadno)</li> <li>1: Pasivno pajkanje (DB2 Event Publishing)</li> </ul>	Ni na voljo.
Target/LastResetTime	Ni na voljo.	Zadnji statistični podatki ponovne nastavitve: <ul style="list-style-type: none"> <li>0: Aktivno pajkanje (Navadno)</li> <li>1: Pasivno pajkanje (DB2 Event Publishing)</li> </ul>	Ni na voljo.

Tabela 24. Podrobne statusne informacije o prostoru za pajkanje za pajke Exchange Server, DB2 Content Manager in Content Edition

Ime elementa in atributa	Pajek strežnika Exchange Server	Pajek DB2 Content Manager	Pajek Content Edition
Target@Name	Ime podmape	Ime tipa postavke	Ime razreda postavke

Tabela 24. Podrobne statusne informacije o prostoru za pajkanje za pajke Exchange Server, DB2 Content Manager in Content Edition (nadaljevanje)

Ime elementa in atributa	Pajek strežnika Exchange Server	Pajek DB2 Content Manager	Pajek Content Edition
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>0: Nepajkan</li> <li>1: Pajkanje</li> <li>2: Končano (brez razporeda)</li> <li>3: Čakanje</li> <li>4: Začasno ustavljen</li> <li>-1: Napaka</li> </ul>
Target/NumberOfRecords	Ni na voljo.	Ni na voljo.	Ni na voljo.
Target/NumberOfCompletedRecords	Število pajkanih dokumentov.	Število pajkanih dokumentov.	Število pajkanih dokumentov.
Target/NumberOfInsertedRecords	Število vstavljenih zapisov.	Število vstavljenih zapisov.	Število vstavljenih zapisov.
Target/NumberOfUpdatedRecords	Ni na voljo.	Število posodobljenih zapisov.	Število posodobljenih zapisov.
Target/NumberOfDeletedRecords	Ni na voljo.	Število izbranih zapisov.	Število izbranih zapisov.
Target/StartTime	Datum in čas, ko se je pajek zadnjič zagnal.	Datum in čas, ko se je pajek zadnjič zagnal.	Datum in čas, ko se je pajek zadnjič zagnal.
Target/EndTime	Datum in čas, ko je bilo pajkanje končano.	Datum in čas, ko je bilo pajkanje končano.	Datum in čas, ko je bilo pajkanje končano.
Target/TotalTime	Čas, ki ga je pajek porabil za pajkanje.	Čas, ki ga je pajek porabil za pajkanje.	Čas, ki ga je pajek porabil za pajkanje.
Target/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.
Target/LastUpdateTime	Ni na voljo.	Ni na voljo.	Ni na voljo.
Target/LastResetTime	Ni na voljo.	Ni na voljo.	Ni na voljo.

Tabela 25. Podrobne statusne informacije o prostoru za pajkanje za pajke QuickPlace, Domino Document Manager, datotečnega sistema UNIX in datotečnega sistema Windows

Ime elementa in atributa	Pajek QuickPlace	Pajek Domino Document Manager	Pajki datotečnih sistemov UNIX in Windows
Target@Name	Ime baze podatkov prostora ali ime baze podatkov sobe	Ime baze podatkov kabineta	Ime podimenika
Target@CrawlType	0	0	0

Tabela 25. Podrobne statusne informacije o prostoru za pajkanje za pajke QuickPlace, Domino Document Manager, datotečnega sistema UNIX in datotečnega sistema Windows (nadaljevanje)

Ime elementa in atributa	Pajek QuickPlace	Pajek Domino Document Manager	Pajki datotečnih sistemov UNIX in Windows
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>
Target/NumberOfRecords	Ni na voljo.	Ni na voljo.	Ni na voljo.
Target/NumberOfCompletedRecords	Število pajkanih dokumentov.	Število pajkanih dokumentov.	Število pajkanih datotek.
Target/NumberOfInsertedRecords	Število vstavljenih zapisov.	Število vstavljenih zapisov.	Število vstavljenih zapisov.
Target/NumberOfUpdatedRecords	Število posodobljenih zapisov.	Število posodobljenih zapisov.	Število posodobljenih zapisov.
Target/NumberOfDeletedRecords	Število izbranih zapisov.	Število izbranih zapisov.	Število izbranih zapisov.
Target/StartTime	Datum in čas, ko se je pajek zadnjič zagnal.	Datum in čas, ko se je pajek zadnjič zagnal.	Datum in čas, ko se je pajek zadnjič zagnal.
Target/EndTime	Datum in čas, ko je bilo pajkanje končano.	Datum in čas, ko je bilo pajkanje končano.	Datum in čas, ko je bilo pajkanje končano.
Target/TotalTime	Čas, ki ga je pajek porabil za pajkanje.	Čas, ki ga je pajek porabil za pajkanje.	Čas, ki ga je pajek porabil za pajkanje.
Target/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.
Target/LastUpdateTime	Ni na voljo.	Ni na voljo.	Ni na voljo.
Target/LastResetTime	Ni na voljo.	Ni na voljo.	Ni na voljo.

Tabela 26. Podrobne statusne informacije o prostoru za pajkanje za pajka WebSphere Portal in Web Content Management

Ime elementa in atributa	Pajek WebSphere Portal	Pajek Web Content Management
Target@Name	Ime strežnika WebSphere Portal	Semenski URL za iskanje, ki predstavlja spletno mesto
Target@CrawlType	0	0

Tabela 26. Podrobne statusne informacije o prostoru za pajkanje za pajka WebSphere Portal in Web Content Management (nadaljevanje)

Ime elementa in atributa	Pajek WebSphere Portal	Pajek Web Content Management
Target/Status	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nepajkan</li> <li>• 1: Pajkanje</li> <li>• 2: Končano (brez razporeda)</li> <li>• 3: Čakanje</li> <li>• 4: Začasno ustavljen</li> <li>• -1: Napaka</li> </ul>
Target/NumberOf Records	Ni na voljo.	Ni na voljo.
Target/NumberOf CompletedRecords	Skupno število pajkanih zapisov.	Skupno število pajkanih zapisov.
Target/NumberOf InsertedRecords	Število vstavljenih zapisov.	Število vstavljenih zapisov.
Target/NumberOf UpdatedRecords	Število posodobljenih zapisov.	Število posodobljenih zapisov.
Target/NumberOf DeletedRecords	Število izbranih zapisov.	Število izbranih zapisov.
Target/StartTime	Datum in čas, ko se je pajek zadnjič zagnal.	Datum in čas, ko se je pajek zadnjič zagnal.
Target/EndTime	Datum in čas, ko je bilo pajkanje končano.	Datum in čas, ko je bilo pajkanje končano.
Target/TotalTime	Čas, ki ga je pajek porabil za pajkanje.	Čas, ki ga je pajek porabil za pajkanje.
Target/AggregationLevel	0: Pajek pajka dokumente v običajnem načinu.	0: Pajek pajka dokumente v običajnem načinu.
Target/LastUpdatedTime	Ni na voljo.	Ni na voljo.
Target/LastResetTime	Ni na voljo.	Ni na voljo.

**Status razčlenjevalnika:** Ko zažene ukaz za pridobitev statusa razčlenjevalnika, ukaz vrne informacije v formatu dokumenta XML. Ukaz za status razčlenjevalnika lahko vrne naslednje informacije:

```
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>

```

Naslednja tabela opisuje elemente XML za informacije, ki jih vrne ukaz za status razčlenjevalnika:

Tabela 27. Elementi za ukaz za status razčlenjevalnika

Element	Opis
Status	<ul style="list-style-type: none"> <li>• 0: Seja razčlenjevalnika za to zbirko je zaustavljena.</li> <li>• 1: Seja razčlenjevalnika za to zbirko se izvaja.</li> </ul>
State	<p>Možna stanja so: Inicializiranje, Mirovanje, Vnovični zagon, Razčlenjevanje, Ustavljen, Začasno ustavljen, Nadaljevanje, StoritveRazčlenjevalnikaNiNaVoljo (NoParserServiceIsAvailable).</p> <p>Status Inicializiranje pomeni, da pajek zaganja in inicializira svoje stanje.</p> <p>Status Mirovanje pomeni, da razčlenjevalnik miruje za <i>N</i> minut, v čakaniu, da prispe več dokumentov od pajkov v tej zbirki. Privzeti čas mirovanja je 300 sekund.</p> <p>Status Vnovični zagon pomeni, da razčlenjevalnik čaka, da se bo JVM za razčlenjevanje/izdelovanje žetonov ponovno zagnal. JVM za razčlenjevanje/izdelovanje žetonov se izvaja v ločeni seji in je prostor končne obdelave dokumentov.</p> <p>Status Razčlenjevanje pomeni, da razčlenjevalnik obdeluje dokumente.</p> <p>Status Začasno ustavljen pomeni, da je razčlenjevalnik začasno ustavila seja gradnje indeksa za to zbirko.</p> <p>Status Nadaljevanje pomeni, da je seja gradnje indeksa za to zbirko spremenila stanje razčlenjevalnika iz Začasno ustavljen na Razčlenjevanje.</p> <p>Status StoritveRazčlenjevalnikaNiNaVoljo pomeni, da ni razpoložljivih JVM-ov za razčlenjevanje/izdelovanje žetonov za obdelavo dokumentov za to zbirko. Ta status pomeni, da vse JVM-e za razčlenjevanje/izdelovanje žetonov uporabljajo ostale zbirke.</p>
NumberOfDocsToBeIndexed	Število dokumentov v shrambi za to zbirko. To število prav tako vključuje dokumente, ki so označeni za izbris iz naslednje gradnje indeksa.
ParseRate	Hitrost razčlenjevanja v dokumentih na sekundo.
ParseRateMBPerHour	Hitrost razčlenjevanja v MB na uro.
NumberOfCpmThreads	Število niti CPM, ki jih JVM za razčlenjevanje/izdelovanje žetonov uporablja za obdelovanje dokumentov za to zbirko.
ParserServiceSession	Ime JVM-a za razčlenjevanje/izdelovanje žetonov, ki obdeluje dokumente za to zbirko. To polje je na voljo samo, če je razčlenjevalnik v stanju Razčlenjevanje.
Name	Ime pajka.
Crawlerid	ID, ki ga je za pajka izdelal sistem.
Type	Tip pajka (Web, NNTP, DB2 itd.)
ParserStatus	<ul style="list-style-type: none"> <li>• 0: Dokumenti tega pajka niso pajkani (seja pajkanja je zaustavljena).</li> <li>• 1: Dokumenti tega pajka so pajkani (seja pajkanja se izvaja).</li> </ul>
NumberOfDocsAlreadyParsed	Število dokumentov tega pajka, ki so bili pajkani.

**Status gradnje indeksa:** Ko zaženete ukaz za pridobitev statusa gradnje indeksa, ukaz vrne informacije v formatu dokumenta XML. Ukaz za status gradnje indeksa lahko vrne naslednje informacije:

```
<?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>
```

Naslednja tabela opisuje vsak element XML za informacije, ki jih vrne ukaz za status gradnje indeksa:

Tabela 28. Elementi za ukaz za status gradnje indeksa

Element	Opis
IndexStatusId	ID statusa indeksa.
StartTime	Čas v sekundah od 1970, ko se je ta gradnja indeksa pričela. Za izračun sedanjega časa, ki ga ta čas predstavlja, uporabite formulo <code>January 1, 1970 %2B StartTime</code> . Za več informacij o času Unix epoch glejte <a href="http://en.wikipedia.org/wiki/Unix_epoch">http://en.wikipedia.org/wiki/Unix_epoch</a> .
Progress	Dokončanje v odstotkih za to gradnjo indeksa.
CurrentPhase	<ul style="list-style-type: none"> <li>1: stopnja prepisovanje shrambe</li> <li>2: stopnja globalne analize</li> <li>3: stopnja gradnje indeksa</li> </ul>
TotalPhase	Število stopenj za to gradnjo indeksa. Ta vrednost je trenutno 3.
IndexCopyProgress	Dokončanje v odstotkih za kopijo indeksa. Proces kopiranja indeksa kopira zgrajen indeks s strežnika gradnje indeksa na iskalne strežnike.
CurrentServer	Iskalni strežnik, na katerega kopija indeksa kopira indeks.
TotalServer	Število iskalnih strežnikov, na katere se kopira indeks.
IndexCopyTime	Skupni čas za kopiranje indeksa na vse iskalne strežnike.
IndexBuildTime	Skupni čas za vse stopnje gradnje indeksa
Status	<ul style="list-style-type: none"> <li>0: gradnja in kopiranje indeksa</li> <li>-1: napaka pri zahtevi gradnje indeksa</li> <li>1: gradnja indeksa, kopiranje indeksa, ali oba, sta v teku</li> </ul>
JobID	Unikatni ID, ki je povezan z vsako zahtevo za gradnjo indeksa.



Tabela 28. Elementi za ukaz za status gradnje indeksa (nadaljevanje)

Element	Opis
MessagesAvailable	Boolova vrednost, ki kaže, ali so sporočila o napakah na voljo (v primeru napake).
StopTime	Končni čas za gradnjo (vse stopnje) in kopiranje indeksa.
TotalTime	Časovno obdobje med pričetkom in ustavitvijo.
NumberOfDocuments	Število dokumentov v indeksu.
CurrentIndexWildcardSupport	Nastavitev univerzalnega znaka za naslednjo gradnjo indeksa. Možne vrednosti so None, QueryExpansion ali IndexExpansion.
ScheduleStatus	<ul style="list-style-type: none"> <li>• 0 če načrtovanje za ta tip zbirke in indeksa ni omogočeno.</li> <li>• 1 če načrtovanje za ta tip zbirke in indeksa je omogočeno.</li> </ul>
ScheduledTimeEnabled	Čas v sekundah od 1970, ko se bo pričela naslednja gradnja indeksa za ta tip zbirke in indeksa. Za izračun sedanjega časa, ki ga ta čas predstavlja, uporabite formulo <code>January 1, 1970 %2B ScheduledTimeEnabled</code> . Za več informacij o času Unix epoch glejte <a href="http://en.wikipedia.org/wiki/Unix_epoch">http://en.wikipedia.org/wiki/Unix_epoch</a> .

**Status iskalnega strežnika:** Ko zaženete ukaz za pridobitev statusa iskalnega strežnika, ukaz vrne informacije v formatu dokumenta XML. Ukaz za status iskalnega strežnika lahko vrne naslednje informacije:

```
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>
```

Naslednja tabela opisuje elemente XML za informacije, ki jih vrne ukaz za status iskalnega strežnika:

Tabela 29. Elementi za ukaz za status iskalnega strežnika

Element	Opis
SearchStatusName	Ime in ID seje upravljalnika iskanja, ki nadzoruje in ohranja iskalni indeks za to zbirko.
HostName	Ime gostitelja strežnika, kjer se izvaja iskalni indeks.
Status	<ul style="list-style-type: none"> <li>• 0 če se iskalni indeks za to zbirko ne izvaja.</li> <li>• 1 če se iskalni indeks za to zbirko izvaja.</li> </ul>

**Podrobni status iskalnega strežnika:** Ukaz za vrnitev v iskalni strežnik lahko vrne naslednje informacije:

```
FFQC5303I Search Manager (node1) (sid: searchmanager.node1)
is already running. PID: 15711
FFQC5314I Result: PID=18390
CacheHits=3
QueryRate=1
Port=44008
SessionId=coll.runtime.node1
CacheHitRate=0.333
ResponseTime=70
Status=1
SessionName=coll.runtime.node1.1
```

Naslednja tabela opisuje postavke v informacijah, ki jih vrne ukaz za podrobni status iskalnega strežnika:

Tabela 30. Postavke za ukaz za podrobni status iskalnega strežnika

Postavka	Opis
CacheHits	Število rezultatov, priklicanih iz iskalnega predpomnilnika.
QueryRate	Število poizvedb, prejetih v zadnjem časovnem intervalu. Po privzetku je časovni interval pet minut.
Port	Številka vrat, ki jo iskalni indeks uporablja za poslušanje ali prejemanje poizvedb.
SessionId	ID seje za iskalni indeks te zbirke.
CacheHitRate	Število rezultatov, priklicanih iz iskalnega predpomnilnika, kot odstotek vseh rezultatov iskanja.
ResponseTime	Povprečni odzivni čas v milisekundah za podani časovni interval. (Po privzetku pet minut.)
Status	<ul style="list-style-type: none"> <li>• 0 če se iskalni indeks za to zbirko ne izvaja.</li> <li>• 1 če se iskalni indeks za to zbirko izvaja.</li> </ul>
SessionName	Ime seje za iskalni indeks te zbirke.

## Povratne kode za ukaze esadmin

Za ukaze **esadmin** so lahko vrnjene naslednje kode:

Tabela 31. Povratne kode za ukaze esadmin

Koda	Name	Opis
0	CODE_ERROR_NONE	Ukaz uspešno končan.
102	CODE_ERROR_INSTANTIATION_EXCEPTION	Med pripravo nadzornika ukazov je prišlo do napake.
103	CODE_ERROR_ACCESS_EXCEPTION	Med pripravo nadzornika ukazov je prišlo do napake zaradi neveljavnega dostopa.
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	

## Pridobivanje ID-jev sej

Za prikaz seznama komponent za iskanje v podjetju in njihovih ustreznih ID-jev seje uporabite ukaz **esadmin check**. Naslednja tabela prikazuje seznam skupnih sej, njihovih ID-jev, strežnik, na katerem se nahajajo, in stanje seje.

Tabela 32. Primeri imen sej, izvornih strežnikov, ID-jev sej in stanj sej

Seja	Strežnik, na katerem se izvaja seja	ID seje	Stanje seje
configmanager	indeksni strežnik	10433	Zagnana
controller	indeksni strežnik	10464	Zagnana

Tabela 32. Primeri imen sej, izvornih strežnikov, ID-jev sej in stanj sej (nadaljevanje)

Seja	Strežnik, na katerem se izvaja seja	ID seje	Stanje seje
customcommunication	indeksni strežnik	Ni na voljo	Ni na voljo
odkritje	indeksni strežnik	10649	Zagnana
nadzornik	indeksni strežnik	10682	Zagnana
parserservice	indeksni strežnik	10718	Zagnana
resource.node1	indeksni strežnik	10759	Zagnana
samplecpp	indeksni strežnik	10827	Zagnana
sampletest	indeksni strežnik	10857	Zagnana
scheduler	indeksni strežnik	10889	Zagnana
searchmanager.node1	indeksni strežnik	10927	Zagnana
utilities.node1	indeksni strežnik	10384	Zagnana

### S tem povezani pojmi

“Nadziranje dejavnosti iskanja v podjetju” na strani 273

“Varnostno kopiranje in obnavljanje sistema iskanja v podjetju” na strani 303



Sporočila za iskanje v podjetju



Sporočila za iskanje v podjetju

### S tem povezana opravila

“Nadziranje pajkov” na strani 275

“Zagon sistema iskanja v podjetju” na strani 267

“Zaustavitev sistema iskanja v podjetju” na strani 269

“Izvajanje skrbništva nad iskalnimi strežniki v samostojnem načinu” na strani 271



---

## Občutljivost na velike in male črke pri iskanju v podjetju

Komponente iskanja v podjetju, kot so skladnja poizvedbe, hitre povezave, imena polj itd., različno obravnavajo velike in male črke.

### Skladnja poizvedbe

Iskanje ne razlikuje med velikimi in malimi črkami, razen v naslednjih primerih:

#### Imena elementov in atributov XML

Razlikujejo med velikimi in malimi črkami. Izrazi in vrednosti atributov razlikujejo o medvelikimi in malimi črkami tudi v poizvedbah XML. Na primer v naslednjem dokumentu:

```
<book>
 <Author>
 <Name>Ferdinand</Name>
 <Contact Type="eMail">ferdi@nand.org</Contact Type>
 <Contact Type="Phone">+1 408 876 4242</Contact Type>
 </Author>
</book>
```

Naslednje poizvedbe ne bodo vrnila dokumenta:

- `@xmlns::'author[Name ftcontains ("Ferdinand")]`
- `@xmlns::'//contact[@type="eMail"]'`
- `@xmlf2::'<author><name>Ferdinand</name></author>`
- `@xmlf2::'<CONTACT TYPE="email">ferdi</contact>`

Naslednje poizvedbe pa bodo vrnila dokument:

- `@xmlns::'Author[Name ftcontains ("ferdinand")]`
- `@xmlns::'//Contact[@Type="email"]'`
- `@xmlf2::'<Author><Name>ferdinand</Author><Name>`
- `@xmlf2::'<Contact Type="email">ferdi</Contact>`

#### Seznami za nadzor dostopa (ACL-i)

Razlikujejo med velikimi in malimi črkami.

#### URL-ji v izrazih docid: in samegroupas:

Razlikujejo med velikimi in malimi črkami. Vendar pa deli URL-ja v izrazih site: in url: ne ločijo med velikimi in malimi črkami. V dokumentu z URL-jem `http://www.here.com/HR/` velja naslednje:

- Poizvedba `docid:http://www.here.com/hr` ne bo vrnila dokumenta.
- Poizvedbe `url:hr` in `url:HERE` bodo vrnila dokument.
- Poizvedba `site:HERE.com` bo vrnila dokument.

#### ID-ji kategorij in taksonomij

Razlikujejo med velikimi in malimi črkami. V poizvedbi `taxonomy_id::category_id` velikost črk pri `taxonomy_id` in `category_id` na primer ni važna. Poizvedba `RuleBased::c42` se ujema s kategorijo `c42` v taksonomiji na osnovi pravil ter se prav tako ujema z `rulebased::C42`. Po imenu kategorije ne morete iskati, lahko pa iščete po ID-ju kategorije.

#### Območja

Ne razlikujejo med velikimi in malimi črkami. Oba dokumenta `Scope:RESEARCH` in `scope::research` bosta na primer vrnila dokumente iz območja z imenom `Research`.

**Univerzalni znaki**

Ne razlikujejo med velikimi in malimi črkami. Tako je izraz Fer\*n\*d enakovreden fer\*n\*d.

**Imena polj**

Ne razlikujejo med velikimi in malimi črkami. Tako so na primer poizvedbe Title:Expenses,TITLE:expense in title:expenses enakovredne. Vsa imena polj ne razlikujejo med velikimi in malimi črkami, tudi če izvirajo iz datoteke za preslikavo XML. Kljub temu lahko zunanji viri obravnavajo imena polj kot občutljiva na velike in male črke glede na svojo semantiko poizvedb.

**Hitre povezave**

Ne razlikujejo med velikimi in malimi črkami.

**Kategorizator na osnovi pravila**

Pravila URL razlikujejo med velikimi in malimi črkami, pravila vsebine dokumenta pa ne.

**Skrčeni URI-ji in definicije pospeševanja, ki temeljijo na vzorcu URI-ja**

URI-ji razlikujejo med velikimi in malimi črkami, skržena skupinska imena URI pa ne. Dveh skrčenih imen skupin URI, ki se razlikujeta samo v velikih in malih črkah, ne morete podati.

**Imena polj**

Ne razlikujejo med velikimi in malimi črkami. To pravilo velja za imena polj v poizvedbah, preslikave XML, definicije razreda pospeševanja in vse ostale skrbniške vmesnike, kjer so podana imena polj. Če v vmesniku konfigurirate ime polja v velikih črkah ali mešanih črkah, lahko sistem te spremeni v male črke, tako da se bo ime ob naslednjem ogledu konfiguracije prikazalo v malih črkah. Zunanji vir lahko interpretira tudi ime polja tako, da zazna velike in male črke.

**Slovarji**

Ne razlikujejo med velikimi in malimi črkami. To pravilo velja za slovarje sinonimov, zaustavitvenih besed, predlog za črkovanje in poviševalnih besed.

**ID-ji in imena zbirke**

Razlikujejo med velikimi in malimi črkami. Če v API-ju skrbništva ali iskanja podate ime ali ID zbirke, se mora natančno ujemati z uporabo velikih in malih črk v zbirki s tem imenom. Čeprav so ID-ji zbirke občutljivi na velike in male črke, ne morete podati dveh ID-jev zbirke, ki se razlikujeta le v velikih in malih črkah. Enaka omejitev velja za ID-je in imena pajka in virov podatkov.

**Iskalne aplikacije**

Razlikujejo med velikimi in malimi črkami. Dveh imen ali ID-jev aplikacij, ki se razlikujeta le v velikih in malih črkah, ne morete podati.

---

## Dokumentacija iskanja v podjetju

Dokumentacija OmniFind Enterprise Edition je na voljo v obliki zapisa PDF ali HTML.

Namestitveni program OmniFind Enterprise Edition samodejno namesti informacijski center, ki zajema različice HTML dokumentacije za iskanje v podjetju. Ko gre za namestitev na več strežnikov, se informacijski center namesti na oba iskalna strežnika. Če ne namestite informacijskega centra, se ob kliku pomoči odpre informacijski center, postavljen na IBM-ovem spletnem mestu.

Če si želite ogledati nameščene različice dokumentov PDF, pojdite v imenik `ES_INSTALL_ROOT/docs/področna_nastavitev/pdf`. Če želite na primer najti angleške dokumente, pojdite v imenik `ES_INSTALL_ROOT/docs/en_US/pdf`.

Če želite priti do različic PDF dokumentacije v vseh jezikih, ki so na voljo, obiščite spletno mesto z dokumentacijo za OmniFind Enterprise Edition, različica 8.5.

Na spletnem mestu s podporo za OmniFind Enterprise Edition so na voljo prenosi in povezave, paketi popravkov, tehnične opombe in informacijski center.

Naslednja tabela prikazuje razpoložljivo dokumentacijo, imena datotek in mesta.

*Tabela 33. Dokumentacija za iskanje v podjetju*

Naslov	Ime datoteke	Mesto
Informacijski center		<a href="http://publib.boulder.ibm.com/infocenter/discover/v8r5/">http://publib.boulder.ibm.com/infocenter/discover/v8r5/</a>
<i>Installation Guide for Enterprise Search</i>	iiysi.pdf	ES_INSTALL_ROOT/docs/področna_nastavitev/pdf/
<i>Quick Start Guide</i> (ta dokument je v natisnjeni obliki na voljo tudi za angleščino, francoščino in japonsščino)	OmniFindEE850_qsg_dvočrkovna_vrednost_področne_nastavitve.pdf	ES_INSTALL_ROOT/docs/področna_nastavitev/pdf/
<i>Administering Enterprise Search</i>	iiysa.pdf	ES_INSTALL_ROOT/docs/področna_nastavitev/pdf/
<i>Programming Guide and API Reference for Enterprise Search</i>	iiysp.pdf	ES_INSTALL_ROOT/docs/en_US/pdf/
<i>Troubleshooting Guide and Messages Reference</i>	iiysm.pdf	ES_INSTALL_ROOT/docs/področna_nastavitev/pdf/
<i>Text Analysis Integration</i>	iiyst.pdf	ES_INSTALL_ROOT/docs/področna_nastavitev/pdf/
<i>Plug-in for Google Desktop Search</i>	iiysg.pdf	ES_INSTALL_ROOT/docs/področna_nastavitev/pdf/





---

## Pripomočki za osebe s posebnimi potrebami

Pripomočki za osebe s posebnimi potrebami so namenjene osebam s fizično omejitvijo, kot je na primer omejena gibljivost ali slabši vid. Nudijo jim pomoč pri uporabi izdelkov s področja informacijske tehnologije.

IBM stremi k temu, da bi bili izdelki na voljo vsakomur, ne glede na starost ali sposobnosti.

### Pripomočki za osebe s posebnimi potrebami

Spodnji seznam zajema najpomembnejše pripomočke za osebe s posebnimi potrebami, ki so na voljo v OmniFind Enterprise Edition:

- Upravljanje samo prek tipkovnice
- Vmesniki, ki jih ponavadi uporabljajo bralniki zaslona

Informacijski center za OmniFind Enterprise Edition in z njim povezane publikacije so prilagojene za osebe s posebnimi potrebami. Pripomočki za osebe s posebnimi potrebami, ki so na voljo v informacijskem centru, so opisani na naslovu [http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/topic/com.ibm.classify.nav.doc/dochome/accessibility\\_info.htm](http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/topic/com.ibm.classify.nav.doc/dochome/accessibility_info.htm).

### Usmerjanje s tipkovnico

Ta izdelek uporablja standardne navigacijske tipke Microsoft Windows.

Po namestitvenem programu za OmniFind Enterprise Edition se lahko pomikate tudi z naslednjimi bližnjicami na tipkovnici.

*Tabela 34. Bližnjice tipkovnice za namestitveni program*

Dejanje	Bližnjica
Označitev izbirnega gumba	Smerna tipka
Izbira izbirnega gumba	Tipka Tab
Označitev gumba	Tipka Tab
Izbira gumba	Tipka Enter
Pomik v naslednje ali prejšnje okno ali preklic	Označite gumb s pritiskom tipke Tab in pritisnite Enter
Pretvorba aktivnega okna v neaktivno	Ctrl + Alt + Esc

### Informacije o vmesniku

Uporabniški vmesniki skrbniške konzole, vzorčne iskalne aplikacije in prilagojevalnika iskalnih aplikacij so na spletnem brskalniku temelječi vmesniki, ki si jih lahko ogledate v brskalnikih Microsoft Internet Explorer ali Mozilla FireFox. Za seznam bližnjic na tipkovnici in pripomočke za osebe s posebnimi potrebami za vaš brskalnik preglejte zaslonsko pomoč za Internet Explorer ali FireFox.

### Sorodne informacije o pripomočkih za osebe s posebnimi potrebami

S programom Adobe Acrobat Reader si lahko ogledate publikacije za OmniFind Enterprise Edition v obliki zapisa PDF (Adobe Portable Document Format). PDF-je najdete na CD-ju, ki

je priložen izdelku, ali na naslovu <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27010938>.

## **IBM in pripomočki za osebe s posebnimi potrebami**

Podrobnejše informacije IBM-ovi obvezi do oseb s posebnimi potrebami najdete v centru IBM Human Ability and Accessibility Center.

---

## Pojmovnik izrazov iskanja v podjetju

Ta pojmovnik definira izraze, uporabljene v vmesnikih in dokumentaciji iskanja v podjetju.

### **seznam za nadzor dostopa**

(access control list) Na področju računalniške zaščite je to seznam, povezan z objektom, ki označuje vse subjekte z dostopom do objekta in njihove pravice za dostop.

### **upravna vloga**

(administrative role) Klasifikacija uporabnika, ki predpisuje dostop.

### **motor za analizo**

(analysis engine) Glejte motor za analizo besedila.

### **rezultati analize**

(analysis results) Informacije, ki jih ustvarijo razlagalniki. Rezultati analize so zapisani v podatkovno strukturo, imenovano struktura splošne analize. Rezultate analize, ki jih ustvarijo motorji za analizo besedila po meri (razlagalniki), lahko omogočite za iskanje z vključitvijo v indeks iskanja v podjetju.

### **razlaga**

(annotation) Informacije o obsegu besedila. Tako lahko na primer razlaga kaže, da predstavlja določen obseg besedila ime podjetja. V arhitekturi UIMA (Unstructured Information Management Architecture) je razlaga posebna vrsta funkcijske strukture.

### **razlagalnik**

(annotator) Komponenta programske opreme, ki izvaja specifične naloge lingvistične analize ter ustvarja in zapisuje razlage. Razlagalnik je logična komponenta analize v motorju za analizo.

### **Iskanje z logičnimi vrednostmi (Boolovo iskanje)**

(boolean search) Iskanje, v katerem enega ali več iskalnih izrazov združite s pomočjo operatorjev kot so AND, NOT in OR.

### **poviševalni razred**

(boost class) Objekt, v katerem so specifikacije, ki lahko vplivajo na relativni rang dokumenta v rezultatih iskanja.

### **poviševalna beseda**

(boost word) Beseda, ki lahko vpliva na relativno uvrstitev dokumenta v rezultatih iskanja. Med obdelavo poizvedbe je lahko pomembnost dokumenta, ki vsebuje poviševalno besedo, zvišana ali znižana glede na točkovanje, ki je vnaprej definirano za besedo.

### **drevo kategorij**

(category tree) Hierarhija kategorij.

### **digitalno potrdilo**

(certificate) Na področju računalniške zaščite je to digitalni dokument, ki poveže javni ključ z identiteto lastnika digitalnega potrdila, s čimer omogoči njegovo overitev. Digitalno potrdilo izda služba za izdajo digitalnih potrdil, ki ga tudi digitalno podpiše.

### **urad za digitalna potrdila**

(certificate authority) Overjena neodvisna organizacija ali podjetje, ki izdaja digitalna potrdila, na podlagi katerih se izdelajo digitalni podpisi in pari javnih in zasebnih ključev. Urad za potrdila jamči za identiteto posameznikov, ki jim dodeli unikatne certifikate.

**normalizacija znakov**

(character normalization) Postopek, v katerem so različne oblike znaka, kot so na primer velike črke ali diakritični znaki, vrnjeni v splošno obliko.

**klitika** (clitic) Beseda, ki sintaktično deluje ločeno, fonetično pa je povezana z drugo besedo. Klitika je lahko zapisana kot povezana ali ločena od besede, s katero je povezana. Splošni primeri klitik v angleščini vključujejo zadnji del okrajšave (*wouldn't* ali *you're*).

**zbirka** (collection) Niz izvorov podatkov in možnosti za pajkanje, razčlenjevanje, indeksiranje in preiskovanje teh izvorov podatkov.

**struktura splošne analize**

(common analysis structure (CAS)) Struktura, ki hrani vsebino in metapodatke dokumenta ter vse rezultate analize, ki jih ustvari motor za analizo besedila. Struktura splošne analize obravnava vso izmenjavo podatkov med analizo dokumenta.

**potrošnik strukture splošne analize**

(CAS consumer) Potrošnik, ki dokončno obdela rezultate analize, shranjene v strukturi splošne analize. Potrošnik na primer indeksira vsebino strukture splošne analize v iskalniku ali poseli relacijsko bazo podatkov s specifičnimi rezultati analize.

**splošna komunikacijska plast**

(common communication layer) Komunikacijska infrastruktura, ki združuje različne komponente (krmilnik, razčlenjevalnik, pajek, indeksni strežnik) izdelka OmniFind Enterprise Edition.

**izdelava pojmovnega izvlečka**

(concept extraction) Funkcija za analizo besedila, ki določa pomembne postavke besednjaka (kot so osebe, mesta ali izdelki) v besedilnih dokumentih in ustvari seznam teh postavk. Glejte tudi izdelava tematskega izvlečka.

**prostor za pajkanje**

(crawl space) Niz izvorov, ki ustrezajo podanim vzorcem (kot so na primer URL-ji (Uniform Resource Locators), imena baz podatkov, poti datotečnega sistema, imena domen in naslovi IP), iz katerih bere pajek, da pridobi postavke za indeksiranje.

**pajek** (crawler) Programska oprema, ki pridobiva dokumente iz izvorov podatkov in zbira informacije, ki jih je mogoče uporabiti za izdelavo iskalnih indeksov.

**poverilnica**

(credential) Podrobne informacije, pridobljene med overjanjem, ki opisujejo uporabnika, morebitne povezave s skupinami in druge z varnostjo povezane attribute identitete. Poverilnice lahko uporabite za izvajanje več storitev, kot so pooblašcanje, beleženje in delegiranje. Informacije, ki jih uporabnik vnese ob prijavi (ID uporabnika in geslo), so na primer poverilnice, ki uporabniku omogočajo dostop do računa.

**motor za analizo besedila po meri**

(custom text analysis engine) Motor za analizo besedila, izdelan s pomočjo kompleta orodij za razvijanje programske opreme (SDK) arhitekture UIMA (Unstructured Information Management Architecture - UIMA), ki ga lahko dodate nizu standardnih motorjev za analizo besedila iskanja v podjetju (imenovanih tudi osnovni razlagalniki iskanja v podjetju). Glejte tudi motor za analizo besedila.

**izvor podatkov**

(data source) Katerikoli repozitorij podatkov, iz katerega lahko pridobivate dokumente, kot so spletne, relacijske in nerelacijske baze podatkov ter sistemi za upravljanje vsebine.

**tip izvora podatkov**

(data source type) Združitev izvorov podatkov glede na protokol, uporabljen za dostopanje do podatkov.

**podatkovna shramba**

(data store) Podatkovna struktura, v kateri so shranjeni dokumenti v njihovi razčlenjeni obliki.

**gradnja delta indeksa**

(delta index build) V sistemu za iskanje v podjetju je to postopek dodajanja novih informacij v obstoječi indeks. Primerjajte z izdelavo glavnega indeksa.

**odstraniti iz čakalne vrste**

(dequeue) Odstraniti postavke iz čakalne vrste.

**diakritični znak**

(diacritic) Znak, ki označuje spremembo v fonetični vrednosti znaka ali kombinaciji znakov.

**odkritelj**

(discoverer) Funkcija pajka, ki določa, kateri izvori podatkov so na voljo pajku za pridobivanje informacij.

**razločevalno ime**

(distinguished name) Ime, ki unikatno določa vnos v imeniku. Razločevalno ime je sestavljeno iz parov atribut:vrednost, ločenih z vejico. Uporabljeni so lahko tudi nizi parov ime-vrednost (kot je CN=ime osebe in C=država ali regija), ki unikatno določajo entiteto v digitalnem certifikatu.

**objektni model dokumentov**

(Document Object Model) Sistem, v katerem je strukturiran dokument, kot je datoteka XML, prikazan kot drevo objektov, do katerih je mogoče programsko dostopati in jih ažurirati.

**kabinet Domino Document Manager**

(cabinet Domino Document Manager) Baza podatkov Domino Document Manager, uporabljena za razvrstitev dokumentov. V kabinetih so shranjene baze podatkov Domino.

**knjižnica Domino Document Manager**

(Domino Document Manager library) Baza podatkov Domino Document Manager, ki je vstopna točka za Domino Document Manager.

**Domino Internet Inter-ORB Protocol (DIIOP)**

(Domino Internet Inter-ORB Protocol (DIIOP)) Strežniška naloga, ki se izvaja na strežniku in deluje s posrednikom zahtev po objektih Domino, da omogoča komunikacije med programčki Java, ki jih izdelate z razredi Notes Java in strežnikom Domino. Uporabniki pregledovalnika in strežniki Domino uporabljajo DIIOP za komuniciranje in za izmenjavo podatkov objektov.

**dinamično rangiranje**

(dynamic ranking) Vrsta rangiranja, v kateri so izrazi poizvedbe analizirani v skladu z dokumenti, v katerih se izvaja iskanje, da se določi rang rezultatov. Glejte tudi rezultati, ki temeljijo na besedilu. Primerjajte s statičnim rangiranjem.

**dinamično povzemanje**

(dynamic summarization) Vrsta povzemanja, v kateri so izrazi iskanja označeni, rezultati iskanja pa vsebujejo fraze, ki najboljše predstavljajo pojme dokumenta, ki jih išče uporabnik. Primerjajte s statičnim povzemanjem.

**uvrstiti v čakalno vrsto**

(enqueue) Postaviti sporočilo ali postavko v čakalno vrsto.

**skrbnik iskanja v podjetju**

(enterprise search administrator) Upravna vloga, ki uporabniku omogoča upravljanje celotnega sistema za iskanje v podjetju.

**osnovni razlagalniki iskanja v podjetju**

(enterprise search base annotators) Niz standardnih motorjev za analizo besedila, uporabljenih pri iskanju v podjetje za privzeto obdelavo analize dokumentov.

**ubežni znak**

(escape character) Znak, ki zadrži ali izbere poseben pomen za enega ali več znakov, ki sledijo.

**zunanji izvor podatkov**

(external data source) Izvor podatkov za združevanje, ki ga OmniFind Enterprise Edition ne pajka, razčleni ali indeksira. Iskanja v zunanjih izvorihih podatkov so poverjena poizvedovalnim aplikacijskim programerskim vmesnikom teh izvorov podatkov.

**pot do funkcije**

(feature path) Pot, uporabljena za dostopanje do vrednosti funkcije v funkcijski strukturi UIMA (Unstructured Information Management Architecture).

**funkcijska struktura**

(feature structure) Podrejena podatkovna struktura, ki predstavlja rezultat analize besedila. Funkcijska struktura je struktura atribut-vrednost. Vsaka funkcijska struktura je določenega tipa, vsak tip pa ime podan niz veljavnih funkcij ali atributov, podobno kot razred Java.

**zvezno iskanje**

(federated search) Zmožnost iskanja, ki omogoča iskanja v različnih iskalnih storitvah in vrne združen seznam rezultatov iskanja.

**federacija**

(federation) Postopek kombiniranja sistemov poimenovanja tako, da lahko združen sistem obdela sestavljena imena, ki se raztezajo čez sisteme poimenovanja.

**polje**

(field) Področje, v katerega se vnese določena kategorija podatkov ali krmilnih informacij.

**iskanje po poljih**

(fielded search) Poizvedba, ki je omejena na določeno polje.

**prosto besedilo**

(free-form text) Nestrukturirano besedilo, sestavljeno iz besed ali stavkov.

**iskanje po prostem besedilu**

(free text search) Iskanje, v katerem je iskalni izraz izražen kot besedilo proste oblike.

**indeks celotnega besedila**

(full text index) Podatkovna struktura, ki se sklicuje na podatkovne postavke, da bi iskanju olajšala iskanje dokumentov z iskanimi izrazi.

**mehko iskanje**

(fuzzy search) Iskanje, ki vrne besede, katerih črkovanje je podobno črkovanju iskalnega izraza.

**mešano iskanje**

(hybrid search) Združeno boolovo iskanje in prosto besedilno iskanje.

**upravljanje identitet**

(identity management) Niz API-jev za iskanje v podjetju, ki nadzirajo dostop do

zaščitenih podatkov in uporabnikom omogočajo iskanje po zbirki, ne da bi ti pri tem morali podati uporabniški ID in geslo za vsak repozitorij v zbirki.

**indeks** (index) Glejte indeks celotnega besedila.

**indeksna čakalna vrsta**

(index queue) Seznam zahtev za obdelavo izdelave glavnega in delta indeksa.

**izdelava izvlečka informacij**

(information extraction) Vrsta izdelave izvlečka pojmov, ki v besedilnih dokumentih samodejno prepozna pomembne postavke besednjaka, kot so imena, besede in izrazi.

**naslov IP**

(IP address) Unikaten naslov naprave ali logične enote v omrežju, ki uporablja standard IP.

**Java Database Connectivity (JDBC)**

(Java Database Connectivity (JDBC)) Industrijski standard za povezljivost, neodvisen od baze podatkov, med platformo Java in obsežnim razponom baz podatkov. Vmesnik JDBC nudi API klicne ravni za dostop do baze podatkov, ki temelji na SQL.

**JavaScript**

(JavaScript) Spletni skriptni jezik, uporabljen v pregledovalnikih in na spletnih strežnikih.

**JavaServer Pages (JSP)**

(JavaServer Pages (JSP)) Skriptna tehnologija strežnika, ki omogoča dinamično vdelovanje kode Java znotraj spletnih stani (datotek HTML) in izvedbo, ko je stran postrežena, z namenom, da je odjemalcu vrnjena dinamična vsebina.

**navidezni računalnik Java (JVM)**

(Java virtual machine (JVM)) Izvedba programske opreme procesorja, ki izvaja prevedeno kodo Java (programčke in aplikacije).

**Katakana**

(Katakana) Nabor znakov, sestavljen iz simbolov, uporabljenih v eni od dveh splošnih japonskih fonetičnih abeced, ki se v glavnem uporablja za fonetično zapisovanje tujih besed.

**Datoteka baze podatkov ključev**

(key database file) Glejte geslo obroč ključev. obroč ključev.

**obroč ključev**

(key ring) Na področju računalniške zaščite je to datoteka, ki vsebuje javne ključe, zasebne ključe, overjene korene in certifikate. Glejte tudi pojem datoteka shrambe ključev.

**datoteka shrambe ključev**

(keystore file) Obroč ključev, ki vsebuje tako javne ključe, shranjene kot certifikate podpisnikov, kot tudi zasebne ključe, shranjene v osebnih certifikatih.

**jezikovna identifikacija**

(language identification) Pri iskanju v podjetju je to funkcija, ki določa jezik dokumenta.

**lema**

(lemma) Osnovna oblika besede. Leme so še posebej pomembne v jezikih z veliko skloni, kot je na primer češčina.

**lematizacija**

(lemmatization) Postopek ugotavljanja korenske oblike in različnih slovničnih oblik

določene besede. Iskanje besede mouse poišče tudi dokumente, ki vsebujejo množinsko obliko mice, iskanje besede go pa najde tudi dokumente, ki vsebujejo oblike going, gone ali went.

**leksična podobnost**

(lexical affinity) Povezava iskalnih besed v dokumentu, katerih pomen je podoben. Leksična podobnost se uporablja za izračun ustreznosti rezultata.

**knjižnica**

(library) Sistemski objekt, ki služi kot imenik za druge objekte. Glejte tudi knjižnica Domino Document Manager.

**vezava** (ligature) Dva ali več povezanih znakov, ki so predstavljeni kot en znak. Ff in ffi so na primer znaki, ki jih je mogoče predstaviti kot vezavo.

**LDAP (Lightweight Directory Access Protocol)**

(LDAP) Odprt protokol, ki s pomočjo TCP/IP nudi dostop do imenikov, ki podpirajo model X.500 in ki zahteva manj virov kot kompleksnejša različica X.500 Directory Access Protocol (DAP). S protokolom LDAP je mogoče na primer v intranetnem ali internetnem imeniku poiskati osebe, organizacije in druge vire.

**lingvistično iskanje**

(linguistic search) Vrsta iskanja, ki preiskuje, pridobiva in indeksira dokument z izrazi, ki so zmanjšani v njihovo osnovno obliko (tako da je na primer izraz *mobi* indeksiran kot *mobilnik*) ali razširjeni z njihovo osnovno obliko (kot pri sestavljenih besedah).

**analiza povezav**

(link analysis) Metoda, ki temelji na analizi hiperpovezav med dokumenti, in se uporablja za določitev, katere strani v zbirki so pomembne za uporabnike.

**lokalni združevalnik**

(local federator) V aplikaciji za iskanje v podjetju je to objekt odjemalca, ki ga izdelajo API-ji za iskanje in indeksiranje, s pomočjo katerih lahko uporabniki preiščejo skupino heterogenih zbirk in pridobijo enoten niz rezultatov iskanja.

**prostor Lotus QuickPlace**

(Lotus QuickPlace place) Spletno shajališče, ki ga nudi Lotus QuickPlace, ki omogoča geografsko razpršenim udeležencem sodelovanje pri projektih in komuniciranje v priključnem načinu v strukturiranem in zaščitenem delovnem mestu.

**soba Lotus QuickPlace**

(Lotus QuickPlace room) Porazdeljeno področje prostora Lotus QuickPlace, omejeno na pooblaščen uporabnike s skupnim interesom in potrebo za skupno delo.

**izdelava glavnega indeksa**

Pri iskanju v podjetju je to postopek izgradnje celotnega indeksa. Primerjajte z izdelavo delta indeksa.

**maskirni znak**

(masking character) Znak, ki predstavlja neobvezne znake na začetku, na sredini ali na koncu iskalnega izraza. Maskirni znaki se običajno uporabljajo za iskanje različic izraza v indeksu. Glejte tudi univerzalni znak.

**tip MIME**

(MIME type) Internetni standard za določitev tipa objekta, ki se prenaša prek interneta.

**nadzornik**

(monitor) Uporabnik iskanja v podjetju, ki ima pooblastilo za nadzorovanje procesov na ravni zbirke.



**znak za novo vrstico**

(newline character) Krmilni znak, ki povzroči pomik položaja tiskanja ali prikaza za eno vrstico navzdol.

**segmentacija n-gram**

(n-gram segmentation) Način analize, ki upošteva prekrivajoča se zaporedja podanega števila znakov kot ene besede, namesto da bi za ločitev besed uporabljal presledke kot v segmentaciji presledkov, ki temelji na Unicode.

**smernica 'ne sledi'**

Smernica na spletni strani, ki robotom (kot je spletni pajek) ukazuje, naj ne sledijo povezavam, ki jih najdejo na tej strani.

**smernica 'ne indeksiraj'**

Smernica na spletni strani, ki robotom (kot je spletni pajek) ukazuje, naj ne vključijo vsebine te strani v indeks.

**klic oddaljeni proceduri Notes**

(Notes remote procedure call (NRPC)) Komunikacijski mehanizem Lotus Notes, uporabljen za vse komunikacije Notes z Notes.

**operater**

(operator) Uporabnik iskanja v podjetju, ki ima pooblastilo za nadzorovanje, zagon in zaustavitev procesov na ravni zbirke.

**parametrično iskanje**

(parametric search) Vrsta iskanja, ki išče objekte, ki vsebujejo znotraj podanega območja številsko vrednost ali atribut, kot so datumi, cela števila ali drugi številski tipi podatkov.

**razčlenjevalnik**

(parser) Program, ki interpretira dokumente, dodane v podatkovno shrambo iskanja v podjetju. Razčlenjevalnik povzame informacije iz dokumentov in jih pripravi za indeksiranje, preiskovanje in pridobivanje.

**gonilnik razčlenjevalnika**

Pri iskanju v podjetju je to storitev, ki storitev razčlenjevalnika oskrbuje z dokumenti. Za vsako zbirko obstaja en gonilnik razčlenjevalnika. Storitev gonilnika razčlenjevalnika zbirke ustreza razčlenjevalniku zbirke na konzoli za upravljanje iskanja v podjetju.

**storitev razčlenjevalnika**

(parser service) Storitev iskanja v podjetju, ki obravnava celotno obdelavo razčlenjevanja dokumentov in analizo besedila v zbirkah dokumentov. Ves čas se izvaja vsaj ena storitev razčlenjevalnika.

**prostor**

(place) Navidezna lokacija, vidna v portalu, kjer se srečujejo posamezniki in skupine z namenom sodelovanja. V portalu ima vsak uporabnik osebno prostor za zasebno delo, posamezniki in skupine pa imajo dostop do številnih prostorov v skupni rabi, ki so lahko javni ali omejeni. Glejte tudi prostor Lotus QuickPlace.

**rangiranje glede na priljubljenost**

(popular ranking) Tip rangiranja, ki poviša obstoječe rangiranje dokumenta glede na njegovo priljubljenost.

**Portal Document Manager (PDM)**

(Portal Document Manager (PDM)) Uporabnikom omogoča osrednji repozitorij dokumentov za timsko sodelovanje. Skrbniki lahko učinkovito upravljajo svoje dokumente in nadzorujejo načine, na katere uporabniki delajo z informacijami.

**arhiv motorja za obdelavo**

(processing engine archive) Arhivska zip datoteka .pear, ki vsebuje motor za analizo UIMA (Unstructured Information Management Architecture) in vse vire, potrebne za njegovo uporabo za analizo po meri pri iskanju v podjetju.

**iskanje glede na bližino**

(proximity search) Besedilno iskanje, ki vrne rezultat, kadar se dva ali več ustreznih izrazov pojavita blizu drug drugemu, na primer v istem stavku ali odstavku.

**strežnik proxy**

(proxy server) Strežnik, ki deluje kot posrednik za spletne zahteve HTTP, ki jih gosti aplikacija ali spletni strežnik. Strežnik proxy deluje kot nadomestek za strežnike vsebine za podjetje.

**hitra povezava**

(quick link) Povezava med URI-jem (Uniform Resource Identifier) in ključnimi besedami ali frazami.

**rangiranje**

(ranking) Dodeljevanje celoštevilске vrednosti vsakemu od dokumentov, vrnjenih v rezultatih iskanja. Vrstni red dokumentov v rezultatih iskanja temelji na ustreznosti s poizvedbo. Višji rang kaže na bližje ujemanje. Glejte tudi dinamično rangiranje in statično rangiranje.

**shramba surovih podatkov**

(raw data store) Podatkovna struktura, v katero so shranjeni dokumenti, v katerih je bilo izvedeno pajkanje, preden so poslani razčlenjevalniku. Pajki pišejo v shrambo surovih podatkov, razčlenjevalnik pa bere iz nje. Ko so dokumenti razčlenjeni, so odstranjeni iz shrambe surovih podatkov. Izraza ne mešajte s podatkovno shrambo.

**razlagalnik običajnih izrazov**

(regular expression annotator) Programska komponenta, ki v besedilnem dokumentu odkrije entitete ali enote informacij, na primer številke izdelkov, na podlagi regularnih izrazov, ki natanko opisujejo vzorce za iskanje v besedilu dokumenta. Če se eden od običajnih izrazov ujema z deli besedila dokumenta, izdelava razlagalnik običajnih izrazov ustrezne razlage, ki obsegajo ujemanje ali njegov del. Ti razloženi izrazi so nato s pomočjo indeksne datoteke preslikav shranjeni v indeks iskanja v podjetju ali s pomočjo datoteke preslikav baze podatkov v bazo podatkov z zmožnostjo za JDBC.

**oddaljeni združevalec**

(remote federator) Združevalec strežnika, ki združuje niz objektov z možnostjo iskanja.

**Robots Exclusion Protocol**

(Robots Exclusion Protocol) Protokol, ki skrbnikom spletnega mesta omogoča, da za obiskujoče robote nakažejo, katerih delov spletnega mesta naj ne obiščejo.

**soba**

(room) Program, ki uporabnikom omogoča, da izdelajo dokumente, ki jih lahko berejo drugi uporabniki, odgovarjajo na opombe drugih uporabnikov ali pregledajo status in rok projektov. Uporabniki lahko tudi klepetajo z drugimi uporabniki, ki so v isti sobi. Glejte tudi soba Lotus QuickPlace.

**kategorija, ki temelji na pravilih**

(rule-based category) Kategorije, izdelane s pravili, ki podajajo, kateri dokumenti so povezani s katerimi kategorijami. Tako lahko na primer definirate pravila za povezovanje dokumentov, ki vsebujejo ali izključujejo določene besede ali ki se ujemajo z vzorcem URI (Uniform Resource Identifier), z določenimi kategorijami.

**iskalna aplikacija**

(search application) Pri iskanju v podjetju je to program, ki obdeluje poizvedbe, preiskuje indeks, vrne rezultate iskanja in priključuje izvorne dokumente.

**iskalni predpomnilnik**

(search cache) Vmesni pomnilnik, v katerem so shranjeni podatki in rezultati prejšnjih iskalnih zahtev.

**iskalnik**

(search engine) Program, ki sprejme iskalno zahtevo in uporabniku vrne seznam dokumentov.

**indeksne datoteke iskanja**

(search index files) Niz datotek, v katerih je shranjen indeks v iskalniku.

**rezultati iskanja**

(search results) Seznam dokumentov, ki ustrezajo iskalni zahtevi.

**plast zaščitnih vtičnic (SSL)**

(Secure Sockets Layer (SSL)) Zaščitni protokol, ki omogoča zasebnost komunikacij. Prek SSL lahko odjemalske/strežniške aplikacije komunicirajo na način, ki je zasnovan tako, da onemogoča prisluškovanje, sabotaže in ponarejanje sporočil.

**žeton zaščite**

(security token) Informacije o identiteti in zaščiti, uporabljene za dodelitev pooblastila za dostop do dokumentov v zbirki. Različne vrste izvorov podatkov podpirajo različne žetone zaščite. Primeri vključujejo uporabniške vloge, ID-je uporabnikov, ID-je skupin in druge informacije, ki jih je mogoče uporabiti za nadzorovanje dostopa do vsebine.

**stran seznama semen**

(seed list page) V portalu WebSphere je to stran XML, ki vsebuje povezave do strani, ki so na voljo v portalu. Pajki uporabljajo seznam semen za določitev dokumentov za pajkanje. Stran seznama semen vsebuje tudi metapodatke, ki so s pajkanimi dokumenti shranjeni v indeksu iskanja v podjetju.

**zagon URL-ja (Uniform Resource Locator)**

(start Uniform Resource Locator (URL)) Začetna točka za pajkanje.

**segmentacija**

(segmentation) Razdelitev besedila v posebne leksične enote. Obdelava, ki ne temelji na slovarju, vključuje segmentacijo presledkov in segmentacijo n-gram, podpora, temelječa na slovarju, pa vključuje segmentacijo besed, stavkov in odstavkov ter lematizacijo.

**semantično iskanje**

(semantic search) Tip iskanja po ključnih besedah, ki vključuje lingvistično in kontekstualno analizo. Glejte tudi pojem besedilna analiza.

**strežniški programček**

(servlet) Program Java, ki se izvaja na spletnem strežniku in razširja funkcionalnost strežnika z izdelavo dinamične vsebine kot odziv na zahteve spletnega odjemalca. Strežniški programčki se običajno uporabljajo za povezovanje baz podatkov s spletom.

**pokrivatek**

(shingle) Niz zaporednih žetonov (besed), vzetih iz stavka. Pokrivki (ali trigrami), sestavljeni iz treh besed iz fraze "This is a very short sentence", so:

This is a  
is a very  
a very short  
very short sentence

Pokrivke je mogoče uporabljati v statistični lingvistiki. Če imata na primer dve različni besedili veliko skupnih pokrивkov, sta najbrž nekako povezani.

#### **stran programskih napak**

(soft error page) Tip spletne strani, na kateri so informacije o tem, zakaj ni bilo mogoče vrniti zahtevane spletne strani. Namesto da strežnik HTTP vrne preprosto statusno kodo, lahko vrne stran, na kateri je statusna koda podrobneje razložena.

#### **statično rangiranje**

(static ranking) Vrsta rangiranja, v kateri faktorji o dokumentih, ki so rangirani, kot so datum, število povezav, ki kažejo na dokumente in tako dalje, povečajo rang. Primerjajte z dinamičnim rangiranjem.

#### **statično povzemanje**

(static summarization) Vrsta povzemanja, v kateri vsebujejo rezultati iskanja podan, shranjen povzetek iz dokumenta. Primerjajte z dinamičnim povzemanjem.

#### **korenjenje**

(stemming) Glejte korenjenje besed.

#### **izključitvena beseda**

(stop word) Pogosto uporabljena beseda, kot je v angleščini na primer *the*, *an* ali *and*, ki jo iskalna aplikacija prezre.

#### **odstranitev izključitvene besede**

(stop word removal) Postopek odstranjevanja zaustavitvenih besed iz poizvedbe, z namenom, da bi bile splošne besede zanemarjene in bi bili vrnjeni ustreznejši rezultati.

#### **povzemanje**

(summarization) Postopek vključevanja neodvečnih stavkov v rezultate iskanja, s katerimi je na kratko opisana vsebina dokumenta. Glejte tudi dinamično povzemanje in statično povzemanje.

#### **slovar sopomenk**

(synonym dictionary) Slovar, ki uporabnikom pri preiskovanju zbirke omogoča iskanje sopomenk za izraze poizvedbe.

#### **taksonomija**

(taxonomy) Klasifikacija objektov v skupine na osnovi podobnosti. Pri iskanju v podjetju taksonomija organizira podatke v kategorije in podkategorije. Glejte tudi drevo kategorij.

#### **analiza besedila**

(text analysis) Postopek povzetja semantike in drugih informacij iz besedila za izboljšanje berljivosti podatkov v zbirki. Glejte tudi pojem semantično iskanje.

#### **motor za analizo besedila**

(text analysis engine) Komponenta programske opreme, ki je odgovorna za iskanje in predstavitev konteksta in semantične vsebine v besedilu.

#### **točkovanje, temelječe na besedilu**

(text-based scoring) Postopek dodelitve celoštevilске vrednosti dokumentu, ki označuje ustreznost dokumenta glede na izraze poizvedbe. Višja celoštevilška vrednost označuje bližje ujemanje s poizvedbo. Glejte tudi dinamično rangiranje.

**segmentacija besedila**

Glejte tudi segmentacija.

**izdelava tematskega izvlečka**

(theme extraction) Vrsta izdelava izvlečka pojmov, ki samodejno prepozna pomembne izraze besednjaka v besedilnih dokumentih, da povzame temo ali predmet dokumenta. Glejte tudi izdelava pojmovnega izvlečka.

**žeton** (token) Osnovne besedilne enote, ki jih indeksira iskanje v podjetju. Žetoni so lahko besede v jeziku ali druge enote besedila, ki so primerne za indeksiranje.

**izdelava žetonov**

(tokenization) Postopek razčlenjevanja vhodnih podatkov v žetone.

**izdelovalec žetonov**

(tokenizer) Program za segmentacijo besedila, ki pregleda besedilo in določi, ali in kdaj je mogoče niz znakov prepoznati kot žeton.

**zaključni znak**

(trailing character) Znak na zadnjem mestu v besedi.

**tipski sistem**

(type system) Tipski sistem definira tipe objektov (funkcijske strukture), ki jih v dokumentu lahko odkrije motor za analizo besedila. Tipski sistem definira vse možne funkcijske strukture v skladu s tipi in funkcijami. V tipskem sistemu lahko definirate poljubno število različnih tipov. Tipski sistem je specifičen za domeno in aplikacijo.

**segmentacija presledkov, temelječa na Unicode**

(Unicode-based white space segmentation) Način izdelave žetonov, ki uporablja za razločevanje med žetoni in ločevalnimi znaki lastnosti znakov Unicode.

**Uniform Resource Identifier (URI)**

(Uniform Resource Identifier (URI)) Strnjen niz znakov, ki določa abstrakten ali fizičen vir.

**Uniform Resource Locator (URL)**

(URL) Unikatni naslov vira informacij, do katerega je mogoče priti v omrežju, kot je internet. URL vsebuje okrajšano ime protokola, uporabljenega za dostop do vira informacij, in informacije, s pomočjo katerih protokol poišče vir informacij.

**Unstructured Information Management Architecture (UIMA)**

(Unstructured Information Management Architecture (UIMA)) IBM-ova arhitektura, ki definira sestav za izvajanje sistemov za analiziranje nestrukturiranih podatkov.

**uporabniški agent**

(user agent) Aplikacija, ki preišče splet in na mestih, ki jih obiše, pusti informacije o sebi. Pri iskanju v podjetju ima funkcijo uporabniškega posrednika spletni pajek.

**spletni pajek**

(Web crawler) Tip pajka, ki raziskuje splet, tako da priključuje spletni dokument in obiše povezave, vsebovane v tem dokumentu.

**utežno iskanje izrazov**

(weighted term search) Poizvedba, v kateri je določenim izrazom dodeljena večja pomembnost.

**univerzalni znak**

(wildcard character) Znak, ki predstavlja neobvezne znake na začetku, na sredini ali na koncu iskalnega izraza.

**korenjenje besed**

(word stemming) Postopek lingvistične normalizacije, v katerem so različne oblike besede vrnjene v običajno obliko. Besede, kot so na primer *mobilniki*, *mobilen* in *mobiliziran* so vrnjene v obliko *mobilnik*.

**XML Path Language (XPath)**

(XPath) Jezik, zasnovan tako, da unikatno označi ali naslovi dele izvornih podatkov XML za uporabo s tehnologijami, povezanimi z XML, kot so na primer razčlenjevalniki XSLT, XQuery in XML. XPath je standard konzorcija World Wide Web Consortium.

---

## Opombe in blagovne znamke

---

### Opombe

Pričujoče informacije so bile razvite za izdelke in storitve, ki so na voljo v ZDA.

IBM izdelkov, storitev ali komponent, ki jih opisuje ta dokument, lahko ne bo nudil v drugih državah. Informacije o tem, kateri izdelki in storitve so trenutno na voljo na vašem področju, lahko dobite pri lokalnem IBM-ovem tržnem predstavniku. Nobena referenca na IBM-ov izdelek, program ali storitev na noben način na pomeni, da lahko uporabite samo ta IBM-ov izdelek, program ali storitev. Namesto njih lahko uporabite katerikoli funkcionalno enakovreden izdelek, program ali storitev, ki ne krši IBM-ovih pravic do intelektualne lastnine. Toda odgovornost uporabnika je, da oceni in preveri delovanje vseh ne-IBM-ovih izdelkov, programov ali storitev.

IBM ima lahko patente ali za patent priglašene aplikacije, ki vključujejo predmet, opisan v tem dokumentu. Posedovanje tega dokumenta vam ne daje nobene licence za te patente. Poizvedbe o licencah lahko v pisni obliki pošljete na naslednji naslov:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive Armonk, NY  
10504-1785  
U.S.A.

Vprašanja v zvezi z licencami za DBCS naslovite na IBM-ov oddelek za intelektualno lastnino v vaši državi ali pošljete poizvedbe v pisni obliki na naslov:

IBM World Trade Asia Corporation Licensing  
2-31 Roppongi 3-chome,  
Minato-ku  
Tokyo 106-0032, Japan

Naslednji odstavek ne velja za Veliko Britanijo ali katero koli drugo državo, v kateri takšni predpisi niso v skladu z zakonom: INTERNATIONAL BUSINESS MACHINES CORPORATION OBJAVLJA TO PUBLIKACIJO "TAKŠNO, KOT JE", BREZ KAKRŠNEGAKOLI JAMSTVA, BODISI IZRECNEGA ALI ZAKONSKEGA, VKLJUČUJOČ, VENDAR NE OMEJENO NA, ZAKONSKA JAMSTVA ZA NEKRŠITEV, PRODAJNOST ALI PRIMERNOST ZA DOLOČEN NAMEN. Nekatere države v določenih transakcijah ne dopuščajo zavrnitve izrecnih ali vključenih jamstev, zato ta stavek morda ne velja za vas.

Te informacije lahko vsebujejo tehnične netočnosti ali tipografske napake. Informacije v tem dokumentu občasno spremenimo. Te spremembe bomo vključili v nove izdaje publikacije. IBM lahko kadarkoli in brez obvestila izboljša in/ali spremeni izdelke in/ali programe, opisane v tej publikaciji.

Vse reference v teh informacijah na spletna mesta, ki niso IBM-ova, smo podali zgolj zaradi priročnosti in na noben način ne pomenijo promoviranja teh spletnih mest. Gradivo na teh spletnih straneh ni del gradiva za ta IBM-ov izdelek in te spletne strani uporabljate na lastno odgovornost.

IBM lahko katerekoli informacije, ki nam jih pošljete, uporablja ali razpečuje na kakršenkoli način, ki se mu zdi primeren, pri čemer nima do vas nobene obveznosti.

Imetniki licenc za ta program, ki potrebujejo informacije, da bi omogočili: (i) izmenjavo informacij med neodvisno izdelanimi programi in drugimi programi (vključno s tem) in (ii) medsebojno uporabo informacij, ki so bile izmenjane, naj pošljejo vprašanja na naslednji naslov:

IBM Corporation  
J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003  
U.S.A.

Takšne informacije so lahko na voljo pod določenimi pogoji in določbami, ki v določenih primerih vključujejo tudi plačilo.

Licenčni program, opisan v tem dokumentu, in ves licenčni material, ki je zanj na voljo, IBM nudi pod pogoji IBM-ove Pogodbe za stranke, IBM-ove Mednarodne licenčne pogodbe ali katerekoli enakovredne pogodbe med nami.

Vse podatke o zmogljivosti, vsebovane v tem dokumentu, smo določili v nadzorovanem okolju. Zato se lahko rezultati, dobljeni v drugih operacijskih okoljih, bistveno razlikujejo. Nekatere meritve so bile opravljene v sistemih na razvojni stopnji in zato ne dajemo nobenega jamstva, da bodo te meritve enake tudi v splošno razpoložljivih sistemih. Poleg tega smo nekatere meritve ocenili z ekstrapolacijo. Dejanski rezultati so lahko drugačni. Uporabniki tega dokumenta naj preverijo ustrezne podatke za svoje specifično okolje.

Informacije, ki se nanašajo na izdelke drugih proizvajalcev, smo pridobili pri njihovih dobaviteljih, v njihovih objavah ali v drugih javno razpoložljivih virih. IBM teh izdelkov ni testiral, zato ne more potrditi točnosti podatkov o njihovi zmogljivosti, združljivosti in drugih lastnosti, povezanih z njimi. Vprašanja v zvezi z zmogljivostjo izdelkov, ki niso IBM-ovi, naslovite na dobavitelje teh izdelkov.

Vse izjave, povezane z IBM-ovo bodočo usmeritvijo ali namenom, lahko brez obvestila spremenimo ali umaknemo, in predstavljajo zgolj cilje in namene.

Vse prikazane IBM-ove cene so predlagane IBM-ove veljavne prodajne cene, ki se lahko spremenijo brez predhodnega obvestila. Cene pri prodajalcih se lahko razlikujejo.

Te informacije so namenjene samo za načrtovanje. Informacije v tem dokumentu se bodo spremenile, še preden bodo opisani izdelki dejansko na voljo.

Te informacije vsebujejo primere podatkov in poročil, uporabljenih v dnevni poslovnih operacijah. Da bi bili čim bolj nazorni, smo v primere vključili imena posameznikov, podjetij, znamk in izdelkov. Vsa ta imena so izmišljena; vsaka podobnost z imeni in naslovi dejanskih poslovnih podjetij je zgolj naključna.

#### LICENCA ZA AVTORSKE PRAVICE:

Te informacije vsebujejo vzorčne uporabniške programe v izvornem jeziku, ki kažejo tehnike programiranja na različnih operacijskih platformah. Te vzorčne programe lahko kopirate, spreminjate in jih distribuirate v namene razvijanja, uporabe, trženja ali distribuiranja aplikacij, ki ustrezajo IBM-ovim vmesnikom za programiranje aplikacij, ne da bi vam bilo treba IBM-u za to plačati. Teh primerov nismo natančno preskusili v vseh pogojih, zato IBM ne more zagotoviti zanesljivosti, uporabnosti ali funkcionalnosti teh programov.



Vsaka kopija ali posamezni del teh vzorčnih programov oziroma dela, ki so izpeljana iz njih, morajo vsebovati naslednje obvestilo o avtorskih pravicah:

© (ime vašega podjetja) (leto). Deli kode so izpeljani iz vzorčnih programov IBM Corp.. © Copyright IBM Corp. \_vnesite leto ali leta\_. Vse pravice pridržane.

Deli tega izdelka so:

- Oracle® Outside In Content Access, Copyright © 1992, 2008, Oracle. Vse pravice pridržane.
- Licenčno gradivo IBM XSLT Processor - Lastnina IBM © Copyright IBM Corp., 1999-2008. Vse pravice so pridržane.

## Blagovne znamke

Informacije o blagovnih znamkah IBM najdete na naslovu <http://www.ibm.com/legal/copytrade.shtml>.

Naslednja imena so blagovne znamke ali registrirane blagovne znamke drugih podjetij:

Adobe, Acrobat, Portable Document Format (PDF), PostScript in vse blagovne znamke Adobe so blagovne znamke ali registrirane blagovne znamke Adobe Systems Incorporated v Združenih državah Amerike, drugih državah ali v obojih.

Intel, logotip Intel, Intel Inside, logotip Intel Inside, Intel Centrino, logotip Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium in Pentium so blagovne znamke ali registrirane blagovne znamke družbe Intel Corporation ali njenih podružnic v Združenih državah Amerike in drugih državah.

Java in vse blagovne znamke, ki temeljijo na Javi, so blagovne znamke Sun Microsystems Inc. v ZDA, drugih državah ali v obojih.

Linux je registrirana blagovna znamka Linusa Torvaldsa v Združenih državah Amerike, drugih državah ali v obojih.

Microsoft, Windows, Windows NT in logotip Windows so blagovne znamke družbe Microsoft Corporation v ZDA, drugih državah ali v obojih.

UNIX je registrirana blagovna znamka The Open Group v Združenih državah Amerike in drugih državah.

Druga imena podjetij, izdelkov in storitev so lahko blagovne ali storitvene znamke drugih.



# Stvarno kazalo

## A

aktivna spletna mesta, nadziranje 276, 278  
analiza besedila  
  motorji za analizo besedila 129, 130  
  preslikovanje elementov XML 131  
  struktura splošne analize 132, 133  
analiza besedila po meri  
  motorji za analizo besedila 129, 130  
  opis 128  
  preslikava rezultatov analize v  
    indeks 132  
  preslikava rezultatov analize v relacijsko  
    bazo podatkov 133  
  preslikava strukture splošne analize v  
    indeks 132  
  preslikava strukture splošne analize v  
    relacijsko bazo podatkov 133  
  preslikovanje elementov XML 131  
analiza sidrnega besedila  
  globalna analiza 247  
  indeksiranje dokumentov 247  
  opis 233  
  zaščita zbirke 247  
API za iskanje in indeksiranje 9, 201  
API-ji  
  iskanje in indeksiranje 9, 201  
  opis 9  
aplikacija ESAdmin  
  onemogočanje zaščite 262  
  prijava v 16  
aplikacija ESSearchApplication  
  datoteka config.properties 203, 219  
  onemogočanje zaščite 262  
  zagon 226  
aplikacija ESSearchServer  
  onemogočanje zaščite 262  
  ponovni zagon 219, 220  
aplikacije podjetja  
  aplikacija ESAdmin 262  
  aplikacija ESSearchApplication 262  
  aplikacija ESSearchServer 262  
arhivske datoteke  
  formati URI 108  
  pajkanje 107  
  podprti formati 107

## B

baze podatkov DB2  
  dostop kot zunanji vir 229  
  dostop s pajki DB2 44  
  dostopanje s pajki baze podatkov  
    JDBC 58  
Baze podatkov Oracle  
  dostop kot zunanji vir 229  
  dostop s pajki DB2 44  
  dostopanje s pajki baze podatkov  
    JDBC 58  
baze podatkov SQL Server  
  dostop s pajki DB2 44

baze podatkov SQL Server (*nadaljevanje*)  
  dostopanje s pajki baze podatkov  
    JDBC 58  
bos.iocp.rte module 73  
brisanje  
  indeksi iz čakalne vrste 285  
  pajki 38  
  zbirke 33

## C

Center za iskanje, WebSphere Portal  
  opis 311  
  različica 6, integracija v sistem iskanja v  
    podjetju 320

## Č

čarovnik za selitev  
  datoteka dnevnika 334  
  na pravih temelječe taksonomije 331  
  opis 331  
  privzete nastavitve pajka 332  
  privzete nastavitve zbirke 332  
  zagon 331  
  zbirke 331  
Čarovnik za zbirke 29

## D

datotečne pripone  
  izključitev iz prostorov za spletno  
    pajkanje 84  
  ki jih podpirajo razčlenjevalniki  
    Stellent 144  
  podprte s strani razčlenjevalnikov  
    zbirke 139, 141  
  podprti razčlenjevalniki Stellent 146  
datoteka AdminLinkBarInclude.jsp 315  
datoteka banner.jspf 322  
datoteka banner\_searchControl.jspf 322  
datoteka ccl.properties 22  
datoteka CCLServer\_date.log 21  
datoteka config.properties 251, 252  
  kloniranje 221  
  opisi lastnosti 203  
  prilaganje 220  
  urejanje 219  
datoteka cookies.ini  
  format 93  
  konfiguriranje 94  
  opis 93  
datoteka crawl.rules 95  
datoteka crawler\_rdb\_plugin.xml 62  
datoteka Default.jsp 315, 322  
datoteka es.cfg 17, 18, 22, 24, 226, 251, 252  
datoteka es.search.provider.jar 318, 326  
datoteka es.security.jar 313, 318, 326  
datoteka es.wp5.install.jar 313  
datoteka es.wp6.install.jar 318, 326

datoteka esapi.jar 313, 318, 326  
datoteka ESPACServer.ear 313, 318, 326  
datoteka ESSearchAdapter.ear 313  
datoteka ESSearchPortlet.war 313, 318, 326  
datoteka  
  ESSearchRegistrationPortlet.war 313  
datoteka followindex.rules  
  konfiguriranje 96  
  opis 96  
datoteka global.rules 95  
datoteka nodes.ini 22, 226  
datoteka  
  parserdriver.collection.properties 142  
datoteka parserTypes.cfg 137, 139, 141  
datoteka runtime-generic.properties 185, 287  
datoteka SearchBarInclude.jsp 315  
datoteka siapi.jar 313  
datoteka stellent.properties 144  
datoteka stellenttypes.cfg 144  
datoteka stellentTypes.cfg 137  
datoteka vbr\_access\_services.jar 42, 43  
datoteka WpsMigratorLog.log 334  
datoteke dnevnika  
  čarovnik za selitev 334  
  dnevnik poizvedb 299  
  filtriranje 300  
  konfiguriranje strežnika SMTP 297  
  konfiguriranje velikosti 299  
  menjava 296  
  možnosti elektronske pošte 298  
  nadziranje 290, 300  
  največja velikost 296  
  opis 293  
  prikazovanje 300  
  prikazovanje odstranjenih  
    dokumentov 290  
  privzeto nahajališče 293  
  ravni resnosti 296  
  za sledenje dokumentov 289  
datoteke EAR  
  aplikacija ESAdmin 262  
  aplikacija ESSearchApplication 262  
  aplikacija ESSearchServer 262  
datoteke robots.txt  
  identifikacija uporabniških  
    posrednikov 81  
  skladnost spletnega pajka 82  
datoteke shrambe ključev 226  
datoteke tar  
  formati URI 108  
  pajkanje 107  
datoteke zip  
  formati URI 108  
  pajkanje 107  
datumi pajkanih dokumentov  
  konfiguriranje spletnih pajkov 97  
delta indeksi  
  odkrivanje sprememb 161  
  opis 6, 157  
  razporejanje 158  
  sočasna gradnja 160

- dinamično povzemanje 184
- dinamično rangiranje 189
- dokumentacija
  - HTML 365
  - iskanje 365
  - PDF 365
- dokumentacija HTML za iskanje v podjetju 365
- dokumentacija PDF za iskanje v podjetju 365
- dokumenti HTML
  - iskanje 126, 127
  - pravila zamenjave 142, 143
  - razčlenjevanje 142, 143
- dokumenti XML
  - iskanje 123
  - iskanje po izvornem XML 135
- domene Lotus Domino 256, 257
- domene Windows 260
- domenska pravila za spletne pajke 84
- drevesa kategorij
  - opis 120
  - selitev s portala WebSphere Portal 331

## E

- elektronska obvestila
  - konfiguriranje strežnika SMTP 297
  - za opozorila 298
  - za sporočila 298
- elementi Dublin Core 127
- elementi XML
  - iskanje 123, 131
  - preslikovanje v analizo splošne strukture 131
  - preslikovanje v iskalna polja 123
- enobajtno kodiranje 153
- ES\_INSTALL\_ROOT, opis 17, 18
- ES\_NODE\_ROOT, opis 17, 18

## F

- Fragmenti XML, iskanje po izvornem XML 135

## G

- geslo skrbnika
  - spreminjanje v enem strežniku 17
  - spreminjanje v več strežnikih 18
- geslo, skrbnik sistema iskanja v podjetju 17, 18
- glavni indeksi
  - odkrivanje sprememb 161
  - opis 6, 157
  - razporejanje 158
  - sočasna gradnja 160
- globalna analiza
  - analiza sidrnega besedila 233, 247
  - opis 6
  - zaznavanje dvojnikov dokumentov 167, 233, 237
- globalna zaščita za WebSphere
  - lastnosti iskalne aplikacije 219
  - onemogočanje 262
  - prilagojevalnik iskalnih aplikacij 220
- globalni prostor za spletno pajkanje 95

- globina poti URL-ja 84
- gonilniki JDBC
  - za pajke baze podatkov JDBC 58
  - za zunanje vire JDBC 229
- gradnja indeksov 157
- gradnje indeksov
  - odkrivanje sprememb 161
  - opis 157
  - paralelno 160
  - razporejanje 159
  - sočasno 160
  - status sistema 285
  - ukaz startIndexBuild 161
  - zagon 285
  - zaustavitev 162, 285
- gruče
  - WebSphere Portal 325
- gruče WebSphere Portal
  - integracijski skripti 326
  - odstranjevanje iskanja v podjetju 329
  - smernice za integracijo 325

## H

- hitre povezave
  - formati URI 108
  - iskanje 186
  - izdelovanje 186
  - opis 186
- hitrost pajkanja, nadziranje 278
- HTTPS, konfiguriranje iskalnih strežnikov 226

## I

- ID zbirke 33
- ID zbirke, pravila skladnje 30
- ID-ji aplikacije 238
- ID-ji sej, iskanje v podjetju 335
- indeksi
  - brisanje iz čakalne vrste 285
  - formati URI 108
  - nadziranje 285
  - območja 163, 171
  - odkrivanje sprememb 161
  - odstranjevanje URI-jev 163, 174
  - omogočanje razporeda 159, 285
  - onemogočanje razporeda 159, 285
  - opis 6, 157
  - opozorila o 294
  - paralelna gradnja 160
  - razporejanje 158
  - sidrno besedilo 247
  - skrčeni URI-ji 163, 172, 173
  - sočasna gradnja 160
  - spreminjanje razporeda 159
  - ukaz startIndexBuild 161
  - univerzalni znaki 163, 168, 170
  - vpliv univerzalnih znakov 169
- indeksna čakalna vrsta 285
- integracija v WebSphere Portal
  - datoteka es.wp5.install.jar 313
  - datoteka es.wp6.install.jar 318, 326
  - gručni sistem 325
  - Lotus Quickr 311
  - namestitveni skripti 312

- integracija v WebSphere Portal (*nadaljevanje*)
  - opis 311
  - skript wp5\_install 313
  - skript wp6\_cluster\_install 326
  - skript wp6\_install 318
  - Web Content Management 311
- intervali za vnovično pajkanje za spletne pajke 88
- IOCP, konfiguriranje pajka 73
- iskalna polja HTML
  - preslikovanje elementov v 126
- Iskalna polja HTML
  - elementi Dublin Core 127
  - izdelovanje 127
  - opis 126
  - preslikovanje elementov v 127
- iskalna polja s popolnim ujemanjem, opis 163
- Iskalna polja XML
  - izdelovanje 123
  - opis 123, 131
  - preslikovanje elementov v 123, 131
- iskalna vrstica, WebSphere Portal
  - različica 5.1, preusmeritev na sistem iskanja v podjetju 315
- Iskalna vrstica, WebSphere Portal
  - različica 6, preusmeritev na sistem iskanja v podjetju 322
- iskalne aplikacije
  - dostopanje 226
  - ID-ji aplikacije 238
  - opis 10
  - po meri 201
  - povezovanje z zbirkami 202
  - povezovanje z zunanjimi viri 231
  - prilagajanje 220
  - vzorec 201, 202
  - zagon 226
  - zaščita na ravni zbirke 238
- iskalni predpomnilnik
  - konfiguriranje 178
  - opis 177
- iskalni strežniki
  - iskalni predpomnilnik 177, 178
  - izračunavanje števila poizvedb 287
  - konfiguriranje HTTPS 226
  - konfiguriranje SSL 226
  - konfiguriranje strežnika proxy 228
  - nadziranje 286, 287
  - nedavne poizvedbe 286, 287
  - opis 7, 177
  - pogoste poizvedbe 286, 287
  - povezovanje s slovarji poviševalnih besed 194
  - povezovanje slovarjev sopomenk 180
  - povezovanje slovarjev zaustavitvenih besed 182
  - slovarji poviševalnih besed 192
  - slovarji sopomenk 178, 180
  - slovarji zaustavitvenih besed 181
  - status sistema 286
  - vnovična razmestitev slovarjev 183
  - zagon 267, 271, 286
  - zaustavitev 269, 271, 286
  - zgodovina odzivnega časa 286
- iskanje
  - dokumenti HTML 126, 127

- iskanje (*nadaljevanje*)
  - dokumenti XML 123, 131
  - hitre povezave 186
  - kategorije 118
  - zbirke 163
- iskanje po izvornem XML 135
- iskanje po poljih
  - opis 163
  - razvrstitev po nizih 163
- iskanje v podjetju
  - API-ji 9
  - datoteke dnevnika 293
  - diagram toka podatkov 11
  - formati URI 108
  - ID-ji sej 335
  - integracija v WebSphere Portal 311
  - integracija z Lotus Notes 309
  - iskalne aplikacije 10
  - iskalni strežniki 7, 177
  - komponente 3
  - konfiguriranje številke vrat 21
  - konzola za upravljanje 8
  - nadziranje 273
  - obnovitev iz varnostne kopije 305
  - povratne kode 335
  - pregled 1
  - razčlenjevalniki 4, 117
  - skrbniške vloge 235
  - skripti izdelovanja varnostnih kopij 303
  - skripti za obnovitev 303
  - strežniki pajka 4, 35
  - strežniki za indeksiranje 6, 157
  - ukazi 335
  - varnostno kopiranje 304
  - zagon iskalnih strežnikov 271
  - zagon strežnikov 267
  - zaščita 233
  - zaščita na ravni dokumenta 238
  - zaščita na ravni zbirke 237
  - zaustavitev strežnikov 267, 269
  - zaustavljanje iskalnih strežnikov 271
- izdelovanje
  - hitre povezave 186
  - Iskalna polja HTML 127
  - Iskalna polja XML 123
  - kategorije na osnovi pravil 121
  - območja 171
  - pajki 36
  - poročila spletnega pajka 279
  - zbirke 29, 30

## J

- japonsščina
  - odstranitev znakov za novo vrstico 154
  - segmentacija n-gram 154
- javanski spojnik za DB2 Content Manager 53, 54
- jeziki
  - dvomestne kode 151
  - iskanje 151
  - podprte 151, 152
  - samodejno zaznavanje 152
- jeziki preprostih besedil 151
- jeziki zapletenih besedil 151
- jezikovna podpora
  - analiza besedila po meri 128

- jezikovna podpora (*nadaljevanje*)
  - iskanje po izvornem XML 135
  - kode jezikov 151
  - odstranitev presledkov 154
  - področne nastavitve 151
  - segmentacija n-gram 154
  - semantično iskanje 128, 135
  - slovarji poviševalnih besed 192
  - slovarji sopomenk 178
  - slovarji zaustavitvenih besed 181
  - zaznavanje jezikov 152
  - zaznavanje kodne strani 153

## K

- kategorije
  - drevesa kategorij 120
  - formati URI 108
  - gnezdenje podkategorij 120
  - iskanje 118
  - izdelovanje 121
  - na osnovi pravil 118, 121
  - opis 118
  - selitev s portala WebSphere Portal 331
  - vrsta kategorizacije 121
- kategorije na osnovi pravil
  - izbira vrste kategorizacije 121
  - izdelovanje 121
  - opis 118
- kitajščina
  - odstranitev znakov za novo vrstico 154
  - segmentacija n-gram 154
- ključne besede v hitrih povezavah 186
- kloniranje
  - iskalne aplikacije 221
  - pajki 35
- kodiranje unicode 153
- kodne strani
  - podprte 153
  - samodejno zaznavanje 153
- konfiguriranje dnevnika poizvedb 299
- konfiguriranje overjenega strežnika 257
- konfiguriranje pomoči za imenike 259
- konfiguriranje strežnika SMTP 297
- konfiguriranje uporabnika Domino, pajki QuickPlace 258
- konfiguriranje vmesnika povratne zanke 24
- konzola za upravljanje
  - opis 8
- korejščina
  - analiza zloženek 135
  - segmentacija n-gram 154

## L

- lastnosti pajka
  - opis 4
  - urejanje 37
- Lotus Notes
  - integracija v sistem iskanja v podjetju 309
  - iskalna vrstica 309
  - namestitve vtičnika 309
  - spletno mesto s posodobitvami vtičnika 309

- Lotus Quickr
  - integracija v WebSphere Portal 311, 323
  - konfiguriranje pajka QuickPlace 75
  - konfiguriranje pajkov seznama semen 78

## M

- motorji za analizo besedila
  - dodajanje v sistem 129
  - opis 128
  - povezovanje z zbirkami 130
  - preslikava rezultatov analize v indeks 132
  - preslikovanje elementov XML 131
  - preslikovanje rezultatov analize v relacijske baze podatkov 133
  - preslikovanje strukture splošne analize v relacijske baze podatkov 133
- možnost removeCjNewlineChars 154
- možnost removeCjNewlineCharsMode 154
- možnosti iskanja
  - iskanje po poljih 163
  - parametrično iskanje 163
  - popolno ujemanje 163
  - prosto besedilno iskanje 163
  - razvrstitvena polja 163
  - vsebina dokumenta 163
  - za rezultate iskanja 163

## N

- na pravilih temelječa taksonomija, selitev s portala WebSphere Portal 331
- nadziranje
  - Aktivna mesta spletnega pajka 278
  - datoteke dnevnika 300
  - hitrost pajkanja spletnega pajka 278
  - iskalni strežniki 286, 287
  - iskanje v podjetju 273
  - nedavne poizvedbe 286, 287
  - odstranjeni dokumenti 289, 290
  - pajki 275
  - Podrobnosti o nitih spletnega pajka 277
  - podrobnosti o URI-ju 274
  - pogoste poizvedbe 286, 287
  - poslušalnik podatkov 288
  - razčlenjevalniki 284
  - Spletni pajki 276
  - zbirke 274
  - zgodovina odzivnega časa 286
- nadzor dostopa
  - onemogočanje za zbirko 264
  - opis 235
  - potrditev trenutnega uporabnika 241
  - upravljanje identitet 241, 244
  - zahteve za datotečne sisteme Windows 260
  - zahteve za Lotus Domino 256, 257
  - zaščita na ravni dokumenta 240
  - zaščita z enotno prijavo 245
- nadzornik
  - konfiguriranje vloge 236
  - opis 235
- najboljši rezultati
  - stolpčni grafikon 223
- najdaljši interval za vnovično pajkanje 88

- najkrajši interval za vnovično pajkanje 88
- namestitveni skripti
  - WebSphere Portal 312
- naslovi IP
  - podpira IPv6 25
  - podpora za dvojne 24
  - vmesnik povratne zanke 24
- naslovi IP, spreminjanje 22
- nedavne poizvedbe
  - izračunavanje 287
- nedavne poizvedbe, nadziranje 286
- nedavno pajkani URL-ji, nadziranje 276
- neposredni način, repozitoriji Content Edition 41
- neznani tipi dokumentov 141
- nit
  - razčlenjevalnik 134
  - spletni pajek 277
- niz konteksta zaščite uporabnika 241
- niz USC 241

## O

- obdelava besedila
  - motorji za analizo besedila 128
  - razlagalniki 128
  - struktura splošne analize 128
- obiskovanje URL-jev takoj, ko je mogoče 89
- objavljanje dogodkov
  - konfiguriranje pajka DB2 48, 50
  - nameščanje v operacijskem okolju Solaris 46
  - nameščanje v operacijskem sistemu AIX 46
  - nameščanje v operacijskem sistemu Linux 46
  - nameščanje v sistemu Windows 47
  - opis 44
- območja
  - formati URI 108
  - iskanje 171
  - izdelovanje 171
  - opis 171
  - vzorci URI-jev 171
- obnovitev iskanja v podjetju 303
- obnovitev sistema 303, 305
- ocenjevanje sistemskih virov 273
- odkritje 4
- odstranitev presledkov 154
- odstranitev znaka za novo vrstico 154
- odstranjeni dokumenti
  - datoteke dnevnika za 290
  - konfiguriranje datoteke dnevnika 289
  - opis 289
  - poročila o 289
- odstranjevanje URI-jev iz indeksa 174
- odzivni čas iskanja
  - nadziranje 286
  - opozorila o 294
- omejevanje prostora za spletno pajkanje 84
- OmniFind Enterprise Edition
  - API-ji 9
  - diagram toka podatkov 11
  - ID-ji sej 335
  - integracija v WebSphere Portal 311
  - integracija z Lotus Notes 309
  - iskalne aplikacije 10

- OmniFind Enterprise Edition (*nadaljevanje*)
  - iskalni strežniki 7
  - komponente 3
  - konfiguriranje številke vrat 21
  - konzola za upravljanje 8
  - podpora dvojnimi IP-jem 24
  - podpora za protokol IPv6 25
  - povratne kode 335
  - pregled 1
  - razčlenjevalniki 4
  - spreminjanje gesla v enem strežniku 17
  - spreminjanje gesla v več strežnikih 18
  - spreminjanje naslovov IP 22
  - strežniki pajka 4
  - strežniki za indeksiranje 6
  - ukazi 335
- omogočanje razporedov indeksiranja 159
- onemogočanje razporedov indeksiranja 159
- operacijski sistem AIX
  - konfiguriranje objavljavanja dogodkov 46
  - konfiguriranje pajka Content Edition 42
  - konfiguriranje pajka DB2 46
  - konfiguriranje pajka DB2 Content Manager 53
  - konfiguriranje pajka Domino Document Manager 69
  - konfiguriranje pajka Notes 69
  - konfiguriranje pajka QuickPlace 69
- operacijski sistem Linux
  - konfiguriranje objavljavanja dogodkov 46
  - konfiguriranje pajka Content Edition 42
  - konfiguriranje pajka DB2 46
  - konfiguriranje pajka DB2 Content Manager 53
  - konfiguriranje pajka Domino Document Manager 69
  - konfiguriranje pajka Notes 69
  - konfiguriranje pajka QuickPlace 69
  - operacijsko okolje Solaris
    - konfiguriranje objavljavanja dogodkov 46
- operacijski sistem Solaris
  - konfiguriranje pajka DB2 46
- operacijski sistem Windows
  - konfiguriranje objavljavanja dogodkov 47
  - konfiguriranje pajka 71
  - konfiguriranje pajka Content Edition 43
  - konfiguriranje pajka DB2 47
  - konfiguriranje pajka DB2 Content Manager 54
- operacijsko okolje Solaris
  - konfiguriranje pajka Content Edition 42
  - konfiguriranje pajka DB2 Content Manager 53
  - konfiguriranje pajka Domino Document Manager 69
  - konfiguriranje pajka Notes 69
  - konfiguriranje pajka QuickPlace 69
- operator
  - konfiguriranje vloge 236
  - opis 235
- opozorila
  - indeksirani dokumenti 294
  - konfiguriranje strežnika SMTP 297
  - možnosti elektronske pošte 294, 295
  - na ravni sistema 295
  - odzivni časi iskanja 294

- opozorila (*nadaljevanje*)
  - omejitve indeksa 273
  - opis 293
  - pajkani dokumenti 294
  - prejemanje elektronske pošte za 298
  - raven zbirke 273, 294
  - razpoložljiv prostor na strežnikih 295
- opozorila o razpoložljivem prostoru 295
- osnovno overjanje HTTP 90, 91
- overjanje
  - onemogočanje za aplikacije podjetja 262
  - opis 235
- overjanje na osnovi obrazcev 90, 91
- overjeni strežniki Lotus Domino 257

## P

- pajki
  - aplikacije poslušalnika podatkov 104
  - arhivske datoteke 107
  - baza podatkov JDBC 58, 60, 62
  - brisanje 38
  - Content Edition 39, 41, 42
  - datotečni sistem UNIX 79
  - Datotečni sistem Windows 102
  - DB2 44
  - DB2 Content Manager 52
  - Domino Document Manager 56
  - Exchange Server 58, 256
  - formati URI 108
  - izdelovanje 36
  - nadziranje 275
  - NNTP 65
  - omogočanje zaščite na ravni dokumentov 35
  - opis 4
  - Opombe 66, 68
  - osnovne vrednosti za 35
  - podpora za zunanje 104
  - privzete nastavitve selitve 332
  - QuickPlace 75
  - razporejanje 35, 39
  - rpregled konfiguriranja 35
  - Seznam semen 78
  - Splet 80
  - status sistema 275
  - urejanje lastnosti pajka 37
  - urejanje prostorov za pajkanje 38
  - vtičniki 105
  - Web Content Management 98, 101
  - WebSphere Portal 100, 101
  - začetne vrednosti za 36
  - zaščita na ravni dokumenta 238
  - združevanje tipov pajkov 35
- pajki baze podatkov JDBC
  - formati URI 108
  - konfiguriranje 58
  - pajkanje več tabel 60, 62
  - podprti gonilniki 58
- Pajki baze podatkov JDBC
  - vtičnik za pajkanje več tabel 60, 62
- pajki Content Edition
  - formati URI 108
  - konfiguriranje 39
  - nameščanje v operacijskem okolju Solaris 42
  - neposredni način 41

- pajki Content Edition *(nadaljevanje)*
  - strežniški način 42
- Pajki Content Edition
  - nameščanje v operacijskem sistemu AIX 42
  - nameščanje v operacijskem sistemu Linux 42
  - nameščanje v sistemu Windows 43
- pajki datotečnega sistema UNIX
  - formati URI 108
  - konfiguriranje 79
- pajki datotečnega sistema Windows
  - formati URI 108
  - konfiguriranje 102
  - konfiguriranje zaščite na ravni dokumenta 260
- pajki DB2
  - formati URI 108
- Pajki DB2
  - konfiguriranje 44
  - konfiguriranje objavljanja dogodkov 46, 47
  - konfiguriranje WebSphere II Event Publisher Edition 48
  - konfiguriranje WebSphere MQ 50
  - namestitve v operacijskem okolju Solaris 46
  - namestitve v operacijskem sistemu AIX 46
  - namestitve v operacijskem sistemu Linux 46
  - namestitve WebSphere MQ na sistemu AIX 46
  - namestitve WebSphere MQ na sistemu Linux 46
  - namestitve WebSphere MQ na sistemu Solaris 46
  - namestitve WebSphere MQ na sistemu Windows 47
  - nameščanje v sistemu Windows 47
  - objavljanje dogodkov 44
  - WebSphere II Classic Federation 51
- pajki DB2 Content Manager
  - formati URI 108
  - nameščanje v operacijskem sistemu Linux 53
  - nameščanje v sistemu Windows 54
- Pajki DB2 Content Manager
  - konfiguriranje 52
  - nameščanje v operacijskem okolju Solaris 53
  - nameščanje v operacijskem sistemu AIX 53
- pajki Domino Document Manager
  - formati URI 108
  - konfiguriranje protokola DIIOP 72
  - nameščanje v operacijskem sistemu AIX 69
- Pajki Domino Document Manager
  - konfiguriranje 56
  - konfiguriranje IOCP 73
  - nameščanje v operacijskem okolju Solaris 69
  - nameščanje v operacijskem sistemu Linux 69
  - nameščanje v sistemu Windows 71
  - protokol NRPC 69, 71
- pajki NNTP, konfiguriranje 65
- pajki Notes
  - formati URI 108
  - konfiguriranje 66
  - konfiguriranje protokola DIIOP 72
  - nasveti za uporabo 68
  - pravila za preslikavo polj 68
- Pajki Notes
  - konfiguriranje IOCP 73
  - konfiguriranje zaščite na ravni dokumenta 256
  - nameščanje v operacijskem okolju Solaris 69
  - nameščanje v operacijskem sistemu AIX 69
  - nameščanje v operacijskem sistemu Linux 69
  - nameščanje v sistemu Windows 71
  - overjeni strežnik Lotus Domino 257
  - potrditev trenutnih poverilnic 257
  - protokol NRPC 69, 71
- pajki QuickPlace
  - formati URI 108
  - konfiguriranje pomoči za imenike 259
  - konfiguriranje protokola DIIOP 72
  - nameščanje v operacijskem okolju Solaris 69
  - nameščanje v operacijskem sistemu Linux 69
- Pajki QuickPlace
  - konfiguriranje 75
  - konfiguriranje IOCP 73
  - konfiguriranje uporabnika Domino 258
  - nameščanje v operacijskem sistemu AIX 69
  - nameščanje v sistemu Windows 71
  - protokol NRPC 69, 71
  - zaščita lokalnega uporabnika 258
- Pajki seznama semen
  - formati URI 108
  - integracija v WebSphere Portal 311, 323
  - konfiguriranje 78
- Pajki strežnika Exchange Server
  - formati URI 108
  - konfiguriranje 58
  - zaščiteni dokumenti 256
- pajki Web Content Management
  - formati URI 108
  - konfiguriranje 98
  - kopiranje URL-jev spletnega mesta 101
- pajki WebSphere Portal
  - formati URI 108
  - konfiguriranje 100
- paralelna gradnja indeksov 160
- parameter distinctRecentQueryCheck 287
- parameter newHtmlTagReplacement 142
- parametrična polja
  - numerična razvrstitev 163
  - opis 163
- piškotki za spletno pajkanje
  - format 93
  - konfiguriranje 94
  - opis 93
- podpora za JavaScript v spletnih pajkih 83
- podrobnosti o niti, nadziranje 276
- podrobnosti o URI-ju
  - nadziranje 274
- podrobnosti o URI-ju *(nadaljevanje)*
  - odstranjeni dokumenti 289
- področne nastavitve
  - iskanje 151
  - razčlenjevanje 151
- pogled Zbirke
  - izdelovanje zbirk 30
- Pogled Zbirke
  - opis 13
- pogoste poizvedbe
  - izračunavanje 287
- pogoste poizvedbe, nadziranje 286
- poizvedbe z nizkim priklicem
  - opis 196
  - privzeti poviševalni faktorji 199
- poizvedbe z visokim priklicem
  - opis 196
  - privzeti poviševalni faktorji 199
- pojmovni jeziki 151
- polja rezultatov iskanja, opis 163
- polja za metapodatki, stolpčni grafikon najboljših rezultatov 223
- polja, preslikava v poviševalne razrede 198
- pomembnost dokumenta
  - obnovitev privzetih vrednosti 191
  - omogočanje za zbirko 30
  - poviševalni razredi 196, 199
  - slovarji poviševalnih besed 194
  - statično 190
  - v preseljenih zbirkah 331
  - vzorci URI-jev 195
- pooblašcanje, opis 235
- poročila o podrobnostih spletnega mesta
  - izdelovanje 279
  - opis 276
- poročila o zgodovini pajka
  - izdelovanje 279
  - opis 276
- Poročilo o statusnih kodah HTTP 279
- Poročilo o strani 279
- portalski programčki
  - integracija v WebSphere Portal 5.1 313
  - integracija v WebSphere Portal 6 318, 323
  - integracija z gručami WebSphere Portal 326
  - iskanje v podjetju 311
  - nastavljanje za vire Lotus Quickr 323
  - odstranjevanje iz gruče portala WebSphere Portal 329
  - odstranjevanje iz WebSphere Portal 5.1 317
  - odstranjevanje iz WebSphere Portal različice 6 324
  - opis 311
- Poslušalnik podatkov
  - konfiguriranje 104
  - nadziranje 288
  - ponovni zagon 104, 288
- potrditev poizvedbe 241
- potrditev trenutnih poverilnic 241, 256, 257, 260
- poviševalni faktorji
  - konfiguriranje poviševalnega razreda 196, 199
  - za poviševalne razrede 198, 199
  - za slovarje poviševalnih besed 192

- poviševalni faktorji (*nadaljevanje*)
  - za vzorce URI-jev 195
- poviševalni razredi
  - konfiguriranje 198, 199
  - opis 196
  - poizvedbe z nizkim priklicem 196, 199
  - poizvedbe z visokim priklicem 196, 199
  - preslikava polj v 198
  - privzete vrednosti 199
  - zaznavanje dvojnikov dokumentov 196
- povratne kode, iskanje v podjetju 335
- povzetek naloge, skrbniška konzola 13
- povzetki
  - dinamično 184
  - prilagajanje 184, 185
- povzetki dokumentov
  - prilagajanje 184
  - urejanje lastnosti za 185
- požarni zidovi, pakiranje dokumentov strežnika
  - Exchange Server 256
- pravila kategorije
  - konfiguriranje 121
  - vsebina dokumenta 118, 121
  - vzorci URI-jev 118, 121
- pravila s predponami za spletne pajke 84
- pravila z naslovi IP za spletne pajke 84
- pravila z naslovi za spletne pajke 84
- pravila zamenjave za HTML 142, 143
- preslikovanje
  - elementov XML v strukturo splošne analize 131
  - iskalna polja HTML 127
  - Iskalna polja XML 123
  - polj v poviševalne razrede 198
  - rezultatov analize v relacijske baze podatkov 133
  - struktur splošne analize v indeks 132
  - strukture splošne analize v relacijske baze podatkov 133
- prijava na skrbniško konzolo 16
- prikaz Nadzornik, opis 13
- prikaz Sistem, opis 13
- prikaz Zaščita, opis 13
- prikazovanje
  - datoteke dnevnika 300
  - datoteke dnevnika za odstranjene dokumente 290
  - podrobnosti o URI-ju 274
- prilagajanje iskalne aplikacije 219, 220
- Prilagojevalnik iskalnih aplikacij
  - datoteka config.properties 220
  - zagon 220
- pripomočki za osebe s posebnimi potrebami, ki so na voljo v tem izdelku 367
- privzeta iskalna aplikacija 219
- prosto besedilno iskanje, opis 163
- prostor pakiranja
  - konfiguriranje spletnega pajka 84
  - opozorila o 294
  - urejanje 38
- prostor za pakiranje
  - opis 4
- protokol DIIOp, konfiguriranje pajka 72
- protokol IPv6 25
- protokol NRPC, konfiguriranje pajka 69, 71

- protokol za izključevanje robotov
  - identifikacija uporabniških posrednikov 81
  - skladnost spletnega pajka 82

## R

- rangiranje dokumentov
  - obnovitev privzetih vrednosti 191
- rangiranje rezultatov iskanja
  - datoteka runtime.properties 191
  - dinamično 189
  - obnovitev privzetih vrednosti 191
  - opis 189
  - poviševalni razredi 196, 198, 199
  - slovarji poviševalnih besed 194
  - statično 190, 191
  - točkovanje, temelječe na besedilu 189
  - vzorci URI-jev 195
- razčlenjevalnik ASCII 141
- razčlenjevalnik Stellent
  - opis 137
  - povezovanje tipov dokumentov 144
  - privzeti tipi dokumentov 146
  - razčlenjevanje tipov dokumentov 139
- razčlenjevalniki
  - analiza zloženek 135
  - datoteke brez pripon 141
  - iskanje po izvornem XML 135
  - izbira tipa razčlenjevalnika 137
  - jezikovna obdelava 151
  - nadziranje 284
  - naloge analiziranja podatkov 4
  - neznani tipi dokumentov 141
  - nit 134
  - odstranitev presledkov 154
  - odstranitev znaka za novo vrstico 154
  - opis 4, 117
  - podprti jeziki 151
  - podprti tipi dokumentov Stellent 146
  - pravila zamenjave za HTML 142, 143
  - razčlenjevalnik ASCII 141
  - razčlenjevanje tipov dokumentov 139
  - segmentacija n-gram 154
  - status sistema 284
  - tipi dokumentov za razčlenjevalnike Stellent 144
  - tipi dokumentov za storitve razčlenjevalnika 139, 141
  - zagon 284
  - zaustavitev 284
  - zaznavanje formata dokumentov 137
  - zaznavanje jezikov 152
  - zaznavanje kodne strani 153
- razlagalniki 128
- razmestitev portalskega programčka Iskanje
  - datoteka es.wp5.install.jar 313
  - datoteka es.wp6.install.jar 318, 326
  - skript wp5\_install 313
  - skript wp5\_uninstall 317
  - skript wp6\_cluster\_install 326
  - skript wp6\_install 318
  - skript wp6\_uninstall 324
  - wp6\_cluster\_uninstall 329
- razporejanje
  - gradnje indeksov 158, 159
  - pajki 35, 39

- razširitev indeksa
  - opis 168
  - vpliv na čas gradnje indeksa 169
  - vpliv na velikost indeksa 169
- razširitev poizvedb
  - opis 168
  - vpliv na čas gradnje indeksa 169
  - vpliv na velikost indeksa 169
- razvrstitevna polja
  - numerična razvrstitev 163
  - razvrstitev po nizih 163
- rezervirano polje
  - es\_special\_field.default\_field 199
- rezervirano polje
  - es\_special\_field.default\_metadata\_field 199
- rezervirano polje
  - es\_special\_field.regular\_text 199
- rezultati iskanja
  - dinamično povzemanje 184
  - dinamično rangiranje 189
  - grupiranje 172, 173
  - konfiguriranje poviševalnega razreda 196, 198, 199
  - konfiguriranje vzorcev URI-jev 195
  - krčenje 172, 173, 265
  - opis 189
  - po-filtriranje 239
  - povzetki 184, 185
  - predfiltriranje 239
  - prilagajanje povzetkov 184, 185
  - rangiranje 195
  - razširitev univerzalnih znakov 170
  - statično rangiranje 190
  - točkovanje, temelječe na besedilu 189
  - univerzalni znaki 168

## S

- samodejno zaznavanje
  - jeziki 152
  - kodne strani 153
- segmentacija n-gram 154
- selitev
  - na pravih temelječa taksonomija 331
  - zbirke 331
- semantično iskanje 128, 131, 135
- SIAPI (API za iskanje in indeksiranje) 9, 201
- sistemski viri
  - ocenjevanje 273
- skladnja poizvedbe XML, izvorni 135
- skrbnik iskanja v podjetju
  - konfiguriranje vloge 236
  - opis 235
  - spreminjanje gesla v enem strežniku 17
  - spreminjanje gesla v več strežnikih 18
- skrbnik zbirke
  - konfiguriranje vloge 236
  - opis 235
- skrbniška konzola
  - povzetek naloge 13
  - prijava 16
  - vmesnik 13
- skrbniške vloge
  - konfiguriranje 236
  - nadzornik 235, 236
  - operater 235, 236
  - opis 235



skrbniške vloge (*nadaljevanje*)  
 skrbnik iskanja v podjetju 235, 236  
 skrbnik zbirke 235, 236

skrčeni rezultati iskanja  
 konfiguriranje 173  
 omejitve zaščite 265  
 opis 172

skrčeni URI-ji  
 konfiguriranje 173  
 omejitve zaščite 265  
 opis 172

skript esbackup.bat 304  
 skript esbackup.sh 304  
 skript eschange.pw 17, 18  
 skript escrm.sh 53  
 skript escrm.vbs 54  
 skript escrdb2.sh 46  
 skript escrdb2.vbs 47  
 skript escrnote.sh 69  
 skript escrnote.vbs 71  
 skript escrvbr.sh 42  
 skript escrvbr.vbs 43  
 skript esrestore.bat 305  
 skript esrestore.sh 305  
 skript startccl 305  
 skript wp5\_install 313  
 skript wp5\_uninstall 317  
 skript wp6\_cluster\_install 326  
 skript wp6\_cluster\_uninstall 329  
 skript wp6\_install 318  
 skript wp6\_uninstall 324

skripti  
 esbackup.bat 304  
 esbackup.sh 304  
 escrm.sh 53  
 escrm.vbs 54  
 escrdb2.sh 46  
 escrdb2.vbs 47  
 escrnote.sh 69  
 escrnote.vbs 71  
 escrvbr.sh 42  
 escrvbr.vbs 43  
 esrestore.bat 305  
 esrestore.sh 305  
 startccl 305

skripti izdelovanja varnostnih kopij  
 izvajanje 304  
 opis 303

skripti za obnovitev  
 izvajanje 305  
 opis 303

sledenje dokumentov  
 datoteke dnevnika 290  
 konfiguriranje datoteke dnevnika 289  
 omogočanje 289  
 onemogočanje 289  
 opis 289  
 poročila 289

slovarji poviševalnih besed  
 dodajanje v sistem 194  
 opis 192  
 povezovanje z zbirko 194  
 vnovična razmestitev 183

slovarji sopomenk  
 dodajanje v sistem 180  
 opis 178  
 povezovanje z zbirko 180

slovarji sopomenk (*nadaljevanje*)  
 vnovična razmestitev 183

slovarji zaustavitvenih besed  
 dodajanje v sistem 182  
 opis 181  
 povezovanje z zbirko 182  
 vnovična razmestitev 183

sočasna gradnja indeksov 160

spletni pajki  
 datumi pajkanih dokumentov 97

Spletni pajki  
 aktivna spletna mesta 276, 278  
 datoteka followindex.rules 96  
 datoteke robots.txt 81, 82  
 format piškotkov 93  
 globalni prostor za pajkanje 95  
 hitrost pajkanja 278  
 intervali za vnovično pajkanje 88  
 izdelava poročil o 279  
 konfiguriranje 80  
 konfiguriranje piškotkov 94  
 nadziranje 276  
 nedavno pajkani URL-ji 276  
 obiskovanje URL-jev takoj, ko je mogoče 89  
 omejevanje prostora za pajkanje 84  
 piškotki 93  
 podpora za JavaScript 83  
 podrobnosti o niti 276, 277  
 podrobnosti spletnega mesta 276  
 pravila za pajkanje 84  
 status sistema 276  
 status URL-ja 276  
 strani s programskimi napakami 89  
 strežniki proxy 92  
 ukazi no-follow 96  
 ukazi no-index 96  
 uporabniški posredniki 81  
 z geslom zaščiteni spletna mesta 90, 91  
 začetni URL-ji 84, 89  
 zgodovina pajka 276

sporočila o napakah  
 konfiguriranje strežnika SMTP 297  
 prejemanje elektronske pošte za 296, 298  
 prikaz datotek dnevnikov 300  
 prikazovanje datotek dnevnika za odstranjene dokumente 290

SSL, konfiguriranje iskalnih strežnikov 226

statično rangiranje  
 omogočanje za zbirko 30  
 opis 190  
 v preseljenih zbirkah 331

status sistema  
 gradnje indeksov 285  
 iskalni strežniki 286  
 pajki 275  
 razčlenjevalniki 284  
 Spletni pajki 276  
 zbirke 274

statusne kode HTTP  
 poročilo spletnega pajka 279  
 prejeli spletni pajki 279

stolpični grafikoni  
 najboljši rezultati 223  
 razredi Java za najboljše rezultate 223

stran Dnevnik, opis 13  
 stran Indeksiranje, opis 13

stran Iskanje, opis 13  
 stran Pajkanje, opis 13  
 stran Razčlenjevanje, opis 13  
 strani s programskimi napakami, spletni pajki 89

strežniki pajka  
 zagon 267, 275  
 zaustavitev 269, 275

strežniki proxy 92  
 konfiguracija iskalnega strežnika 228

strežniki proxy HTTP 92

strežniki razčlenjevalnika  
 konfiguriranje niti 134

strežniki za indeksiranje  
 zagon 267  
 zaustavitev 269

strežniki za iskanje v podjetju  
 podpora dvojnimi IP-jem 24  
 podpora za protokol IPv6 25  
 spreminjanje naslovov IP 22

strežniški način, repozitoriji Content Edition 42

struktura splošne analize  
 opis 128  
 preslikava v indeks 132  
 preslikovanje elementov XML v 131  
 preslikovanje v relacijske baze podatkov 133

## Š

številka vrat, iskanje v podjetju 21

## T

taksonomije, selitev s portala WebSphere Portal 331

tipi dokumentov  
 podprte v razčlenjevalnikih Stellent 146  
 razčlenjevanje 139  
 za razčlenjevalnike Stellent 144  
 za storitve razčlenjevalnika 139, 141  
 zaznavanje 137

tipi virov podatkov  
 ki jih podpira iskanje v podjetju 2  
 podpora za zunanje 2, 9

tipi MIME, vključitev v prostore za spletno pajkanje 84

tipi pajkov  
 osnovne vrednosti za 35  
 združevanje v zbirki 35

tipi virov podatkov  
 baza podatkov DB2 for iSeries 44  
 baze podatkov CA-Datcom 44  
 baze podatkov DB2 44, 58, 229  
 baze podatkov Domino Document Manager 56  
 baze podatkov IMS 44  
 baze podatkov Informix 44  
 baze podatkov JDBC 58, 60, 62, 229  
 baze podatkov Notes 66, 68  
 Baze podatkov Oracle 44, 58, 229  
 baze podatkov QuickPlace 75  
 baze podatkov Software AG Adabas 44  
 baze podatkov SQL Server 44, 58  
 baze podatkov Sybase 44

- tipi virov podatkov (*nadaljevanje*)
  - baze podatkov VSAM 44
  - datotečni sistemi UNIX 79
  - Datotečni sistemi Windows 102
  - DB2 for z/OS 44
  - javne mape strežnika Exchange Server 58
  - Lotus Quickr for Domino 75
  - Lotus Quickr for WebSphere Portal 78
  - novičarske skupine NNTP 65
  - relacijske baze podatkov 44
  - repozitoriji Content Edition 39, 41, 42
  - Spletna mesta 80
  - spletna mesta Web Content Management 98
  - Spletna mesta WebSphere Portal 100
  - tipi postavk DB2 Content Manager 52
- točkovanje, temelječe na besedilu 189
- tok podatkov, sistem iskanja v podjetju 11

## U

- UIMA
  - dodajanje motorjev za analizo besedila v sistem 129
  - opis 128
  - povezovanje z zbirkami 130
  - preslikava rezultatov analize v indeks 132
  - preslikava strukture splošne analize v indeks 132
  - preslikovanje elementov XML 131
  - preslikovanje rezultatov analize v relacijske baze podatkov 133
  - preslikovanje strukture splošne analize v relacijske baze podatkov 133
  - struktura splošne analize 132, 133
- ukaz esadmin 335
- ukaz esadmin startSearch 271
- ukaz esadmin stopIndex 162
- ukaz esadmin stopSearch 271
- ukaz esadmin system startall 335
- ukaz esadmin system stopall 335
- ukaz eschangeproxywp 228
- ukaz eschangetrustwp 226
- ukaz eschangewaspw
  - konfiguracija z enim strežnikom 251
  - konfiguracija z več strežniki 252
- ukaz startIndexBuild 161
- ukazi no-follow
  - konfiguriranje 96
  - opis 96
- ukazi no-index
  - konfiguriranje 96
  - opis 96
- ukazi, iskanje v podjetju 335
- univerzalni znaki
  - razširitev indeksa 168, 169, 170
  - razširitev poizvedb 168, 170
  - v poizvedbah 168
- uporabnik strežnika WebSphere Application Server
  - geslo za konfiguracijo z enim strežnikom 251
  - geslo za konfiguracijo z več strežniki 252
- uporabniški posredniki 81
- uporabniški profili
  - konfiguriranje 246

- uporabniški profili (*nadaljevanje*)
  - opis 244
- uporabniški register LDAP 249
- upravljanje identitet
  - konfiguriranje 246
  - kontekst zaščite uporabnika 241
  - niz poizvedbe XML 241
  - onemogočanje 241
  - opis 241
  - podpora za enotno prijavo 245
  - povzetek skupine 241
  - uporabniški profili 244
- urejanje
  - aplikacije poslušalnika podatkov 104
  - lastnosti iskalne aplikacije 203, 219
  - lastnosti pajka 37
  - prostori 38
  - zbirke 32
- URI-ji
  - formati v iskanju v podjetju 108
  - hitre povezave 186
  - območja 171
  - odstranjevanje iz indeksa 174
  - pravila kategorije 118, 121
  - prikaz podrobnosti o 274
  - skrčeni v rezultatih iskanja 172, 173
  - vplivanje na statično oceno 195

## V

- V/I zaključna vrata, konfiguriranje pajka 73
- varnostna kopija sistema 303, 304
- varnostno kopiranje iskanja v podjetju 303, 304
- večbajtno kodiranje 153
- vnovični obisk URL-jev takoj, ko je mogoče 89
- vrsta kategorizacije
  - izbiranje 30, 121
  - na osnovi pravil 118
- vsebina dokumenta, opis 163
- vtičnik za Lotus Notes
  - namestitvev 309
  - spletno mesto s posodobitvami 309
- vtičnik za pajkanje več tabel 60, 62
- vtičniki
  - pajkanje več strukturiranih tabel 60, 62
  - pajki baze podatkov JDBC 60, 62
- vtičniki pajkov 105
- vtičniki, za pajke 105
- vzorčna iskalna aplikacija
  - datoteka config.properties 203, 219
  - iskalne funkcije 201, 202
  - kloniranje 221
  - onemogočanje zaščite 262
  - opis 10, 202
  - privzeta razmestitev 219
  - uveljavitev HTTPS 226
  - uveljavitev SSL 226

## W

- Web Content Management
  - integracija v WebSphere Portal 311
- WebSphere II Classic Federation 51

- WebSphere II Event Publisher Edition, konfiguriranje pajka DB2 48
- WebSphere MQ, konfiguriranje pajka DB2 50
- WebSphere MQ, konfiguriranje strežnika pajkanja 46, 47
- WebSphere Portal
  - Center za iskanje, opis 311
  - gručni sistem 325
  - integracija v sistem iskanja v podjetju 311
  - Iskalna vrstica, opis 311
  - namestitveni skripti za sistem iskanja v podjetju 312
  - privzete nastavitve selitve 332
  - različica 5.1, konfiguriranje iskalne vrstice 315
  - različica 5.1, odstranjevanje sistema iskanja v podjetju 317
  - različica 5.1, skripti za integracijo 313
  - različica 6, konfiguriranje Centra za iskanje 320
  - različica 6, konfiguriranje iskalne vrstice 322
  - različica 6, odstranjevanje sistema iskanja v podjetju 324
  - različica 6, skripti za integracijo 318
  - selitev drevesa kategorij 331
  - selitev taksonomije 331
  - selitev zbirke 331
- Windows
  - podpora za protokol IPv6 25

## X

- XPath, izvirmi XML search 135

## Z

- z geslom zaščiten spletna mesta 90
  - osnovno overjanje HTTP 91
  - overjanje na osnovi obrazcev 91
- začetni URL-ji za spletne pajke 84, 89
- zagon
  - čarovnik za selitev 331
  - gradnje indeksov 285
  - iskalne aplikacije 226
  - iskalni strežniki 271, 286
  - Poslušalnik podatkov 288
  - Prilagojevalnik iskalnih aplikacij 220
  - razčlenjevalniki 284
  - strežniki pajka 275
  - strežniki za iskanje v podjetju 267
- zaščita
  - analiza sidnega besedila 247
  - dokumenti Lotus Domino 256, 257
  - domene Windows 260
  - globalna zaščita za WebSphere 262
  - globalna, aplikacijski strežnik WebSphere Application Server 248, 249
  - ID-ji iskalne aplikacije 238
  - izognitev nadzoru dostopa na ravni dokumenta 264
  - konfiguriranje HTTPS za iskanje 226
  - konfiguriranje SSL za iskanje 226

- zaščita (*nadaljevanje*)
  - na ravni dokumenta 238, 239, 240, 241, 246, 264
  - nadzor dostopa 235
  - namestitev na en strežnik 251
  - namestitev na več strežnikov 252
  - omogočanje za zbirko 30, 233
  - omogočitev za iskanje v podjetju 248
  - onemogočanje za aplikacijo podjetja 262
  - opis 233
  - overjanje 235, 262
  - podpora za enotno prijavo 245
  - raven zbirke 237, 264
  - skrbniške vloge 236
  - skrčeni rezultati iskanja 265
  - uporabniški profili 244
  - uporabniški register LDAP 249
  - upravljanje identitet 241, 246
  - vtičniki pajkov 105
  - zaznavanje dvojnikov dokumentov 237
- zaščita lokalnega uporabnika, pajki QuickPlace 258
- zaščita na ravni dokumenta
  - dokumenti Lotus Domino 257
  - indeksiran nadzor dostopa 240
  - konfiguriranje pajka 35
  - opis 233, 238
  - podpora za enotno prijavo 245
  - potrditev trenutnih poverilnic 241
  - potrditev v dejanskem času 241
  - rezultati po-filtriranja 239
  - rezultati predfiltriranja 239
  - uporabniški profili 244
  - upravljanje identitet 241, 246
  - vtičniki pajkov 105
  - za datotečne sisteme Windows 260
  - za dokumente Lotus Domino 256
  - žetoni zaščite 240
- zaščita na ravni zbirke
  - analiza sidrnega besedila 247
  - ID-ji aplikacije 238
  - omogočanje 30
  - opis 233, 237
  - zaznavanje dvojnikov dokumentov 237
- zaščita z enotno prijavo
  - konfiguriranje 246
  - upravljanje identitet 245
- zaustavitev
  - gradnje indeksov 285
  - iskalni strežniki 271, 286
  - razčlenjevalniki 284
  - strežniki pajka 275
  - strežniki za iskanje v podjetju 267, 269
- zaznavanje dvojnikov dokumentov
  - globalna analiza 167, 237
  - konfiguriranje poviševalnega razreda 196
  - omogočanje zaščite 237
  - opis 167, 233
- zbirke
  - brisanje 33
  - datoteka MigrationWizard.log 331
  - federacija 29
  - iskalni strežniki 177
  - iskanje 163
  - izdelovanje s čarovnikom za zbirke 29
  - izdelovanje s pogledom Zbirke 30
- zbirke (*nadaljevanje*)
  - izognitev nadzoru dostopa na ravni dokumenta 264
  - načini za izdelavo 29
  - nadziranje 274
  - ocenjevanje velikosti 30
  - ocenjevanje virov 273
  - opis 3
  - osnutek 29
  - povezovanje z iskalnimi aplikacijami 202
  - privzete nastavitve selitve 332
  - razčlenjevanje 117
  - selitev s portala WebSphere Portal 331
  - status sistema 274
  - ugotavljanje ID-ja 33
  - urejanje 32
  - zaščita 237
  - zaščita dvojnikov dokumentov 237
  - zaščita ID-ja aplikacije 238
  - zaščita sidrnega besedila 247
  - zaznavanje dvojnikov dokumentov 167
- združene zbirke 29
- zgodovina odzivnega časa, nadziranje 286
- zloženke, razčlenjevanje 135
- zunanjji pajki
  - aplikacije poslušalnika podatkov 104
  - konfiguriranje 104
- zunanjji viri
  - konfiguriranje 229
  - opis 229
  - povezovanje z iskalnimi aplikacijami 231
  - zaščita ID-ja aplikacije 238
- zunanjji viri JDBC
  - brisanje 229
  - gonilniki JDBC 229
  - konfiguriranje 229
  - urejanje 229
- zunanjji viri LDAP
  - brisanje 229
  - konfiguriranje 229
  - urejanje 229

## Ž

- žetoni zaščite
  - konfiguriranje pajka 240
  - onemogočanje za zbirko 264
  - zaščita na ravni dokumenta 240





**IBM**



**Java**<sup>™</sup>  
**COMPATIBLE**

SC18-9283-04



Spine information:

OmniFind Enterprise Edition **Različica 8.5**

**Izvajanje skrbništva nad iskanjem v podjetju**

