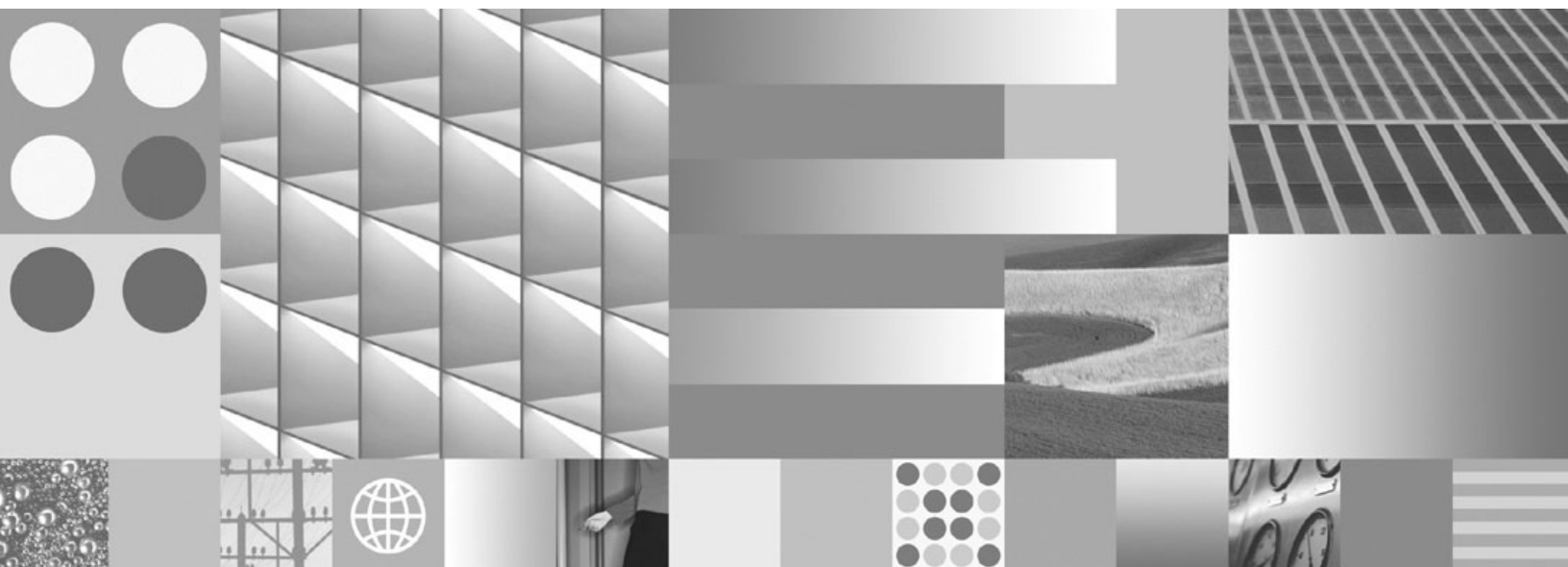


## Správa podnikového vyhledávání





## Správa podnikového vyhledávání

**Poznámka**

Před použitím těchto informací a produktu, k jehož podpoře se vztahují, si přečtěte informace v části “Doložky a ochranné známky” na stránce 395.

**Poznámka k vydání**

Toto vydání se vztahuje na verzi 8, vydání 5, modifikaci 0 produktu IBM OmniFind Enterprise Edition (číslo produktu 5724-C74) a všechny následující vydání a modifikace, dokud nebude v nových vydáních uvedeno jinak.

Odesláním informací společnosti IBM udělujete společnosti IBM nevýhradní právo k použití nebo distribuci jakýchkoli informací, které jí sdělíte, libovolným způsobem, který společnost považuje za odpovídající, bez vyžádání vašeho svolení.

© Copyright International Business Machines Corporation 2004, 2008. Všechna práva vyhrazena.

# Obsah

## ibm.com a související zdroje informací . . . . . vii

Odeslání vašich komentářů . . . . . vii

Kontaktování společnosti IBM . . . . . vii

## Co je podnikové vyhledávání? . . . . . 1

Typy zdrojů dat podporované podnikovým vyhledáváním . . . . . 2

Přehled komponent podnikového vyhledávání . . . . . 3

    Prolézací moduly podnikového vyhledávání . . . . . 4

    Analyzátoři podnikového vyhledávání . . . . . 4

    Indexy podnikového vyhledávání . . . . . 6

    Vyhledávací servery podnikového vyhledávání . . . . . 7

    Konzola pro správu podnikového vyhledávání . . . . . 8

    Monitorování systému podnikového vyhledávání . . . . . 9

    Soubory žurnálu podnikového vyhledávání . . . . . 9

    Přízpůsobení podnikového vyhledávání . . . . . 10

    Ukázková vyhledávací aplikace pro podnikové vyhledávání . . . . . 10

Datový tok podnikového vyhledávání . . . . . 11

## Správa systému podnikového vyhledávání . . . . . 13

Přihlášení ke konzole pro správu . . . . . 16

Změna hesla administrátora podnikového vyhledávání v konfiguraci s jedním serverem . . . . . 17

Změna hesla administrátora podnikového vyhledávání v konfiguraci s více servery . . . . . 18

Čísla portů TCP používaná pro podnikové vyhledávání . . . . . 20

    Změna čísla portu pro systém podnikového vyhledávání . . . . . 21

Změna adres IP nebo názvů hostitele serveru podnikového vyhledávání . . . . . 22

Konfigurace podpory pro duální adresy IP . . . . . 24

Povolení podpory protokolu IPv6 . . . . . 25

## Kolekce podnikového vyhledávání . . . . . 29

Vytvoření kolekce pomocí Průvodce kolekcí . . . . . 29

Vytvoření kolekce s použitím pohledu Kolekce . . . . . 30

Úprava kolekce . . . . . 32

Odstranění kolekce . . . . . 33

Určení ID kolekce . . . . . 33

## Administrace prolézacích modulů . . . . . 35

Vytvoření prolézacího modulu . . . . . 36

Úprava vlastností prolézacího modulu . . . . . 37

Úprava prolézaného prostoru . . . . . 38

Odstranění prolézacího modulu . . . . . 38

Časový plán prolézacího modulu . . . . . 39

Prolézací moduly Content Edition . . . . . 39

    Přístup k úložištím Content Edition v přímém režimu . . . . . 41

    Přístup k úložištím WebSphere II Content Edition v režimu serveru . . . . . 42

Prolézací moduly DB2 . . . . . 44

    Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2 . . . . . 46

    Konfigurování prolézacího serveru v systému Windows pro prolézací moduly DB2 . . . . . 47

    Konfigurování produktu WebSphere Information Integrator Event Publisher Edition pro prolézací moduly DB2 . . . . . 48

    Konfigurování produktu WebSphere MQ pro prolézací moduly DB2 . . . . . 50

    Prolézání databázi DB2 na klasickém serveru zdroje dat . . . . . 52

    Prolézací moduly DB2 Content Manager . . . . . 53

    Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2 Content Manager . . . . . 54

    Konfigurování prolézacího serveru v systému Windows pro produkt DB2 Content Manager . . . . . 55

    Prolézací moduly Domino Document Manager . . . . . 57

    Prolézací moduly Exchange Server . . . . . 59

    Prolézací moduly databáze JDBC . . . . . 60

    Mapy vztahů pro databáze JDBC . . . . . 62

    Prolézání více strukturovaných tabulek databáze JDBC . . . . . 64

    Prolézací moduly NNTP . . . . . 68

    Prolézací moduly Notes . . . . . 68

    Rady pro prolézání databázi Lotus Domino . . . . . 71

    Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino . . . . . 72

    Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino . . . . . 73

    Konfigurování serverů používajících protokol DIIOP . . . . . 75

    Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino . . . . . 76

    Prolézací moduly QuickPlace . . . . . 78

    Prolézací moduly počátečního seznamu . . . . . 81

    Prolézací moduly souborového systému UNIX . . . . . 82

    Webové prolézací moduly . . . . . 83

    Konfigurace uživatelského agenta . . . . . 84

    Způsob využití protokolu Robots Exclusion ve webovém prolézacím modulu . . . . . 85

    Podpora skriptů JavaScript . . . . . 87

    Pravidla omezující prolézaný webový prostor . . . . . 87

    Testování spojení adres URL s webovým prolézacím modulem . . . . . 92

    Nastavení intervalu opakovaného prolézání ve webovém prolézacím modulu . . . . . 92

    Volby pro návštěvy adres URL webovým prolézacím modulem . . . . . 93

    Způsob práce webového prolézacího modulu se stránkami méně závažných chyb . . . . . 93

    Podpora prolézání zabezpečených webových serverů . . . . . 94

    Webové servery obsluhované servery proxy . . . . . 97

    Správa souborů cookie . . . . . 97

    Konfigurace globálního prolézaného webového prostoru . . . . . 99

    Direktivy no-follow a no-index . . . . . 100

    Potlačení direktiv no-follow a no-index na webových stránkách . . . . . 101

    Konfigurování data používaného webovým prolézacím modulem pro prolézané dokumenty . . . . . 102

    Prolézací moduly Web Content Management . . . . . 103

Prolézací moduly WebSphere Portal . . . . .	105
Kopírování adresy URL pro prolézání ze systému WebSphere Portal . . . . .	106
Prolézací moduly souborového systému Windows . . . . .	107
Konfigurování podpory pro aplikace modulu pro příjem dat . . . . .	109
Vlastní moduly plug-in pro prolézací moduly . . . . .	110
Podpora prolézání souborů archivu . . . . .	112
Formáty identifikátorů URI v indexu podnikového vyhledávání . . . . .	113

## **Administrace analyzátorů . . . . . 123**

Práce s kategoriemi . . . . .	124
Kategorie založené na pravidlech . . . . .	124
Stromy kategorií . . . . .	126
Výběr typu kategorizace . . . . .	127
Konfigurování kategorií . . . . .	127
Práce s vyhledávacími poli XML . . . . .	128
Vyhledávací pole XML . . . . .	129
Mapování prvků XML na vyhledávací pole . . . . .	129
Práce s vyhledávacími poli HTML . . . . .	132
Vyhledávací pole HTML . . . . .	132
Mapování prvků metadat HTML na vyhledávací pole . . . . .	133
Vlastní zpracování textu . . . . .	134
Přidávání strojů pro analýzu textu do systému . . . . .	135
Asociování stroje pro analýzu textu s kolekcí . . . . .	136
Mapování prvků XML na obecnou strukturu analýzy . . . . .	137
Mapování obecné struktury analýzy na index . . . . .	138
Mapování obecné struktury analýzy na relační databázi . . . . .	139
Konfigurování podprocesů pro službu analyzátoru . . . . .	140
Povolení rozšířené analýzy složených výrazů . . . . .	141
Povolení podpory nativního vyhledávání XML . . . . .	142

## **Detekce formátu dokumentu . . . . . 143**

Výchozí podporované typy dokumentů . . . . .	144
Typy dokumentů přiřazené k analyzátorům kolekcí a k analyzátorům Stellent . . . . .	144
Přiřazení typů dokumentů k analyzátoru kolekce . . . . .	145
Výchozí pravidla služby analyzátoru kolekce . . . . .	147
Analýza neznámých typů dokumentů . . . . .	147
Změna pravidel náhrady pro některé značky HTML . . . . .	148
Výchozí pravidla náhrady kódu HTML . . . . .	149
Přiřazení typů dokumentů k analyzátoru Stellent . . . . .	150
Výchozí pravidla analýzy pro analyzátor Stellent . . . . .	153

## **Podpora jazyků a kódových stránek 157**

Automatická detekce jazyka . . . . .	158
Automatická detekce kódové stránky . . . . .	159
Lingvistická analýza dokumentů v čínštině, japonštině a korejštině . . . . .	160
N-gramová segmentace . . . . .	160
Odebrání mezerových znaků z textu . . . . .	160

## **Administrace indexů . . . . . 163**

Plánování sestavení indexu . . . . .	164
Změna časového plánu indexu . . . . .	165
Povolení a zákaz časových plánů indexování . . . . .	165
Konfigurování souběžných sestavení indexu . . . . .	166
Sestavení indexů pouze při detekci změn . . . . .	167

Zastavení sestavení indexu . . . . .	168
Volby ovlivňující prohledatelné zobrazení indexu . . . . .	169
Volby vytváření indexu pro prohledávání dokumentů . . . . .	169
Detekce duplicitních dokumentů . . . . .	173
Zástupné znaky v dotazech . . . . .	174
Obory . . . . .	177
Konfigurování oborů . . . . .	177
Sbalené identifikátory URI . . . . .	178
Sbalení identifikátorů URI ve výsledcích vyhledávání . . . . .	179
Odebrání identifikátorů URI z indexu . . . . .	180

## **Administrace vyhledávacího serveru 183**

Mezipaměti pro vyhledávání . . . . .	183
Konfigurování mezipaměti vyhledávání . . . . .	184
Vlastní slovníky synonym . . . . .	184
Přidávání slovníků synonym do systému . . . . .	186
Asociování slovníku synonym s kolekcí . . . . .	186
Vlastní slovníky zakázaných slov . . . . .	187
Přidávání slovníků zakázaných slov do systému . . . . .	188
Asociování slovníku zakázaných slov s kolekcí . . . . .	188
Opakovaná implementace vlastních slovníků . . . . .	189
Dynamické shrnutí . . . . .	190
Přizpůsobení shrnutí dokumentů v konzole pro správu . . . . .	190
Přizpůsobení souhrnů dokumentů provedením úprav vlastností . . . . .	191
Práce s rychlými odkazy . . . . .	192
Rychlé odkazy . . . . .	192
Konfigurování rychlých odkazů . . . . .	192

## **Ohodnocení důležitosti dokumentů 195**

Hodnocení na základě textu . . . . .	195
Statické řazení . . . . .	196
Obnovení výchozích hodnot pro statické ohodnocení důležitosti dokumentů . . . . .	197
Vlastní slovníky slov pro upřednostnění . . . . .	198
Přidávání slovníků slov pro upřednostnění do systému . . . . .	200
Asociování slovníku slov pro upřednostnění s kolekcí . . . . .	200
Hodnocení důležitosti dokumentů na základě vzorů identifikátoru URI . . . . .	201
Ovlivňování skóre dokumentů shodujících se se vzory identifikátorů URI . . . . .	201
Hodnocení důležitosti dokumentů na základě tříd upřednostnění . . . . .	202
Mapování polí na třídy upřednostnění . . . . .	204
Konfigurování faktorů upřednostnění pro třídy upřednostnění . . . . .	205
Výchozí hodnoty tříd upřednostnění . . . . .	205

## **Vyhledávací aplikace podnikového vyhledávání . . . . . 209**

Asociace vyhledávacích aplikací s kolekcemi . . . . .	210
Funkce ukázkové vyhledávací aplikace . . . . .	210
Vlastností vyhledávací aplikace . . . . .	211
Úpravy vlastností ukázkové vyhledávací aplikace . . . . .	227
Přizpůsobení podnikových aplikací . . . . .	228
Klonování ukázkové vyhledávací aplikace . . . . .	230
Analýza nejlepších výsledků . . . . .	231
Přístup k vyhledávacím aplikacím . . . . .	234
Konfigurování vyhledávacích serverů tak, aby přijímaly pouze zabezpečené požadavky (SSL) . . . . .	235

Konfigurování vyhledávacích serverů tak, aby přijímaly požadavky prostřednictvím serveru proxy . . . . . 236

## **Podpora externích zdrojů . . . . . 239**

Přidávání externích zdrojů do systému . . . . . 239  
Asociování vyhledávacích aplikací s externími zdroji . . . . . 241

## **Zabezpečení podnikového vyhledávání 243**

Zabezpečení instalace . . . . . 244  
Ověřování a řízení přístupu . . . . . 245  
Administrativní role . . . . . 246  
    Konfigurování administrativních uživatelů . . . . . 246  
Zabezpečení na úrovni kolekce . . . . . 247  
    Analýza duplicitních dokumentů a zabezpečení kolekce . . . . . 247  
    Identifikátory vyhledávací aplikace . . . . . 248  
Zabezpečení na úrovni dokumentu . . . . . 248  
    Předběžné a dodatečné filtrování výsledků vyhledávání . . . . . 249  
    Ověřování pomocí uložených prvků zabezpečení . . . . . 250  
    Ověřování aktuálního pověření během zpracování dotazu . . . . . 251  
    Analýza kotvícího textu . . . . . 257  
Povolení zabezpečení pro podnikové vyhledávání . . . . . 259  
    Konfigurace globálního zabezpečení a registru uživatelů LDAP na serveru WebSphere Application Server . . . . . 260  
    Povolení zabezpečení pro systém podnikového vyhledávání s jediným serverem . . . . . 262  
    Povolení zabezpečení pro systém podnikového vyhledávání s více servery . . . . . 262  
Požadavky nastavení prolézacího modulu pro podporu zabezpečení . . . . . 263  
    Ověřování přístupu k zabezpečeným dokumentům serveru Exchange Server . . . . . 267  
    Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino . . . . . 267  
    Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows . . . . . 271  
Zákaz zabezpečení pro podnikové vyhledávání . . . . . 273  
    Zákaz zabezpečení pro podnikové aplikace na aplikačním serveru WebSphere . . . . . 274  
    Zakázání zabezpečení na úrovni dokumentu . . . . . 275  
    Zákaz zabezpečení pro sbalené výsledky vyhledávání 276

## **Spouštění a zastavování systému podnikového vyhledávání . . . . . 279**

Spuštění systému podnikového vyhledávání . . . . . 279  
Zastavení systému podnikového vyhledávání . . . . . 281  
Řízení spouštěných a zastavovaných komponent . . . . . 282  
Správa vyhledávacích serverů v samostatném režimu . . . . . 283

## **Monitorování aktivity podnikového vyhledávání . . . . . 285**

Odhad počtu dokumentů v kolekci . . . . . 285  
Monitorování kolekce . . . . . 286  
Zobrazení podrobností o identifikátoru URI . . . . . 286  
Monitorování prolézacích modulů . . . . . 288  
Zobrazení podrobných informací o aktivitě webového prolézacího modulu . . . . . 289

Podrobnosti o podprocesech webového prolézacího modulu. . . . . 290  
Aktivní weby webového prolézacího modulu . . . . . 290  
Frekvence prolézání webového prolézacího modulu . . . . . 291  
Vytváření sestav webového prolézacího modulu . . . . . 292  
Stavové kódy HTTP vrácené webovému prolézacímu modulu. . . . . 293  
Monitorování modulu analýzy . . . . . 297  
Monitorování aktivity indexu pro kolekci . . . . . 298  
Monitorování fronty indexů podnikového vyhledávání . . . . . 298  
Monitorování vyhledávacích serverů . . . . . 299  
    Změna způsobu výpočtu statistik dotazů . . . . . 300  
Monitorování modulu pro příjem dat . . . . . 301  
Sledování dokumentů . . . . . 302  
    Konfigurování souborů žurnálu pro sledování dokumentů . . . . . 302  
    Zobrazení sestav odstraněných dokumentů . . . . . 303  
    Zobrazení souborů žurnálu odstraněných dokumentů 304

## **Soubory žurnálu a výstrahy . . . . . 305**

Výstrahy . . . . . 305  
    Konfigurování výstrah na úrovni kolekce . . . . . 306  
    Konfigurování výstrah na úrovni systému . . . . . 307  
Konfigurování souborů žurnálu . . . . . 308  
    Konfigurování informací o serveru SMTP . . . . . 309  
    Příjem e-mailů o protokolovaných zprávách . . . . . 310  
    Změna velikostí protokolu dotazů . . . . . 311  
Zobrazení souborů žurnálu . . . . . 312

## **Zálohování a obnovování systému podnikového vyhledávání . . . . . 315**

Zálohování systému podnikového vyhledávání . . . . . 316  
Obnovení systému podnikového vyhledávání . . . . . 317  
Export a import konfigurací kolekce . . . . . 318

## **Integrace s produktem Lotus Notes verze 8 . . . . . 321**

Vytváření webu aktualizace modulu plug-in podnikového vyhledávání . . . . . 321  
Instalace modulu plug-in podnikového vyhledávání v klientovi produktu Lotus Notes verze 8 . . . . . 321

## **Integrace s produktem WebSphere Portal . . . . . 323**

Instalační skripty integrace podnikového vyhledávání do portálu WebSphere Portal . . . . . 324  
Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 5.1 . . . . . 325  
    Konfigurování vyhledávacího panelu produktu WebSphere Portal verze 5.1 pro práci s podnikovým vyhledáváním . . . . . 327  
    Odebrání podnikového vyhledávání z produktu WebSphere Portal verze 5.1 . . . . . 329  
Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 6 . . . . . 330  
    Konfigurování vyhledávacího centra produktu WebSphere Portal verze 6 pro podnikové vyhledávání . 332

Konfigurování vyhledávacího panelu produktu WebSphere Portal verze 6 pro práci s podnikovým vyhledáváním . . . . .	334
Nastavení portletu podnikového vyhledávání pro produkt Lotus Quickr . . . . .	335
Odebrání podnikového vyhledávání z produktu WebSphere Portal verze 6 . . . . .	336
Integrace podnikového vyhledávání s klastrovanými systémy produktu WebSphere Portal . . . . .	337
Nastavení podnikového vyhledávání v klastrovém systému WebSphere Portal . . . . .	338
Odebrání podnikového vyhledávání z klastrovaného systému WebSphere Portal . . . . .	341

**Migrace z produktu WebSphere Portal do podnikového vyhledávání . . . . . 343**

Migrace kolekce z produktu WebSphere Portal . . . . .	343
Migrované nastavení kolekce . . . . .	344
Soubor žurnálu průvodce migrací . . . . .	346

**Příkazy podnikového vyhledávání, návratové kódy a ID relací . . . . . 347**

**Rozlišování malých a velkých písmen v podnikovém vyhledávání. . . . . 377**

**Dokumentace k podnikovému vyhledávání . . . . . 379**

**Funkce usnadnění přístupu . . . . . 381**

**Slovníček pojmů z oblasti podnikového vyhledávání . . . . . 383**

**Doložky a ochranné známky . . . . . 395**

Poznámky . . . . .	395
Ochranné známky . . . . .	397

**Rejstřík . . . . . 399**



---

## ibm.com a související zdroje informací

Podpora a dokumentace k produktu je k dispozici na webu [ibm.com](http://www.ibm.com).

### Podpora

Podpora produktu je k dispozici na webu.

#### 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>

### Informační centrum

Dokumentaci k produktu můžete zobrazit pomocí webového prohlížeče v Informačním centru založeném na technologii Eclipse. Informační centrum je k dispozici na adrese <http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/>.

### Příručky ve formátu PDF

Soubory PDF lze zobrazit online pomocí programu Adobe Acrobat Reader pro příslušný operační systém. Pokud nemáte program Acrobat Reader instalován, můžete jej stáhnout z webu společnosti Adobe na adrese <http://www.adobe.com>.

Příručky ve formátu PDF naleznete na následujících webech:

Produkt	Adresa webu
OmniFind Enterprise Edition, verze 8.5	<a href="http://www.ibm.com/support/docview.wss?rs=63&amp;uid=swg27010938">http://www.ibm.com/support/docview.wss?rs=63&amp;uid=swg27010938</a>
OmniFind Discovery Edition, verze 8.4	<a href="http://www.ibm.com/support/docview.wss?rs=3035&amp;uid=swg27008552">http://www.ibm.com/support/docview.wss?rs=3035&amp;uid=swg27008552</a>
OmniFind Yahoo! Edition, verze 8.4	<a href="http://www.ibm.com/support/docview.wss?rs=3193&amp;uid=swg27008932">http://www.ibm.com/support/docview.wss?rs=3193&amp;uid=swg27008932</a>

---

## Odeslání vašich komentářů

Váš názor je důležitý v naší snaze poskytovat co nejpřesnější a nejkvalitnější informace.

Odešlete své komentáře pomocí formuláře pro komentáře čtenářů na adrese [https://www14.software.ibm.com/webapp/iwm/web/signup.do?lang=en\\_US&source=swg-rcf](https://www14.software.ibm.com/webapp/iwm/web/signup.do?lang=en_US&source=swg-rcf).

---

## Kontaktování společnosti IBM

Chcete-li kontaktovat zákaznický servis společnosti IBM v USA nebo v Kanadě, zavolejte na číslo 1-800-IBM-SERV (1-800-426-7378).

Chcete-li získat další informace o servisu k dispozici, volejte na jedno z následujících telefonních čísel:

- V USA: 1-888-426-4343
- V Kanadě: 1-800-465-9600

Více informací o možnostech kontaktování společnosti IBM naleznete na webové stránce Kontaktujte IBM na adrese <http://www.ibm.com/contact/cz/>.

---

## Co je podnikové vyhledávání?

System podnikového vyhledávání poskytuje rozsáhlé možnosti prohledávání libovolného počtu strukturovaných i nestruturovaných zdrojů dat jediným dotazem. Díky rychlé odezvě na dotazy a konsolidovaným výsledným sadám založeným na rozsáhlé analýze textu nemusíte pouze vyhledávat dokumenty, které vás zajímají, ale také extrahovat části obsahu dokumentu podle významu.

Komponenty podnikového vyhledávání instalované spolu s produktem IBM OmniFind Enterprise Edition shromažďují informace z celého podniku. Zadáním dotazu do webového prohlížeče můžete současně prohledávat lokální i vzdálené databáze, systémy spolupráce, systémy správy obsahu, souborové systémy a interní i externí webové stránky.

System podnikového vyhledávání, navržený tak, aby byl schopen bezproblémové integrace s existujícími systémy, zajišťuje veškeré organizační operace potřebné k shromažďování dat z různých zdrojů a indexování těchto dat pro rychlé načítání. Díky použití lingvistické analýzy a dalších typů analýzy dat může podnikové vyhledávání poskytovat vysoce využitelné výsledky vyhledávání. Uživatel, který potřebuje prohledávat více typů datových úložišť, se navíc nemusí seznamovat s různými rozhraními.

Můžete přidat podporu pro prohledávání zdrojů dat, které nechcete zahrnout do indexu podnikového vyhledávání. Díky možnosti federovaného vyhledávání v rámci podnikového vyhledávání můžete tyto externí zdroje prohledávat v době prohledávání indexovaných zdrojů dat.

### Kvalita vyhledávání

Aby produkt OmniFind Enterprise Edition mohl zaručit, že uživatelé najdou hledané informace, podporuje architekturu UIMA (Unstructured Information Management Architecture) společnosti IBM. Architektura UIMA je otevřená struktura, která definuje standardní společné rozhraní pro moduly analyzující text. Při rozsáhlé sémantické analýze může podnikové vyhledávání identifikovat pojmy, skryté významy, vztahy, fakta a další relevantní data, která jsou v nestruturovaném textu často skryta. Informace extrahované během analýzy lze využít ke zvýšení kvality výsledků vyhledávání nebo ke zvýšení kvality jiných aplikací, například obchodní inteligence nebo dolování dat.

### Zabezpečení

Nedílnou součástí podnikového vyhledávání je zabezpečení. Administraci systému mohou provádět pouze uživatelé s příslušnou autorizací. Mechanismy zabezpečení dostupné u aplikačního serveru IBM WebSphere Application Server vám umožňují konfigurovat administrativní role a řídit, kteří uživatelé budou mít přístup k různým funkcím správy.

Pomocí speciálních voleb můžete přiřadit prvky zabezpečení ke shromažďovaným datům. Pokud vaše vyhledávací aplikace podporují zabezpečení, můžete pomocí těchto prvků, které jsou ukládány spolu s dokumenty v indexu, vynutit řízení přístupu a omezit možnost zadávání dotazů na data a prohlížení výsledků vyhledávání pouze na uživatele s patřičným pověřením.

Pro určité typy zdrojů dat lze konfigurovat volby ověřování přihlašovacích údajů pověřením uživatele s použitím aktuálního řízení přístupu během zpracování dotazu. Tato dodatečná vrstva zabezpečení zaručuje ověření uživatelských oprávnění s nativním zdrojem dat v reálném čase. Tato volba může chránit před instancemi, v nichž se údaje pověřením uživatele změni po provedení indexace dokumentu a příslušných prvků zabezpečení.

## Výukový program k produktu

Výukový program online je k dispozici na webu <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html>. Výukový program popisuje jednotlivé kroky instalace a konfigurace, zobrazuje způsob prohledávání různých typů zdrojů dat a uvádí, jak lze použít rozhraní API produktu k rozšíření funkcí podnikového vyhledávání. Výukový program je určen pro starší verzi produktu OmniFind Enterprise Edition, ale většinu koncepcí a postupů lze stále použít.

### Související pojmy

“Přehled komponent podnikového vyhledávání” na stránce 3

“Datový tok podnikového vyhledávání” na stránce 11

Zabezpečení podnikového vyhledávání



Integrace vlastní analýzy textu



Základní koncepce používané při zpracování analýzy textu

### Související odkazy

“Typy zdrojů dat podporované podnikovým vyhledáváním”

---

## Typy zdrojů dat podporované podnikovým vyhledáváním

K dispozici je předdefinovaná podpora prohledávání zdrojů dat mnoha různých typů.

Po instalaci produktu IBM OmniFind Enterprise Edition můžete zahájit shromažďování dat ze zdrojů dat následujících typů:

- typy položek aplikace IBM DB2 Content Manager (dokumenty, prostředky a položky),
- databáze IBM DB2,
- databáze IBM Domino Document Manager (dříve Domino.Doc),
- databáze IBM Lotus Notes,
- databáze IBM Lotus QuickPlace,
- knihovny obsahu IBM Lotus Quickr,
- úložiště produktu IBM WebSphere Information Integrator Content Edition, včetně produktů Documentum, FileNet Panagon Content Services, FileNet P8 Content Manager, Hummingbird Document Management (DM), Microsoft SharePoint, OpenText Livelink Enterprise Server a WebSphere Portal Document Manager (PDM),
- tabulky přezdívek produktu IBM WebSphere Information Integrator pro řadu typů databázových systémů, včetně systémů IBM DB2 for z/OS, IBM Informix, Microsoft SQL Server, Oracle a Sybase,
- weby IBM WebSphere Portal,
- weby IBM Workplace Web Content Management,
- veřejné složky na serveru Microsoft Exchange Server,
- databáze Microsoft SQL Server,
- souborové systémy Microsoft Windows,
- diskusní skupiny NNTP (Network News Transfer Protocol),
- databáze Oracle,
- souborové systémy UNIX,
- weby v síti Internet nebo v místním intranetu.

Můžete také přidat podporu pro prohledávání následujících typů externích zdrojů, aniž by dokumenty z těchto zdrojů byly přidávány do indexu podnikového vyhledávání:

- Databáze podporující protokol JDBC (Java Database Connectivity) (pouze databázové systémy DB2 a Oracle). Pro každou tabulku, pro kterou je povoleno vyhledávání, je vytvořen samostatný externí zdroj.
- Servery LDAP (Lightweight Directory Access Protocol). Pro každý server LDAP je vytvořen jeden externí zdroj.

Nejnovější informace o podporovaných typech zdrojů dat a podporovaných verzích produktů naleznete na stránce se systémovými požadavky na webu podpory produktu OmniFind Enterprise Edition.

#### **Související pojmy**

Podpora externích zdrojů

“Datový tok podnikového vyhledávání” na stránce 11

---

## **Přehled komponent podnikového vyhledávání**

Komponenty podnikového vyhledávání shromažďují data z celého podniku, analyzují a kategorizují získané informace a vytvářejí indexy, které mohou prohledávat uživatelé.

*Kolekce* podnikového vyhledávání reprezentuje množinu zdrojů, které mohou uživatelé prohledávat pomocí jediného dotazu. Při vytvoření kolekce určíte, které zdroje do ní mají být zahrnuty, a nastavíte možnosti prohledávání indexovaných dat uživateli.

Můžete vytvořit více kolekcí, přičemž každá kolekce může obsahovat data z různých zdrojů dat. Vytvořená kolekce může zahrnovat například dokumenty z databázi IBM DB2 Universal Database, IBM Lotus Notes a IBM DB2 Content Manager. Výsledky prohledávání této kolekce mohou zahrnovat dokumenty ze všech těchto zdrojů dat.

Podpora federovaného vyhledávání umožňuje uživatelům prohledávat jediným dotazem více než jednu kolekci. Výsledky vyhledávání pak mohou zahrnovat dokumenty ze všech kolekcí a externích zdrojů v systému podnikového vyhledávání.

Vytvoření a správa kolekce zahrnuje následující činnosti:

#### **Sběr dat**

Komponenty nazývané *prolézací moduly* shromažďují dokumenty obsažené ve zdrojích dat, a to nepřetržitě nebo podle plánu, který nastavíte. Časté prolézání zaručuje uživatelům trvalý přístup k nejaktuálnějším informacím.

#### **Analýza dat**

Komponenty nazývané *analyzátoři* extrahují text z dokumentů a provádějí lingvistickou analýzu a další typy analýzy všech dokumentů připravených prolézacím modulem. Podrobná analýza obsahu zvyšuje kvalitu výsledků vyhledávání.

#### **Indexování dat**

Komponenty *indexu*, spouštěné podle plánu v pravidelných intervalech, přidávají do indexu informace o nových a změněných dokumentech. Komponenty indexu provádějí rovněž globální analýzu dokumentů v kolekci, což dále zvyšuje kvalitu výsledků vyhledávání.

#### **Vyhledávání dat**

*Vyhledávací* komponenty prohledávají index a ve spolupráci s vyhledávacími aplikacemi zpracovávají vyhledávací požadavky a vracejí výsledky vyhledávání.

Další komponenty produktu OmniFind Enterprise Edition vám umožňují zadat předvolby zabezpečení, monitorovat aktivitu systému a řešit vzniklé problémy. Součástí produktu je také funkční ukázková vyhledávací aplikace, kterou můžete použít jako šablonu při vytváření vlastních vyhledávacích aplikací.

#### **Související pojmy**

Administrace prolézacích modulů

Administrace analyzátorů

Administrace indexů

Administrace vyhledávacího serveru

“Datový tok podnikového vyhledávání” na stránce 11

## **Prolézací moduly podnikového vyhledávání**

Prolézací moduly podnikového vyhledávání shromažďují dokumenty obsažené ve zdrojích dat a připravují je pro analýzu, indexování a prohledávání.

Komponenta prolézacího modulu dodávaná s produktem OmniFind Enterprise Edition plní následující funkce:

- Při konfigurování prolézacího modulu shromáždí *zjišťovací* procesy informace o zdrojích dostupných pro prolézání, například názvy všech pohledů a složek v databázi Lotus Notes nebo názvy všech souborových systémů na serveru UNIX.
- Po výběru zdrojů, které chcete prolézat, a spuštění prolézacího modulu shromažďují komponenty prolézacího modulu data obsažená ve zdrojích a připravují je pro analýzu a indexování.

V jedné kolekci může existovat více prolézacích modulů a každý z těchto prolézacích modulů je uzpůsoben k shromažďování dat z určitého typu zdrojů dat. Chcete-li například v jedné kolekci zkombinovat data ze souborových systémů, z databází Notes a z relačních databází, vytvoříte tři prolézací moduly. Také můžete vytvořit více prolézacích modulů stejného typu a nastavit pro ně různé plány prolézání podle toho, jak často se mění data zpracovávaná jednotlivými prolézacími moduly.

Webový prolézací modul po spuštění pracuje nepřetržitě. Poté, co zadáte adresy URL, které chcete prolézat, se prolézací modul k těmto zdrojům pravidelně vrací a hledá nová nebo změněná data. Další typy prolézacích modulů můžete spouštět a zastavovat ručně nebo pro ně nastavit plány prolézání. Naplánováním prolézacího modulu určíte, kdy bude poprvé spuštěn a jak často musí znovu navštěvovat zdroje dat a prolézat nové a změněné dokumenty.

*Vlastnosti prolézacího modulu* jsou sady pravidel, která určují chování určitého prolézacího modulu při jeho činnosti. Můžete například nastavit pravidla určující způsob, jakým prolézací modul využívá systémové prostředky. Množina zdrojů určených k prolézání tvoří *prolézáný prostor* daného prolézacího modulu. Po vytvoření prolézacího modulu můžete jeho vlastnosti kdykoli upravit a změnit tak způsob, jakým prolézací modul shromažďuje data. Kromě toho můžete upravit prolézáný prostor a změnit plán prolézání, přidat nové zdroje nebo odebrat zdroje, které již nechcete prohledávat.

#### **Související pojmy**

Administrace prolézacích modulů

#### **Související úlohy**

Monitorování prolézacích modulů

## **Analyzátoři podnikového vyhledávání**

Analyzátor podnikového vyhledávání provádí analýzu dokumentů shromážděných prolézacím modulem a připravuje je pro indexování.

Komponenta analyzátoru dodávaná s produktem OmniFind Enterprise Edition analyzuje obsah dokumentů a metadata v dokumentech. Výsledky analýzy ukládá do datového skladu, kde k nim má přístup komponenta indexu. Analyzátor provádí následující úlohy:

- Extrahuje text z formátu, ve kterém je v dokumentu uložen. Může například extrahovat text ze značek v dokumentech XML a HTML. Díky využití technologie Outside In Content Access společnosti Oracle (dříve Stellent) pro produkt IBM OmniFind Enterprise Edition dokáže analyzátor extrahovat text také z binárních formátů, které používají například dokumenty aplikace Microsoft Word a dokumenty PDF aplikace Adobe Acrobat.
- Detekuje kódování znakové sady použité v jednotlivých dokumentech. Tyto informace použije ještě před zahájením lingvistické analýzy k převedení veškerého textu do kódování Unicode.
- Detekuje zdrojový jazyk jednotlivých dokumentů.
- Extrahuje text a přidává prvky zlepšující možnosti načtení dat. V této fázi provádí analyzátor následující operace:
  - Normalizace znaků, například normalizace velkých a malých písmen a diakritických značek, jako je například čárka, háček nebo kroužek.
  - Analýza struktury odstavců, vět, slov a mezer. Pomocí prostředků lingvistické analýzy analyzátor rozloží složená slova a nastaví přiřazení prvků dovolující vyhledávat ve slovnících a synonymech.
- Aplikuje pravidla analýzy, která jste pro kolekci nastavili. Při konfigurování analyzátoru můžete konfigurovat následující aktivity analýzy:

#### **Pravidla mapování polí pro dokumenty XML a HTML**

Tato volba umožňuje uživatelům prohledávat strukturovaný i nestruturovaný obsah dokumentů XML a HTML. Namapujete-li prvky XML nebo prvky metadat HTML na vyhledávací pole v indexu podnikového vyhledávání, uživatelé budou moci v dotazu zadat názvy polí a prohledávat konkrétní části dokumentů XML a HTML. Dotazy, které prohledávají konkrétní pole, mohou vracet přesnější výsledky než volné textové dotazy prohledávající celý obsah dokumentu.

#### **Categories**

Tato volba umožňuje uživatelům prohledávat dokumenty podle kategorií, do nichž tyto dokumenty patří. Uživatelé mohou také vybrat kategorie ve výsledcích vyhledávání a procházet pouze dokumenty, které patří do stejné kategorie.

Při vytváření kolekce můžete zvolit, zda chcete používat kategorizaci. V případě používání kategorií *založených na pravidlech* jsou dokumenty přiřazeny ke kategoriím na základě pravidel, která definujete. Kategorie založené na pravidlech lze konfigurovat v kolekcích podnikového vyhledávání, které sami vytvoříte, a v kolekcích, které migrujete z produktu IBM WebSphere Portal.

#### **Vlastní analýza textu**

Vývojáři aplikací mohou vytvořit vlastní analytické programy, které budou provádět složitou lingvistickou analýzu dat určených k prohledávání. Tyto programy lze zapojit do systému podnikového vyhledávání a používat k anotaci obsahu kolekcí. Indexováním anotací umožníte sémantické prohledávání kolekcí.

Uživatelé mohou vyhledávat například dotazovací výrazy, které se v textu vyskytují blízko sebe nebo v jedné větě, nebo vztahy mezi výrazy v dotazu. Uživatelé mohou například požadovat dokumenty zabývající se prodejcem společnosti IBM, který se jmenuje Smith, nikoli technickým pracovníkem společnosti IBM stejného jména.

#### **Podpora n-gramové segmentace**

Chcete-li zlepšit možnosti načítání dokumentů napsaných v čínštině, japonštině či korejštině, můžete povolit metodu lexikální analýzy nazývanou n-gramová

segmentace. Tato analytická metoda nepoužívá k oddělování slov prázdný prostor. Po vytvoření kolekce již nelze metodu segmentace změnit.

### **Podpora prohledávání dokumentů XML s použitím nativního vyhledávání XML**

Nativní vyhledávání XML může poskytovat přesnější výsledky vyhledávání díky prohledávání struktury značek XML. V dotazu může být například uvedeno, že se slovo musí vyskytovat v určitém prvku jazyka XML.

### **Třídy pro upřednostnění skóre relativní důležitosti polí**

Mapováním polí na třídy upřednostnění můžete ovlivnit řazení dokumentů ve výsledcích vyhledávání. Může být například vhodné zvýšit skóre polí s názvem a zajistit tak, že v případě výskytu výrazu dotazu v názvu budou mít dokumenty, které příslušný výraz obsahují v názvu, ve výsledcích vyhledávání vyšší ohodnocení důležitosti.

### **Související pojmy**

Práce s kategoriemi

vyhledávací pole XML

vyhledávací pole HTML



Integrace vlastní analýzy textu



Analýza textu integrovaná v podnikovém vyhledávání

### **Související úlohy**

Monitorování modulu analýzy

## **Indexy podnikového vyhledávání**

Indexovací komponenty podnikového vyhledávání, spouštěné podle plánu v pravidelných intervalech, přidávají do indexu informace o nových a změněných dokumentech.

Aby byl zajištěn trvalý přístup uživatelů k nejnovějším informacím z prohledávaných zdrojů, probíhá sestavení indexu ve dvou fázích:

### **Sestavování hlavního indexu**

Během sestavování hlavního indexu je celý index sestaven znovu, takže jeho struktura je po reorganizaci optimální. Procesy indexování načtou všechna data shromážděná prolézacími moduly a analyzovaná modulem analýzy.

### **Sestavování rozdílových indexů**

Při sestavování rozdílového indexu jsou do indexu přidány informace získané prolézáním od posledního sestavení hlavního indexu.

Při konfigurování voleb indexu pro kolekci můžete určit časové plány sestavování hlavních a rozdílových indexů. Četnost sestavování indexu závisí na systémových prostředcích a na tom, zda indexované zdroje obsahují statická nebo dynamická data.

V zájmu zajištění dostupnosti nových informací naplánujte co nejčastější sestavování rozdílového indexu. Plán pravidelného sestavování hlavního indexu zajistí konsolidaci všech nových informací, analýzu nového obsahu a optimalizaci výkonu indexu.

Procesy indexování lze spustit také neplánovaně. Změníte-li například některá pravidla analýzy a chcete-li tyto změny zpřístupnit vyhledávacím aplikacím, můžete spustit sestavování hlavního indexu ihned po novém prolezení a analýze dat bez čekání na nejbližší spuštění sestavování indexu v naplánovaném čase.

Při řízení využití prostředků určujete, kolik kolekcí může sdílet procesy indexování a souběžně odesílat požadavky na sestavení indexu. Při souběžném sestavování indexů se



sníží nebezpečí, že sestavování velmi rozsáhlého hlavního indexu zablokuje sestavování rozdílových indexů pro jiné kolekce. Sestavování indexu může klást vysoké nároky na systémové prostředky, takže u velkých systémů je nutné monitorovat zatížení systému a podle potřeby upravovat časové plány sestavování hlavního a rozdílového indexu.

Při sestavování indexu provádějí indexovací procesy globální analýzu dokumentů. V této fázi je použit algoritmus, který identifikuje duplicitní dokumenty, analyzuje strukturu odkazů v dokumentech a provádí speciální zpracování kotvicích textů (textů, které popisují cílovou stránku v hypertextovém odkazu) ve webových dokumentech.

Můžete nastavit volby pro následující aktivity indexování:

- Chcete-li uživatelům povolit zadávání zástupných znaků, můžete podporu rozšiřování výrazů v dotazu zabudovat do indexu nebo určit, že dotazovací výrazy budou doplňovány během zpracování dotazu. Při rozhodování o tom, kterou variantu použijete, je třeba brát ohled na protichůdné požadavky na využití prostředků a dobu odezvy dotazů.
- V případě potřeby můžete konfigurovat obory. *Obor* umožňuje omezit v kolekci rozsah dat, která jsou viditelná pro uživatele. Můžete například vytvořit jeden obor, který obsahuje identifikátory URI dokumentů oddělení technické podpory, a jiný obor zahrnující identifikátory URI dokumentů personálního oddělení. Pokud vyhledávací aplikace podporuje obory, mohou uživatelé prohledávat a načítat dokumenty omezené na dané části kolekce.
- Pro dokumenty se stejnou předponou identifikátoru URI lze nastavit volby sbalení výsledků vyhledávání. Můžete rovněž zadat název skupiny a ve výsledcích vyhledávání společně sbalit výsledné dokumenty s více různými předponami identifikátoru URI.
- Po sestavení indexu můžete odebrat identifikátory URI, v jejichž prohledávání chcete uživatelům zabránit.

#### **Související pojmy**

Administrace indexů

Zástupné znaky v dotazech

Obory

Sbalené identifikátory URI

Hodnocení důležitosti dokumentů na základě vzorů identifikátoru URI

#### **Související úlohy**

Plánování sestavení indexu

Konfigurování souběžných sestavení indexu

Odebrání identifikátorů URI z indexu

Monitorování aktivity indexu pro kolekci

Monitorování fronty indexů podnikového vyhledávání

## **Vyhledávací servery podnikového vyhledávání**

Vyhledávací servery podnikového vyhledávání ve spolupráci s vyhledávacími aplikacemi zpracovávají dotazy, prohledávají index a vracejí výsledky vyhledávání.

Vyhledávací servery podnikového vyhledávání se instalují spolu s produktem OmniFind Enterprise Edition. Při konfigurování vyhledávacích serverů pro kolekci můžete zadat volby ovlivňující způsob prohledávání kolekce:

- Často se vyskytující vyhledávací dotazy mohou být ukládány do mezipaměti vyhledávání. Tato mezipaměť může zlepšit výkon při hledání a načítání dat.
- Můžete zadat výchozí jazyk pro prohledávání dokumentů v kolekci.
- Pokud vývojáři aplikací vytvoří vlastní slovníky, můžete tyto slovníky asociovat s jednotlivými kolekcemi.

- Když uživatel zadá dotaz na kolekci, která používá *slovník synonym*, budou do výsledků vyhledávání zahrnuty dokumenty obsahující synonyma výrazů uvedených v dotazu.
- Když uživatel zadá dotaz na kolekci, která používá *slovník zakázaných slov*, budou z dotazu před zpracováním zakázaná slova odebrána.
- Když uživatel zadá dotaz na kolekci, která používá *slovník slov pro upřednostnění*, bude důležitost dokumentů obsahujících slova pro upřednostnění zvýšena nebo snížena v závislosti na faktoru upřednostnění asociovaném s příslušným slovem ve slovníku.
- Jestliže předem vytipujete konkrétní dokumenty, které mají význam pro určité dotazy, můžete konfigurovat rychlé odkazy. *Rychlý odkaz* přiřazuje k určitým klíčovým slovům a frázím konkrétní identifikátor URI. Obsahuje-li dotaz některé z klíčových slov nebo frází uvedených v definici rychlého odkazu, bude ve výsledcích vyhledávání automaticky vrácen přiřazený identifikátor URI.

U konfigurace s více servery je ochrana před selháním k dispozici nejen na úrovni serveru, ale také na úrovni kolekcí. Přestane-li být kolekce z nějakého důvodu dostupná na jednom vyhledávacím serveru, budou dotazy na tuto kolekci automaticky směrovány na druhý vyhledávací server.

#### **Související pojmy**

Vyhledávací aplikace podnikového vyhledávání

Mezipaměti pro vyhledávání

Vlastní slovníky synonym

Vlastní slovníky zakázaných slov

Vlastní slovníky slov pro upřednostnění

Rychlé odkazy

#### **Související úlohy**

Monitorování vyhledávacích serverů

## **Konzola pro správu podnikového vyhledávání**

Konzola pro správu podnikového vyhledávání pracuje v prohlížeči, takže administrátoři k ní mají přístup kdykoli a z libovolného místa. Mechanismy zabezpečení zaručují, že administrační funkce mohou používat pouze autorizovaní uživatelé.

Konzola pro správu podnikového vyhledávání se instaluje na vyhledávací servery při instalaci produktu OmniFind Enterprise Edition.

Administrační konzola obsahuje průvodce, kteří vám pomohou s některými základními úlohami správy. Průvodce kolekcí vám například pomůže vytvořit kolekci a umožní vám uložit rozpracované změny v režimu konceptu. Průvodci prolézacími moduly se specializují na různé typy zdrojů dat a pomohou vám vybrat zdroje, jejichž prohledávání chcete uživatelům umožnit.

U ostatních úloh správy můžete vybrat jednotlivé položky, které chcete spravovat. Při úpravách kolekce můžete například vybrat stránku Index a změnit plán indexu nebo vybrat stránku Analýza a upravit pravidlo analýzy dokumentů XML.

#### **Související pojmy**

Správa systému podnikového vyhledávání

Administrativní role

#### **Související úlohy**

Přihlášení ke konzole pro správu

## Monitorování systému podnikového vyhledávání

Pomocí konzoly pro správu podnikového vyhledávání můžete monitorovat aktivity systému a upravit jeho provozní parametry podle potřeby.

Po instalaci produktu OmniFind Enterprise Edition a vytvoření alespoň jedné kolekce lze zobrazit podrobnou statistiku pro každý základní okruh aktivit (prolézání, analýza, indexování a hledání). Zobrazené informace zahrnují průměrnou dobu odezvy a informace o průběhu, například počet dokumentů, které byly prolezeny nebo indexovány v průběhu konkrétní relace prolézání nebo sestavování indexu.

Většinu aktivit lze zastavit a spustit. Můžete například pozastavit aktivitu, změnit její konfiguraci nebo vyřešit problém, a znovu zpracování spustit, když jste připraveni povolit pokračování aktivity.

K dispozici je rovněž možnost konfigurování výstrah, tj. e-mailů s informacemi o určitých aktivitách, zasílaných při každém výskytu monitorované události. Výstrahu můžete obdržet například v případě, že doba odezvy vyhledávání překročí nastavený práh.

Pokud byl dokument odstraněn ze systému podnikového vyhledávání, můžete jej sledovat a určit, kdy, kde a proč byl dokument odstraněn. Může se například stát, že prolézací modul nemůže dokument prolézt nebo že administrátor odebere dokument z indexu.

### Související pojmy

Monitorování aktivity podnikového vyhledávání

Spouštění a zastavování systému podnikového vyhledávání

## Soubory žurnálu podnikového vyhledávání

Soubory žurnálu jsou vytvářeny pro jednotlivé kolekce a pro relace na úrovni systému.

Při konfigurování voleb žurnálu pro kolekce podnikového vyhledávání nebo pro systém zadáváte typy zpráv, které chcete zapisovat do žurnálu, například chybové zprávy a varovné zprávy. Můžete rovněž určit, jak často má systém znovu používat staré soubory žurnálu a uvolňovat tak místo pro novější zprávy. Můžete vybrat volby pro zasílání e-mailů informujících o výskytu určitých zpráv (včetně výstrah) nebo všech chybových zpráv.

Při prohlížení souborů žurnálu vyberete soubor žurnálu, který chcete zobrazit. Název souboru obsahuje informace o tom, kdy byl soubor vytvořen a která komponenta zprávy vygenerovala. Můžete rovněž specifikovat filtry zobrazení. Tímto způsobem lze zobrazit například pouze chybové zprávy nebo pouze zprávy vygenerované určitou relací podnikového vyhledávání.

### Související pojmy

Soubory žurnálu a výstrahy

Výstrahy

 Zprávy podnikového vyhledávání

### Související úlohy

Konfigurování souborů žurnálu

Konfigurování informací o serveru SMTP

Příjem e-mailů o protokolovaných zprávách

Zobrazení souborů žurnálu

## Přizpůsobení podnikového vyhledávání

Rozhraní API pro podnikové vyhledávání vám umožňují vytvářet vlastní vyhledávací aplikace, vlastní aplikace pro aktualizaci obsahu kolekcí, vlastní programy pro analýzu textu a vlastní slovníky synonym, zakázaných slov a slov pro upřednostnění.

Po instalaci produktu OmniFind Enterprise Edition máte k dispozici následující skupiny rozhraní API, pomocí nichž můžete rozšiřovat kolekce podnikového vyhledávání:

### Rozhraní API pro vyhledávání a indexování (SI-API)

Toto rozhraní API slouží k sestavování vlastních vyhledávacích aplikací a vlastního rozhraní pro správu.

### Moduly plug-in prolézacího modulu

Prostřednictvím rozhraní API modulu plug-in můžete do dokumentů při jejich prolézání přidávat metadata nebo s nimi asociovat prvky zabezpečení, které vynucují obchodní pravidla a pravidla zabezpečení příslušné organizace.

Možnosti načítání informací lze rozšířit integrováním vlastních programů pro lingvistickou analýzu do kolekcí podnikového vyhledávání. Po přidání vlastních strojů pro analýzu textu do systému můžete tyto stroje asociovat s kolekcemi. Když uživatelé zadávají dotazy na kolekce, mohou využívat asociace mezi slovy, které do indexu zařadí vlastní programy. Uživatelé tak mohou například vyhledávat koncepty a vztahy mezi výrazy, nejen samotné výrazy.

Dostupnost informací lze zlepšit také integrací vlastních slovníků, které odrážejí například zkratky a odborné výrazy spadající do oboru vašeho podnikání. Po přidání slovníků do systému můžete slovníky asociovat s jednotlivými kolekcemi. Když uživatelé zadávají dotazy na kolekce, jsou slovníky užitečné následujícím způsobem:

- Pokud dotaz obsahuje výrazy, které jsou definované jako synonyma, budou ve výsledcích vyhledávání zahrnuty dokumenty obsahující synonyma výrazů uvedených v dotazu.
- Pokud dotaz obsahuje zakázaná slova, budou zakázaná slova z dotazu odebrána, aby ve výsledcích vyhledávání nebyly vráceny irelevantní dokumenty.
- Pokud dotaz obsahuje slova pro upřednostnění, budou mít dokumenty, které obsahují slova pro upřednostnění, ve výsledcích vyhledávání vyšší nebo nižší ohodnocení důležitosti v závislosti na hodnotě pro upřednostnění asociované s příslušným slovem ve slovníku.

#### Související pojmy

Vyhledávací aplikace podnikového vyhledávání

Vlastní slovníky synonym

Vlastní slovníky zakázaných slov

Vlastní slovníky slov pro upřednostnění



Vyhledávací a indexové rozhraní API - přehled



Moduly plug-in prolézacího modulu

## Ukázková vyhledávací aplikace pro podnikové vyhledávání

Ukázkovou vyhledávací aplikaci pro podnikové vyhledávání můžete použít jako šablonu při vývoji vlastních vyhledávacích aplikací.

Ukázková vyhledávací aplikace se instaluje současně s produktem OmniFind Enterprise Edition. Tato aplikace demonstruje většinu funkcí vyhledávání a načítání, které jsou v prostředí podnikového vyhledávání k dispozici. Zároveň tato aplikace tvoří funkční příklad, který vám umožňuje prohledávat všechny aktivní kolekce a externí zdroje v systému podnikového vyhledávání. Pomocí ukázkové aplikace můžete otestovat nové kolekce a externí zdroje ještě předtím, než je zpřístupníte uživatelům.

Ukázková vyhledávací aplikace demonstruje podporu federovaného vyhledávání tím, že vám umožňuje prohledávat jednu nebo více kolekcí nebo externích zdrojů současně.

Pro určité typy prolézacích modulů můžete prostřednictvím komponenty pro správu identit podnikového vyhledávání ověřovat aktuální údaje pověření při přístupu uživatelů k vyhledávací aplikaci. Pokud je prohledávaná doména chráněna zabezpečením SSO (single sign-on), lze k ověřování uživatele v rámci relace vyhledávání používat mechanismy zabezpečení SSO. V opačném případě může komponenta pro správu identit šifrovat a ukládat údaje pověření uživatelů v profilu a používat je pro vyloučení zakázaných dokumentů z výsledků vyhledávání.

Chcete-li přizpůsobit ukázkovou vyhledávací aplikaci, můžete použít modul pro přizpůsobení vyhledávacích aplikací. Jde o grafické uživatelské rozhraní, které uživatelům umožňují zobrazovat účinek prováděných změn. Vyhledávací aplikaci lze rovněž přizpůsobit úpravou konfiguračního souboru pro aplikaci.

Chcete-li vytvořit vlastní vyhledávací aplikaci, použijte vyhledávací a indexové rozhraní API pro podnikové vyhledávání.

#### **Související pojmy**

Vyhledávací aplikace podnikového vyhledávání

Funkce ukázkové vyhledávací aplikace

 Vyhledávací a indexové rozhraní API - přehled

#### **Související úlohy**

Přístup k vyhledávacím aplikacím

Úpravy vlastností ukázkové vyhledávací aplikace

Přizpůsobení podnikových aplikací

---

## **Datový tok podnikového vyhledávání**

Komponenty podnikového vyhledávání, které instalujete spolu s produktem OmniFind Enterprise Edition, navzájem úzce spolupracují na zajištění datového toku uvnitř systému.

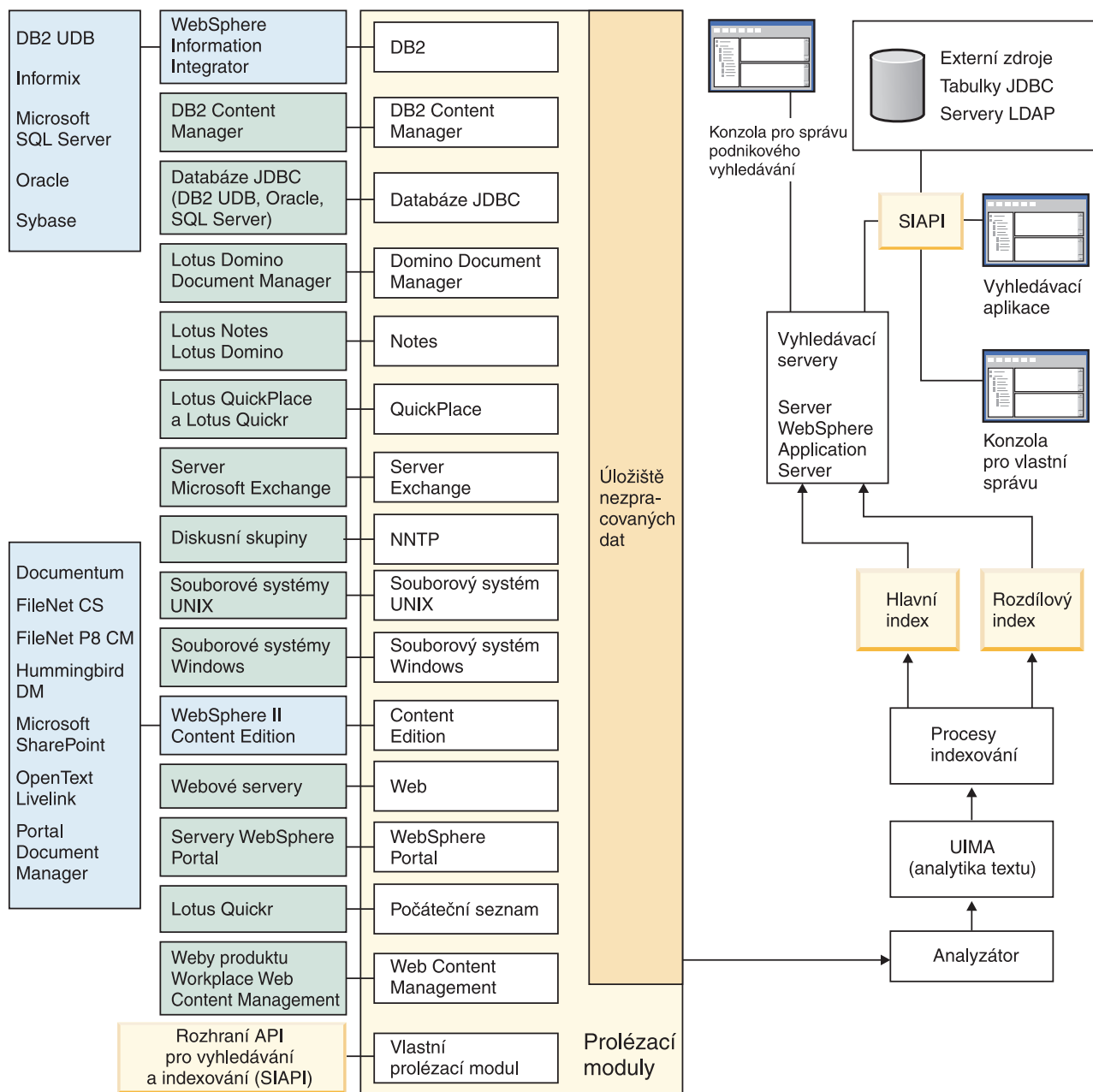
Prolézací moduly shromažďují dokumenty ze zdrojů dat v celém podniku. Analyzátor extrahuje z prolezených dokumentů užitečné informace a generuje prvky, které mohou například přiřazovat dokumenty ke kategoriím a pomáhat při určování důležitosti dokumentu z hlediska výrazů obsažených ve vyhledávacím požadavku. V indexu jsou data uložena tak, aby je bylo možné efektivně načítat.

Uživatelé prohledávají indexované kolekce a externí zdroje pomocí webového prohlížeče a vyhledávací aplikace. Vyhledávací aplikace může v prohlížeči zobrazit seznam výsledků, na něž mohou uživatelé klepnout. Složitější aplikace mohou vracet dynamicky generovaný obsah založený na informacích z různých zdrojů.

Aplikace prohledávání katalogu může například přizpůsobit zobrazení produktů vyhovujících vyhledávacímu požadavku. Jediný dotaz může prohledávat dokumenty z různých typů zdrojů dat, například kombinaci dokumentů produktu IBM DB2 Content Manager a dokumentů v úložištích IBM DB2 Content Manager a Lotus Notes.

Administrátoři určují, jaká data mají být shromažďována a jak má probíhat jejich prolézání, analýza, indexování a prohledávání. Při monitorování aktivity systému mohou administrátoři kromě toho také upravit nastavení tak, aby dosáhli optimální propustnosti dat.

Následující schéma znázorňuje tok informací v systému podnikového vyhledávání.



Obrázek 1. Tok dat v systému podnikového vyhledávání

### Související pojmy

“Co je podnikové vyhledávání?” na stránce 1

“Přehled komponent podnikového vyhledávání” na stránce 3

### Související odkazy

“Typy zdrojů dat podporované podnikovým vyhledáváním” na stránce 2

---

## Správa systému podnikového vyhledávání

Konzola pro správu podnikového vyhledávání umožňuje vytvářet a spravovat kolekce a externí zdroje, spouštět a zastavovat komponenty, monitorovat aktivity systému a soubory žurnálu, konfigurovat administrační uživatele, přiřazovat vyhledávací aplikace ke kolekcím a externím zdrojům a zadávat informace potřebné k zajištění zabezpečení.

**Tip:** Výukový program online je k dispozici na webu <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html>. Výukový program popisuje jednotlivé kroky instalace a konfigurace, zobrazuje způsob prohledávání různých typů zdrojů dat a uvádí, jak lze použít rozhraní API produktu k rozšíření funkcí podnikového vyhledávání. Výukový program je určen pro starší verzi produktu OmniFind Enterprise Edition, ale většinu koncepcí a postupů lze stále použít.


Podrobné příklady konfigurace prolézacích modulů a povolení zabezpečení v prostředí malých, středních i velkých organizací naleznete v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

### Pohled Kolekce

Podle následujících pokynů můžete vytvořit svou první kolekci a spravovat systém.

#### Přihlášení

Přihlaste se ke konzole pro správu podnikového vyhledávání. Pohled Kolekce představuje vstupní bod pro vytváření a správu kolekcí.

**Tip:** Potřebujete-li pomoc při práci s konzolou pro správu, můžete kdykoli klepnout na volbu **Nápověda** na panelu nástrojů nebo na volbu **Nápověda pro tuto stránku**. Jsou-li pro příslušné zprávy k dispozici podrobná vysvětlení a opravné akce, můžete klepnutím na ikonu  **Další informace** na konci zprávy zobrazit podrobnosti.


#### Vytvoření kolekce

Zvolte jeden z následujících přístupů při vytváření nové, prázdné kolekce:

- Chcete-li vytvořit kolekci pomocí průvodce kolekcí, klepněte na tlačítko **Průvodce kolekcí** a postupujte podle pokynů průvodce.
- Chcete-li vytvořit kolekci pomocí pohledu Kolekce, klepněte na tlačítko **Vytvořit kolekci**, vyplňte pole na stránce Vytvořit kolekci a poté klepněte na tlačítko **OK**.

#### Konfigurování kolekce

Novou prázdnou kolekci musíte upravit, vložit do ní obsah, nastavit volby prolézání dat a způsob, kterým chcete umožnit prohledávání dat.

Klepněte na tlačítko  **Upravit** u nové kolekce a poté vyberte stránku s volbami kolekce.

**Upozornění:** Chcete-li se vrátit k dříve zobrazené stránce nebo aktualizovat informace v konzole pro správu, klepněte na tlačítko **Předchozí** nebo **Aktualizovat** v konzole pro správu podnikového vyhledávání. Klepnete-li na tlačítko **Zpět** nebo **Aktualizovat** ve webovém prohlížeči, mohou ve výsledcích vzniknout nekonzistence a hrozí i ztráta dat.

- Na stránce **Obecné** můžete zadat volby, které budou platit pro celou kolekci:
  - Obecné volby upravte tehdy, chcete-li změnit název či popis kolekce nebo odhad její velikosti.

- Na této stránce se kromě toho zobrazují informace o kolekci, které nelze změnit, například ID kolekce nebo metoda statického hodnocení dokumentů ve výsledcích vyhledávání.
- Pokud bylo pro kolekci při jejím vytváření povoleno zabezpečení, lze zapnout nebo vypnout volby zabezpečení na úrovni dokumentu.
- Na stránce Prolézání konfiguruje alespoň jeden prolézací modul.  
Každá kolekce může obsahovat data pocházející z různých zdrojů dat. Pro každý typ zdroje dat, který chcete zahrnout, je nutné konfigurovat alespoň jeden prolézací modul. S konfigurací prolézacího modulu, který vytvoříte, vám pomůže průvodce určený speciálně pro zvolený typ prolézáných dat.
- Na stránce Analýza můžete nastavit konfiguraci voleb pro analýzu prolézáných dat tak, aby jejich prohledávání bylo efektivní:
  - Můžete určit, zda mají být dokumenty XML analyzovány tak, aby je bylo možné prohledávat pomocí nativního vyhledávání XML.
  - Dokumenty můžete přiřadit ke kategoriím, čímž umožníte uživatelům prohledávat podmnožiny kolekcí nebo procházet dokumenty ve výsledcích vyhledávání podle kategorií, do nichž patří.
  - Prvky XML a prvky metadat HTML můžete namapovat na vyhledávací pole v indexu, takže uživatelé budou moci v dotazu zadat názvy polí a prohledávat konkrétní části dokumentů.
  - Pokud jste do systému podnikového vyhledávání přidali vlastní stroje pro analýzu textu, můžete určit, který z nich chcete používat pro danou kolekci, a poté zadat volby zpracování textu, které zlepšují možnosti načítání informací a zajistí podporu sémantického vyhledávání.
  - Můžete asociovat pole s třídami upřednostnění a ovlivňovat tak ohodnocení polí, která odpovídají výrazům dotazu, ve výsledcích vyhledávání.
- Na stránce Index nastavte konfiguraci plánů sestavení indexu. Sestavování indexu naplánujte s dostatečnou frekvencí, aby uživatelé měli vždy přístup k nejnovějším informacím. Kromě toho můžete provádět následující volitelné akce:
  - Povolit uživatelům zadávání zástupných znaků ve dotazovacích výrazech.
  - Konfigurovat obory a umožnit tak uživatelům prohledávat pouze omezenou část kolekce namísto všech dokumentů v indexu.
  - Sbalit výsledky vyhledávání, takže dokumenty pocházející ze stejného zdroje budou ve výsledcích vyhledávání sbalené.
  - Přiřadit faktory upřednostnění a ovlivnit ohodnocení důležitosti dokumentů odpovídajících vzoru identifikátoru URI ve výsledcích vyhledávání.
  - Odebrání některých identifikátorů URI z indexu. Tímto způsobem můžete po vytvoření kolekce například bránit uživatelům v zobrazování některých dokumentů.
- Na stránce Vyhledávání můžete zadat volby prohledávání dokumentů v kolekci:
  - Pro výsledky vyhledávání můžete vyhradit prostor v mezipaměti. Dále můžete změnit výchozí jazyk kolekce.
  - Pokud jste do systému podnikového vyhledávání přidali vlastní slovníky synonym, zakázaných slov a slov pro upřednostnění, můžete vybrat slovníky, které mají být použity při prohledávání této kolekce uživateli.
  - Můžete zadat délku souhrnných údajů o dokumentech zobrazovaných ve výsledcích vyhledávání.
  - Chcete-li, aby se ve výsledcích vyhledávání automaticky objevily určité identifikátory URI vždy, když se v dotazu vyskytnou určitá klíčová slova nebo fráze, můžete konfigurovat rychlé odkazy.
- Na stránce Žurnál můžete provádět následující činnosti:



- Zadat volby pro typy zpráv, které chcete ukládat do žurnálu, a pro interval přepisování starých souborů žurnálu.
- Zadat volby pro odesílání výstrah při určitých aktivitách kolekce. Například vás může výstraha informovat, že průměrná doba odezvy při vyhledávání překračuje zadané omezení.
- Zadat volby pro odeslání e-mailu vždy, když se v žurnálu objeví určité zprávy nebo typy zpráv.
- Určit volby pro protokolování informací, na základě kterých lze určit čas, místo a důvod odstranění dokumentu ze systému podnikového vyhledávání.

## Spuštění komponent

Po zadání prolézáných zdrojů dat a voleb pro shromažďování a prohledávání dat můžete spustit proces sestavení kolekce. Při spouštění komponent je třeba dodržovat stanovené pořadí. Nejprve musí prolézací moduly zpracovat data, poté lze tato data analyzovat a indexovat. Stejně tak je nutné nejprve sestavit hlavní index a teprve poté mohou vyhledávací servery začít zpracovávat požadavky na vyhledávání.

## Pohled Externí zdroje

Chcete-li prohledávat zdroje dat, aniž by je bylo třeba prolézat či indexovat, můžete po klepnutí na volbu **Externí zdroje** na panelu nástrojů určit volby pro povolení prohledávání pro příslušné zdroje dat. Musíte zadat informace, které zpřístupní databáze JDBC (Java Database Connectivity) a servery LDAP (Lightweight Directory Access Protocol) pro podnikové vyhledávání. Po asociování externích zdrojů s vyhledávacími aplikacemi mohou uživatelé tyto zdroje prohledávat souběžně s prohledáváním kolekcí obsahujících data, která byla prolezena, analyzována a indexována.

## Pohled Systém

Jestliže jste členem role administrátora podnikového vyhledávání, můžete po klepnutí na tlačítko **Systém** provádět následující činnosti. Administrátoři kolekcí, operátoři a osoby pověřené monitorováním mají k tomuto pohledu přístup jen tehdy, pokud jim administrátor podnikového vyhledávání udělí příslušné oprávnění.

- Přidávat do systému stroje pro analýzu textu.
- Přidávat do systému vlastní slovníky synonym, zakázaných slov a slov pro upřednostnění.
- Nastavte počet kolekcí, pro které smějí být souběžně sestavovány indexy, a určete, zda může sestavování hlavního indexu pro jednu kolekci běžet současně se sestavováním rozdílových indexů.
- Konfigurovat výstrahy pro události na úrovni systému.
- Určovat volby protokolování zpráv generovaných relacemi na úrovni systému.
- Zadávat údaje o svém poštovním serveru, abyste mohli dostávat e-maily s informacemi o aktivitách podnikového vyhledávání.

## Pohled Zabezpečení


Jestliže jste členem role administrátora podnikového vyhledávání, můžete klepnout na tlačítko **Zabezpečení** a zadat volby zabezpečení. Administrátoři kolekcí, operátoři a osoby pověřené monitorováním nemají k tomuto pohledu přístup.

Pokud povolíte zabezpečení na aplikačním serveru WebSphere, můžete pomocí pohledu Zabezpečení konfigurovat administrativní role. Konfigurováním administrativních rolí můžete správu systému zpřístupnit více uživatelům, avšak omezit přístup jednotlivých uživatelů pouze na určité funkce a kolekce.

V pohledu Zabezpečení můžete také konfigurovat volby správy identit. Můžete například zadat volby pro ukládání pověřovacích údajů uživatelů do profilů, které lze použít k ověření aktuálních pověřovacích údajů uživatele během zpracování dotazu. Pokud je zdroj, který má být prohledán, chráněn zabezpečením SSO (single sign-on), můžete nastavit také volby použití ověřovacích metod SSO k ověření aktuálních pověřovacích údajů uživatele během zpracování dotazu.

Dokud nevytvoříte vlastní vyhledávací aplikace, můžete k prohledávání všech kolekcí a externích zdrojů používat ukázkovou vyhledávací aplikaci. Po vytvoření vlastní vyhledávací aplikace použijte pohled Zabezpečení k přiřazení aplikace ke kolekcím a externím zdrojům, které smí prohledávat.

## Pohled Monitorování

V kterémkoli okamžiku můžete klepnout na ikonu  **Monitor** a zahájit monitorování komponent systému nebo kolekce. Pokud to dovoluje vaše administrativní role, můžete také spouštět a zastavovat procesy komponent, které monitorujete.

### Související pojmy

“Administrace prolézacích modulů” na stránce 35

“Monitorování aktivity podnikového vyhledávání” na stránce 285

### Související úlohy

“Spuštění systému podnikového vyhledávání” na stránce 279

“Správa vyhledávacích serverů v samostatném režimu” na stránce 283

“Zastavení systému podnikového vyhledávání” na stránce 281

“Vytvoření kolekce pomocí Průvodce kolekcí” na stránce 29

“Vytvoření kolekce s použitím pohledu Kolekce” na stránce 30

---

## Přihlášení ke konzole pro správu

Chcete-li spravovat systém podnikového vyhledávání, musíte nejprve zadat adresu URL ve webovém prohlížeči a poté se přihlásit ke konzole pro správu.

### Než začnete

Musíte se přihlásit pod jménem uživatele, kterému je uděleno oprávnění pro přístup ke konzole pro správu podnikového vyhledávání:

- Pokud nepovolíte globální zabezpečení na aplikačním serveru WebSphere, má přístup ke konzole pro správu pouze administrátor podnikového vyhledávání, který byl určen při instalaci produktu OmniFind Enterprise Edition.
- Pokud povolíte globální zabezpečení na serveru WebSphere Application Server, můžete použít konzolu pro správu podnikového vyhledávání ke konfiguraci administrativních rolí. Konfigurovaná jména uživatelů musí existovat v registru uživatelů na serveru WebSphere Application Server. Konfigurováním administrativních rolí umožníte dalším uživatelům přihlásit se ke konzole pro správu, můžete však určit, ke kterým funkcím a kolekcím mají mít jednotliví administrativní uživatelé přístup.

### Postup

Chcete-li se přihlásit ke konzole pro správu podnikového vyhledávání, postupujte takto:

1. Zadejte adresu URL konzoly pro správu ve webovém prohlížeči. Příklad:

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

*SearchServer.com* je hostitelský název vyhledávacího serveru podnikového vyhledávání.

V závislosti na konfiguraci webového serveru může být rovněž třeba zadat číslo portu.

Příklad:

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

2. Na úvodní stránce zadejte své jméno uživatele a heslo a klepněte na tlačítko **Přihlásit**.

Zobrazí se pohled Kolekce tvořící vstupní bod ke správě systému a kolekci. Pokud používáte administrativní role, závisí akce, které můžete provádět, a kolekce, které uvidíte, na vaší administrativní roli.

Pokud je vaše relace po určitou dobu neaktivní, systém vás automaticky odhlásí. Chcete-li pokračovat ve správě systému, znovu se přihlaste.

Po dokončení správy kolekci se od konzoly můžete odhlásit klepnutím na tlačítko **Odhlásit**. Poté se můžete přihlásit s jiným ID a heslem nebo zavřít webový prohlížeč, čímž ukončíte práci konzoly pro správu.

#### **Související pojmy**

“Administrativní role” na stránce 246

---

## **Změna hesla administrátora podnikového vyhledávání v konfiguraci s jedním serverem**

Heslo administrátora podnikového vyhledávání je uloženo v šifrovaném formátu. Chcete-li toto heslo změnit, použijte skript **eschangepw**.

#### **Omezení**

Hesla mohou obsahovat následující speciální znaky:

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

Zadáte-li v systému AIX, Linux nebo Solaris heslo obsahující speciální znaky, musíte celé heslo uvést v apostrofech.

Příklad: 'mojehsl@\$\$%'

Zadáte-li heslo obsahující speciální znaky v systému Windows, musíte celé heslo uvést v uvozovkách.

Příklad: "moje?+!hsl"

#### **O této úloze**

Heslo pro počáteční jméno administrátora podnikového vyhledávání se nastavuje při instalaci produktu OmniFind Enterprise Edition.

Chcete-li toto heslo změnit, musíte spustit skript **eschangepw**, který zajistí rozšíření informace o změně do celého systému podnikového vyhledávání. Instalační program vytváří dvě proměnné prostředí, které můžete využít v kombinaci se skriptem **eschangepw**:

#### **ES\_INSTALL\_ROOT**

Instalační adresář podnikového vyhledávání.

#### **ES\_NODE\_ROOT**

Datový adresář podnikového vyhledávání. Heslo pro jméno administrátora podnikového vyhledávání se ukládá do souboru **es.cfg** v tomto adresáři.

## Postup

Chcete-li změnit heslo administrátora podnikového vyhledávání v prostředí konfigurace s jedním serverem, postupujte takto:

1. Přihlaste se jako administrátor podnikového vyhledávání a zastavte systém podnikového vyhledávání:

```
esadmin system stopall
```

**Důležité:** V době, kdy je systém zastaven, nemohou uživatelé zadávat vyhledávací požadavky.

2. Změňte systémové heslo odpovídající uživatelskému jménu administrátora podnikového vyhledávání pomocí příkazů operačního systému (AIX, Linux, Solaris) nebo pomocí prostředků operačního systému pro změnu hesla (Windows).
3. Spusťte následující skript, přičemž parametr *nová\_hodnota* označuje heslo zadané v kroku 2:

**Systém AIX, Linux nebo Solaris**

```
eschangepw.sh nová_hodnota
```

**Windows**

```
eschangepw nová_hodnota
```

4. Restartujte systém podnikového vyhledávání:

```
esadmin system startall
```

**Související odkazy**

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 347

---

## Změna hesla administrátora podnikového vyhledávání v konfiguraci s více servery

Heslo administrátora podnikového vyhledávání je uloženo v šifrovaném formátu. Chcete-li toto heslo změnit, použijte skript **eschangepw**, který je změni ve všech počítačích zapojených do systému podnikového vyhledávání.

### Omezení

Hesla mohou obsahovat následující speciální znaky:

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

Zadáte-li v systému AIX, Linux nebo Solaris heslo obsahující speciální znaky, musíte celé heslo uvést v apostrofěch.

Příklad: 'mojehsl@\$%'

Zadáte-li heslo obsahující speciální znaky v systému Windows, musíte celé heslo uvést v uvozovkách.

Příklad: "moje?+!hsl"

### O této úloze

Heslo pod ID administrátora podnikového vyhledávání, které se poprvé nastavuje při instalaci produktu OmniFind Enterprise Edition, musí být na všech serverech podnikového vyhledávání shodné.

Chcete-li toto heslo změnit a rozšířit informaci o změně do celého systému podnikového vyhledávání, musíte spustit skript **eschangepw** ve všech počítačích, které využíváte pro podnikové vyhledávání. Následující postup popisuje doporučené pořadí změn hesel na jednotlivých serverech. Toto pořadí nemusíte dodržet, musíte však provést všechny kroky vyžadované pro jednotlivé typy serverů.

Instalační program vytváří dvě proměnné prostředí, které můžete využít v kombinaci se skriptem **eschangepw**:

#### **ES\_INSTALL\_ROOT**

Instalační adresář podnikového vyhledávání.

#### **ES\_NODE\_ROOT**

Datový adresář podnikového vyhledávání. Heslo pro jméno administrátora podnikového vyhledávání se ukládá do souboru **es.cfg** v tomto adresáři.

### **Postup**

Chcete-li změnit heslo administrátora podnikového vyhledávání v prostředí konfigurace s více servery, postupujte takto:

1. Přihlaste se k indexovému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání a zastavte systém podnikového vyhledávání:  
`esadmin system stopall`

**Důležité:** V době, kdy je systém zastaven, nemohou uživatelé zadávat vyhledávací požadavky.

- a. Změňte systémové heslo odpovídající uživatelskému jménu administrátora podnikového vyhledávání pomocí příkazů operačního systému (AIX, Linux, Solaris) nebo pomocí prostředků operačního systému pro změnu hesla (Windows).
- b. Spusťte následující skript, přičemž parametr *nová\_hodnota* označuje heslo zadané v kroku 1a:

#### **Systém AIX, Linux nebo Solaris**

```
eschangepw.sh nová_hodnota
```

#### **Windows**

```
eschangepw nová_hodnota
```

2. V ostatních počítačích zapojených do systému podnikového vyhledávání postupujte takto:
  - a. Přihlaste se jako administrátor podnikového vyhledávání.
  - b. Zastavte službu CCL (Common Communication Layer) podnikového vyhledávání:

#### **Systém AIX, Linux nebo Solaris**

```
stopccl.sh
```

#### **Příkazový řádek systému Windows**

```
stopccl
```

#### **Nástroj pro správu Služby systému Windows**

- 1) Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
- 2) Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Zastavit**.
- c. Změňte systémové heslo odpovídající uživatelskému jménu administrátora podnikového vyhledávání pomocí příkazů operačního systému (AIX, Linux, Solaris) nebo pomocí prostředků operačního systému pro změnu hesla (Windows). Toto heslo se musí shodovat s heslem, které jste zadali v kroku 1a.

- d. Spusíte následující skript, přičemž parametr *nová\_hodnota* označuje heslo zadané v kroku 1a na stránce 19:

**System AIX, Linux nebo Solaris**  
`eschangepw.sh nová_hodnota`

**Windows**  
`eschangepw nová_hodnota`

- e. Znovu spusíte službu CCL:

**System AIX, Linux nebo Solaris**  
`startccl.sh -bg`


**Příkazový řádek systému Windows**  
`startccl`

**Nástroj pro správu Služby systému Windows**

Chcete-li spustit službu CCL na pozadí, postupujte takto:

- 1) Spusíte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - 2) Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Vlastnosti**.
  - 3) Klepněte na kartu **Přihlášení**.
  - 4) Změňte heslo zadáním jeho nové hodnoty a poté klepněte na tlačítko **OK**.
  - 5) Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.
3. Přihlaste se k indexovému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání a znovu spusíte systém podnikového vyhledávání:  
`esadmin system startall`

**Související odkazy**

-  [Nastavení stejného šifrovaného hesla administrátora na všech serverech](#)
- [“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 347](#)

## Čísla portů TCP používaná pro podnikové vyhledávání

Zkontrolujte, zda jsou v systému podnikového vyhledávání použita výchozí čísla portů, aby nedocházelo ke konfliktům portů při konfigurování prostředků a přiřazování čísel portů dalším aplikacím.

Při konfigurování brány firewall musíte explicitně povolit přístup ke konkrétním číslům portů. Dále musíte zajistit umístění všech serverů podnikového vyhledávání dovnitř brány firewall.

*Tabulka 1. Konfigurace čísel portů pro podnikové vyhledávání*

Název portu	Čísla portů	Místo nastavení
CCL (Common Communication Layer)	6002	ES_NODE_ROOT/nodeinfo/es.cfg a ES_NODE_ROOT/master_config/nodes.ini na všech serverech podnikového vyhledávání
Port HTTP na vyhledávacích serverech	80	HTTP_SERVER_ROOT/conf/http.conf na vyhledávacích serverech
Konzola pro správu produktu WebSphere Application Server verze 6	9060	Na vyhledávacích serverech

Tabulka 1. Konfigurace čísel portů pro podnikové vyhledávání (pokračování)

Název portu	Čísla portů	Místo nastavení
Konzola pro správu produktu WebSphere Application Server verze 5.1	9090	Na vyhledávacích serverech
Prolézací modul DB2	6000, 6001, 6002, 60003, 50000	Na prolézacím serveru
Informační centrum	8888	Na vyhledávacích serverech
Anonymní nebo dynamické porty pro CCL, přenos souborů (ESFTP) a kopírování indexu	49152 až 65535	Na všech serverech podnikového vyhledávání
Síťový server Apache Derby	1527	Na prolézacím serveru
Vlastní komunikace	8890	Na prolézacím serveru
Připojení vzdálených klientů k serveru DB2 (používá se pouze u vydání starších než verze 8.4 produktu OmniFind Enterprise Edition	50000	Na prolézacím serveru
Databáze FastObjects produktu WebSphere Information Integrator Content Edition	6001 (alternativně 6002)	V konzole pro správu produktu WebSphere Information Integrator Content Edition
Konektor proxy pro vzdálené volání metod (RMI) produktu WebSphere Information Integrator Content Edition	1251 (port RMI)	V konzole pro správu produktu WebSphere Information Integrator Content Edition

## Změna čísla portu pro systém podnikového vyhledávání

Pokud je číslo portu používaného systémem podnikového vyhledávání pro komunikaci v konfliktu s číslem portu používaného jiným produktem, musíte změnit číslo portu podnikového vyhledávání.

### O této úloze

Číslo portu pro systém podnikového vyhledávání je určeno při instalaci produktu OmniFind Enterprise Edition. V případě konfigurace s více servery je pro všechny servery určeno stejné číslo portu.

Pokud číslo portu nelze použít (číslo portu je například přiřazeno jinému produktu na stejném serveru), je výsledkem konfliktu následující chybová zpráva v souboru `CCLServer_datum.log`, kde parametr `datum` určuje datum vytvoření souboru žurnálu):

```
FFQ00273W Vyskytlo se interní varování - zpráva o výjimce: {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)
```

### Postup

Chcete-li změnit číslo portu používaného podnikovým vyhledáváním, postupujte takto:

1. Přejděte k počítači, v němž je třeba změnit číslo portu, přihlaste se jako administrátor podnikového vyhledávání a zastavte systém podnikového vyhledávání:

esadmin system stopall

**Důležité:** V době, kdy je systém zastaven, nemohou uživatelé zadávat vyhledávací požadavky.

2. Otevřete soubor ES\_NODE\_ROOT/nodeinfo/es.cfg pro úpravy, vyhledejte následující vlastnost, zadejte novou hodnotu čísla portu a poté soubor uložte a zavřete jej.

`CCLPort=nové_číslo_portu`

3. Restartujte službu CCL (Common Communication Layer) podnikového vyhledávání:

**Systém AIX, Linux, nebo Solaris**

startccl.sh

**Příkazový řádek systému Windows**

startccl

**Nástroj pro správu Služby systému Windows**

Chcete-li spustit vrstvu CCL na pozadí, postupujte takto:

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.
4. Přejděte k indexovému serveru, přihlaste se jako administrátor podnikového vyhledávání a zastavte službu CCL:

**Systém AIX, Linux, nebo Solaris**

stopccl.sh

**Příkazový řádek systému Windows**

stopccl

**Nástroj pro správu Služby systému Windows**

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Zastavit**.
5. Upravte soubor ES\_NODE\_ROOT/nodeinfo/es.cfg:

- a. Vyhledejte následující vlastnost, přičemž parametr *název\_počítače* určuje název počítače, pro který jste upravili číslo portu v kroku 2. Číslo *N* ve vlastnosti *nodeN* je číslo identifikující server podnikového vyhledávání.

`nodeN.destination=název_počítače`

- b. Vyhledejte následující dílčí vlastnost, zadejte stejné číslo portu, jako jste zadali pro server v kroku 2, a poté soubor uložte a zavřete:

`nodeN.port=nové_číslo_portu`

6. Restartujte systém podnikového vyhledávání:

esadmin system startall

**Související odkazy**

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 347

“Číslo portů TCP používaná pro podnikové vyhledávání” na stránce 20

---

## Změna adres IP nebo názvů hostitele serveru podnikového vyhledávání

Je možné změnit názvy hostitelů a adresy IP, k jejichž používání jsou servery podnikového vyhledávání konfigurovány.



Změnit adresy IP může být vhodné v případě, že používáte několik karet NIC (Network Interface Card) na každém serveru podnikového vyhledávání a zjistíte, že indexový server je konfigurován k používání pomalé sítě. Změnou nastavení v konfiguračních souborech můžete indexovému serveru umožnit používání rychlejší sítě.

**Tip:** Pokud nechcete upravovat konfigurační soubory, můžete přeinstalovat produkt OmniFind Enterprise Edition a zadat nové názvy hostitelů nebo adresy IP při spuštění instalačního programu.

#### **Postup:**

Chcete-li změnit názvy hostitelů nebo adresy IP, postupujte takto:

1. Přihlaste se jako administrátor podnikového vyhledávání. Používáte-li konfiguraci s více servery, přihlaste se k libovolnému serveru podnikového vyhledávání.
2. Zastavte relaci podnikového vyhledávání:  
`esadmin system stopall`
3. Zastavte vrstvu CCL (Common Communication Layer). Používáte-li konfiguraci s více servery, použijte k zastavení vrstvy CCL na všech serverech podnikového vyhledávání jednu z následujících metod:

##### **System AIX, Linux, nebo Solaris**

```
stopccl.sh
```

##### **Příkazový řádek systému Windows**

```
stopccl
```

##### **Nástroj pro správu Služby systému Windows**

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Zastavit**.
4. Na každém serveru podnikového vyhledávání proveďte následující kroky:
    - a. Upravte soubor `ES_INSTALL_ROOT/configurations/ccl.properties` a zadejte název hostitele nebo adresu IP, které chcete používat pro daný server, do parametru `es_server_hostName`.
    - b. Upravte soubor `ES_NODE_ROOT/nodeinfo/es.cfg` a zadejte název hostitele nebo adresu IP, které chcete používat pro daný server, do parametru `LocalHostName`.
    - c. Upravte soubor `ES_NODE_ROOT/master_config/nodes.ini` a nahraďte všechny výskyty parametru `nodeN.destination` názvem hostitele nebo adresou IP, které chcete používat pro tento server.
    - d. Upravte soubor `ES_NODE_ROOT/config/nodes.ini` a nahraďte všechny výskyty parametru `nodeN.destination` názvem hostitele nebo adresou IP, které chcete používat pro tento server.
  5. Znovu spusťte službu CCL. Používáte-li konfiguraci s více servery, použijte k restartování vrstvy CCL na všech serverech podnikového vyhledávání jednu z následujících metod:

##### **System AIX, Linux nebo Solaris**

```
startccl.sh -bg
```

##### **Příkazový řádek systému Windows**

```
startccl
```

##### **Nástroj pro správu Služby systému Windows**

Chcete-li spustit službu CCL na pozadí, postupujte takto:

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.

- b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.
6. Restartujte relace podnikového vyhledávání:  
esadmin system startall

---

## Konfigurace podpory pro duální adresy IP

Pokud jsou servery, na které instalujete podnikové vyhledávání, konfigurovány k podpoře duálních adres IP, je nutné ručně konfigurovat servery podnikového vyhledávání, aby je bylo možné spustit v tomto prostředí.

Například administrátor pracovní plochy může instalovat produkt Microsoft Loopback Adapter (adaptér zpětné smyčky) k vytvoření virtuální sítě, která podporuje požadavky na síť pro některé produkty, jako je například produkt Microsoft SQL Server.

### Postup:

Chcete-li konfigurovat systém podnikového vyhledávání, aby jej bylo možné spustit na serverech s podporou duálních adres IP, postupujte takto:

1. Přihlaste se jako administrátor podnikového vyhledávání. Používáte-li konfiguraci s více servery, přihlaste se k libovolnému serveru podnikového vyhledávání.
2. Zastavte relaci podnikového vyhledávání:  
esadmin system stopall
3. Zastavte vrstvu CCL (Common Communication Layer). Používáte-li konfiguraci s více servery, použijte k zastavení vrstvy CCL na všech serverech podnikového vyhledávání jednu z následujících metod:

#### Systém AIX, Linux, nebo Solaris

```
stopccl.sh
```

#### Příkazový řádek systému Windows

```
stopccl
```

#### Nástroj pro správu Služby systému Windows

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Zastavit**.
4. Na každém serveru podnikového vyhledávání, která podporuje duální adresy IP, upravte soubor ES\_NODE\_ROOT/nodeinfo/es.cfg a přidejte parametr s názvem **LocalIPAddress**. Hodnotě přiřaďte adresu IP, kterou lze zjistit pomocí modulu DNS.
5. Znovu spusťte službu CCL. Používáte-li konfiguraci s více servery, použijte k restartování vrstvy CCL na všech serverech podnikového vyhledávání jednu z následujících metod:

#### Systém AIX, Linux nebo Solaris

```
startccl.sh -bg
```

#### Příkazový řádek systému Windows

```
startccl
```

#### Nástroj pro správu Služby systému Windows

Chcete-li spustit službu CCL na pozadí, postupujte takto:

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.

6. Restartujte relace podnikového vyhledávání:  
esadmin system startall

---

## Povolení podpory protokolu IPv6

Chcete-li povolit podporu adres URL, které používají protokol IPv6 (IP verze 6), je nutné definovat proměnnou prostředí, která instruuje systém podnikového vyhledávání k používání pouze adres soketu IPv6.

### Než začnete

1. Ověřte, že jsou vlastnosti **LocalHostName** a **LocalIPAddress** (je-li zadána) v souboru %ES\_CFG% názvem hostitele nebo platnou adresou protokolu IPv6 v lokálním počítači. Tyto hodnoty můžete v případě potřeby změnit ručně a poté uložit soubor %ES\_CFG%.
2. Ověřte, že je hodnota vlastnosti **DerbyServerHostName** v souboru %ES\_CFG% názvem hostitele. Tato hodnota nesmí být adresou protokolu IPv4 nebo IPv6. Pokud tuto vlastnost změníte, uložte soubor %ES\_CFG%.
3. Ověřte, že jsou hodnoty vlastností **destination** a **serverhost** v souboru ES\_NODE\_ROOT/master\_config/nodes.ini na indexovém serveru platnou adresou protokolu IPv6 nebo platným názvem hostitele. Tyto hodnoty nesmí být adresami protokolu IPv4. Pokud tento soubor změníte, uložte změny.

### Omezení

Podpora protokolu IPv6 je k dispozici pouze v systémech podnikového vyhledávání instalovaných na servery Windows 2003 Server. Po povolení podpory pro adresy protokolu IPv6 již podnikové vyhledávání nebude používat adresy protokolu IPv4 pro žádnou soketovou komunikaci.

### Podporované prohlížeče

Prohlížeče Internet Explorer a Mozilla Firefox zpracovávají adresy protokolu IPv6 různým způsobem.

#### Mozilla Firefox

Chcete-li spustit konzolu pro správu podnikového vyhledávání nebo vyhledávací aplikaci, můžete zadat adresu protokolu IPv6 nebo název hostitele jako adresu URL. Příklad:

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

#### Internet Explorer

Chcete-li spustit konzolu pro správu podnikového vyhledávání nebo vyhledávací aplikaci, nelze zadat adresu protokolu IPv6 jako adresu URL. Je nutné použít následující formát a zajistit, aby byl název hostitele mapován na adresu protokolu IPv6 v modulu DNS, nebo použít soubor c:\windows\system32\etc\hosts:

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

Při zadání adresy protokolu IPv6 jako adresy URL se zobrazí zpráva **Chyba syntaxe**. Další informace o tomto omezení naleznete v tématu <http://support.microsoft.com/kb/325414>.

### Podpora zdrojů dat protokolu IPv4

Zdroje dat spuštěné na serveru s protokolem IPv4 jsou podporovány prolézacími moduly podnikového vyhledávání. Při konfiguraci webových prolézacích modulů a zadání počátečních adres URL ve formátu adresy protokolu IPv6 se ujistěte, že jsou adresy URL uvedeny v závorkách. Příklad:

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

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

## Postup

Postup při povolení podpory protokolu IPv6:

1. Přihlaste se k indexovému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání a zastavte systém podnikového vyhledávání:  
`esadmin system stopall`

**Důležité:** V době, kdy je systém zastaven, nemohou uživatelé zadávat vyhledávací požadavky.

2. Zastavte službu CCL (Common Communication Layer) na všech serverech podnikového vyhledávání:
  - a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Zastavit**.
3. Definujte volbu **ES\_IPV6=1** v systémových proměnných prostředí. V případě instalace podnikového vyhledávání na více serverů proveďte tento krok na vyhledávacích serverech.
4. Aktualizujte přizpůsobené vlastnosti prostředí JVM (Java Virtual Machine) na serveru WebSphere Application Server. V případě instalace podnikového vyhledávání na více serverů proveďte tyto kroky na vyhledávacích serverech.
  - a. Je-li tento server spuštěn, zastavte aplikační server ESSearchServer na serveru WebSphere Application Server.
  - b. Otevřete konzolu pro správu serveru WebSphere Application Server a přejděte na panel přizpůsobených vlastností prostředí JVM (Java Virtual Machine). Vyberte volbu **Servery** → **Aplikační servery** → **ESSearchServer** → **Správa procesů a prostředí Java** → **Definice procesu** → **Prostředí JVM (Java Virtual Machine)** → **Přizpůsobené vlastnosti**.
  - c. Konfigurujte následující vlastnosti:

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

Pokud přizpůsobená vlastnost není ještě uvedena v seznamu, vytvořte novou vlastnost a zadejte název vlastnosti do pole **Název** a platnou hodnotu do pole **Hodnota**.

- d. Pokud používáte produkt WebSphere Application Server verze 6.1, přidejte do souboru `ES_NODE_ROOT/nodeinfo/es.cfg` vlastnost **LocalIPAddress** obsahující adresu IPv6 tohoto vyhledávacího serveru.
  - e. Restartujte aplikaci ESSearchServer.
5. Tyto kroky proveďte při spuštění serveru WebSphere Application Server verze 6.1. V případě instalace podnikového vyhledávání na více serverů proveďte tyto kroky na vyhledávacích serverech.
    - a. Upravte soubor `httpd.conf` serveru IBM HTTP Server.
    - b. Odebráním znaku `#` ze začátku řádku zrušte umístění pokynů v komentáři, a umožněte tak serveru IBM HTTP Server přijímat informace z připojení protokolu IPv6 prostřednictvím portu 80:

```
# Listen [::]:80
```

- c. Restartujte server IBM HTTP Server.
6. Restartujte službu CCL na všech serverech podnikového vyhledávání:
  - a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.
7. Restartujte systém podnikového vyhledávání:  
esadmin system startall
8. Zkontrolujte soubor ES\_NODE\_ROOT/node/logs/ccl0.log a ověřte, že došlo ke spuštění serveru CCL bez chyb. Při správném povolení protokolu IPv6 dojde k protokolování zpráv podobných následujícím zprávám:

```
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. Chcete-li ověřit, že vyhledávací aplikace a aplikace správy pracují správně, postupujte podle následujících kroků:
  - a. Spusťte vyhledávací servery pro libovolnou kolekci a poté otevřete ukázkovou vyhledávací aplikaci a zadejte dotaz. Tento krok ověřuje, že zpracování odesláni dotazů funguje i v případě, že v indexu nejsou k dispozici žádné dokumenty k vyhledání.
  - b. Otevřete konzolu pro správu a ověřte, že se můžete bez problémů přihlásit.

### Zakázání podpory protokolu IPv6

1. Zobrazte sekci **Než začnete** a pro každou instanci adresy protokolu IPv6 nahraďte hodnotu názvem hostitele nebo adresou protokolu IPv4.
2. Pokračujte v proceduře povolení podpory protokolu IPv6 s následujícími výjimkami:
  - V kroku 3 na stránce 26 odeberte proměnnou prostředí **ES\_IPV6**.
  - V kroku 4 na stránce 26 odeberte přízpůsobené vlastnosti prostředí JVM nastavené pro aplikaci ESSearchServer v produktu WebSphere Application Server.
3. Pokud vlastnost **LocalIPAddress** v souboru ES\_NODE\_ROOT/nodeinfo/es.cfg obsahuje adresu ve formátu IPv6, odeberte ji.
4. Pokračujte v proceduře od kroku 9 a ověřte vyhledávací aplikace a aplikace správy.



---

## Kolekce podnikového vyhledávání

Kolekce podnikového vyhledávání obsahuje celou sadu zdrojů, které mohou uživatelé prohledávat v rámci jediného dotazu. Díky využití principu federování mohou uživatelé prohledávat více kolekcí jediným dotazem.

Při vytváření kolekce zadáváte volby platné pro celou kolekci. Kolekce je prázdná, dokud do ní nepřidáte obsah.

Kolekce můžete do systému podnikového vyhledávání přidávat dvěma způsoby:

- Pokud neovládáte práci s konzolou pro správu podnikového vyhledávání nebo se teprve seznamujete s vzájemnými vztahy jednotlivých komponent kolekce, můžete k vytvoření kolekce použít Průvodce kolekcí. Průvodce kolekcí vám pomůže postupně projít jednotlivé úlohy a vytvářené kolekce průběžně ukládat jako koncepty.
- Jestliže jste se již s konzolou pro správu dostatečně seznámili, můžete dát přednost vytváření kolekcí výběrem konkrétních stránek, které chcete spravovat, v pohledu Kolekce.

Po vytvoření kolekce slouží ovládací prvky v pohledu Kolekce k zobrazení a monitorování kolekce, systému podnikového vyhledávání a voleb zabezpečení.

### Federování kolekcí

Pokud vyhledávací aplikace obsahuje podporu federování, mohou uživatelé prohledávat více kolekcí současně. Federování vám rovněž umožňuje obejít omezení velikosti kolekce (20 000 000 dokumentů na kolekci). Uživatelé mohou například prohledávat dvě kolekce, z nichž každá může obsahovat 20 000 000 dokumentů.

Kvalita vyhledávání závisí na ohodnoceních generovaných jednotlivými kolekcemi, jejichž sloučením vzniká konečná sada výsledků. Výsledky jsou stejné jako při odeslání dvou samostatných vyhledávacích požadavků a následném sloučení a seřazení výsledků.

#### Související úlohy

“Monitorování kolekce” na stránce 286

---

## Vytvoření kolekce pomocí Průvodce kolekcí

Pokud jste novými uživateli podnikového vyhledávání, může vám s vytvářením kolekce pomoci průvodce. Průvodce poskytuje podrobné údaje o jednotlivých krocích procesu a umožňuje průběžně ukládat nastavení.

### Než začnete

Chcete-li vytvořit kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

Chcete-li přidat obsah do kolekce nebo určit volby pro analýzu, indexování nebo vyhledávání obsahu v kolekci, musíte být administrátor podnikového vyhledávání nebo administrátor kolekce pro danou kolekci.

### Omezení

Pomocí Průvodce kolekcí můžete vytvořit následující typy prolézacích modulů:

- Content Edition


- DB2
- DB2 Content Manager
- Exchange Server
- Notes
- Souborový systém UNIX
- webový prolézací modul
- Souborový systém Windows

### O této úloze

Při vytváření můžete kolekci uložit ve stavu konceptu. Pokud je kolekce ve stavu konceptu, může v ní každý administrátor s oprávněním ke správě kolekce provádět změny. Můžete například chtít, aby administrátor kolekce, který má zkušenosti se zdroji Lotus Notes, konfiguroval prolézací modul Notes. Později může administrátor kolekce, který má zkušenosti se systémy UNIX, upravit koncept kolekce a konfigurovat prolézací modul Souborový systém UNIX.

### Postup

Vytvoření kolekce pomocí Průvodce kolekcí:

1. Klepnutím na volbu **Kolekce** otevřete pohled Kolekce.
2. Klepněte na volbu **Průvodce kolekcí**.
3. Pomocí pokynů v průvodci vytvoříte prázdnou kolekci a přidejte do ní obsah.  
Je nutné konfigurovat obecné informace o kolekci a vytvořit alespoň jeden prolézací modul. U zbývajících voleb konfigurace můžete přijmout výchozí hodnoty nebo můžete určit volby pro novou kolekci.
4. Chcete-li kolekci uložit před jejím dokončením, klepněte na volbu **Uložit jako koncept**. Kolekce bude uvedena s ostatními koncepty kolekcí v pohledu Kolekce. Pokud pro kolekci povolíte zabezpečení, bude vedle názvu kolekce zobrazena ikona  **Je povoleno zabezpečení kolekce**.
5. Chcete-li se vrátit k vytvářené kolekci, klepněte na volbu **Zpět do průvodce** v pohledu Kolekce.
6. Klepnutím na tlačítko **Dokončit** vytvoříte kolekci.  
Nová kolekce bude uvedena s ostatními kolekcemi v pohledu Kolekce.

Po vytvoření kolekce je nutné spustit procesy pro prolézání, analýzu, indexování a prohledávání kolekce. Dokud nebudete připraveni asociovat kolekci s vyhledávacími aplikacemi, které by ji mohly prohledávat, můžete novou kolekci prohledávat pomocí vzorové vyhledávací aplikace (označené jako Výchozí).

---

## Vytvoření kolekce s použitím pohledu Kolekce

Prostřednictvím pohledu Kolekce vytvoříte prázdnou kolekci. Tuto kolekci potom můžete upravit určením voleb pro přidávání dat do kolekce a umožnit tak její prohledávání.

### Než začnete

Chcete-li vytvořit kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

Chcete-li přidat obsah do kolekce nebo určit volby pro analýzu, indexování nebo vyhledávání obsahu v kolekci, musíte být administrátor podnikového vyhledávání nebo administrátor kolekce pro danou kolekci.



## O této úloze


Informace o hodnotách, které můžete zadat pro novou kolekci, získáte po klepnutí na tlačítko **Nápověda** při vytváření kolekce.

## Postup

Vytvoření kolekce z pohledu Kolekce:

1. V pohledu Kolekce klepněte na volbu **Vytvořit kolekci**.
2. Na stránce Vytvořit kolekci zadejte informace nebo proveďte výběr v následujících polích:
  - **Název kolekce.** Zadejte popisný název označující obsah nebo účel kolekce.
  - **Zabezpečení kolekce.** Určete, zda chcete pro kolekci povolit zabezpečení. Po vytvoření kolekce nelze toto nastavení změnit. Pokud je zabezpečení kolekce povoleno, můžete později určit volby pro vynucení řízení přístupu na úrovni dokumentů.
  - **Důležitost dokumentů (statický model určení pořadí).** Určete strategii pro přiřazování statických tříd důležitosti, která bude použita k řazení dokumentů ve výsledcích vyhledávání. Tuto hodnotu nelze po vytvoření kolekce změnit.
  - **Typ kategorizace.** Určete, zda chcete umožnit vyhledávání dokumentů podle kategorií, do nichž náleží.
  - **Používaný jazyk.** Určete výchozí jazyk pro vyhledávání dokumentů v dané kolekci.
3. U následujících polí potvrďte výchozí hodnoty nebo určete volby, které chcete použít pro danou kolekci:
  - **Popis.** Při výchozím nastavení není vytvořen žádný popis.
  - **Odhadovaný počet dokumentů.** Výchozí odhadovaná velikost kolekce je 1 000 000 dokumentů. Systém používá tuto hodnotu k odhadu prostředků paměti a disků pro danou kolekci, nikoli k omezení její velikosti.
  - **Umístění pro data kolekce.** Výchozí umístění pro soubory související s kolekcí je na indexovém serveru. Tuto hodnotu nelze po vytvoření kolekce změnit.
  - **ID kolekce.** Výchozí ID kolekce je založeno na názvu kolekce. Tuto hodnotu nelze po vytvoření kolekce změnit. Pokud zadáte vlastní ID kolekce, bude vyhledávací aplikace volat kolekci s tímto identifikátorem, nikoli s identifikátorem, který byl vytvořen systémem a může být kryptický.
  - **N-gramová segmentace.** Výchozí metoda segmentace je segmentace ve formátu Unicode podle mezerových znaků. Volbu použití n-gramové segmentace vyberte pouze v případě, že kolekce zahrnuje dokument v čínštině, japonštině nebo korejštině a chcete, aby modul analýzy používal k oddělování slov raději n-gramovou segmentaci. Tuto hodnotu nelze po vytvoření kolekce změnit.

Další informace o konfiguraci podpory pro úplnou n-gramovou analýzu a převod na prvky v kolekcích podnikového vyhledávání najdete v dokumentu <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088>.
4. Klepněte na tlačítko **OK**.

V pohledu Kolekce bude nová kolekce uvedena spolu s ostatními kolekcemi v systému podnikového vyhledávání. Pokud pro kolekci povolíte zabezpečení, bude vedle názvu kolekce zobrazena ikona  **Je povoleno zabezpečení kolekce**.

Kolekce je prázdná, dokud do ní nepřidáte obsah. Chcete-li do nové kolekce přidat obsah, vyberte kolekci v pohledu Kolekce, upravte ji, vytvořte alespoň jeden prolézací modul a určete volby pro analýzu, indexování a vyhledávání dat.

Poté je nutné spustit procesy pro prolézání, indexování a prohledávání kolekce. K prohledávání nové kolekce můžete používat vzorovou vyhledávací aplikaci, dokud nebudete připraveni používat vlastní vyhledávací aplikace.

---

## Úprava kolekce

Při úpravách kolekce můžete zadat informace o dokumentech, které chcete zahrnout do kolekce.

### Než začnete


Chcete-li upravovat kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

Při úpravách kolekce můžete určit volby pro prolézání zdrojů dat, analýzu dokumentů, sestavování indexu, vyhledávání indexovaného obsahu a protokolování chybových zpráv. V rámci úprav kolekce po jejím vytváření musíte přidat do kolekce obsah. Při pozdějších úpravách kolekce můžete aktualizovat obsah nebo změnit způsob prolézání, analýzy, indexování, prohledávání nebo protokolování informací.

### Postup

Úpravy kolekce:

1. Klepnutím na volbu **Kolekce** na panelu nástrojů otevřete pohled Kolekce.
2. V seznamu kolekcí vyhledejte kolekci, kterou chcete upravit, a klepněte na volbu  **Upravit**.
3. Proveďte změny na následujících stránkách:

#### Prolézání

Určete zdroje dat, které chcete prolézat, a volby pro způsob prolézání obsahu. Každá kolekce musí obsahovat alespoň jeden prolézací modul; jedna kolekce může obsahovat data z několika typů zdrojů dat. Pro každý typ zdroje dat, který chcete zahrnout do kolekce, je nutné konfigurovat alespoň jeden prolézací modul.

#### Analýza

Určete volby pro způsob, jakým chcete analyzovat prolezené dokumenty. Můžete konfigurovat kategorie, které uživatelům umožní prohledávat podmnožiny kolekcí, a pravidla, která umožní prohledávat určité části dokumentů XML a HTML. Pokud do systému podnikového vyhledávání přidáte vlastní stroje pro analýzu textu, můžete určit, který z nich má být použit pro analýzu a anotace obsahu v této kolekci. Také můžete asociovat pole s třídami upřednostnění a ovlivňovat tak ohodnocení důležitosti dokumentů ve výsledcích vyhledávání.

#### Index

Určete časové plány pro sestavování celého indexu a aktualizaci indexu s použitím nového a změněného obsahu. Lze také konfigurovat volby pro používání zástupných znaků v dotazech, omezení zobrazení kolekce na určitý rozsah identifikátorů URI, sbalení výsledků vyhledávání ze stejného webu a odebrání identifikátorů URI z indexu.

#### Vyhledávání

Určete volby pro procesy vyhledávání, například pro konfigurování mezipaměti vyhledávání a výběr jazyka vyhledávání. Můžete také konfigurovat rychlé odkazy, což je funkce zajišťující vrácení předem určených identifikátorů URI, jestliže uživatel v dotazu uvede určitá slova nebo fráze. Pokud jste do systému

podnikového vyhledávání přidali vlastní slovníky, můžete vybrat slovníky, které chcete použít pro prohledávání této kolekce.

**Žurnál** Určete typy zpráv, které chcete protokolovat, a volby pro vytváření a cyklické použití souborů žurnálu. Můžete konfigurovat výstrahy, abyste mohli být upozorněni na určité události, a určit volby pro příjem e-mailů, kdykoli jsou zaznamenány určité zprávy nebo určité typy zpráv. Můžete také určit volby protokolování informací o dokumentech odstraněných ze systému podnikového vyhledávání.

#### **Obecné**

Zadejte obecné informace o kolekci a nastavení pohledu, které nelze změnit. Pokud bylo pro kolekci při jejím vytváření povoleno zabezpečení, lze konfigurovat volby zabezpečení na úrovni dokumentu.

---

## Odstranění kolekce

Při odstranění kolekce jsou ze systému podnikového vyhledávání zcela odebrány všechny informace o dané kolekci.

#### **Než začnete**

Chcete-li odstranit kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

Před odstraněním je nutné zastavit všechny procesy přiřazené ke kolekci.


#### **O této úloze**

Odstranění kolekce může trvat delší dobu. Jakmile potvrdíte, že chcete danou kolekci odstranit, systém odstraní všechna data související s kolekcí.

**Tip:** Zpráva o vypršení časového limitu požadované operace se může zobrazit i v případě, že proces stále ještě běží na pozadí. Chcete-li zjistit, zda byla úloha dokončena, klepněte na tlačítko **Aktualizovat** v konzole správy (nepoužívejte tlačítko **Aktualizovat** webového prohlížeče). Proces odstranění byl dokončen, jestliže se název kolekce již nezobrazuje v seznamu kolekcí.

#### **Postup**

Odstranění kolekce:

1. Klepnutím na volbu **Kolekce** otevřete pohled Kolekce.
2. V seznamu kolekcí vyhledejte kolekci, kterou chcete odstranit, a klepněte na volbu  **Odstranit**.

---

## Určení ID kolekce

Pro řadu úloh správy je nutné znát ID kolekce.

#### **Než začnete**

Chcete-li zobrazit ID kolekce, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

#### **O této úloze**

Při vytváření kolekce můžete hodnotu ID zadat nebo můžete umožnit systému, aby ID přiřadil za vás. Chcete-li určit ID kolekce po vytvoření kolekce, můžete použít konzolu pro správu nebo zobrazit konfigurační soubor.

### Postup

1. Chcete-li určit ID kolekce pomocí konzoly pro správu, postupujte takto:
  - a. Klepnutím na volbu **Kolekce** otevřete pohled Kolekce.
  - b. Na stránce Obecné klepněte na tlačítko **Zobrazit nastavení kolekce**.  
Na stránce Nastavení kolekce je zobrazeno ID kolekce, úplná cesta k umístění dat kolekce a statický model ohodnocení důležitosti používaný pro dokumenty této kolekce.
2. Chcete-li určit ID kolekce pomocí zobrazení konfiguračního souboru, postupujte takto:

- a. Otevřete soubor `ES_NODE_ROOT/master_config/collections.ini`. Zjednodušte zobrazení seřazením dat tohoto souboru. V následujícím ukázkovém výstupu je použito ID kolekce `coll1`:

```
% 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
...
```

---

## Administrace prolézacích modulů

Prolézací moduly je třeba konfigurovat pro různé typy dat, které chcete zahrnout do kolekce. Jedna kolekce může obsahovat libovolný počet prolézacích modulů.

**Tip:** Výukový program online je k dispozici na webu <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html>. Výukový program popisuje jednotlivé kroky instalace a konfigurace, zobrazuje způsob prohledávání různých typů zdrojů dat a uvádí, jak lze použít rozhraní API produktu k rozšíření funkcí podnikového vyhledávání. Výukový program je určen pro starší verzi produktu OmniFind Enterprise Edition, ale většinu koncepcí a postupů lze stále použít.

Podrobné příklady konfigurace prolézacích modulů a povolení zabezpečení v prostředí malých, středních i velkých organizací naleznete v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

### Konfigurování prolézacích modulů

K vytváření, úpravám a odstraňování prolézacích modulů slouží konzola pro správu podnikového vyhledávání. Konfiguraci prolézacího modulu obvykle provádí odborník na daný typ prolézáných dat. Například při konfigurování modulu pro prolézání zdrojů serveru Lotus Notes by měl být administrátor kolekce buď administrátorem serveru Notes, nebo úzce spolupracovat s někým, kdo rozumí prolézáným databázím.

Se zadáním vlastností určujících způsob, jakým prolézací modul, který vytvoříte, využívá systémové prostředky, vám pomůže průvodce určený pro zvolený typ prolézáných dat. Pomocí průvodce můžete také vybrat zdroje, které chcete prohledávat.

V existujících prolézacích modulech můžete kdykoli provést změny. Podle potřeby můžete upravit vlastnosti prolézacího modulu nebo části prolézaného prostoru. Také s těmito změnami vám mohou pomoci průvodci prolézacími moduly.

### Naplnění nového prolézacího modulu základními hodnotami

Prolézací modul můžete vytvořit použitím výchozích systémových hodnot nebo zkopírováním hodnot zadaných pro jiný prolézací modul stejného typu. Pokud jako základ nového prolézacího modulu použijete existující prolézací modul, můžete rychle vytvořit více prolézacích modulů s podobnými vlastnostmi a poté je konfigurovat například pro prolézání různých zdrojů nebo pro práci s různými plány prolézání.

Po zkopírování prolézacího modulu můžete rozdělit zátěž spojenou s prolézáním mezi více prolézacích modulů používajících stejná pravidla prolézání. Můžete například zkopírovat prolézací modul Notes, protože chcete použít stejné vlastnosti a pravidla prolézání polí u jiného serveru Lotus Notes. Jediným rozdílem mohou být databáze, které každý z těchto prolézacích modulů zpracovává, a nastavení zabezpečení na úrovni dokumentu.

### Kombinování typů prolézacích modulů v kolekci

Prolézací moduly podnikového vyhledávání jsou určeny k shromažďování informací pocházejících z různých typů zdrojů dat. Při konfigurování prolézacích modulů v kolekci se musíte rozhodnout, jakým způsobem tyto různé typy zdrojů dat zkombinujete, aby uživatelé mohli snadno prohledávat podniková data. Chcete-li například uživatelům umožnit prohledávat jediným dotazem souborové systémy Microsoft Windows a současně veřejné

složky na serveru Microsoft Exchange, vytvořte kolekci zahrnující prolézací moduly Souborový systém Windows a Exchange Server.

Pokud v jedné kolekci zkombinujete několik typů prolézacích modulů, zajistěte, aby všechny prolézací moduly používaly stejnou metodu statického řazení. (Metoda statického řazení se určuje při vytvoření kolekce.) Pokud například kombinujete webové zdroje (které jako třídu důležitosti používají počet odkazů na dokument) a zdroje NNTP (které jako třídu důležitosti používají data dokumentu), může být kvalita výsledků vyhledávání snížena.

## Konfigurace zabezpečení na úrovni dokumentu

Pokud při vytvoření kolekce povolíte její zabezpečení, můžete konfigurovat volby zabezpečení na úrovni dokumentů. Každý prolézací modul může přiřadit k prolézáným dokumentům prvky zabezpečení. Pokud při konfigurování prolézacího modulu určíte, že má být použito zabezpečení na úrovni dokumentů, prolézací modul přiřadí k jednotlivým dokumentům zadané prvky zabezpečení a přidá je do indexu spolu s dokumenty.

Jestliže povolíte zabezpečení ve vlastních aplikacích podnikového vyhledávání, mohou vaše aplikace používat prvky zabezpečení, které prolézací moduly přiřadili k dokumentům, k ověřování uživatelů. Tato funkce vám umožňuje omezit přístup k některým dokumentům v kolekci a prohledávání jiných dokumentů povolit všem uživatelům. V určité kolekci můžete například povolit všem uživatelům přístup ke všem dokumentům ve veřejných složkách na serveru Microsoft Exchange, zatímco přístup k dokumentům v databázích Lotus Notes umožníte pouze uživatelům s určitými uživatelskými jmény.

Na základě použití vlastních obchodních pravidel můžete určit hodnotu prvků zabezpečení zakódováním pravidel ve třídě Java. Při konfigurování vlastností prolézacího modulu můžete určit název modulu plug-in, který má prolézací modul při prolézání dokumentů používat. Prvky zabezpečení přidané příslušným modulem plug-in jsou uchovávány v indexu a jejich prostřednictvím lze řídit přístup k dokumentům.

Při konfigurování určitých typů prolézacích modulů můžete určit další ovládací prvky zabezpečení. Můžete například určit, že chcete během zpracování dotazu ověřovat uživatele. Povolíte-li tuto volbu, budou údaje pověření uživatele porovnávány s aktuálními seznamy řízení přístupu spravovaných zdrojů dat, které mají být prohledávány. Tímto ověřováním aktuálních údajů pověření lze nahradit ověřování založené na použití prvků zabezpečení v indexu podnikového vyhledávání nebo mohou být oba způsoby ověřování použity společně.

### Související pojmy

“Zabezpečení na úrovni dokumentu” na stránce 248

### Související úlohy

“Monitorování prolézacích modulů” na stránce 288

### Související odkazy

“Požadavky nastavení prolézacího modulu pro podporu zabezpečení” na stránce 263

---

## Vytvoření prolézacího modulu

Při vytváření prolézacího modulu určujete typ prolézacího modulu, který chcete vytvořit. Informace o datech, které chcete zahrnout do kolekce, můžete zadávat pomocí průvodce.

### Než začnete

Chcete-li vytvořit prolézací modul, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

## O této úloze

Pro kolekci je nutné vytvořit alespoň jeden prolézací modul. Typ vytvářeného prolézacího modulu závisí na typu dat, která chcete zahrnout do kolekce. Určování voleb pro prolézací modul vám usnadní průvodce pro typ vytvářeného prolézacího modulu. Průvodce vám například může usnadnit určování voleb způsobu využívání systémových prostředků prolézacím modulem. Pomocí průvodce můžete také vybrat zdroje dat, které chcete zahrnout do kolekce.

## Postup

Vytvoření prolézacího modulu:

1. Upravte kolekci, vyberte stránku Prolézání a klepněte na volbu **Vytvořit prolézací modul**.
2. Vyberte typ prolézacího modulu a základní hodnoty pro jeho nastavení:
  - a. Vyberte typ prolézacího modulu podporující typ dat, která chcete prolézat, například webové servery, databáze Lotus Notes nebo souborové systémy UNIX.  
Po výběru typu prolézacího modulu se zobrazí volby pro jeho vytvoření.
  - b. Vyberte základní hodnoty pro prolézací modul:

### Použití výchozí hodnoty systému pro nový prolézací modul

Naplní počáteční nastavení prolézacího modulu výchozími instalačními hodnotami.

Po výběru této volby zahajete konfiguraci nového prolézacího modulu klepnutím na tlačítko **Další**.

### Klonovat hodnoty existujícího prolézacího modulu pro nový prolézací modul

Naplní počáteční nastavení prolézacího modulu hodnotami, které byly použity ke konfigurování jiného prolézacího modulu tohoto typu.

Při výběru této volby se zobrazí seznam prolézacích modulů, které vyhovují tomuto typu. Vyberte prolézací modul, který chcete použít pro nový prolézací modul, a klepnutím na tlačítko **Další** zahajete konfiguraci nového prolézacího modulu.

Otevře se průvodce pro typ vytvářeného prolézacího modulu. Podle pokynů v průvodci vytvoříte prolézací modul. Další informace o volbách, které můžete určit pro daný typ prolézacího modulu, zobrazíte klepnutím na volbu **Nápověda** na jednotlivých stránkách průvodce.

Nový prolézací modul bude uveden na stránce Prolézání mezi ostatními prolézacími moduly, které patří do dané kolekce. Pokaždé, když bude nutné provést změny prolézacího modulu, klepněte na příslušné volby pro úpravy vlastností prolézacího modulu a prolézaného prostoru.

---

## Úprava vlastností prolézacího modulu

Můžete změnit informace o prolézacím modulu a způsob, jakým prolézá data. Můžete například změnit způsob, jakým prolézací modul využívá systémové prostředky.

### Než začnete


Chcete-li upravovat vlastnosti prolézacího modulu, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

## O této úloze

Klepnutím na volbu **Nápověda** při úpravách vlastností prolézacího modulu zobrazíte informace o typech změn, které můžete provést. Vlastností, které lze upravovat, závisí na typu prolézacího modulu.

### Postup

Úpravy vlastností prolézacího modulu:

1. Upravte kolekci, vyberte stránku Prolézání , vyhledejte prolézací modul, který chcete upravit, a klepněte na volbu  **Vlastnosti prolézacího modulu**.
2. Změňte vlastnosti prolézacího modulu a klepněte na tlačítko **OK**.
3. Změny se projeví po zastavení a restartu prolézacího modulu. (Změníte-li pouze popis prolézacího modulu, není třeba restartovat jej.)

---

## Úprava prolézaného prostoru

Můžete změnit informace o zdrojích dat, které prolézací modul prolézá. Můžete například přidat nebo odebrat zdroje dat, změnit časový plán prolézání a změnit pravidla pro prolézání dokumentů v konkrétním zdroji dat.

### Než začnete


Chcete-li upravovat prolézaný prostor, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

### O této úloze

Informace o změnách, které můžete provádět pro daný typ prolézacího modulu, který spravujete, zobrazíte klepnutím na volbu **Nápověda** při úpravách prolézaného prostoru.

### Postup

Úpravy prolézaného prostoru:

1. Upravte kolekci, vyberte stránku Prolézání, vyhledejte prolézací modul, který chcete upravit, a klepněte na volbu  **Prolézaný prostor**.
2. Vyberte požadované volby a změňte prolézaný prostor.  
Volby, které jsou k dispozici pro výběr, závisí na typu prolézacího modulu. U některých voleb, například pro přidávání zdrojů dat do kolekce, se otevře průvodce pro daný typ prolézacího modulu, který vám usnadní změnu prolézaného prostoru.
3. Změny se projeví po zastavení a restartu prolézacího modulu.

---

## Odstranění prolézacího modulu

Při odstranění prolézacího modulu jsou ze systému podnikového vyhledávání odebrány všechny informace o daném modulu. Informace prolézané tímto modulem však zůstávají v indexu, dokud není vytvořeno hlavní sestavení indexu.

### Než začnete

Chcete-li odstranit prolézací modul, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

### O této úloze




Odstranění prolézacího modulu může trvat delší dobu. Jakmile potvrdíte, že chcete prolézací modul odstranit, systém odstraní všechna data související s tímto prolézacím modulem.

**Tip:** Provedení této úlohy trvá nějakou dobu, a proto se může zobrazit zpráva o vypršení časového limitu operace, a to i v případě, že je proces na pozadí stále spuštěn. Chcete-li určit, zda byla úloha dokončena, občas klepněte na tlačítko **Aktualizovat** v konzole pro správu (nikoli na tlačítko **Aktualizovat** ve webovém prohlížeči). Jestliže název prolézacího modulu již není zobrazen v seznamu prolézacích modulů, je proces odstranění dokončen.

### Postup

Odstranění prolézacího modulu:

1. Upravte kolekci a vyberte stránku Prolézání.
2. Vyhledejte prolézací modul, který chcete odstranit, a klepněte na volbu  **Odstranit**.

---

## Časový plán prolézacího modulu

Prolézací moduly, které vytvoříte pro webové zdroje pracují nepřetržitě. Spustíte-li webový prohlížeč, obvykle je nemusíte zastavovat, pokud nepotřebujete změnit konfiguraci. Při konfigurování všech ostatních typů prolézacích modulů je třeba zadat plán prolézání.

U některých typů zdrojů dat existuje jediný společný plán zpracování všech zdrojů dat v prolézaném prostoru daným prolézacím modulem. U jiných zdrojů dat můžete zadat individuální plány pro konkrétní zdroje dat. Individuální plán prolézání lze zadat například pro každou databázi Lotus Notes zpracovávanou prolézacím modulem.

Při konfigurování časového plánu můžete určit typ prolézání, které má být prováděno. Můžete naplánovat úplné prolézání všech dokumentů v prolézaném prostoru, prolézání zahrnující všechny aktualizace prolézaného prostoru (nové dokumenty, upravené dokumenty a odstraněné dokumenty) nebo prolézání zahrnující pouze nové a upravené dokumenty. Úplné prolézání trvá nejdéle. Prolézání s odebíráním odstraněných dokumentů trvá déle než prolézání, při kterém jsou odstraněné dokumenty ignorovány.

V rámci úprav prolézaného prostoru prolézacího modulu můžete určit druhý časový plán prolézání. Můžete například konfigurovat jeden časový plán pro prolézání všech dokumentů v prolézaném prostoru každou sobotu v noci a druhý časový plán pro častější spuštění a prolézání pouze nových a upravených dokumentů.

Vytvoříte-li více plánů prolézání, získáte přesnější kontrolu nad tím, kdy prolézací modul navštěvuje cílové zdroje. Při prolézání databází umístěných v různých časových pásmech můžete například naplánovat spuštění prolézacího modulu na čas po pracovní době většiny uživatelů.

---

## Prolézací moduly Content Edition

Chcete-li zahrnout úložiště IBM WebSphere Information Integrator Content Edition do kolekce podnikového vyhledávání, musíte konfigurovat prolézací modul Content Edition.

Prolézací modul Content Edition můžete využít k prolézání následujících typů úložišť:

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

- OpenText Livelink
- Portal Document Manager (PDM)

Při konfiguraci prolézacího modulu určujete volby ovlivňující, jakým způsobem bude prolézací modul prolézat všechna úložiště v prolézaném prostoru. Také určujete třídy položek, které chcete v jednotlivých úložištích prolézat.

### Tip:

Podrobné příklady konfigurace konektorů a zabezpečeného prolézacího modulu Content Edition najdete ve scénáři pro střední organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

## Konfigurace prolézacího serveru

Způsob přípravy úložišť pro prolézání závisí na tom, zda máte v úmyslu použít pro přístup k prolézaným datům přímý režim nebo režim serveru. Pokud používáte přímý režim, musíte konfigurovat konektor v produktu WebSphere Information Integrator Content Edition. Pokud používáte režim serveru, musíte na prolézacím serveru spustit skript. Tento skript se dodává s produktem OmniFind Enterprise Edition a umožňuje prolézacímu modulu Content Edition komunikovat se servery WebSphere Information Integrator Content Edition.

Pokud používáte režim serveru, proveďte před vytvořením prolézacího modulu Content Edition úlohu odpovídající vašemu prostředí:

- “Konfigurování prolézacího serveru v systému UNIX pro produkt WebSphere II Content Edition” na stránce 42.
- “Konfigurování prolézacího serveru v systému Windows pro produkt WebSphere II Content Edition” na stránce 43.

Podrobné pokyny pro konfigurování systému podnikového vyhledávání pro prohledávání úložišť WebSphere Information Integrator Content Edition naleznete v článku *Správa dokumentů vyhledávacího portálu WebSphere pomocí produktu WebSphere Information Integrator OmniFind Edition* ze série IBM developerWorks na adrese <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0606lee/>.

## Přehled konfigurace

Chcete-li vytvořit nebo změnit prolézací modul Content Edition, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechna úložiště v prolézaném prostoru.
- Určete, zda prolézací modul používá pro přístup k úložištím přímý režim nebo režim serveru. V případě režimu serveru musíte rovněž zadat informace, které prolézacímu modulu umožní přistupovat k webovému aplikačnímu serveru.
- Vyberte úložiště, která chcete prolézat.
- Zadat jména uživatelů a hesla, která prolézacímu modulu umožňují přístup k obsahu ve vybraných úložištích.
- Nastavit plán prolézání úložišť.
- Vybrat třídy položek, které chcete prolézat v jednotlivých úložištích.

- Zadat volby umožňující prohledávání vlastností tříd položek. Můžete například vyloučit z prolézaného prostoru určité typy dokumentů nebo zadat konkrétní verzi úložiště, která má být prolézána.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na seznamech přístupových práv nebo prvcích zabezpečení.

Pro třídy položek Documentum, FileNet Panagon Content Services, Hummingbird DM, Portal Document Manager a SharePoint můžete rovněž vybrat volbu ověřování údajů pověření uživatelů při odeslání dotazu. V tomto případě systém neporovnává údaje pověření uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat. Tento typ ověřování aktuálních údajů pověření není pro jiné typy úložišť k dispozici.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

## Přístup k úložištím Content Edition v přímém režimu

Prolézací modul Content Edition lze konfigurovat pro přístup k úložištím produktu WebSphere Information Integrator Content Edition v přímém režimu.

### O této úloze

V přímém režimu prolézací modul používá konektor WebSphere Information Integrator Content Edition instalovaný na prolézací server při instalaci produktu OmniFind Enterprise Edition. Prolézací modul používá pro přímé připojení k prolézaným úložištím rozhraní API pro integraci obsahu. Pracuje-li server pro integraci obsahu v přímém režimu, nejsou k dispozici všechny funkce tohoto serveru. Informace o provozování serveru pro integraci obsahu v přímém režimu a o rozdílech dostupných funkcí ve srovnání se serverem pro integraci obsahu provozovaným v režimu serveru naleznete v dokumentaci k produktu WebSphere Information Integrator Content Edition.

Tento postup shrnuje kroky potřebné pro nastavení přístupu v přímém režimu. Podrobné pokyny naleznete v článku *Správa dokumentů vyhledávacího portálu WebSphere pomocí produktu WebSphere Information Integrator OmniFind Edition* ze série IBM developerWorks na adrese <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0606lee/>.

### Postup

Chcete-li konfigurovat systém tak, aby prolézací modul mohl přistupovat k úložištím v přímém režimu, postupujte takto:

1. Potvrďte, že proměnné prostředí VBR\_HOME a JAVA\_HOME v souboru *kořenový\_adresář\_instalace\_produkту\_iice/bin/config.sh* (v systému UNIX) nebo *kořenový\_adresář\_instalace\_produkту\_iice\bin\config.bat* (v systému Microsoft Windows) určují správný adresář.
2. Chcete-li konfigurovat konzolu pro správu produktu WebSphere Information Integrator Content Edition tak, aby mohla pracovat v přímém režimu, přidejte do souboru *kořenový\_adresář\_instalace\_produkту\_iice/bin/Admin.sh* (v systému UNIX) nebo *kořenový\_adresář\_instalace\_produkту\_iice\bin\Admin.bat* (v systému Windows) systémovou vlastnost Java **-Dvbr.as.operationMode=direct**. Příklad:

#### Soubor Admin.sh

```
java -classpath \
"$VBR_CLASSPATH" \
-Dvbr.home="$VBR_HOME" \
```

```
-Dvbr.as.operationMode=direct \  
-Dlog4j.category.com.venetica.vbr.tools.admin=WARN \  
com.venetica.vbr.tools.admin.AdminFrame $1 $2 $3 $4
```

#### Soubor Admin.bat

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

3. Spusíte konzolu pro správu produktu WebSphere Information Integrator Content Edition v přímém režimu a konfiguruje konektor pro prolézací server OmniFind Enterprise Edition.
4. Používáte-li ke konfigurování prolézacího modulu Content Edition konzolu pro správu podnikového vyhledávání, vyberte volbu přímého režimu.

## Přístup k úložištím WebSphere II Content Edition v režimu serveru

Prolézací modul Content Edition lze konfigurovat pro přístup k úložištím v režimu serveru.

V režimu serveru je konektor WebSphere Information Integrator Content Edition používaný prolézacím modulem pro přístup k datům instalován jako podniková aplikace na serveru WebSphere Application Server a prolézací modul přistupujete k úložištím prostřednictvím serveru. Tento přístup vám umožňuje využívat prostředí aplikačního serveru J2EE.

Před konfigurováním prolézacího modulu pro přístup k úložištím WebSphere Information Integrator Content Edition v režimu serveru musíte na prolézacím serveru spustit skript. Tento skript se dodává s produktem OmniFind Enterprise Edition a umožňuje prolézacímu modulu Content Edition přistupovat k úložišti na serveru.

Proveďte úlohu odpovídající vašemu prostředí:

- “Konfigurování prolézacího serveru v systému UNIX pro produkt WebSphere II Content Edition”.
- “Konfigurování prolézacího serveru v systému Windows pro produkt WebSphere II Content Edition” na stránce 43.

### Konfigurování prolézacího serveru v systému UNIX pro produkt WebSphere II Content Edition

Pokud jste instalovali produkt OmniFind Enterprise Edition do počítače se systémem IBM AIX či Linux nebo s provozním prostředím Solaris a konfigurovali prolézací modul Content Edition tak, aby při přístupu k úložištím používal režim serveru, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožňuje prolézacímu modulu Content Edition přistupovat k úložištím WebSphere Information Integrator Content Edition.

#### O této úloze

Prolézací modul Content Edition používá knihovnyJava produktu WebSphere Information Integrator Content Edition jako klienta Java. V režimu serveru tyto knihovny Java vyžadují knihovny produktu WebSphere Application Server související se sadou. Chcete-li se ujistit, že prolézací modul Content Edition může pracovat s knihovnami Java, musíte spusťte po instalaci produktu WebSphere Application Server spustit instalační skript, který je k dispozici na prolézacím serveru produktu OmniFind Enterprise Edition.

Produkt WebSphere Information Integrator Content Edition je instalován na prolézací server při instalaci produktu OmniFind Enterprise Edition. Chcete-li mít možnost používat prolézací modul Content Edition v režimu serveru, musíte zkopírovat soubor `vbr_access_services.jar` ze serveru WebSphere Information Integrator Content Edition na prolézací server.

## Postup

Chcete-li konfigurovat prolézací server, aby mohl prolézat úložiště WebSphere Information Integrator Content Edition, postupujte takto:

1. Je-li produkt OmniFind Enterprise Edition instalován v konfiguraci s více servery, instalujte a svažte knihovny Java produktu WebSphere Application Server.
2. Na prolézacím serveru spusíte skript pro nastavení prolézacího modulu Content Edition:
  - a. Přihlašte se jako administrátor podnikového vyhledávání.
  - b. Spusíte následující skript instalovaný v adresáři \$ES\_INSTALL\_ROOT/bin a odpovězte na výzvy, které se zobrazí:  
escrvbr.sh
3. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

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

4. Zkopírujte soubor vbr\_access\_services.jar ze serveru WebSphere Information Integrator Content Edition na prolézací server.

### Kopírovat z:

Soubor vbr\_access\_services.jar je uložen v následujícím výchozím umístění:

```
kořenový_instalační_adresář_was/installedApps/název_serveru/  
název_aplikace
```

Adresář *kořenový\_instalační\_adresář\_was* je instalační adresář aplikačního serveru WebSphere, *název\_serveru* je název serveru, který jste zadali, a *název\_aplikace* je zadaný název aplikace WebSphere Information Integrator Content Edition na serveru WebSphere Application Server.

### Kopírovat do:

Cílový adresář na prolézacím serveru je *kořenový\_instalační\_adresář\_iice/lib*, kde *kořenový\_instalační\_adresář\_iice* je instalační adresář produktu WebSphere Information Integrator Content Edition na prolézacím serveru.

## Konfigurování prolézacího serveru v systému Windows pro produkt WebSphere II Content Edition

Pokud instalujete produkt OmniFind Enterprise Edition do počítače se systémem Microsoft Windows a konfiguruje prolézací modul Content Edition tak, aby při přístupu k úložištím používal režim serveru, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožňuje prolézacímu modulu Content Edition přistupovat k úložištím WebSphere Information Integrator Content Edition.

### O této úloze

Prolézací modul Content Edition používá knihovnyJava produktu WebSphere Information Integrator Content Edition jako klienta Java. V režimu serveru tyto knihovny Java vyžadují knihovny produktu WebSphere Application Server související se sadou. Chcete-li se ujistit, že prolézací modul Content Edition může pracovat s knihovnami Java, musíte spusíte po instalaci produktu WebSphere Application Server spustit instalační skript, který je k dispozici na prolézacím serveru produktu OmniFind Enterprise Edition.

Produkt WebSphere Information Integrator Content Edition je instalován na prolézací server při instalaci produktu OmniFind Enterprise Edition. Chcete-li mít možnost používat prolézací modul Content Edition v režimu serveru, musíte zkopírovat soubor vbr\_access\_services.jar ze serveru WebSphere Information Integrator Content Edition na prolézací server.

## Postup

Chcete-li konfigurovat prolézací server, aby mohl prolézat úložiště WebSphere Information Integrator Content Edition, postupujte takto:

1. Je-li produkt OmniFind Enterprise Edition instalován v konfiguraci s více servery, instalujte a svažte knihovny Java produktu WebSphere Application Server.
2. Na prolézacím serveru spusťte skript pro nastavení prolézacího modulu Content Edition:
  - a. Přihlašte se pod jménem administrátora podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition).
  - b. Spusťte následující skript instalovaný v adresáři %ES\_INSTALL\_ROOT%\bin a odpovězte na výzvy, které se zobrazí:  
escrvbr.vbs
3. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

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

4. Zkopírujte soubor vbr\_access\_services.jar ze serveru WebSphere Information Integrator Content Edition na prolézací server.

### Kopírovat z:

Soubor vbr\_access\_services.jar je uložen v následujícím výchozím umístění:

```
kořenový_instalační_adresář_was\installedApps\název_serveru\  
název_aplikace
```

Adresář *kořenový\_instalační\_adresář\_was* je instalační adresář aplikačního serveru WebSphere, *název\_serveru* je název serveru, který jste zadali, a *název\_aplikace* je zadaný název aplikace WebSphere Information Integrator Content Edition na serveru WebSphere Application Server.

### Kopírovat do:

Cílový adresář na prolézacím serveru je *kořenový\_instalační\_adresář\_iice\lib*, kde *kořenový\_instalační\_adresář\_iice* je instalační adresář produktu WebSphere Information Integrator Content Edition na prolézacím serveru.

---

## Prolézací moduly DB2

Pomocí prolézacího modulu DB2 lze do kolekce zahrnout databáze IBM DB2.

Pokud použijete nástroj IBM WebSphere Information Integrator k federování a vytvoření přezdívek tabulek pro následující typy databázových systémů, můžete tyto tabulky zpracovávat prolézacími moduly DB2 prostřednictvím definovaných přezdívek:

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

Pro každý databázový server, který chcete prolézat, musíte konfigurovat samostatný prolézací modul. Při konfigurování prolézacího modulu zadáváte volby určující, jakým způsobem má

prolézací modul zpracovávat všechny databáze na daném serveru. Můžete také vybrat konkrétní tabulky, které chcete v jednotlivých databázích prolézat.

Tabulkami vybranými pro prolézání musí být databázové tabulky, přezdívkové tabulky nebo pohledy. Prolézací modul DB2 nepodporuje spojené tabulky.

**Tip:**

Podrobné příklady konfigurace zabezpečeného prolézacího modulu DB2 najdete ve scénáři pro velké organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

## Konfigurace prolézacího serveru

Před prolézáním databázových tabulek musíte na prolézací server nainstalovat klienta DB2 Administration Client. Poté musíte na prolézacím serveru spustit skript. Tento skript se dodává s produktem OmniFind Enterprise Edition a umožňuje prolézacímu modulu DB2 komunikovat se servery.

Před konfigurováním prolézacího modulu DB2 prostřednictvím konzoly pro správu podnikového vyhledávání proveďte úlohu, která odpovídá vašemu prostředí:

- “Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2” na stránce 46.
- “Konfigurování prolézacího serveru v systému Windows pro prolézací moduly DB2” na stránce 47.

## Publikování událostí

Pokud používáte produkt WebSphere Information Integrator Event Publisher Edition a asociujete databáze, které chcete prolézat, s mapováním fronty publikování, může prolézací modul DB2 používat tato mapování k prolézání aktualizací databázových tabulek.

Mapování fronty publikování identifikuje frontu produktu WebSphere MQ, která přijímá zprávy XML při publikování aktualizací databázové tabulky. Prolézací modul přijme z fronty informace o příslušných publikovaných událostech a při aktualizaci tabulek aktualizuje prolézáný prostor (při prvním prolézání tabulky prolézací modul prolézá všechny dokumenty).

Publikování událostí umožňuje zpřístupnit nové a změněné dokumenty pro prohledávání rychleji než dokumenty, které prolézací modul prolézá na základě příslušného časového plánu.

Pokud jsou některé nebo všechny tabulky konfigurovány pro použití služby publikování událostí, můžete zadat při konfigurování prolézacího modulu zadat informace, které prolézacímu modulu umožní přístup k produktu WebSphere MQ a k mapování fronty publikování.

Musíte rovněž zajistit, že jsou na serveru, který má být prolézán, konfigurovány produkty WebSphere MQ a WebSphere Information Integrator Event Publisher Edition a že je na prolézacím serveru konfigurován klientský modul WebSphere MQ. Chcete-li používat službu publikování událostí v kombinaci s prolézacím modulem DB2, proveďte následující úlohy:

- “Konfigurování produktu WebSphere MQ pro prolézací moduly DB2” na stránce 50.
- “Konfigurování produktu WebSphere Information Integrator Event Publisher Edition pro prolézací moduly DB2” na stránce 48.

## Přehled konfigurace

Chcete-li vytvořit nebo změnit prolézací modul DB2, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu stanovují způsob prolézání všech databází na daném databázovém serveru.

- Zadat informace o typech databází, které chcete prolézat.

Hodláte-li prolézat vzdálené databáze, které nejsou katalogizovány na lokálním databázovém serveru, musíte dříve, než budete moci použít prolézací modul DB2 k prolézání těchto databází, spustit na vzdáleném serveru produkt DB2 Administration Server. Při konfigurování prolézacího modulu musíte navíc zadat hostitelský název a port vzdáleného databázového serveru.

- Určit databáze, které chcete prolézat.
- Zadat jména uživatelů a hesla, která prolézacímu modulu umožní přístup k databázím využívajícím řízení přístupu.
- Nastavit časový plán pro prolézání databáze.
- Vybrat tabulky, které chcete v jednotlivých databázích procházet.

**Upozornění:** Chcete-li optimalizovat výkon zjišťovacích procesů (a zabránit vypršení časového limitu procesu konfigurace prolézacího modulu), vyberte prolézání všech tabulek pouze v případě, že databáze neobsahuje příliš mnoho tabulek nebo že jednotlivé tabulky obsahují pouze malý počet polí. Vyberete-li nějaké tabulky k prolézání již nyní, budete moci později upravit prolézáný prostor a přidat do kolekce další tabulky.

- Vybrat tabulky, které mají být prolézány, jsou-li pro ně publikovány aktualizace ve frontě publikování aktualizací, a zadejte informace, které prolézacímu modulu umožní přístup k frontě publikování událostí.
- Zadat volby umožňující prolézání sloupců v konkrétních tabulkách. Můžete například povolit použití určitých sloupců v parametrických dotazech nebo rozhodnout o tom, které sloupce smí být vráceny ve výsledcích vyhledávání.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

## Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2

Pokud jste instalovali produkt OmniFind Enterprise Edition do počítače se systémem IBM AIX či Linux nebo s provozním prostředím Solaris, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožní prolézacímu modulu DB2 komunikaci s databázovými servery. Pokud používáte publikování událostí, skript rovněž umožní prolézacímu modulu přístup ke správcům front a frontám produktu WebSphere MQ.

**O této úloze**



Chcete-li prolézacímu modulu DB2 umožnit prolézání databázových tabulek, spusťte instalační skript `escrdb2.sh`, který je v produktu OmniFind Enterprise Edition k dispozici na prolézacím serveru.

Než skript spustíte, musíte se ujistit, zda je na prolézacím serveru nainstalován klient DB2 Administration Client.

Pokud používáte publikování událostí, musíte na prolézací server nainstalovat moduly WebSphere MQ 5.3 pro službu Java Messaging, aby měl prolézací modul DB2 přístup ke správcům front a frontám WebSphere MQ. Instalační skript `escrdb2.sh` musíte spustit až po instalaci modulů WebSphere MQ.

### Postup

Chcete-li konfigurovat prolézací server pro podporu prolézacích modulů DB2, postupujte takto:

1. Volitelné: Pokud máte v úmyslu používat publikování událostí, nainstalujte na prolézací server moduly produktu WebSphere MQ 5.3 pro službu Java Messaging:
  - a. Přihlaste se jako uživatel root a zadejte následující příkaz:

```
export LD_ASSUME_KERNEL=2.4.19
```
  - b. Vložte disk CD produktu WebSphere MQ.
  - c. Přejděte do adresáře obsahujícího moduly MQ pro službu Java Messaging.
  - d. Instalujte moduly zadáním následujícího příkazu:

```
rpm -i MQSeriesJava-5.3.0-1.i386.rpm
```
2. Na prolézacím serveru spusťte skript pro nastavení prolézacího modulu DB2:
  - a. Přihlaste se jako administrátor podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition).
  - b. Spusťte následující skript instalovaný v adresáři `$ES_INSTALL_ROOT/bin` a odpovězte na výzvy, které se zobrazí:

```
escrdb2.sh
```
3. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

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

## Konfigurování prolézacího serveru v systému Windows pro prolézací moduly DB2

Pokud jste instalovali produkt OmniFind Enterprise Edition do počítače se systémem Microsoft Windows, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožní prolézacímu modulu DB2 komunikaci s databázovými servery. Pokud používáte publikování událostí, skript rovněž umožní prolézacímu modulu přístup ke správcům front a frontám produktu WebSphere MQ.

### O této úloze

Chcete-li prolézacímu modulu DB2 umožnit prolézání databázových tabulek, spusťte instalační skript `escrdb2.vbs`, který je v produktu OmniFind Enterprise Edition k dispozici na prolézacím serveru.

Než skript spustíte, musíte se ujistit, zda je na prolézacím serveru nainstalován klient DB2 Administration Client.

Pokud používáte publikování událostí, musíte na prolézací server nainstalovat moduly WebSphere MQ 5.3 pro službu Java Messaging, aby měl prolézací modul DB2 přístup ke správcům front a frontám WebSphere MQ. Instalační skript `escrdb2.vbs` musíte spustit až po instalaci modulů WebSphere MQ.

### Postup

Chcete-li konfigurovat prolézací server pro podporu prolézacích modulů DB2, postupujte takto:

1. Volitelné: Pokud máte v úmyslu používat publikování událostí, nainstalujte na prolézací server moduly produktu WebSphere MQ 5.3 pro službu Java Messaging:
  - a. Vložte disk CD produktu WebSphere MQ.
  - b. Spusťte instalační program produktu WebSphere MQ.
  - c. V okně Choose Product Features vyberte volbu instalace **Java Messaging**.
2. Na prolézacím serveru spusťte skript pro nastavení prolézacího modulu DB2:
  - a. Přihlašte se pod jménem administrátora podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition).
  - b. Spusťte následující skript instalovaný v adresáři `%ES_INSTALL_ROOT%\bin` a odpovzte na výzvy, které se zobrazí:  
`escrdb2.vbs`
3. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

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

## Konfigurování produktu WebSphere Information Integrator Event Publisher Edition pro prolézací moduly DB2

Před konfigurováním prolézacího modulu DB2 pro používání služby publikování událostí ověřte, že je na serveru, ze kterého bude prolézací modul přijímat data, konfigurován produkt IBM WebSphere Information Integrator Event Publisher Edition.

### O této úloze

Při konfigurování produktu WebSphere Information Integrator Event Publisher Edition pro použití s prolézacím modulem DB2 postupujte podle následujících pokynů:

- Změněné i nezměněné sloupce ve zdrojových tabulkách musí být vybrány pro publikování.
- Odstraněné řádky ve zdrojových tabulkách musí být vybrány pro publikování.
- Fronta publikování událostí nemůže být sdílena více databázemi.
- Jedna databáze může mít přiřazeno více map front a více front.
- K tabulce by měla být přiřazena jedna položka publikování ve formátu XML asociovaná s jednou mapou fronty publikování. (K tabulce by neměla být přiřazena více než jedna položka publikování ve formátu XML asociovaná s jednou mapou fronty publikování. K tabulce může být přiřazena více než jedna položka publikování ve formátu XML, je-li každá položka publikování ve formátu XML asociovaná s jinou mapou fronty publikování.)

### Postup

Provedením následujících kroků konfiguruje databázový server, aby mohl prolézací modul DB2 přistupovat k aktualizacím tabulek publikovaným ve frontě publikování událostí. (Potřebujete-li pomoc při provádění těchto kroků, naleznete příslušné informace v dokumentaci k produktu WebSphere Information Integrator Publisher Edition.)

1. Instalujte produkt WebSphere Information Integrator Event Publisher Edition na databázovém serveru, který má být prolézán.
2. Spusťte příruční panel replikačního centra:

**Systém AIX, Linux, nebo Solaris**  
db2rc

#### **Příkazový řádek systému Windows**

Klepněte na příkaz **Start → IBM DB2 Replication Center.**

3. Vytvořte řídicí tabulky Q-zachycení:
  - a. Jako zobrazení příručního panelu vyberte volbu **Publikování událostí**, vyberte volbu **Vytvořit řídicí tabulky pro Q-zachycení** a poté klepněte na tlačítko **Další**.
  - b. V poli **Server pro Q-zachycení** vyberte ze seznamu databázových serverů k dispozici server, který chcete používat jako server pro Q-zachycení, a klepněte na tlačítko **OK**.
  - c. Zadejte jméno uživatele a heslo s oprávněním pro přístup k vybranému serveru pro Q-zachycení. Změňte schéma Q-zachycení nebo potvrďte výchozí název schématu a klepněte na tlačítko **Další**.
  - d. Zadejte názvy správce fronty, fronty administrace a fronty restartu zadané při konfigurování produktu WebSphere MQ na tomto databázovém serveru a klepněte na tlačítko **Další**.
  - e. Klepněte na tlačítko **Dokončit**. Po zobrazení stránky se zprávami a skripty SQL klepněte na tlačítko **Zavřít**.
  - f. Jako volbu zpracování vyberte hodnotu **Spustit nyní** a klepněte na tlačítko **OK**. Po zobrazení zprávy s informací, že zpracování skriptů SQL bylo dokončeno, klepněte na tlačítko **Zavřít**.
4. Vytvořte publikování ve formátu XML:
  - a. V příručním panelu replikačního centra vyberte jako zobrazení příručního panelu volbu **Publikování událostí**, vyberte volbu **Vytvořit publikování XML** a poté klepněte na tlačítko **Další**.
  - b. Na stránce Spuštění klepněte na tlačítko **Další**.
  - c. Na stránce Server a mapa fronty potvrďte, že je server pro Q-zachycení a schéma Q-zachycení správné, klepněte na volbu u pole **Mapa fronty publikování** a klepnutím na příkaz **Nové** vytvořte mapu fronty publikování.
  - d. Na stránce Obecné zadejte název pro mapu fronty.
  - e. Na stránce Vlastnosti zadejte název fronty odesílání (například název datové fronty, který jste zadali při konfigurování produktu WebSphere MQ na tomto serveru), jako typ obsahu zprávy vyberte volbu **Řádková operace** nebo **Transakce**, zrušte zaškrtnutí políček pro odesílání synchronizačních zpráv a přidávání hlaviček zpráv JMS a klepněte na tlačítko **OK**.
  - f. Po zobrazení stránky se zprávami a skripty SQL klepněte na tlačítko **Zavřít**.
  - g. Jako volbu zpracování vyberte hodnotu **Spustit nyní** a klepněte na tlačítko **OK**. Po zobrazení zprávy s informací, že zpracování skriptů SQL bylo dokončeno, klepněte na tlačítko **Zavřít**.
  - h. Na stránce Vybrat mapu fronty publikování vyberte mapu fronty, kterou jste vytvořili, a klepněte na tlačítko **OK**.
  - i. Na stránce Server a mapa fronty potvrďte, že je název mapy fronty správný, a klepněte na tlačítko **Další**.
  - j. Na stránce Zdrojová tabulka klepněte na volbu **Přidat**, poté klepněte na volbu **Načíst vše**, vyberte tabulku, pro kterou chcete povolit publikování událostí, klepněte na tlačítko **OK** a poté klepněte na tlačítko **Další**.
  - k. Na stránce Sloupce a řádky vyberte sloupce, které má prolézací modul DB2 prolézat (nebo všechny sloupce), a vyberte sloupce klíče. Na stránce pro výběr řádků, které

mají být prolézány (nebo všech řádků), vyberte volbu publikování odstranění ze zdrojové tabulky. Po dokončení konfigurování uvedených voleb klepněte na tlačítko **Další**.

- l. Na stránce Obsah zpráv vyberte volbu zahrnutí dat změněných i nezměněných sloupců a vyberte volbu pouze nových hodnot dat. Ověřte, že je políčko automatického spuštění publikování XML zaškrtnuto, a klepněte na tlačítko **Další**.
  - m. Na stránce Zkontrolovat a dokončit publikování XML klepněte na tlačítko **Další**.
  - n. Na stránce Souhrn klepněte na tlačítko **Dokončit**. Po zobrazení stránky se zprávami a skripty SQL klepněte na tlačítko **Zavřít**.
  - o. Jako volbu zpracování vyberte hodnotu **Spustit nyní** a klepněte na tlačítko **OK**. Po zobrazení zprávy s informací, že zpracování skriptů SQL bylo dokončeno, klepněte na tlačítko **Zavřít**.
5. Spusťte server pro Q-zachycení:
- a. Zavřete příruční panel replikačního centra a spusťte replikační centrum.
  - b. Ve stromu objektů klepněte na položky **Q-replikace** → **Definice** → **Servery pro Q-zachycení**.
  - c. Klepněte pravým tlačítkem myši na ikonu serveru pro Q-zachycení, který jste konfigurovali, a vyberte volbu **Zpřístupnit databázi pro Q-replikaci**.
  - d. Po zobrazení varování klepněte na tlačítko **OK**.
  - e. Po zobrazení stránky se zprávami produktu DB2 klepněte na tlačítko **Zavřít**.
  - f. Ve stromu objektů klepněte pravým tlačítkem myši na ikonu serveru pro Q-zachycení a vyberte volbu **Spustit program Q-Capture**.
  - g. Jako volbu zpracování vyberte hodnotu **Spustit nyní**, zadejte název systému, jméno uživatele a heslo pro uživatele produktu DB2, cestu k adresáři, do kterého jsou ukládány žurnály, a název instance DB2. Poté klepněte na tlačítko **OK**.
  - h. Po zobrazení zprávy s informací, že byl požadavek odeslán, klepněte na tlačítko **Zavřít**.
  - i. Ve stromu objektů klepněte pravým tlačítkem myši na ikonu serveru pro Q-zachycení a vyberte volbu **Zkontrolovat stav**.
- Zobrazí se stav serveru pro Q-zachycení. Došlo-li k chybám, je ve stavové zprávě uvedeno, že server zřejmě neběží. Pokud si chcete prohlédnout žurnály a určit příčinu jakýchkoli chyb, zadejte na příkazový řádek následující příkaz:
- ```
asnqcap Capture_Server=název_serveru_pro_zachycení LOGSTDOUTOUT=y
```

## Konfigurování produktu WebSphere MQ pro prolézací moduly DB2

Před konfigurováním prolézacího modulu DB2 pro používání služby publikování událostí ověřte, že je na serveru, ze kterého bude prolézací modul přijímat data, konfigurován produkt IBM WebSphere MQ.

### Než začnete

Ověřte, že jsou na cílovém databázovém serveru instalovány produkty DB2, WebSphere Information Integrator Event Publisher Edition a WebSphere MQ.

### Omezení

Je-li cílový databázový server instalován v počítači se systémem Linux, všichni uživatelé produktu DB2, uživatelé produktu WebSphere MQ a uživatelé produktu OmniFind Enterprise Edition musí nastavit následující proměnnou prostředí:

```
export LD_ASSUME_KERNEL=2.4.19
```

Tato proměnná prostředí umožňuje exportování implementací vláken LinuxThread z libovolného shellu, ve kterém je prováděna instalace, spouštěny řídicí příkazy produktu WebSphere MQ nebo spouštěny aplikace WebSphere MQ. Produkt WebSphere MQ vyžaduje exportování této proměnné prostředí.

Prolézací moduly DB2, které využívají publikování událostí, se připojují k frontám produktu WebSphere MQ pomocí klientského připojení. Chcete-li povolit klientská připojení, přihlaste se jako administrátor produktu WebSphere MQ a spuštěním následujícího příkazu nastavte identifikátor CCSID správce fronty na hodnotu 819:

```
runmqsc  
název_správce_fronty  
ALTER_QMGR CCSID(819)  
END
```

### O této úloze

Prolézací modul DB2 podporuje režim připojení klienta k serveru WebSphere MQ. Prolézací modul přijímá zprávy ve formátu XML publikované ve frontě publikování událostí. Prolézací modul nemůže přijímat zprávy ve formátu XML přenášené prostřednictvím více než jedné fronty.

Po konfigurování produktu WebSphere MQ prolézací modul DB2 pomocí názvu správce front, názvu fronty, názvu hostitele serveru, čísla portu serveru a názvu kanálu serveru získává z fronty publikování zprávy ve formátu XML. Prolézací modul zprávy analyzuje a aktualizuje prolézáný prostor informacemi o aktualizovaných tabulkách.

### Postup

Provedením následujících kroků konfiguruje databázový server, aby mohl prolézací modul DB2 přijímat data z fronty publikování událostí. (Potřebujete-li pomoc při provádění těchto kroků, naleznete příslušné informace v dokumentaci k produktu WebSphere MQ.)

1. Přihlaste se s použitím role administrátora produktu WebSphere MQ a zadáním následujících příkazů vytvořte správce front a fronty.
  - a. Na příkazový řádek zadejte následující příkaz:  
crtmqm QM1
  - b. Po zobrazení zprávy Setup completed (Instalace dokončena) zadejte následující příkaz:  
strmqm QM1
  - c. Po zobrazení zprávy 'QM1' started (Správce front QM1 byl spuštěn) zadejte následující příkaz:  
runmqsc QM1
  - d. Po zobrazení zprávy Starting MQSC for queue manager QM1 (Probíhá spouštění modulu MQSC pro správce front QM1) vytvořte frontu správy zadáním následujícího příkazu:  
DEFINE QLOCAL('ASN.QM1.ADMINQ')
  - e. Po zobrazení zprávy WebSphere MQ queue created (Fronta produktu WebSphere MQ byla vytvořena) restartujte frontu zadáním následujícího příkazu:  
DEFINE QLOCAL('ASN.QM1.RESTARTQ')
  - f. Po opětovném zobrazení zprávy WebSphere MQ queue created (Fronta produktu WebSphere MQ byla vytvořena) vytvořte datovou frontu zadáním následujícího příkazu:  
DEFINE QLOCAL('ASN.QM1.DATAQ')

- g. Po opětém zobrazení zprávy **WebSphere MQ queue created** (Fronta produktu WebSphere MQ byla vytvořena) ukončete práci zadáním následujícího příkazu:
- ```
end
```
2. Zadáním následujícího příkazu spusíte modul pro příjem dat produktu MQ na databázovém serveru (modul pro příjem dat produktu MQ musí být při vytváření prolézacího modulu DB2, který využívá službu publikování událostí, spuštěn). V tomto příkladu hodnota 1414 odpovídá číslu portu serveru a je použit výchozí kanál SYSTEM.DEF.SVRCONN:
- ```
runmqtsr -m QM1 -t TCP -p 1414 &
```
3. Zadáním následujících příkazů udělte uživateli produktu DB2 oprávnění pro přístup ke správci front a k frontám prostřednictvím rozhraní MQI (Message Queuing Interface) pro publikování událostí (v tomto příkladu je použito jméno uživatele db2inst1):
- ```
setmqaut -m QM1 -t qmgr -p db2inst1 +allmqi
setmqaut -m QM1 -t queue -n ASN.QM1.DATQAQ -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. Zadejte následující příkazy pro jméno uživatele použité k vytvoření a spuštění prolézacího modulu DB2 s použitím služby publikování událostí. Tyto příkazy udělí příslušnému jménu uživatele oprávnění pro přístup ke správci front a k frontám prostřednictvím rozhraní MQI (Message Queuing Interface) pro publikování událostí. V tomto příkladu je použito jméno uživatele esuser:
- ```
setmqaut -m ASN.QM1.QM2 -t qmgr -p esuser +allmqi
setmqaut -m ASN.QM1.QM2 -t queue -n ASN.QM1.DATQAQ -p esuser +allmqi
```

## Prolézání databází DB2 na klasickém serveru zdroje dat

Prolézací modul DB2 dokáže prolézat databázi DB2 umístěnou na klasickém serveru zdroje dat prostřednictvím produktu WebSphere Information Integrator Classic Federation.

### O této úloze

Aby bylo možné prolézat databázi DB2 na klasickém serveru zdroje dat, musí být tato databáze sdružena (federována) s databází DB2 umístěnou na jiném než klasickém serveru zdroje dat pomocí modulu wrapper ODBC produktu WebSphere Information Integrator Classic Federation.

### Postup

Chcete-li provést federování databáze, postupujte takto:

1. Na klasický server zdroje dat nainstalujte produkt WebSphere Information Integrator Classic Federation.
2. Na server zdroje dat jiného než klasického typu, na němž je umístěna databáze, která bude použita pro federování s databází na klasickém serveru zdroje dat, nainstalujte modul klienta WebSphere Information Integrator Classic Federation.
3. Nakonfigurujte ovladač ODBC produktu WebSphere Information Integrator Classic Federation pro připojení ke klasickému serveru zdroje dat.
4. Připojte se k federující databázi a vytvořte modul wrapper ODBC pro federování s databází na klasickém serveru zdroje dat.
5. Nakonfigurujte prolézací modul DB2 pro prolézání federující databáze na serveru zdroje dat jiného než klasického typu. Tento postup umožňuje prolézat databázi DB2 umístěnou na klasickém serveru zdroje dat prostřednictvím produktu WebSphere Information Integrator Classic Federation.

---

## Prolézací moduly DB2 Content Manager

Chcete-li do kolekce podnikového vyhledávání zahrnout typy položek produktu IBM DB2 Content Manager, musíte konfigurovat prolézací modul DB2 Content Manager.

### Konfigurace prolézacího serveru

Chcete-li prolézat server DB2 Content Manager, musíte nejprve spustit skript na prolézacím serveru. Tento skript se dodává s produktem OmniFind Enterprise Edition a umožňuje prolézacímu modulu DB2 Content Manager komunikovat se servery DB2 Content Manager.

Před konfigurováním prolézacího modulu DB2 Content Manager prostřednictvím konzoly pro správu podnikového vyhledávání proveďte úlohu, která odpovídá vašemu prostředí:

- “Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2 Content Manager” na stránce 54.
- “Konfigurování prolézacího serveru v systému Windows pro produkt DB2 Content Manager” na stránce 55.

#### Tip:

Podrobné příklady konfigurace zabezpečeného prolézacího modulu DB2 Content Manager najdete ve scénáři pro velké organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

### Přehled konfigurace

Prolézací modul DB2 Content Manager lze použít k prolézání libovolného počtu serverů DB2 Content Manager. Při konfiguraci prolézacího modulu určujete volby, jakým způsobem bude prolézací modul prolézat všechny servery DB2 Content Manager v prolézaném prostoru. Také určujete typy položek, které chcete prolézat na každém serveru.

Chcete-li vytvořit nebo změnit prolézací modul DB2 Content Manager, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem prolézací modul prolézá všechny typy položek všech serverů DB2 Content Manager v prolézaném prostoru.
- Vybrat servery DB2 Content Manager, které chcete prolézat.
- Zadat jméno uživatele a heslo, která umožní prolézacímu modulu přistupovat k obsahu serverů DB2 Content Manager.
- Nastavit časový plán pro prolézání serverů.
- Vybrat typy položek, které chcete prolézat na každém serveru DB2 Content Manager.
- Určit volby, které nastaví atributy některých typů položek jako vyhledatelné. Můžete například vyloučit z prolézaného prostoru určité typy dokumentů a zadat, které atributy budou vraceny ve výsledcích vyhledávání.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

## Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2 Content Manager

Pokud jste instalovali produkt OmniFind Enterprise Edition do počítače se systémem IBM AIX či Linux nebo s provozním prostředím Solaris, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript dovoluje prolézacímu modulu DB2 Content Manager komunikovat se servery IBM DB2 Content Manager.

### O této úloze

Prolézací modul DB2 Content Manager používá konektor Java pro produkt DB2 Content Manager verze 8 pro přístup k serverům DB2 Content Manager. Tento konektor je instalován při instalaci jednoho z následujících produktů na prolézací server:

- IBM DB2 Information Integrator for Content, verze 8.3 pro systém AIX, Solaris nebo Linux,
- IBM DB2 Information Integrator for Content, verze 8.2 pro systém AIX nebo Solaris,
- IBM DB2 Content Manager Toolkit, verze 8.2 pro systém Linux.

Chcete-li ověřit, že prolézací modul DB2 Content Manager dokáže pracovat s produktem DB2 Content Manager, spusíte po instalaci konektoru skript pro nastavení, který je k dispozici v rámci produktu OmniFind Enterprise Edition na prolézacím serveru.

### Postup

Chcete-li konfigurovat prolézací server, aby mohl prolézat servery DB2 Content Manager, postupujte takto:

1. Instalujte na prolézací server konektor Java pro produkt DB2 Content Manager verze 8:
  - a. Přihlašte se k prolézacímu serveru jako uživatel root:

```
su - root
```
  - b. Spusíte soubor db2profile. Příklad:

```
./home/db2inst/sqllib/db2profile
```
  - c. Exportujte proměnnou prostředí JAVAHOME. Příklad:

```
export JAVAHOME=/usr/IBMJava2-141
```
  - d. Přidejte adresář konektoru Java do proměnné prostředí PATH:

```
export PATH=$PATH:$JAVAHOME/bin
```
  - e. Vložte instalační disk CD produktu DB2 Information Integrator for Content Manager a spusíte průvodce instalací.
  - f. V okně Výběr komponent proveďte následující akce. (Pracujete-li s produktem DB2 Information Integrator for Content verze 8.3, lze okno Výběr komponent zobrazit prostřednictvím volby Vlastní instalace.)
    - 1) V seznamu **Komponenty** vyberte položku **Lokální konektory** a v seznamu **Dílní komponenty** položku **Konektor Content Manager V8**.
    - 2) V seznamu **Komponenty** vyberte položku **Sady nástrojů a ukázky konektorů** a v seznamu **Dílní komponenty** položku **Konektor Content Manager V8**.



- g. Zadejte název databáze, jméno uživatele a heslo pro knihovnu DB2 Content Manager a potvrďte výchozí nastavení pro zbývající okna.
2. Přihlaste se k prolézacímu serveru s použitím jména uživatele, které patří do skupiny administrátorů systému DB2.
3. Katalogizujte vzdálenou databázi serveru knihovny DB2 Content Manager a zkontrolujte, zda se prolézací server může připojit k serveru DB2 Content Manager:

```
db2 catalog tcpip node název_uzlu remote název_hostitele server port
db2 catalog database název_databáze as alias at node
název_uzlu
```

kde:

*název\_uzlu*

Zkrácený název hostitele serveru DB2 Content Manager (například *ibmes*).

*název\_hostitele*

Úplný název hostitele serveru DB2 Content Manager (například *ibmes.ibm.com*).

*port*

Číslo portu serveru DB2 Content Manager.

*název\_databáze*

Název databáze DB2 Content Manager (například *ICMNLSDDB*).

*alias*

Alias databáze DB2 Content Manager (například *CMSVR*).

4. Volitelné: Přihlaste se jako uživatel *root* a vyzkoušejte připojení k databázi:

```
. instalační_adresář_produkту_Information_Integrator_for
_Content/bin/cmbenv81.sh
cd instalační_adresář_produkту_Information_Integrator_for
_Content/samples/java/icm
javac *.java
java SConnectDisconnectICM název_databáze_ICM ID_administrátora
_CM heslo_administrátora_CM
```

5. Na prolézacím serveru spusíte skript pro nastavení prolézacího modulu DB2 Content Manager:
  - a. Změňte adresář *ES\_INSTALL\_ROOT/bin*:

```
cd $ES_INSTALL_ROOT/bin
```
  - b. Spusíte následující skript a odpovězte na výzvy:

```
escrcm.sh
```
6. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

```
esadmin system stopall
```

```
esadmin system startall
```

## Konfigurování prolézacího serveru v systému Windows pro produkt DB2 Content Manager

Pokud jste instalovali produkt OmniFind Enterprise Edition do počítače se systémem Microsoft Windows, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožní prolézacímu modulu DB2 Content Manager komunikovat se serverem IBM DB2 Content Manager.

### O této úloze

Prolézací modul DB2 Content Manager používá konektor Java pro produkt DB2 Content Manager verze 8 pro přístup k serverům DB2 Content Manager. Tento konektor je instalován při instalaci produktu IBM DB2 Information Integrator for Content verze 8.2 nebo verze 8.3 pro systém Windows na prolézací server. Chcete-li ověřit, že prolézací modul DB2 Content Manager dokáže pracovat s produktem DB2 Content Manager, spusťte po instalaci konektoru skript pro nastavení, který je k dispozici v rámci produktu OmniFind Enterprise Edition na prolézacím serveru.

## Postup

Chcete-li konfigurovat prolézací server, aby mohl prolézat servery DB2 Content Manager, postupujte takto:

1. Instalujte na prolézací server konektor Java pro produkt DB2 Content Manager verze 8:
  - a. Vložte instalační disk CD produktu DB2 Information Integrator for Content Manager. Instalační program se spustí automaticky.  
Otevře se průvodce instalací podnikového informačního portálu DB2 Content Manager.
  - b. V okně Výběr komponent proveďte následující akce. (Pracujete-li s produktem DB2 Information Integrator for Content verze 8.3, lze okno Výběr komponent zobrazit prostřednictvím volby Vlastní instalace.)
    - 1) V seznamu **Komponenty** vyberte položku **Lokální konektory** a v seznamu **Díličí komponenty** položku **Konektor Content Manager V8**.
    - 2) V seznamu **Komponenty** vyberte položku **Sady nástrojů a ukázky konektorů** a v seznamu **Díličí komponenty** položku **Konektor Content Manager V8**.
  - c. Zadejte název databáze, jméno uživatele a heslo pro knihovnu DB2 Content Manager a potvrďte výchozí nastavení pro zbývající okna.
2. Katalogizujte vzdálenou databázi serveru knihovny DB2 Content Manager a zkontrolujte, zda se prolézací server může připojit k serveru DB2 Content Manager. Na příkazový řádek prolézacího serveru zadejte následující příkazy:

```
db2 catalog tcpip node název_uzlu remote název_hostitele server port  
db2 catalog database název_databáze as alias at node  
název_uzlu
```

kde:

*název\_uzlu*

Zkrácený název hostitele serveru DB2 Content Manager (například ibmes).

*název\_hostitele*

Úplný název hostitele serveru DB2 Content Manager (například ibmes.ibm.com).

*port*

Číslo portu serveru DB2 Content Manager.

*název\_databáze*

Název databáze DB2 Content Manager (například ICMNLSDB).

*alias*

Alias databáze DB2 Content Manager (například CMSVR).

3. Volitelné: Připojení k databázi můžete otestovat otevřením příkazového řádku a zadáním následujících příkazů:

```
cmbenv81.bat  
cd instalační_adresář_produkту_Information_Integrator_for  
_Content\\samples\\java\\icm
```

```
javac *.java
java SConnectDisconnectICM název_databáze_ICM ID_administrátora
   _CM heslo_administrátora_CM
```

4. Na prolézacím serveru spusíte skript pro nastavení prolézacího modulu DB2 Content Manager:
  - a. Přejděte do adresáře ES\_INSTALL\_ROOT\bin:

```
cd %ES_INSTALL_ROOT%\bin
```
  - b. Spusíte následující skript a odpovíte na výzvy:

```
escrcm.vbs
```
5. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

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

---

## Prolézací moduly Domino Document Manager

Chcete-li do kolekce podnikového vyhledávání zahrnout knihovny a kabinety Domino Document Manager, musíte konfigurovat prolézací modul Domino Document Manager.

### Konfigurace prolézacího serveru

Pokud server Domino Document Manager, který máte v úmyslu prolézat, používá protokol NRPC (Notes Remote Procedure Call), je nutné spustit na prolézacím serveru instalační skript. Tento skript se dodává s produktem OmniFind Enterprise Edition a umožňuje prolézacímu modulu Domino Document Manager komunikovat se servery využívajícími protokol NRPC.

Pokud server Domino Document Manager, který máte v úmyslu prolézat, používá protokol DIIOP (Domino Internet Inter-ORB Protocol), nemusíte na prolézacím serveru spouštět žádný instalační skript. Je však nezbytné nastavit server Domino Document Manager tak, aby k němu měl prolézací modul Domino Document Manager přístup.

**Důležité:** V případě, že server Domino Document Manager používá protokol DIIOP a prolézací modul je konfigurován tak, aby používal protokol HTTPS nebo DIIOP přes SSL z důvodu šifrování přenosu mezi prolézacím modulem a serverem, musíte zkopírovat soubor `TrustedCerts.class` (například `c:\certs` nebo `/data/certs`) ze serveru Domino Document Manager na prolézací server. V případě konfigurace se dvěma nebo čtyřmi servery musíte také zkopírovat soubor `TrustedCerts.class` na servery, na nichž je instalována komponenta pro vyhledávání. Je třeba zajistit, aby se tento soubor nacházel na prolézacím serveru a na vyhledávacích serverech ve stejném umístění. Cesta k adresáři se souborem `TrustedCerts.class` je zadána při konfiguraci prolézacího modulu.

Je-li produkt OmniFind Enterprise Edition instalován v systému IBM AIX, musíte zaručit, že je na prolézacím serveru instalován modul I/O Completion Port a že je tento modul k dispozici.

Před konfigurováním prolézacího modulu Domino Document Manager prostřednictvím konzoly pro správu podnikového vyhledávání, proveďte úlohy, které odpovídají vašemu prostředí:

- “Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino” na stránce 72.
- “Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 73.
- “Konfigurování serverů používajících protokol DIIOP” na stránce 75.

- “Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 76.

## Zabezpečení na úrovni dokumentu

Pokud je povoleno zabezpečení kolekce a server, který máte v úmyslu prolézat, používá protokol NRPC, je nutné konfigurovat na prolézacím serveru produkt Lotus Domino Trusted Server. Prostřednictvím produktu Trusted Server lze prosazovat řízení přístupu na úrovni dokumentu. Před zpřístupněním kolekce pro prohledávání uživateli proveďte následující úlohy:

- Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů
- Povolení globálního zabezpečení na serveru WebSphere Application Server a konfigurování vyhledávací aplikace pro použití zabezpečení. Tento krok zajistí ověřování přihlašovacích údajů při pokusu uživatele o použití vyhledávací aplikace. Vyhledávací servery používají tyto pověřovací údaje k ověření oprávnění každého uživatele pro přístup k dokumentům Lotus Domino.

## Přehled konfigurace

Prolézací modul Domino Document Manager lze použít k prolézání libovolného knihoven Domino Document Manager. Při vytváření prolézacího modulu vyberete knihovny na jednom serveru Domino Document Manager, které chcete prolézat. Při pozdějších úpravách prolézaného prostoru můžete přidat dokumenty z dalšího serveru Domino Document Manager, které chcete zahrnout do stejného prolézaného prostoru. Při vytváření nebo úpravě prolézacího modulu můžete určit, zda chcete prolézat všechny kabinety v knihovnách vybraných pro prolézání nebo zda chcete prolézat konkrétní kabinety.

Chcete-li vytvořit nebo změnit prolézací modul Domino Document Manager, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny dokumenty v prolézaném prostoru.
- Identifikovat server Domino Document Manager a komunikační protokol.
- Pokud je server konfigurován pro použití protokolu DIOP, můžete určit, jakým způsobem se má prolézací modul připojovat k objektům Domino. Můžete například zadat volby pro šifrování komunikace pomocí protokolu HTTPS nebo Secure Sockets Layer (SSL).
- Vybrat knihovny, které chcete prolézat.
- Nastavit časový plán pro prolézání knihoven.
- Vybrat dokumenty, které chcete prolézat. Prolézací modul může prolézat všechny kabinety v knihovně nebo může prolézat pouze dokumenty ve vybraných kabinetech.
- Zadat volby umožňující prohledávání polí v různých knihovnách a kabinetech. Můžete například vyloučit z prolézaného prostoru některá pole a zadat volby pro prohledávání příloh.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

#### **Související pojmy**

“Rady pro prolézání databází Lotus Domino” na stránce 71

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino” na stránce 267

#### **Související úlohy**

“Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino” na stránce 72

“Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 73

“Konfigurování serverů používajících protokol DIIOP” na stránce 75

“Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 76

“Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů” na stránce 268

---

## **Prolézací moduly Exchange Server**

Chcete-li do kolekce podnikového vyhledávání zahrnout veřejné složky na serveru Microsoft Exchange, musíte konfigurovat prolézací modul Exchange Server.

Prolézací modul Exchange Server lze použít k prolézání libovolného počtu složek a podsložek na serverech veřejných složek Exchange Server. Při vytvoření prolézacího modulu vyberte obsah, který chcete na serveru veřejných složek prolézat. Později můžete prolézáný prostor upravit a přidat do něj obsah z jiného serveru veřejných složek.

Chcete-li vytvořit nebo změnit prolézací modul Exchange Server, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat každou z podsložek na všech serverech v prolézáném prostoru.
- Zadat informace o serveru veřejných složek Exchange Server, který chcete prolézat. Musíte zadat jméno uživatele a heslo, které prolézacímu modulu umožní přístup k obsahu na serveru. Pokud server používá protokol SSL (Secure Sockets Layer), můžete zadat volby, které prolézacímu modulu umožní přístup k souboru úložiště klíčů na prolézacím serveru.
- Nastavit plán prolézání serveru veřejných složek.
- Vybrat prolézané složky a podsložky.
- Zadat volby umožňující prohledávání dokumentů v podsložkách. Můžete například vyloučit z prolézáného prostoru určité typy dokumentů.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o

zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

#### Související úlohy

“Ověřování přístupu k zabezpečeným dokumentům serveru Exchange Server” na stránce 267

---

## Prolézací moduly databáze JDBC

Prolézací modul Databáze JDBC slouží k zahrnutí databázi, k nimž lze získat přístup pomocí protokolu JDBC (Java Database Connectivity), do kolekce podnikového vyhledávání.

Pro každý typ databázového systému, který chcete prolézat, musíte konfigurovat samostatný prolézací modul. Při vytvoření prolézacího modulu určujete volby prolézání jedné databáze. Později můžete do prolézaného prostoru přidat další databáze stejného typu.

Každý řádek v tabulce databáze je považován za dokument a hodnoty sloupců databáze jsou analyzovány a indexovány jako pole s možností prohledávání. Přiřazením modulu plug-in k prolézacímu modulu při konfiguraci vlastností prolézacího modulu můžete prolézací modul konfigurovat k prolézání více strukturovaných tabulek. Při použití tohoto modulu plug-in lze řádky více tabulek v relační databázi se stejnými klíčovými poli spojit a pracovat s nimi jako s jedním dokumentem. Pokud uživatel prohledává databázi, dojde v případě zobrazení dokumentu ve výsledcích vyhledávání k zobrazení dat ze spojených tabulek jako dalšího pole.

### Podporované databázové systémy a ovladače

Chcete-li pro prolézání tabulek v databázi použít protokol JDBC, musí na prolézacím serveru existovat příslušný ovladač JDBC. Prolézací modul Databáze JDBC podporuje následující databázové systémy a ovladače JDBC typu 4:

| Databázový systém                                                                                                    | Název ovladače JDBC typu 4                                                                                          | Standardní cesta ke třídě ovladače JDBC                                                                                                                                           |
|----------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| IBM DB2 Universal Database verze 8.2 a IBM DB2 Enterprise Server Edition verze 9.1 pro systémy Linux, UNIX a Windows | com.ibm.db2.jcc.DB2Driver                                                                                           | <i>kořenový_adresář_instalace_produkту_db2</i><br>/java/db2jcc.jar<br><i>kořenový_adresář_instalace_produkту_db2</i><br>/java/db2jcc_license_cu.jar                               |
| Oracle 9i a 10g                                                                                                      | Oracle.jdbc.driver.OracleDriver                                                                                     | <i>domovský_adresář_oracle/jdbc/lib/ojdbc14.jar</i>                                                                                                                               |
| Microsoft SQL Server 2000                                                                                            | com.microsoft.jdbc.sqlserver.SQLServerDriver                                                                        | <i>mssql_jdbc_home/lib/mssqlserver.jar</i><br><i>mssql_jdbc_home/lib/msbase.jar</i><br><i>mssql_jdbc_home/lib/msutil.jar</i>                                                      |
| Microsoft SQL Server 2005                                                                                            | com.microsoft.sqlserver.jdbc.SQLServerDriver<br>(Ovladač JDBC pro SQL Server 2005 není podporován v systémech AIX.) | <i>instalační_adresář/sqljdbc_1.0/lokalita/sqljdbc.jar</i><br>kde text <i>lokalita</i> reprezentuje vaši lokalitu, např.<br><i>instalační_adresář/sqljdbc_1.0/enu/sqljdbc.jar</i> |

## Srovnání prolézacích modulů Databáze JDBC a DB2

Pokud v současné době používáte prolézací modul DB2, pravděpodobně jej budete chtít používat i nadále. Uložená data prolézacího modulu DB2 nelze převést do prolézacího modulu Databáze JDBC.

Prolézací modul DB2 použijte namísto prolézacího modulu Databáze JDBC v následujících situacích:

- Chcete-li prolézat databáze DB2 s ovladačem JDBC typu 2.
- Chcete-li prolézat databáze Oracle a SQL Server federované s databází DB2. Prolézací modul DB2 vám umožňuje přístup ke všem těmto typům databází prostřednictvím přezdívky.
- Chcete-li prolézat databáze DB2 for z/OS, DB2 for iSeries, Informix, Sybase, VSAM, IMS, CA-Datcom nebo Software AG Adabas. Tyto typy databází byste měli federovat s databází DB2 a přistupovat k nim pomocí prolézacího modulu DB2 prostřednictvím přezdívky.
- Chcete-li používat publikování událostí k aktualizaci indexu podnikového vyhledávání při publikování aktualizací databází.

### Přehled konfigurace

Chcete-li vytvořit nebo změnit prolézací modul Databáze JDBC, přihlaste se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny databáze v prolézaném prostoru.
- Určit typ databáze, kterou chcete prolézat.
- Vybrat databázi, kterou chcete prolézat, a v případě nutnosti zadat ID uživatele a heslo, které prolézacímu modulu umožní přístup k databázi.
- Nastavit plán prolézání databáze.
- Vybrat tabulky, které chcete prolézat.

**Upozornění:** Chcete-li optimalizovat výkon zjišťovacích procesů a zabránit vypršení časového limitu procesu konfigurace prolézacího modulu, vyberte prolézání všech tabulek pouze v případě, že databáze neobsahuje příliš mnoho tabulek nebo že jednotlivé tabulky obsahují pouze malý počet polí. Vyberete-li nějaké tabulky k prolézání již nyní, budete moci později upravit prolézaný prostor a přidat do kolekce další tabulky.

- Zadat volby umožňující prolézání sloupců v konkrétních tabulkách. Můžete například povolit použití určitých sloupců v parametrických dotazech nebo rozhodnout o tom, které sloupce smí být vráceny ve výsledcích vyhledávání.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

## Mapy vztahů pro databáze JDBC

Při vytváření pravidel pro modul plug-in, který prolézá více strukturovaných tabulek databáze JDBC, zadejte informace o kořenové tabulce a způsobu spojení nadřazených a podřazených tabulek.

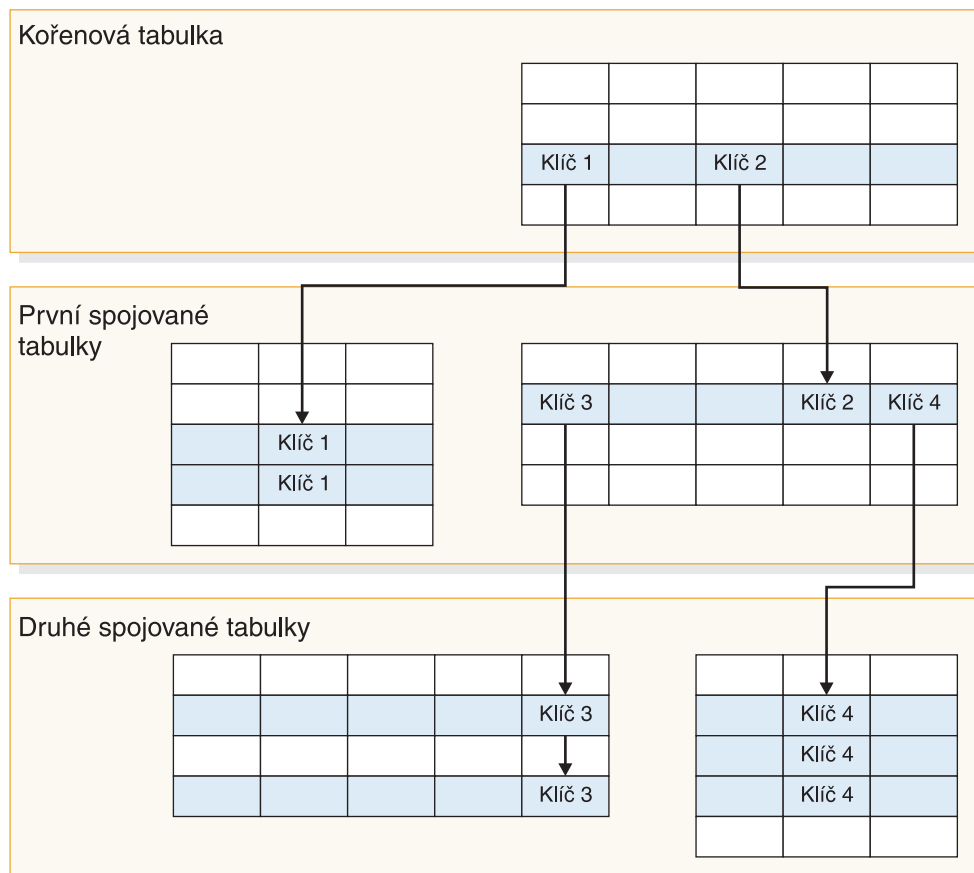
Modul plug-in poskytovaný pro podnikové vyhledávání umožňuje prolézacímu modulu Databáze JDBC spojit více strukturovaných tabulek. Tento modul plug-in můžete vytvořit zadáním pravidel v souboru `ES_INSTALL_ROOT/default_config/crawler_rdb_plugin.xml`. Po konfiguraci prolézacího modulu k používání modulu plug-in dojde ke sloučení řádků z tabulek se stejnými klíčovými poli a je s nimi pracováno jako s jedním dokumentem. Pokud uživatel prohledává databázi, dojde v případě zobrazení dokumentu ve výsledcích vyhledávání k zobrazení dat ze spojených tabulek jako dalšího pole.

### Spojení tabulek pomocí klíčových sloupců

V následujícím obrázku je uveden postup vytvoření mapy vztahů pro více tabulek. Prolézací modul Databáze JDBC skenuje kořenovou tabulku v databázi. Některé sloupce v této tabulce jsou klíčovými poli, které lze použít k spojení tabulek. Sloupce ve spojených tabulkách lze poté použít jako klíče ke spojování dalších tabulek. Řádky více tabulek jsou v prolézaném prostoru považovány za jeden dokument. Hodnoty sloupců jsou považovány za metadata dokumentu. Kořenová tabulka je v tomto vztahu *nadřazeným prvkem* a spojená tabulka na první úrovni je *podřazeným prvkem*. Podřazené tabulky na první úrovni mohou být nadřazenými tabulkami spojenými na druhé úrovni.

V tomto příkladu slouží sloupce Klíč 1 a Klíč 2 v kořenové (nadřazené) tabulce jako klíčová pole umožňující spojení této tabulky s pomocnými tabulkami, které také obsahují sloupce Klíč 1 a Klíč 2. Tabulka spojená na první úrovni obsahuje klíčová pole, Klíč 3 a Klíč 4, které jí umožňují spojení s dalšími tabulkami.



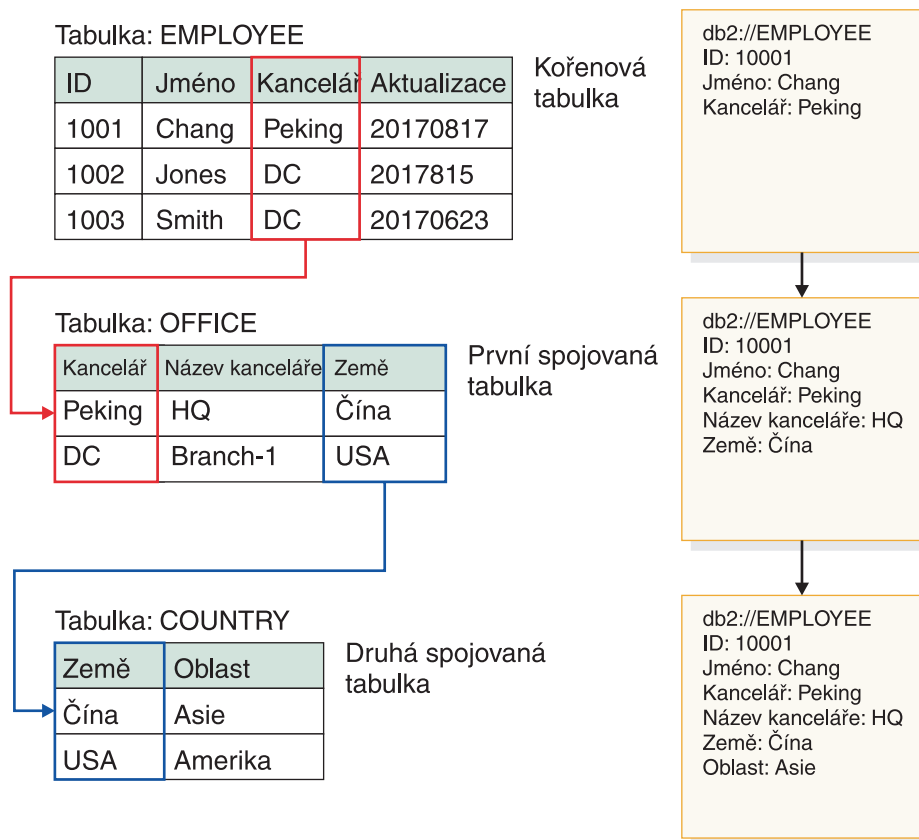


Obrázek 2. Tabulky databáze JDBC spojené pomocí klíčových polí

## Zobrazení výsledků vyhledávání

Na následujícím obrázku je uveden způsob, jakým jsou data z více strukturovaných tabulek prezentována jako jeden dokument ve výsledcích vyhledávání. Bez použití modulu plug-in by se uživatel, který prohledává tabulku ZAMĚSTNANEC, zobrazil ve výsledcích vyhledávání řádek z kořenové tabulky pouze s hodnotami ze sloupců tabulky ZAMĚSTNANEC (ID, Jméno a Kancelář).

Díky modulu plug-in může prolézací modul používat sloupec Kancelář jako klíč ke spojení tabulky ZAMĚSTNANEC s tabulkou KANCELÁŘ. Sloupec Země v tabulce KANCELÁŘ slouží jako klíč ke spojení této tabulky s tabulkou ZEMĚ. Po spojení těchto tabulek se uživateli, který prohledává tabulku ZAMĚSTNANEC, zobrazí hodnoty ze sloupců v tabulkách KANCELÁŘ a ZEMĚ jako další pole ve výsledcích vyhledávání.



Obrázek 3. Hodnoty ze spojených tabulek databáze JDBC se zobrazí ve výsledcích vyhledávání

## Prolézání více strukturovaných tabulek databáze JDBC

Prolézací modul Databáze JDBC můžete konfigurovat ke spojení více strukturovaných tabulek se stejnými klíčovými poli.

### Než začnete

K provedení této úlohy je vyžadováno oprávnění k přihlášení jako administrátor podnikového vyhledávání.

### O této úloze

Při konfiguraci vlastností prolézacího modulu Databáze JDBC můžete zadat modul plug-in pro prolézání více strukturovaných tabulek, které jsou spojeny pomocí klíčových polí. Bez použití tohoto modulu plug-in by byly řádky v databázové tabulce považovány za jednotlivé dokumenty a hodnoty sloupců databáze by bylo možné vyhledávat jako jednotlivá pole. Při použití tohoto modulu plug-in jsou řádky více tabulek v relační databázi se stejnými klíčovými poli spojeny a pracuje se s nimi jako s jedním dokumentem. Prolézací modul přidává data získaná ze spojených tabulek jako metadata původního řádku databázové tabulky. Pokud uživatel prohledává databázi, zobrazí se tato data jako další pole při zobrazení dokumentu ve výsledcích vyhledávání.

### Omezení

### **Datové typy, které nelze prolézat**

Prolézací modul nemůže prolézat pole ve spojených tabulkách, které obsahují následující binární datové typy:

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

### **Omezení oboru prolézaného prostoru**

Spojované tabulky se musí nacházet ve stejné relační databázi. Nelze spojovat tabulky mezi různými databázemi.

Je-li tabulka v databázi konfigurována ke spojení s jinými tabulkami, je toto nastavení univerzální pro všechny prolézací moduly v kolekci, které mohou používat modul plug-in. Je však možné vytvořit více kolekcí a konfigurovat oddělené prolézací moduly k prolézání různých kořenových tabulek a ke spojování různých tabulek.

### **Omezení použití jiných modulů plug-in**

Pokud konfiguruje prolézací modul tak, aby používal modul plug-in k prolézání více strukturovaných tabulek, nelze mu přiřadit jiný modul plug-in. Například není možné zadat vlastní modul plug-in pro používání obchodních pravidel a pravidel zabezpečení. Prolézacímu modulu nelze přiřadit více než jeden modul plug-in.

### **Omezení počtu tabulek, řádků, polí a klíčů**

Maximální počet spojených tabulek v databázi je pět a součet jejich řádků nesmí přesahovat jeden milion. Maximální počet polí, které lze načíst z tabulky, je 10. Ke spojování tabulek je používán klíčový pár. Není tedy možné spojit tabulky pomocí více klíčů.

### **Zajištění prolézání změn ve spojených tabulkách**

Pokud se řádky v kořenové tabulce mezi jednotlivými prolézáními nezmění a prolézací modul není konfigurován k provádění úplného prolézání, ignoruje tento modul nezměněné řádky. Pokud se změní řádky v tabulce, která je spojená s kořenovou tabulkou, a řádky v kořenové tabulce se nezmění, je nutné pomocí některé z následujících akcí zajistit, aby byly změny zjištěny a prolezeny:

- Kořenová tabulka v cílové databázi by měla obsahovat pole časové značky. Konfigurujte cílovou databázi, aby obsahovala pole časové značky, které je aktualizováno při změně řádku v kořenové tabulce nebo při změně libovolné ze spojených pomocných tabulek. Při nastavení prolézacího modulu Databáze JDBC se ujistěte, že jste zadali pole časové značky jako pole, které by měl prolézací modul používat k určení, zda došlo v tabulce ke změnám.
- Při konfiguraci časového plánu prolézacího modulu určete, že prolézací modul má provádět úplné prolézání. Tato volba zajišťuje, že došlo k prolezení všech tabulek bez ohledu na změny.

### **Postup**

Chcete-li nastavit prolézací modul Databáze JDBC k prolézání více strukturovaných tabulek, postupujte takto:

1. Přihlaste se jako administrátor podnikového vyhledávání k prolézacímu serveru a zkopírováním souboru `ES_INSTALL_ROOT/default_config/crawler_rdb_plugin.xml` vytvořte soubor `ES_NODE_ROOT/master_config/crawler_rdb_plugin.xml`.
2. Upravte soubor `ES_NODE_ROOT/master_config/crawler_rdb_plugin.xml` pomocí textového editoru, který podporuje kódování UTF-8.
  - a. Upravte prvek `<Server DBURL="jdbc:db2://db_server_url:50000/SAMPLE">` a nahraďte položku `jdbc:db2://db_server_url:50000/SAMPLE` adresou URL databáze


JDBC, která má být prolézána. Při konfiguraci prolézacího modulu se ujistěte, že jste zadali stejnou adresu URL jako pro databázi, která má být prolézána.

- b. Pokud databáze, která má být prolézána, není databází DB2, upravte prvek `<JDBCdriver>com.ibm.db2.jcc.DB2Driver</JDBCdriver>` a nahraďte ovladač `com.ibm.db2.jcc.DB2Driver` vhodným ovladačem JDBC. Při konfiguraci prolézacího modulu se ujistěte, že jste zadali stejný ovladač jako pro databázi, která má být prolézána.
- c. Upravte prvek `<User>jméno_uživatele</User>` a nahraďte položku `jméno_uživatele` jménem uživatele, který má oprávnění k přístupu k databázi, která má být prolézána.
- d. Upravte prvek `<Password Encryption="True">šifrované_heslo</Password>` a nahraďte položku `šifrované_heslo` šifrovaným heslem pro dané jméno uživatele. Šifrované heslo můžete zkopírovat ze souboru `ES_NODE_ROOT/master_config/col_název_kolekce.JDBC_název_prolézacího_modulu/jdbccrawler.xml` a vložit je zde. Pokud heslo nevyžaduje šifrování, nahraďte položku `Encryption="True"` položkou `Encryption="False"` a položku `šifrované_heslo` nahraďte textově zadaným heslem.
- e. Pokud ponecháte nastavení prvku `<Delimiters Use="True">`, bude více termínů ve sloupci odděleno znaky (,) definovanými v prvku `<Delimiter>`. Sady termínů v tabulce jsou odděleny znaky (;) definovanými v prvku `<SecondDelimiter>`. Nastavíte-li prvek `<Delimiters Use="True">` na hodnotu `<Delimiters Use="False">`, nebudou oddělovače použity a více polí metadat bude přidáno se stejným názvem pole jako metadata dokumentu.
- f. Používáte-li prvek `<Delimiters Use="True">`, zadejte pomocí úpravy prvků `<Delimiter>`, `</Delimiter>` a `<SecondDelimiter>`; `</SecondDelimiter>` znaky, které budou používány k oddělování hodnot.
- g. Upravte prvek `<RelationMap Root="DB2INST1.TABLE_0">` a nahraďte položku `DB2INST1.TABLE_0` názvem kořenové tabulky, která má být prolézána.
- h. Upravte prvek `<Relation Parent="DB2INST1.TABLE_0" ParentAlias="T0" ParentKey="ID" Child="DB2INST1.TABLE_1" ChildAlias="T1" ChildKey="ID"/>`.
  - Nahraďte položku `Parent="DB2INST1.TABLE_0"` názvem tabulky, která je v daném vztahu nadřizená.
  - Nahraďte položku `ParentKey="T0"` aliasem nadřizené tabulky. Tento alias by měl být v rámci souboru `crawler_rdb_plugin.xml` jedinečný.
  - Nahraďte položku `ParentKey="ID"` názvem sloupce, který je v daném vztahu použit jako klíčové pole.
  - Nahraďte položku `Child="DB2INST1.TABLE_1" ChildAlias="T1" ChildKey="ID"` informacemi o podřizené tabulce, která má být prolézána.

Tato struktura definuje způsob spojování tabulek. Například následující mapa vztahů určuje, že má být prolézána kořenová tabulka s názvem `DB2INST1.TABLE_A`. Tabulky `DB2INST1.TABLE_B` a `DB2INST1.TABLE_C` jsou spojeny pod podmínkou `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. Zopakujte krok 2h a vytvořte prvky `<Relation>` pro všechny vztahy, které spojují tabulky z kořenové tabulky.
- j. Upravte prvek `<Target TableAlias="T1">` a nahraďte položku `TableAlias="T1"` hodnotou `ChildAlias`, která je definována v kroku 2h.

- k. Upravte prvek `<Field Name="ID" FieldName="ID_1" Enabling="True" Searchable="True" FieldSearchable="True" IsContent="True"/>`.
- Nahraďte položku `Name="ID"` názvem sloupce v dokumentech, které mají být prolézány.
  - Nahraďte položku `FieldName="ID_1"` názvem pole metadat v dokumentech, které mají být prolézány. Tato hodnota je používána jako zobrazovaný název sloupce v konzole pro správu podnikového vyhledávání a ve výsledcích vyhledávání.
  - Pokud tento sloupec nemá být zahrnut v metadatech dokumentu, nahraďte položku `Enabling="True"` hodnotou `"False"`.
  - Nahrazením položky `Searchable="True"` hodnotou `"False"` zabráníte uživatelům prohledávat tento sloupec pomocí volného textového dotazu.
  - Nahrazením položky `FieldSearchable="True"` hodnotou `"False"` zabráníte uživatelům prohledávat tento sloupec pomocí názvu sloupce.
  - Nahrazením položky `IsContent="True"` hodnotou `"False"` indikujete, že sloupec neobsahuje obsah s možností prohledávání. Zadáte-li volby `Searchable="True"` a `IsContent="True"`, bude hodnota sloupce použita k detekci duplicitních dokumentů a stane se součástí dynamického souhrnu dokumentu ve výsledcích vyhledávání.
- l. Opakováním kroku 2k vytvořte prvky `<Field>` pro všechny sloupce, které mají být prolézány.
- m. Opakováním kroků 2j na stránce 66 a 2k vytvořte prvky `<Target>` a `<Field>` pro všechny podřízené tabulky, na které je odkazováno v mapě vztahů (`<RelationMap>`).
- n. Opakováním kroků 2g na stránce 66 až 2m vytvořte mapy vztahů pro více kořenových tabulek.
- o. Opakováním kroků 2a na stránce 65 až 2n konfigurujte mapu vztahů pro jinou databázi.
3. Konfigurace prolézacího modulu k používání modulu plug-in:
- a. Otevřete konzolu pro správu podnikového vyhledávání, upravte kolekci a vyberte stránku Prolézání.
  - b. Vytvořte prolézací modul Databáze JDBC nebo vyhledejte prolézací modul, který chcete změnit, a klepněte na tlačítko  **Vlastností prolézacího modulu**.
  - c. Do pole **Název třídy modulu plug-in** zadejte název modulu plug-in používaného k prolézání více strukturovaných tabulek:  
com.ibm.es.plugin.rdb.RDBPlugin
  - d. Do pole **Cesta ke třídě modulu plug-in** zadejte úplné cesty k modulu plug-in a ovladačům JDBC, které modul plug-in používá. Příkladem cesty k ovladači JDBC databáze DB2 v systému Windows může být následující cesta:  
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. Buď klepněte na tlačítko **Další** a pokračujte ve vytváření prolézacího modulu, nebo klepnutím na tlačítko **OK** uložte změny.
4. Chcete-li implementovat soubor crawler\_rdb\_plugin.xml do konfigurace systému, restartujte systém podnikového vyhledávání:

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

---

## Prolézací moduly NNTP

Chcete-li do kolekce podnikového vyhledávání zahrnout články z diskusních skupin NNTP, musíte konfigurovat prolézací modul NNTP.

Prolézací modul NNTP lze použít k prolézání libovolného počtu serverů NNTP. Při konfigurování prolézacího modulu vyberete diskusní skupiny, které chcete prolézat, na jednom serveru NNTP. Při pozdějších úpravách prolézaného prostoru můžete přidat další servery NNTP, které má prolézací modul prolézat.

Při určování diskusních skupin, které mají být prolézány, můžete vybrat skupiny, které mají být do prolézaného prostoru zahrnuty, a skupiny, které z něj mají být vyloučeny. Tímto způsobem můžete prolézacímu modulu snadno povolit prolézání většiny diskusních skupin na serveru a zakázat prolézání několika diskusních skupin, jejichž prohledávání nechcete uživatelům umožnit.

Můžete například zadat pravidla zahrnující všechny diskusní skupiny na určitém serveru NNTP a poté určit, že mají být vyloučeny všechny diskusní skupiny na tomto serveru, jejichž názvy obsahují řetězec `private`.

Chcete-li vytvořit nebo změnit prolézací modul NNTP, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny diskusní skupiny v prolézaném prostoru.
- Určit server NNTP, který má být prolézán. Pokud je server chráněn heslem, musíte zadat ID uživatele a heslo, které prolézací modul použije pro přístup k diskusním skupinám na serveru.
- Nastavit plán prolézání serveru.
- Zadat vzory pro zahrnutí diskusních skupin a vzory, které z prolézaného prostoru některé diskusní skupiny vyloučí.
- Určit, zda má prolézací modul automaticky detekovat jazyk a kódovou stránku prolézáných článků, nebo zda má použít konkrétní jazyk a kódovou stránku.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

---

## Prolézací moduly Notes

Chcete-li do kolekce podnikového vyhledávání zahrnout databáze IBM Lotus Notes, musíte konfigurovat prolézací modul Notes.

### Tip:

Podrobné příklady konfigurace zabezpečeného prolézacího modulu Notes najdete ve scénáři pro velké organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

## Konfigurace prolézacího serveru

Pokud server Lotus Notes, který máte v úmyslu prolézat, používá protokol NRPC (Notes Remote Procedure Call), je nutné spustit na prolézacím serveru instalační skript. Tento skript se dodává s produktem OmniFind Enterprise Edition a umožňuje prolézacímu modulu Notes komunikovat se servery využívajícími protokol NRPC.

Pokud server Lotus Notes, který máte v úmyslu prolézat, používá protokol DIIOP (Domino Internet Inter-ORB Protocol), nemusíte na prolézacím serveru spouštět žádný instalační skript. Je však nezbytné nastavit server Lotus Notes tak, aby k němu měl prolézací modul Notes přístup.

**Důležité:** V případě, že server Lotus Notes používá protokol DIIOP a prolézací modul je konfigurován tak, aby používal protokol HTTPS nebo DIIOP přes SSL z důvodu šifrování přenosu mezi prolézacím modulem a serverem, musíte zkopírovat soubor `TrustedCerts.class` (například `c:\certs` nebo `/data/certs`) ze serveru Lotus Notes na prolézací server. V případě konfigurace se dvěma nebo čtyřmi servery musíte také zkopírovat soubor `TrustedCerts.class` na servery, na nichž je instalována komponenta pro vyhledávání. Je třeba zajistit, aby se tento soubor nacházel na prolézacím serveru a na vyhledávacích serverech ve stejném umístění. Cesta k adresáři se souborem `TrustedCerts.class` je zadána při konfiguraci prolézacího modulu.

Je-li produkt OmniFind Enterprise Edition nainstalován v systému IBM AIX, musíte zaručit, že je na prolézacím serveru nainstalován modul I/O Completion Port a že je tento modul k dispozici.

Před konfigurováním prolézacího modulu Notes prostřednictvím konzoly pro správu podnikového vyhledávání, proveďte úlohy, které odpovídají vašemu prostředí:

- “Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino” na stránce 72.
- “Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 73.
- “Konfigurování serverů používajících protokol DIIOP” na stránce 75.
- “Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 76.

## Zabezpečení na úrovni dokumentu

Pokud je povoleno zabezpečení kolekce a server, který máte v úmyslu prolézat, používá protokol NRPC, je nutné konfigurovat na prolézacím serveru produkt Lotus Domino Trusted Server. Prostřednictvím produktu Trusted Server lze prosazovat řízení přístupu na úrovni dokumentu. Před zpřístupněním kolekce pro prohledávání uživateli proveďte následující úlohy:

- Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů
- Povolení globálního zabezpečení na serveru WebSphere Application Server a konfigurování vyhledávací aplikace pro použití zabezpečení. Tento krok zajistí ověřování přihlašovacích údajů při pokusu uživatele o použití vyhledávací aplikace. Vyhledávací servery používají tyto ověřovací údaje k ověření oprávnění každého uživatele pro přístup k dokumentům Lotus Domino.

## Přehled konfigurace

Prolézací modul Notes lze použít k prolézání libovolného počtu standardních databází Lotus Notes (soubory NSF). Při vytváření prolézacího modulu vyberete databáze nebo adresáře na jednom serveru Lotus Notes, které chcete prolézat. Při pozdějších úpravách prolézaného prostoru můžete přidat dokumenty z dalšího serveru Lotus Notes, které chcete zahrnout do stejného prolézaného prostoru. Při vytvoření nebo úpravě prolézacího modulu můžete určit, zda chcete prolézat všechny databáze či adresáře na serveru nebo pouze konkrétní databáze, pohledy a složky.

Chcete-li vytvořit nebo změnit prolézací modul Notes, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny dokumenty v prolézaném prostoru.
- Zadat hostitelský název, port a komunikační protokol serveru Lotus Notes.
- Pokud je server konfigurován pro použití protokolu DIIOP, můžete určit, jakým způsobem se má prolézací modul připojovat k objektům Domino. Můžete například zadat volby pro šifrování komunikace pomocí protokolu HTTPS nebo Secure Sockets Layer (SSL).
- Vybrat databáze nebo adresáře, které chcete prolézat. Pokud prolézáte adresáře, můžete zadat vzorek určující zahrnuté a vyloučené databáze, čímž lze rozdělit úlohu prolézání rozsáhlých adresářů mezi několik prolézacích modulů.
- Nastavit plán prolézání databází nebo adresářů.
- Vybrat dokumenty, které chcete prolézat. Prolézat můžete všechny dokumenty v adresáři, všechny dokumenty v databázi nebo dokumenty ve vybraných pohledech a složkách databáze.
- Zadat volby umožňující prohledávání polí v různých databázích, pohledech a složkách. Můžete například vyloučit z prolézaného prostoru některá pole a zadat volby pro prohledávání příloh.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení. Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

### Související pojmy

“Rady pro prolézání databází Lotus Domino” na stránce 71

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino” na stránce 267

### Související úlohy

“Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino” na stránce 72

“Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 73



“Konfigurování serverů používajících protokol DIIOP” na stránce 75

“Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 76

“Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů” na stránce 268

## Rady pro prolézání databází Lotus Domino

Před konfigurováním prolézacího modulu Notes zkontrolujte pokyny pro prolézání databází Lotus Domino.

- Databáze Notes založené na standardních šablonách (například databáze diskusí) jsou nejlepším typem databází pro prolézání.
- Prolézací modul Notes použije následující pravidla pro mapování polí:
  - Nejprve jsou registrovány názvy hlavních polí ze standardních šablon Domino.
  - Hodnoty z polí Notes určených v tabulce pravidel mapování jsou použity jako souhrny dokumentů ve výsledcích vyhledávání.
  - Hodnoty z polí Notes, která nejsou určena v tabulce pravidel mapování, nejsou použity v souhrnech dokumentů.
  - Hodnoty z polí Notes mapovaných na pole Title jsou ve výsledcích vyhledávání použity jako nadpis dokumentu.
  - Pole v následující tabulce jsou při výchozím nastavení mapována na názvy vyhledávacích polí:

Tabulka 2. Výchozí pravidla pro mapování polí

Název pole databáze Notes	Název vyhledávacího pole
Title	Title
EventTitle	Title
Subject	Title
Body	Body
Mission	Body
From	Creator
Author	Creator
Keywords	Categories
Categories	Categories
TeamRoomName	Organization
TeamName	Organization
Department	Organization

- Prolézací modul Notes může prolézat všechny typy polí s výjimkou polí vypočtených pro zobrazení.
- Statický text a obrázky umístěné do formuláře Notes nejsou prolézány.
- Při konfigurování prolézacího modulu zaškrtněte políčko **Vše** v části **Prolézat**, chcete-li prolézat všechna pole a maximalizovat data polí, která mají být prolézána (prostřednictvím pole **Prolézat všechna pole kromě** můžete omezit pole, která mají být prolézána).  
Chcete-li minimalizovat prolézání polí, která nejsou potřebná, zrušte zaškrtnutí políčka **Prolézat** pro všechna pole s výjimkou těch, která jsou mapována na vyhledávací pole.

### Související pojmy

“Prolézací moduly Domino Document Manager” na stránce 57

“Prolézací moduly Notes” na stránce 68

## Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino

Pokud instalujete produkt OmniFind Enterprise Edition do počítače se systémem IBM AIX či Linux nebo s provozním prostředím Solaris a máte v úmyslu prolézat servery využívající protokol NRPC (Notes Remote Procedure Call), musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožní produktům Notes, QuickPlace a Domino Document Manager komunikaci s databázovými servery.

### Omezení

Server Domino Server nemůže být spuštěn ve stejném počítači současně s prolézacím modulem Notes, QuickPlace nebo Domino Document Manager, který je konfigurován pro použití protokolu NRPC. Pokud se spustí některý z uvedených prolézacích modulů v době, kdy je spuštěn server Domino Server, dojde k chybě a prolézací modul bude zastaven.

### O této úloze

Prolézací moduly využívající protokol NRPC používají knihovny Domino jako klienta. Tyto knihovny nainstalujete instalací serveru Lotus Domino na prolézací server. Chcete-li se ujistit, že prolézací moduly jsou schopny pracovat s knihovnami Domino, spusťte po instalaci knihoven Domino instalační skript, který je k dispozici na prolézacím serveru produktu OmniFind Enterprise Edition.

### Postup

Chcete-li konfigurovat prolézací server tak, aby mohl prolézat servery Lotus Notes, Lotus QuickPlace a Domino Document Manager, postupujte takto:

1. Vytvořte na prolézacím serveru uživatele `server` a skupinu `notes`:
  - a. Přihlašte se jako uživatel `root`:  
`su - root`
  - b. Přidejte uživatele:  
`useradd server`
  - c. Přidejte heslo pro tohoto uživatele:  
`passwd server`  
Budete vyzváni ke změně hesla.
2. Instalujte na prolézací server produkt Lotus Domino Server:
  - a. Vložte disk CD produktu Domino Server a připojte jej. (Pokud nemáte disk CD k dispozici, můžete si stáhnout instalační obraz).
  - b. Přejděte do složky odpovídající použitému operačnímu systému.  
  
AIX: `cd /mnt/cdrom/aix`  
Linux: `cd /mnt/cdrom/linux`  
Solaris: `cd /mnt/cdrom/solaris`
  - c. Spusťte instalační program:  
`./install`
  - d. Odpovězte na výzvy a přijměte výchozí hodnoty nebo určete vlastní upřednostňované nastavení instalace (například cesty k instalačnímu a datovému adresáři).  
Pokud potřebujete pomoci s instalací serveru Domino, vyhledejte potřebné informace v dokumentaci k produktu Domino.

- e. Ověřte, zda ID administrátora podnikového vyhledávání má oprávnění pro přístup k adresáři `home/server`. Toto ID administrátora se určuje během instalace produktu OmniFind Enterprise Edition.
3. Na prolézacím serveru spusíte instalační skript dodaný v rámci produktu OmniFind Enterprise Edition:
    - a. Přihlašte se jako administrátor podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition).
    - b. Spusíte následující skript instalovaný v adresáři `$ES_INSTALL_ROOT/bin`:  
`escrnote.sh`
    - c. Odpovězte na výzvy:
      - Na následující výzvu odpovězte **A**, je-li server Domino nainstalován do výchozího adresáře, nebo **N**, je-li nainstalován do jiného adresáře:  
Byla nalezena cesta k adresáři Lotus Notes  
`/opt/lotus/notes/latest/linux`.  
Jedná se o správnou cestu k adresáři Lotus Notes?  
  
Výchozí cesta pro provozní prostředí AIX je `/opt/lotus/notes/latest/ibmpow`.  
Výchozí cesta pro systém Linux je `/opt/lotus/notes/latest/linux`.  
Výchozí cesta pro provozní prostředí Solaris je `/opt/lotus/notes/latest/sunspa`.
      - Není-li server Domino nainstalován do výchozího adresáře na prolézacím serveru, odpovězte na následující výzvu zadáním instalačního adresáře serveru Domino:  
Zadejte cestu k adresáři Lotus Notes.  
  
V počítači se systémem Linux můžete například zadat hodnotu  
`/opt/lotus/notes/latest/linux`.
      - Na následující výzvu odpovězte **A**, je-li datový adresář serveru Domino nainstalován do výchozího adresáře, nebo **N**, je-li nainstalován do jiného adresáře:  
Byla nalezena cesta k datovému adresáři Lotus Notes  
`/local/notesdata`.  
Jde o správnou cestu k datovému adresáři Lotus Notes?  
  
Výchozí cesta je `/local/notesdata`.
      - Není-li datový adresář serveru Domino umístěn ve výchozím umístění na prolézacím serveru, odpovězte na následující výzvu zadáním datové cesty produktu Domino:  
Zadejte cestu k datovému adresáři Lotus Notes.
  4. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

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

#### Související pojmy

“Prolézací moduly Domino Document Manager” na stránce 57

“Prolézací moduly Notes” na stránce 68

“Prolézací moduly QuickPlace” na stránce 78

## Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino

Pokud instalujete produkt OmniFind Enterprise Edition do počítače se systémem Microsoft Windows a máte-li v úmyslu prolézat servery využívající protokol NRPC (Notes Remote

Procedure Call), musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožní produktům Notes, QuickPlace a Domino Document Manager komunikaci s databázovými servery.

### Omezení

Server Lotus Domino Server a klient Lotus Notes nemohou být spuštěny současně ve stejném počítači s prolézacím modulem Notes, QuickPlace nebo Domino Document Manager, který je konfigurován pro použití protokolu NRPC. Pokud se spustí některý z uvedených prolézacích modulů v době, kdy je spuštěn server Domino Server, dojde k chybě a prolézací modul bude zastaven.

### O této úloze

Prolézací moduly používající protokol NRPC používají knihovny klienta Lotus Domino. Tyto knihovny nainstalujete instalací serveru Lotus Domino na prolézací server. Chcete-li se ujistit, že prolézací moduly jsou schopny pracovat s knihovnami Domino, spusťte po instalaci knihoven Domino instalační skript, který je k dispozici na prolézacím serveru produktu OmniFind Enterprise Edition.

### Postup

Chcete-li konfigurovat prolézací server tak, aby mohl prolézat servery Lotus Notes, Lotus QuickPlace a Domino Document Manager, postupujte takto:

1. Přihlaste se k prolézacímu serveru s použitím jména uživatele, které je členem skupiny administrátorů. Ujistěte se, že použité jméno uživatele má oprávnění k instalaci produktu Lotus Notes.
2. Instalujte produkt Lotus Notes:
  - a. Vložte disk CD produktu Domino Server. (Pokud nemáte disk CD k dispozici, můžete si stáhnout instalační obraz).
  - b. Spusťte instalační program: **setup.exe**
  - c. Odpovězte na výzvy a přijměte výchozí hodnoty nebo určete vlastní upřednostňované nastavení instalace (například cesty k instalačnímu a datovému adresáři).  
Pokud potřebujete pomoc, vyhledejte potřebné informace v dokumentaci k produktu Lotus Domino.
  - d. Ověřte, zda ID administrátora podnikového vyhledávání má oprávnění pro přístup k datovému adresáři serveru Domino. Toto ID administrátora se určuje během instalace produktu OmniFind Enterprise Edition.
3. Na prolézacím serveru spusťte instalační skript, který byl dodán v rámci produktu OmniFind Enterprise Edition:
  - a. Přihlaste se pod jménem administrátora podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition).
  - b. Spusťte následující skript instalovaný v adresáři `%ES_INSTALL_ROOT%\bin:`  
`escrnote.vbs`
  - c. Odpovězte na výzvy:
    - Na následující výzvu odpovězte **A**, je-li produkt Lotus Notes nainstalován do výchozího adresáře, nebo **N**, je-li nainstalován do jiného adresáře:  
Byla nalezena cesta k adresáři Lotus Notes `c:\lotus\notes`.  
Jedná se o správnou cestu k adresáři Lotus Notes?

Typická instalační cesta v počítači se systémem Windows je `c:\lotus\notes` nebo `c:\lotus\domino`.

- Není-li produkt Lotus Notes nainstalován do výchozího adresáře na prolézacím serveru, odpovězte na následující výzvu zadáním instalačního adresáře produktu Lotus Notes:

Zadejte cestu k adresáři Lotus Notes.

- Na následující výzvu odpovězte **A**, je-li datový adresář produktu Lotus Notes implementován ve výchozím umístění, nebo **N**, je-li implementován v jiném umístění:

Byla nalezena cesta k datovému adresáři

Lotus Notes c:\lotus\notes\data.

Jde o správnou cestu k datovému adresáři Lotus Notes?

Typická cesta v počítači se systémem Windows je c:\lotus\notes\data nebo c:\lotus\domino\data.

- Není-li datový adresář produktu Lotus Notes implementován ve výchozím umístění na prolézacím serveru, odpovězte na následující výzvu zadáním cesty k datovému adresáři:

Zadejte cestu k

datovému adresáři Lotus Notes.

4. Na prolézacím serveru zastavte a restartujte systém podnikového vyhledávání:

```
esadmin system stopall
```

```
esadmin system startall
```

#### Související pojmy

“Prolézací moduly Domino Document Manager” na stránce 57

“Prolézací moduly Notes” na stránce 68

“Prolézací moduly QuickPlace” na stránce 78

## Konfigurování serverů používajících protokol DIIOP

Chcete-li prolézat servery využívající protokol DIIOP (Domino Internet Inter-ORB Protocol), musíte server konfigurovat tak, aby prolézací moduly Notes, QuickPlace a Domino Document Manager mohly protokol používat.

### Než začnete

Na serveru, který chcete prolézat, musí být spuštěny úlohy DIIOP a HTTP.

### Postup

Konfigurování serverů využívajících protokol DIIOP:

1. Konfigurujte dokument server:
  - a. Otevřete dokument **server** na serveru Lotus Notes, Lotus QuickPlace nebo Domino Document Manager, který chcete prolézat. Tento dokument je uložen v adresáři produktu Domino.
  - b. Na stránce Konfigurace rozbalte sekci **server**.
  - c. Na stránce Zabezpečení v oblasti **Programová omezení** zadejte bezpečnostní omezení platná ve vašem prostředí do následujících polí:
    - **Spouštět omezené agenty Lotus Script/Java**
    - **Spouštět omezené procesy Java/Javascript/COM**
    - **Spouštět neomezené procesy Java/Javascript/COM**

Můžete například zadat hvězdičku (\*) a povolit tak neomezený přístup agentům Lotus Script/Java a do pole omezení procesů Java/Javascript/COM zadat jména uživatelů registrovaných v adresáři Domino.

**Důležité:** Prolézací modul, který konfiguruje pro prolézání tohoto serveru s protokolem DIIOP, musí mít možnost používat jména uživatelů, která zadáte v těchto polích.

- d. Otevřete stránku Internetový protokol, poté otevřete stránku HTTP a nastavte volbu **Povolit klientům HTTP procházení databáze** na hodnotu **Ano**.
2. Konfigurujete dokument user:
    - a. Otevřete dokument **user** na serveru Lotus Notes, Lotus QuickPlace nebo Domino Document Manager, který chcete prolézat. Tento dokument je uložen v adresáři produktu Domino.
    - b. Na stránce Základy zadejte heslo do pole **Internetové heslo**.  
Při konfigurování voleb pro prolézání tohoto serveru pomocí konzoly pro správu podnikového vyhledávání zadejte toto jméno uživatele a heslo na stránce s určením serveru, který má být prolézán. Klient používá tyto údaje pověření pro přístup k serveru.
  3. Restartujte úlohu DIIOP na serveru.

#### Související pojmy

“Prolézací moduly Domino Document Manager” na stránce 57

“Prolézací moduly Notes” na stránce 68

“Prolézací moduly QuickPlace” na stránce 78

## Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino

Předtím, než budete moci používat prolézací moduly Notes, QuickPlace nebo Domino Document Manager v systému IBM AIX, musíte instalovat modul IOCP (I/O Completion Port) a konfigurovat jej pro používání prolézacím modulem.

### O této úloze

Bez modulu IOCP dojde selhání procesů zjišťování při pokusu o vytvoření prolézacího modulu. Zobrazí se následující chybová zpráva:

```
FFQM0105E Byla přijata chyba serveru -  
Zpráva: FFQG0024E An unexpected exception was caught: discover
```

Do souboru `$ES_NODE_ROOT/logs/system_yyyymmdd.log` bude zapsána následující zpráva obsahující chybu ENOEXEC. (Část textu zprávy je rozdělena na více řádků pro lepší čitelnost.)

```
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)

```

## Postup

Chcete-li instalovat modul IOCP a ověřit, že je na prolézacím serveru instalován správně, postupujte takto:

Je třeba provést následující akce:

1. Instalujte modul IOCP (**bos.iocp.rte**) z disku CD produktu AIX na prolézací server.  
Po instalaci modulu IOCP a před vytvořením prolézacího modulu Notes, QuickPlace nebo Domino Document Manager instalujte pro modul softwarovou opravu. Pokyny naleznete na následující adrese:

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

2. Zadáním následujícího příkazu se ujistěte, že je na prolézacím serveru nainstalován modul IOCP:

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

Výstup příkazu `lspp` by měl být podobný následujícímu příkladu:

Fileset	Level	State	Description
-----			
Path: /usr/lib/objrepos			
bos.iocp.rte	5.2.0.10	COMMITTED	I/O Completion Ports API
Path: /etc/objrepos			
bos.iocp.rte	5.2.0.10	COMMITTED	I/O Completion Ports API

3. Zadáním následujícího příkazu se ujistěte, že je port IOCP ve stavu **Available** (dostupný):

```
$ lsdev -Cc iocp
```

Výstup příkazu `lsdev` by měl odpovídat následujícímu příkladu:

```
iocp0 Available I/O Completion Ports
```

4. Pokud je port IOCP ve stavu **Defined** (definovaný), změňte jej na stav **Available** (dostupný):
  - a. Přihlaste se k prolézacímu serveru jako uživatel `root` a zadejte následující příkaz:

```
# smit iocp
```
  - b. Vyberte příkaz **Change / Show Characteristics of I/O Completion Ports** (Změna/zobrazení vlastností portů I/O Completion) a změňte parametr **STAV při spuštění systému** z hodnoty **Defined** (definovaný) na hodnotu **Available** (dostupný).
  - c. Znovu zaveďte systém prolézacího serveru.
  - d. Znovu zadejte příkaz `lsdev` a ujistěte se, že se stav portu IOCP změnil na stav **Available** (dostupný).

### Související pojmy

“Prolézací moduly Domino Document Manager” na stránce 57

“Prolézací moduly Notes” na stránce 68

“Prolézací moduly QuickPlace” na stránce 78

---

## Prolézací moduly QuickPlace

Chcete-li zahrnout místa a místnosti Lotus QuickPlace do kolekce podnikového vyhledávání, musíte konfigurovat prolézací modul QuickPlace.

Dále můžete prolézací modul QuickPlace použít k prolézání míst, která jsou spravována pomocí produktu Lotus Quickr services for Lotus Domino. Používáte-li produkt Lotus Quickr services for WebSphere Portal, můžete prolézací modul Počáteční seznam použít k přidání dokumentů uložených v knihovnách produktu Lotus Quickr do kolekce podnikového vyhledávání.

### Tip:

Podrobné příklady konfigurace zabezpečeného prolézacího modulu QuickPlace najdete ve scénáři pro malé organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

## Konfigurace prolézacího serveru

Pokud server, který máte v úmyslu prolézat, používá protokol NRPC (Notes Remote Procedure Call), je nutné spustit na prolézacím serveru instalační skript. Tento skript se dodává s produktem OmniFind Enterprise Edition a umožňuje prolézacímu modulu QuickPlace komunikovat se servery využívajícími protokol NRPC.

Pokud server, který máte v úmyslu prolézat, používá protokol DIIOP (Domino Internet Inter-ORB Protocol), nemusíte na prolézacím serveru spouštět žádný instalační skript. Je však nezbytné nastavit cílový server tak, aby k němu měl prolézací modul QuickPlace přístup.

Pokud server, který chcete prolézat, používá server LDAP (Lightweight Directory Access Protocol), je třeba konfigurovat cílový server tak, aby používal protokol DIIOP (prolézací modul QuickPlace nemůže k prolézání dat LDAP používat protokol NRPC). Musíte rovněž konfigurovat databázi služby Directory Assistance a konfigurovat cílový server, aby používal server LDAP jako sekundární server Domino.

**Důležité:** V případě, že cílový server používá protokol DIIOP a prolézací modul je konfigurován tak, aby používal protokol HTTPS nebo DIIOP přes SSL z důvodu šifrování přenosu mezi prolézacím modulem a serverem, musíte zkopírovat soubor `TrustedCerts.class` (například `c:\certs` nebo `/data/certs`) z cílového serveru na prolézací server. V případě konfigurace se dvěma nebo čtyřmi servery musíte také zkopírovat soubor `TrustedCerts.class` na servery, na nichž je instalována komponenta pro vyhledávání. Je třeba zajistit, aby se tento soubor nacházel na prolézacím serveru a na vyhledávacích serverech ve stejném umístění. Cesta k adresáři se souborem `TrustedCerts.class` je zadána při konfiguraci prolézacího modulu.

Při konfiguraci prolézacího modulu a zadávání ID uživatele, které bude prolézací modul používat, je třeba zadat ID dostatečná oprávnění pro přístup ke všem místům QuickPlace v doméně Domino. Chcete-li to zajistit, přiřaďte ID uživatele do vyhrazené skupiny s názvem `QuickPlaceAdministratorsSUGroup`.

Je-li produkt OmniFind Enterprise Edition instalován v systému IBM AIX, musíte zaručit, že je na prolézacím serveru instalován modul I/O Completion Port a že je tento modul k dispozici.

Před konfigurováním prolézacího modulu QuickPlace prostřednictvím konzoly pro správu podnikového vyhledávání, proveďte úlohy, které odpovídají vašemu prostředí:



- “Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino” na stránce 72.
- “Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 73.
- “Konfigurování serverů používajících protokol DIIOP” na stránce 75.
- “Konfigurování serveru QuickPlace pro používání zabezpečení lokálních uživatelů” na stránce 269.
- “Konfigurování služby Directory Assistance na serveru QuickPlace” na stránce 270.
- “Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 76.

## Zabezpečení na úrovni dokumentu

Pokud je povoleno zabezpečení kolekce a server, který máte v úmyslu prolézat, používá protokol NRPC, je nutné konfigurovat na prolézacím serveru produkt Lotus Domino Trusted Server. Prostřednictvím produktu Trusted Server lze prosazovat řízení přístupu na úrovni dokumentu. Před zpřístupněním kolekce pro prohledávání uživateli proveďte následující úlohy:

- Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů
- Povolení globálního zabezpečení na serveru WebSphere Application Server a konfigurování vyhledávací aplikace pro použití zabezpečení. Tento krok zajistí ověřování přihlašovacích údajů při pokusu uživatele o použití vyhledávací aplikace. Vyhledávací servery používají tyto pověřovací údaje k ověření oprávnění každého uživatele pro přístup k dokumentům Lotus Domino.

## Prolézání příloh

V produktu Lotus QuickPlace můžete importovat a publikovat dokumenty sady Microsoft Office (volby vybrané při importu zahrnují typy Importovaná stránka, Stránka aplikace Microsoft Word, Stránka aplikace Microsoft Excel, Stránka aplikace Microsoft PowerPoint a Více importovaných stránek). Prolézací modul QuickPlace může prolézat tyto typy importovaných dokumentů jako přílohy pouze za následujících podmínek:

- Server používá protokol DIIOP.
- Při konfigurování voleb prolézacího modulu jste povolili prolézání příloh.
- Prolézací modul jste nakonfigurovali pro prolézání pole "\$FILE" nebo všech polí.

## Přehled konfigurace

Prolézací modul QuickPlace lze použít k prolézání libovolného počtu míst QuickPlace. Při vytváření prolézacího modulu vyberete místa na jednom serveru QuickPlace, která chcete prolézat. Při pozdějších úpravách prolézaného prostoru můžete přidat dokumenty z dalšího serveru QuickPlace, které chcete zahrnout do stejného prolézaného prostoru. Při vytváření nebo úpravě prolézacího modulu můžete určit, zda chcete prolézat všechna místa na serveru nebo pouze zadaná místa a zda chcete prolézat všechny místnosti v místech, která mají být prolézána, nebo pouze vybrané místnosti.

**Omezení:** Při zadávání ID uživatele, které bude prolézací modul používat, se ujistěte, že má zadané ID dostatečná oprávnění pro přístup ke všem místům QuickPlace v doméně Domino. Splnění tohoto požadavku můžete zajistit tím, že nakonfigurujete server QuickPlace a přiřadíte ID uživatele rezervované skupině s názvem QuickPlaceAdministratorsSUGroup. Pověšimněte si, že název skupiny neobsahuje mezery a rozlišují se v něm velká a malá písmena.

Chcete-li vytvořit nebo změnit prolézací modul QuickPlace, přihlaste se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny dokumenty v prolézáném prostoru.
- Identifikovat server QuickPlace a komunikační protokol.
- Pokud je server konfigurován pro použití protokolu DIIOP, můžete určit, jakým způsobem se má prolézací modul připojovat k objektům Domino. Můžete například zadat volby pro šifrování komunikace pomocí protokolu HTTPS nebo Secure Sockets Layer (SSL).
- Zadat informace o uživatelském adresáři asociovaném se serverem (prolézací modul tyto informace potřebuje, aby mohlo být vynucováno řízení přístupu při prohledávání kolekce uživateli).
- Vybrat místa, která chcete prolézat.
- Nastavit časový plán pro prolézání míst.
- Vybrat dokumenty, které chcete prolézat. Prolézací modul může prolézat všechny místnosti v místě nebo může prolézat pouze dokumenty ve vybraných místnostech.
- Zadat volby umožňující prohledávání polí v různých místech a místnostech. Můžete například vyloučit z prolézáného prostoru některá pole a zadat volby pro prohledávání příloh.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznámech přístupových práv nebo prvcích zabezpečení. Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

#### **Související pojmy**

“Rady pro prolézání databází Lotus Domino” na stránce 71

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino”  
na stránce 267

#### **Související úlohy**

“Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino”  
na stránce 72

“Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 73

“Konfigurování serverů používajících protokol DIIOP” na stránce 75

“Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 76

“Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů” na stránce 268

“Konfigurování serveru QuickPlace pro používání zabezpečení lokálních uživatelů”  
na stránce 269

“Konfigurování služby Directory Assistance na serveru QuickPlace” na stránce 270

---

## Prolézací moduly počátečního seznamu

Používáte-li produkt IBM Lotus Quickr services for WebSphere Portal, můžete prolézací modul Počáteční seznam použít k přidání dokumentů v knihovnách produktu Lotus Quickr do kolekce podnikového vyhledávání.

Knihovna Lotus Quickr je kontejnerem pro soubory dokumentů. Prolézací modul Počáteční seznam nepodporuje prolézání webového obsahu, jako jsou například wikiweby nebo blogy.

Používáte-li produkt Lotus Quickr services for Lotus Domino, můžete použít prolézací modul QuickPlace k přidání dokumentů produktu Lotus Quickr do kolekce.

### Konfigurace serveru WebSphere Portal

Při instalaci produktu Lotus Quickr na server WebSphere Portal verze 6 můžete použít prolézací modul Počáteční seznam k prolézání dokumentů knihoven produktu Lotus Quickr. Volby pro prolézání těchto dokumentů můžete nastavit odděleně od voleb, které zadáte pro portálové weby prolézané prolézacím modulem WebSphere Portal.

Před vytvořením prolézacího modulu Počáteční seznam musíte provést postup, který nastaví podnikové vyhledávání v portálu WebSphere. K nastavení prostředí podnikového vyhledávání spusíte skript (`wp6_install.sh` v systémech AIX, Linux a Solaris a `wp6_install.bat` v systémech Windows), který se instaluje na vyhledávací servery společně s produktem OmniFind Enterprise Edition.

Řetězec uživatelského agenta identifikuje, který prohlížeč nebo robot přistupuje k serveru. Při prolézání serveru Lotus Quickr prolézací modul Počáteční seznam používá řetězec uživatelského agenta OmniFind SeedlistCrawler/1.0.

### Přehled konfigurace

Prolézací modul Počáteční seznam lze použít k prolézání libovolného počtu dokumentů produktu Lotus Quickr. Při konfiguraci prolézacího modulu zadejte server, který má být prolézán. Prolézací modul poté prolézá všechny dokumenty v knihovnách produktu Lotus Quickr na daném serveru.

Prolézané dokumenty musí být dostupné pod stejným ID a heslem administrátora produktu Lotus Quickr. Chcete-li prolézat weby s použitím jiných pověřovacích údajů, musíte k tomuto účelu konfigurovat samostatný prolézací modul Počáteční seznam.

Chcete-li vytvořit nebo změnit prolézací modul Počáteční seznam, přihlaste se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu stanovují způsob prolézání dokumentů Lotus Quickr.
- Určete server, který má být prolézán, a informace, které umožňují připojení prolézacího modulu k serveru.

Pokud vytvoříte nebo upravíte prolézací modul, můžete otestovat, zda se tento modul dokáže úspěšně připojit k dokumentům, které mají být prolézány. Před spuštěním prolézacího modulu tak můžete zobrazit zprávy informující o tom, zda má prolézací modul přístup k dokumentům, které chcete prolézat.

- Nastavit zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit kontrolu přístupu založenou na uložených seznámech přístupových práv nebo prvcích zabezpečení.

Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat.

**Důležité:** Chcete-li prohledávat zabezpečené dokumenty produktu Lotus Quickr, musíte vyhledávací požadavky zadávat pomocí vyhledávacího portletu pro podnikové vyhledávání v rámci portálu WebSphere. Vyhledávací požadavky odeslané z ukázkové vyhledávací aplikace ESSearchApplication nemají k dispozici potřebné pověřovací údaje a nemohou tedy ověřit oprávnění uživatele pro přístup k dokumentům.

- Zadat informace umožňující prolézacímu modulu komunikaci se serverem proxy, pokud se při poskytování stránek využívá server proxy.
- Pokud k ochraně serveru WebSphere Portal a dokumentů produktu Lotus Quickr používáte jiný produkt (například IBM Tivoli Access Manager WebSEAL nebo CA SiteMinder SSO Agent for PeopleSoft), zadejte údaje pověření pro jednotné přihlášení, které prolézacímu modulu umožní přístup k dokumentům na serveru.
- Zadat informace o souboru úložiště klíčů, aby prolézací modul mohl pro připojení k serveru použít protokol SSL (Secure Sockets Layer).
- Zadat jazyk a kódovou stránku prolézáných dokumentů.
- Nastavit volby pro prolézání a vyhledávání metadat v dokumentech Lotus Quickr.
- Nastavit plán prolézání dokumentů Lotus Quickr.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

#### **Související pojmy**

“Integrace s produktem WebSphere Portal” na stránce 323

#### **Související úlohy**

“Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 6” na stránce 330

“Nastavení portletu podnikového vyhledávání pro produkt Lotus Quickr” na stránce 335

---

## **Prolézací moduly souborového systému UNIX**

Chcete-li do kolekce podnikového vyhledávání zahrnout dokumenty uložené v souborových systémech AIX, Linux a Solaris, musíte konfigurovat prolézací modul Souborový systém UNIX.

Prolézací modul Souborový systém UNIX lze použít k prolézání libovolného počtu souborových systémů. Při konfigurování prolézacího modulu vyberete lokální a vzdálené adresáře a podadresáře, které chcete prolézat.

Pokud instalujete prolézací server do počítače se systémem Windows, nemůžete tento server použít k prolézání zdrojů v souborových systémech AIX, Linux a Solaris (v seznamu dostupných typů prolézacích modulů se neobjeví prolézací modul Souborový systém UNIX).

Prolézací modul Souborový systém UNIX při prolézání dokumentů využívá oprávnění ke čtení nastavené pro administrátora podnikového vyhledávání.

Chcete-li vytvořit nebo změnit prolézací modul Souborový systém UNIX, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny podadresáře v prolézaném prostoru.
- Nastavit časový plán pro prolézání souborových systémů.
- Vybrat podadresáře a úrovně podadresářů, které má daný prolézací modul zpracovávat.
- Zadat volby umožňující prohledávání dokumentů v podadresářích. Můžete například vyloučit z prolézaného prostoru určité typy dokumentů.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

---

## Webové prolézací moduly

Chcete-li do kolekce podnikového vyhledávání zahrnout stránky z webových serverů, musíte konfigurovat webový prolézací modul.

Webový prolézací modul lze použít k prolézání libovolného počtu serverů HTTP (Hypertext Transfer Protocol) a HTTPS (Secure HTTP). Prolézací modul navštíví webový server a načte z něj data. Poté použije odkazy v dokumentech k prolézání dalších dokumentů. Webový prolézací modul je schopen prolézat a extrahovat odkazy z jednotlivých stránek i ze *sad rámců* (stránky vytvořené s použitím rámců HTML).

Prolézaná data mohou být v kterémkoli běžném formátu a pocházet z různých zdrojů v intranetu nebo Internetu. K běžným formátům patří formáty HTML, PDF, Microsoft Word, Lotus WordPro, XML (Extensible Markup Language) atd.

### Tip:

Podrobné příklady konfigurace webového prolézacího modulu najdete ve scénáři pro střední organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Chcete-li vytvořit nebo změnit webový prolézací modul, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít rovněž přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce, která prolézací modul vlastní.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny webové stránky v prolézaném prostoru.
- Zadat pravidla povolující a zakazující návštěvu webových serverů. Určíte-li pravidla prolézání, můžete pravidla otestovat a ověřit, že prolézací model může přistupovat k webům, které chcete zahrnout do prolézacího prostoru.

- Určit volby pro zahrnutí určitých typů souborů a vyloučení souborů s určitými příponami.
- Zadat pravidla pro způsob, jakým má webový prolézací modul pracovat se stránkami méně závažných chyb.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.
- Zadat volby pro prolézání webových serverů chráněných heslem (prolézané webové servery musí pro zobrazování výzev k zadání hesla používat základní ověřování HTTP nebo formuláře HTML).
- Zadat volby pro prolézání webových serverů obsluhovaných serverem proxy.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

## Konfigurace uživatelského agenta

Chcete-li prolézat webový server využívající protokol Robots Exclusion, zkontrolujte, zda je v souboru `robots.txt` na webovém serveru povolen název uživatelského agenta, kterého jste konfigurovali pro přístup webového prolézacího modulu k webovému serveru.

Při spuštění systému podnikového vyhledávání webový prolézací modul načte název uživatelského agenta, který jste pro něj konfigurovali. Před stažením stránky z dosud (nebo delší dobu) nenavštíveného webového serveru se prolézací modul nejprve pokusí stáhnout soubor s názvem `robots.txt`. Tento soubor je uložen v kořenovém adresáři webového serveru.

Pokud soubor `robots.txt` neexistuje, je webový server otevřen pro neomezené prolézání. Pokud tento soubor existuje, obsahuje informace o tom, ke kterým oblastem serveru (adresářům) nemají prolézací moduly přístup. Prolézací moduly, kterým se udělují oprávnění pro prolézání, jsou v souboru `robots.txt` identifikovány pomocí názvu uživatelského agenta.

Použití protokolu Robots Exclusion není povinné, webový prolézací modul podnikového vyhledávání se jej však pokouší podporovat:

- Pokud soubor `robots.txt` obsahuje položku s názvem uživatelského agenta konfigurovaného pro daný webový prolézací modul, tento modul se podřídí omezením stanoveným pro příslušného uživatelského agenta.
- Není-li v souboru `robots.txt` hledaný název uživatelského agenta uveden, ale poslední položka obsahuje text `User-agent: *` (tj. libovolný uživatelský agent) a omezení je nastaveno na hodnotu `Disallow: /` (tj. nepovolit žádné prolézání počínaje kořenem webového serveru), webový prolézací modul nebude takový server prolézat.
- Není-li v souboru `robots.txt` uveden hledaný název uživatelského agenta, ale poslední položka obsahuje text `User-agent: *` s omezením `Allow: /`, smí webový prolézací modul takový server prolézat.

Administrátoři webových serverů často používají poslední položku zakazující přístup všem prolézacím modulům, kterým není udělen explicitně. Pokud konfiguruje nový webový prolézací modul a více, že některé z webových serverů, které budete prolézat, používají protokol Robots Exclusion, požádejte administrátory těchto webových serverů o přidání položky pro váš prolézací modul do souboru `robots.txt`.

Dbejte na to, aby byl ve vlastnostech webového prolézacího modulu a ve všech souborech `robots.txt`, které patří ke sledovaným webovým serverům, uveden stejný název uživatelského agenta.

V případě, že protokol Robots Exclusion nepoužívá žádný z webových serverů, které chcete prolézat, obvykle nezáleží na tom, jakou hodnotu vlastnosti uživatelského agenta zadáte. Některé aplikační servery, moduly JSP a servlety však podle názvu uživatelského agenta upravují své odpovědi. Tyto různé varianty odpovědi se používají například k překonání nekompatibility prohlížečů. V těchto situacích může mít zadaný název uživatelského agenta určitý význam bez ohledu na to, zda je použit protokol Robots Exclusion. Potřebujete-li prolézat tyto typy serverů, obraťte se na jejich administrátory a ujistěte se, že je na ně povolen přístup webovému prolézacímu modulu.

## Způsob využití protokolu Robots Exclusion ve webovém prolézacím modulu

Webový prolézací modul se snaží dodržovat protokol Robots Exclusion a neprolézat webové servery, pokud to zakazují pravidla obsažená v souboru `robots.txt` umístěném na daném serveru.

K úspěšnému stažení dojde v případě, že se webovému prolézacímu modulu podaří načíst z webového serveru soubor `robots.txt` nebo s určitostí zjistit, že soubor `robots.txt` neexistuje. Pokud se prolézacímu modulu nepodaří pravidla načíst nebo zjistit, zda soubor `robots.txt` existuje, je stažení považováno za neúspěšné.

Úspěšné stažení neznamena, že je prolézací modul oprávněn server prolézat, protože prolézání může být zakázáno pravidly v souboru `robots.txt`. Po neúspěšném pokusu o stažení bude prolézání dočasně zakázáno, protože prolézací modul nemůže zjistit, zda jsou nějaká pravidla definována a jaká.

Při pokusu o stažení souboru `robots.txt` postupuje prolézací modul takto:

1. Když prolézací modul nalezne nový web, pokusí se získat adresu IP příslušného serveru. Jestliže tento pokus skončí neúspěchem, prolézání není možné.
2. Pokud je k dispozici alespoň jedna adresa IP, prolézací modul se pokusí stáhnout soubor `robots.txt` pomocí příkazu **HTTP (nebo HTTPS) GET**.
3. Pokud dojde k vypršení časového limitu připojení socketu, k přerušení tohoto připojení nebo k jiné chybě na nízké úrovni (např. problémy s certifikátem SSL), prolézací modul zaznamená tento problém do protokolu a opakuje pokus pro všechny známé adresy cílového serveru.
4. Pokud není připojení navázáno ani po vyčerpání všech známých adres, prolézací modul dvě sekundy vyčká a poté znovu opakuje pokus postupně se všemi adresami.
5. Pokud se podaří navázat připojení a provést výměnu záhlaví HTTP, proběhne kontrola návratového stavu. Pokud je nalezen stavový kód 500 nebo vyšší, prolézací modul interpretuje připojení jako nefunkční a pokračuje v navazování připojení přes další adresy IP. Pokud je zjištěn jakýkoli jiný stav, prolézací modul ukončí navazování připojení přes alternativní adresy IP a pokračuje v souladu s hodnotou stavového kódu.

Jakmile prolézací modul přijme stavový kód HTTP nižší než 500 nebo projde dvakrát po sobě všechny adresy IP, pokračuje takto:

1. Pokud nebyl přijat žádný stav HTTP s hodnotou nižší než 500, je daný web dočasně označen jako nedostupný.
2. Pokud byl přijat stav HTTP s hodnotou 400, 404 nebo 410, je web označen jako vhodný pro prolézání bez jakýchkoli pravidel.
3. Pokud byl přijat stav HTTP v intervalu od 200 do 299, řídí se další činnost následujícími podmínkami:
  - Pokud byl obsah zkrácen, je daný web dočasně označen jako nedostupný.
  - Pokud analýza proběhla bez chyby, je server označen jako vhodný pro prolézání při dodržení nalezených pravidel.

- Pokud analýza proběhla s chybami, je server označen jako vhodný pro prolézání bez jakýchkoli pravidel.
4. Pokud byl přijat jakýkoli jiný stav HTTP, je daný web dočasně označen jako nedostupný.

Když se prolézací modul pokusí stáhnout soubor `robots.txt` pro daný web, aktualizuje trvalou časovou značku nastavenou pro tento web a označovanou jako datum robota. V případě, že je web označen jako nedostupný, protože nejsou k dispozici informace o souboru `robots.txt`, je zvýšena trvale udržovaná hodnota udávající počet selhání robotů.

Po uplynutí intervalu opakování se prolézací modul znovu pokusí načíst informace o souboru `robots.txt` z webu, u kterého předtím došlo k selhání. V případě, že celkový počet po sobě jdoucích selhání dosáhne maximálního povoleného počtu selhání, prolézací modul se přestane pokoušet o načtení souboru `robots.txt` z tohoto webu a označí web jako nedostupný pro prolézání.

V případě, že je web dostupný pro prolézání (kontrola pravidel v souboru `robots.txt` úspěje), je počet selhání vynulován. Prolézací modul bude načtená pravidla používat, dokud neuplyne interval kontroly pravidel. Po uplynutí tohoto intervalu je nutné znovu zkontrolovat dostupnost webu pro prolézání.

#### **Tip:**

- Pokud server vrátí obsah se syntaktickými chybami, pokud používá protokol pro práci s roboty v jiné verzi než 1994 nebo pokud je obsah tvořen jinými daty než pravidly pro roboty (například stránkou méně závažné chyby), prolézací modul považuje daný web za web bez souboru pravidel a označí jej jako dostupný pro prolézání. Tento postup je obvykle v pořádku, protože administrátoři kolekci nekontrolují obsah serveru nebo jeho výchozí chování. Pokud chce administrátor webového serveru zabránit prolézání webu a nechce přitom instalovat standardní soubor pravidel, může administrátor kolekce zablokovat přístup webových prolézacích modulů na daný web uvedením jeho domény, adresy IP nebo předpony HTTP v pravidlech prolézacího modulu.
- Vrátí-li server stavový kód 302 nebo jiný typ kódu přesměrování, prolézací modul tento kód interpretuje jako informaci o tom, že na webu je umístěn soubor `robots.txt`, který je třeba použít, nenachází se však na standardním místě (v kořenovém adresáři webu). Administrátor webového serveru musí soubor přesunout na správné místo, aby se mohl webový prolézací modul řídit pravidly, která jsou v tomto souboru obsažena.
- Nastanou-li problémy s certifikátem (např. prošlý certifikát, nedůvěryhodný původce certifikátu nebo použití certifikátu s vlastním podpisem v případě, že konfigurace prolézacího modulu nedovoluje tyto certifikáty přijímat), prolézací modul interpretuje situaci jako selhání pokusu o připojení a označí web jako nedostupný pro prolézání. V prolézání dalších stránek stejného webu by modulu pravděpodobně zabránily stejné problémy. Pokud chce administrátor kolekce umožnit prolézání daného webu, musí povolit použití certifikátů s vlastním podpisem, přidat původce certifikátu do souboru důvěryhodného úložiště klíčů nebo požádat administrátora webového serveru, aby si obstaral aktuální certifikát.
- Webový prolézací modul lze konfigurovat pro použití základního ověřování HTTP (včetně základního ověřování HTTP serverů proxy). Pokud je konfigurace správně nastavena, je ověřování vyžadováno i při stahování souborů `robots.txt`. Stavový kód 403, 407 nebo jiná reakce vyžadující ověření naznačuje problémy s autorizací a prolézací modul v takovém případě označí web jako nedostupný pro prolézání. (Podporováno je pouze základní ověřování HTTP.)
- Pokud délka souboru `robots.txt` na webu překračuje povolené maximum pro stránku robots, administrátor kolekce může toto maximum zvýšit (výchozí hodnota jednoho miliónu bajtů by však měla postačovat).



Jako pomůcku pro řešení problémů si můžete při monitorování webového prolézacího modulu vyžádat zprávu o webu. Vyberte volby pro zobrazení obsahu souboru robots.txt (abyste zjistili, zda webovému prolézacímu modulu brání v přístupu k webu pravidla definovaná v tomto souboru), data a času, kdy se prolézací modul naposledy pokusil o stažení souboru robots.txt (prolézací modul se nepokusí soubor znovu stáhnout, dokud neuplyne interval opakování), a počtu po sobě jdoucích neúspěšných pokusů prolézacího modulu o stažení souboru robots.txt. Klepnutím na tlačítko **Nápověda** můžete při monitorování webového prolézacího modulu zobrazit další informace o těchto volbách zprávy o webu a o tom, jak interpretovat výslednou zprávu.

Další informace o protokolu Robots Exclusion naleznete na následující adrese:  
<http://www.robotstxt.org/wc/exclusion.html>

## Podpora skriptů JavaScript

Webový prolézací modul pro podnikové vyhledávání může najít odkazy (adresy URL) obsažené ve skriptech JavaScript, které jsou součástí webových dokumentů.

Webový prolézací modul se může setkat s relativními i absolutními odkazy. Pokud dokument HTML obsahuje prvek BASE, prolézací modul tento prvek použije k doplnění relativních odkazů na absolutní. V opačném případě prolézací modul k tomuto účelu použije vlastní adresu URL dokumentu.

Podpora skriptů JavaScript se omezuje pouze na extrakci odkazů. Prolézací modul neprovádí analýzu skriptů JavaScript, nesestavuje model DOM (Document Object Model) a neinterpretuje ani neprovádí příkazy skriptů JavaScript. Prolézací modul vyhledává v obsahu dokumentu (mimo jiné v částech tvořených skripty JavaScript) řetězce, u nichž je pravděpodobné, že se jedná o adresy URL v příkazech jazyka JavaScript. Z toho vyplývají dvě skutečnosti:

- Prolézací modul nalezne některé adresy URL, které by přesnější analyzátor kódu HTML ignoroval. Prolézací modul odmítne všechny řetězce, které nemají platnou syntaxi adresy URL, některé z platných adres URL vrácených v tomto kroku však budou z hlediska vyhledávání nezajímavé.
- Obsah dokumentu generovaný skriptem JavaScript například poté, co uživatel otevře stránku v prohlížeči a prohlížeč zpracuje skript JavaScript, který je součástí této stránky, nelze detekovat pomocí webového prolézacího modulu a proto jej nebude možné indexovat.

Vzhledem k tomu, že webový prolézací modul neanalyzuje skript JavaScript v souborech HTML, nejsou prolézány adresy URL ve skriptu JavaScript. Chcete-li webovému prolézacímu modulu umožnit prolézání adres URL ve skriptu JavaScript, proveďte některou z následujících akcí:

- V konzole pro správu podnikového vyhledávání upravte webový prolézací modul a na stránce Prolézáný webový prostor přidejte adresy URL do seznamu adres URL, které prolézací modul používá jako výchozí bod pro přidávání adres URL do kolekce (**Počáteční adresy URL**). Změny se projeví po restartování webového prolézacího modulu (není nutné zahájit úplné prolézání).
- Použijte kotvicí značku (<a href="..">) k zadání adres URL jako hypertextových odkazů v souboru HTML.

## Pravidla omezující prolézáný webový prostor

Chcete-li zajistit, aby měli uživatelé přístup pouze k webovým serverům, na nichž jim bude povoleno vyhledávání, zadejte pravidla omezující rozsah činnosti webového prolézacího modulu.

Webový prolézací modul při prolézání webové stránky vyhledává odkazy na jiné stránky a tyto odkazy zařazuje do fronty stránek čekajících na prolezení. Prolézání a vyhledávání odkazů lze opakovat tolikrát, kolikrát to dovolí čas a paměťové prostředky. Při konfigurování webového prolézacího modulu určíte, odkud má prolézací modul zahájit prolézání. Z těchto výchozích adres URL (nazývaných *počáteční adresy URL*) se může webový prolézací modul dostat k libovolnému dokumentu na webu, který je propojen přímými nebo nepřímými odkazy.

Chcete-li prolézání prostor omezit, konfigurujte webový prolézací modul tak, aby některé adresy URL zpracovával detailněji a ignoroval odkazy směřující mimo oblast zájmu. Vzhledem k tomu, že prolézací modul standardně přijímá ke zpracování všechny nalezené adresy URL, musíte zadat pravidla určující, které adresy URL chcete do kolekce zahrnout, a eliminující ostatní stránky.

Výběr stránek, které webový prolézací modul má a nemá zpracovávat, lze určit několika způsoby. Můžete nastavit:

- seznam počátečních adres URL, od nichž má prolézací modul zahájit prolézání,
- tři typy pravidel prolézání (doménová pravidla, pravidla pro adresy IP a pravidla pro předpony adres URL),
- seznam typů MIME pro dokumenty, které chcete zahrnout,
- seznam přípon souborů s dokumenty, které chcete vyloučit,
- maximální počet adresářů v cestě URL.

Pravidla prolézání mají následující formát:

```
akce typ cíl
```

Parametr *akce* může mít hodnotu *forbid* (zakázat) nebo *allow* (povolit); *typ* je *domain* (doména), *IP address* (adresa IP) nebo *předpona adresy URL* (HTTP nebo HTTPS) a *cíl* závisí na hodnotě *typ*. S určitými omezeními lze používat hvězdičku (\*) jako zástupný znak a zadat tak cíle, které odpovídají určitému vzoru.

## Doménová pravidla

Cílem doménového pravidla je název domény DNS. Můžete například povolit prolézání celé domény `www.ibm.com`:

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

Zadáte-li hvězdičku ve funkci zástupného znaku, bude pravidlo použito pro všechny názvy hostitelů, které odpovídají vzoru zadanému za hvězdičkou. Můžete například stanovit, že nemají být prolézány žádné domény, jejichž názvy začínají slovem `server` a končí řetězcem `ibm.com`:

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

Při porovnávání názvů hostitelů se rozlišují velká a malá písmena bez ohledu na to, zda zadáte explicitní název domény nebo vzor názvu domény. Vzor `*.user.ibm.com` například odpovídá názvům `joe.user.ibm.com` a `mary.smith.user.ibm.com`, nikoli však názvu `joe.user.IBM.com`.

Pravidlo domény, v němž není určeno číslo portu, platí pro všechny porty v dané doméně. V následujícím příkladu jsou v doméně `sales` povoleny všechny porty:

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

Pokud je v pravidlu domény uvedeno číslo portu, vztahuje se toto pravidlo pouze k danému portu. V následujícím příkladu je v doméně `sales` povolen pouze port 443:

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

## Pravidla pro předpony

Pravidlo pro předpony řídí prolézání adres URL, které začínají určitým řetězcem. Cílem je jediná adresa URL, která obvykle obsahuje jednu nebo více hvězdiček definujících proměnná místa ve vzoru. Hvězdička se často používá například jako poslední znak řetězce předpony.

Pravidlo pro předpony umožňuje prolézat celý web nebo pouze jeho část. Můžete zadat cestu k adresáři nebo vzor cesty a poté povolit nebo zakázat prolézání veškerých dat umístěných za daným bodem ve stromu adresářů. Následující kombinace pravidel například povoluje prolézacímu modulu prolézat veškerá data v adresáři `public` na serveru `sales.ibm.com`, avšak zakazuje prolézacímu modulu přístup k jakýmkoli jiným stránkám na tomto serveru:

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

Při zadávání pravidel pro předpony můžete použít i více než jednu hvězdičku. Hvězdičky se mohou v řetězci předpony vyskytovat kdekoli, ne nutně jen na poslední pozici. Následující pravidlo například zakazuje prolézacímu modulu prolézání všech dokumentů v adresáři nejvyšší úrovně na serveru `sales.ibm.com`, jestliže název adresáře končí písmeny `fs`. (Může se jednat například o přípojný bod souborového systému, které neobsahují informace použitelné ve vyhledávacím indexu.)

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

## Pravidla pro adresy

Pravidlo pro adresy vám umožňuje řídit prolézání celých hostitelů nebo sítí zadáním adresy IP a masky podsítě jako cíle. Příklad:

```
IPv4 allow address 9.0.0.0 255.0.0.0
```

**IPv6** Pokud spouštíte podnikové vyhledávání na serveru Windows 2003 Server a povolíte, aby systém podnikového vyhledávání používal protokol IP verze 6 (IPv6), je třeba uzavřít adresu do hranatých závorek.

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

Pomocí síťové masky můžete zadat vzory pro porovnání. Pro kandidátskou adresu IP je použito pouze pravidlo adresy obsahující adresu IP totožnou s kandidátskou adresou IP s výjimkou bitů, které jsou v masce podsítě vynulovány. Pravidlo adresy definuje vzor a maska podsítě definuje významné bity ve vzoru adresy. Nula v masce podsítě má stejný účinek jako zástupný znak a označuje, že za shodnou bude považována libovolná hodnota uvedená na dané bitové pozici v adrese.

Pravidlo `allow` v předchozím příkladu bude použito pro všechny adresy IP s hodnotou 9 v prvním oktetu a s libovolnými hodnotami ve zbývajících třech oktetech.

Na konci seznamu pravidel je vhodné uvést následující pravidlo. Toto pravidlo platí pro všechny adresy IP, protože podle masky podsítě jsou všechny bity nevýznamné (pravidlo zakazuje všechny adresy, které nejsou povoleny některým z předcházejících pravidel v seznamu).

```
IPv4 forbid address 0.0.0.0 0.0.0.0
```

**IPv6**

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

**Omezení pro servery proxy:** Máte-li v úmyslu prolézat weby obsluhované serverem proxy, nezadávejte pravidla pro adresy IP. Server proxy je obvykle používán v případě, že uživatelský agent (prohlížeč nebo prolézací modul) nemá přímý přístup k sítím, v nichž jsou umístěny webové servery. Server proxy HTTP například předává požadavky HTTP z prolézacího modulu na webový server a odpovědi doručuje zpět prolézacímu modulu.

Pokud webový prolézací modul používá server proxy, je adresa IP serveru proxy jedinou adresou IP, kterou má prolézací modul k dispozici pro jiného hostitele. Pokud je prolézací modul omezen pomocí pravidel pro adresy IP na podsíť adres IP, omezení způsobí, že téměř všechny identifikátory URL budou klasifikovány s použitím stavového kódu 760 (tento kód určuje, že jsou zakázány webovým prostorem).

## Pořadí pravidel prolézání

Prolézací modul používá pravidla prolézání v různých fázích procesu vyhledávání a prolézání adres URL. Na pořadí pravidel záleží, avšak pouze v rámci pravidel téhož typu. Je rozdíl v tom, zda je pravidlo pro adresy uvedeno před jiným pravidlem pro adresy nebo za ním, ale není rozdíl v tom, zda je pravidlo adresy uvedeno před nebo za pravidlem pro předpony, protože prolézací modul tato pravidla nepoužívá ve stejné situaci.

V rámci sady pravidel stejného typu prolézací modul porovnává kandidátskou doménu, adresu IP nebo adresu URL s jednotlivými pravidly od prvního do posledního, dokud nenalezne pravidlo, které vyhovuje. Poté je provedena akce uvedená v prvním vyhovujícím pravidlu.

Ze závislosti na pořadí vyplývá typická struktura většiny pravidel prolézání:

- Sada doménových pravidel obvykle začíná pravidly forbid, které z prolézaného prostoru eliminují jednotlivé domény. Administrátor kolekce může například zjistit, že některé domény neobsahují užitečné informace.
- Za seznamem pravidel forbid obvykle následuje posloupnost pravidel allow (se zástupnými znaky) povolujícími prolézacímu modulu navštívit všechny domény, jejichž názvy končí názvem některé domény vyšší úrovně definující podnikový intranet (například \*.ibm.com nebo \*.lotus.com).

Ukončete sadu doménových pravidel následujícím výchozím pravidlem vylučujícím domény, které nebyly explicitně povoleny některým z předcházejících pravidel:

```
forbid domain *
```

Toto závěrečné pravidlo je velmi důležité, protože brání rozšíření prolézaného prostoru na celý Internet.

- Sada pravidel pro adresy obvykle začíná malým počtem pravidel allow, která prolézacímu modulu povolují prolézat síť vyšších úrovní (třídy A, B a C) v rámci podnikového intranetu.

Pročtěte si předchozí diskusi o pravidlech pro adresy, kde jsou uvedeny příklady uvedení závěrečného pravidla v seznamu pravidel, které prolézacímu modulu brání v přístupu k webovým serverům mimo podnikovou síť.

- Sada pravidel pro předpony bývá nejrozsáhlejší, protože obsahuje podrobné specifikace povolených a zakázaných oblastí zadané ve formě stromů a podstromů. Osvědčuje se povolit nebo zakázat nejprve úzce lokalizované oblasti a poté zadat opačné pravidlo s obecnějším vzorem, které povoluje nebo zakazuje všechny ostatní oblasti.

Sekke předpon zpravidla nekončí uzavíracím pravidlem. Výše zmiňovaná uzavírací pravidla pro domény a adresy sama o sobě zajistí, že prolézací modul nebude pracovat s oblastmi mimo podnikovou síť, a to efektivněji, než by bylo možné docílit testováním předpon adres URL.

Prolézací modul je schopen aplikovat pravidla pro předpony efektivněji, pokud je seskupíte podle akcí (forbid a allow). Namísto střídavého zadávání krátkých posloupností pravidel allow a forbid zadejte dlouhou posloupnost pravidel provádějících jednu akci a poté dlouhou posloupnost pravidel provádějících druhou akci. Střídání pravidel allow a forbid je opodstatněné, pokud jim směřujete k určité cílové struktuře prolézaného prostoru. Seskupení pravidel allow do jedné skupiny a pravidel forbid do druhé skupiny však může zvýšit výkon prolézacího modulu.

## Přípony souborů, typy MIME a maximální hloubka prolézání

Tyto volby představují další způsoby stanovení obsahu prolézaného prostoru. Můžete z něj vyloučit určité typy dokumentů na základě přípon souborů s dokumenty a zahrnout do něj určité typy dokumentů na základě typu MIME daného dokumentu. Při zadávání typů MIME, které má prolézací modul zpracovávat, počítejte s tím, že typy MIME jsou ve webových dokumentech často nastaveny nesprávně.

Maximální hloubka prolézání určuje maximální počet lomítek v adrese URL od kořenu daného serveru. Tato volba vám umožňuje zabránit prolézacímu modulu v prolézání rekurzivních struktur souborového systému o nekonečné hloubce. Hloubka prolézání nesouvisí s počtem odkazů, které prolézací modul použije při přecházení z jednoho dokumentu do druhého.

## Počáteční adresy URL

Počáteční adresy URL jsou adresy URL, z nichž prolézací modul při prolézání vychází. Tyto adresy URL jsou do prolézání vloženy při každém spuštění prolézacího modulu. Pokud již u počátečních adres URL proběhlo zjišťování, nebudou tyto adresy prolezeny ani opětovně prolezeny dříve než jiné webové servery povolené v pravidlech prolézání.

Počáteční adresa URL je důležitá při prvním spuštění webového prolézacího modulu s prázdným prolézaným prostorem. Další situace, v níž jsou počáteční adresy URL důležité, je přidávání nových adres URL, u nichž dosud neproběhlo zjišťování, do prolézaného prostoru.

Jako počáteční adresy URL lze použít pouze úplné adresy URL, nikoli pouze názvy domén. Je nutné zadat protokol a také číslo portu, pokud není použit port 80.

Dále jsou uvedeny příklady platných počátečních adres URL:

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

Příklad neplatné počáteční adresy URL:

```
www.ibm.com
```

Počáteční adresy URL je nutné zahrnout do pravidel prolézání. Prolézací modul například nemůže zahájit prolézání zadané počáteční adresy URL, pokud pravidla prolézání zakazují přístup k této adrese URL.

**Podpora adres IPv6:** Pokud spouštíte podnikové vyhledávání na serveru Windows 2003 Server a povolíte, aby systém podnikového vyhledávání používal protokol IP verze 6 (IPv6), je třeba uzavřít počáteční adresy URL do hranatých závorek. Příklad:

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

### Související úlohy

“Povolení podpory protokolu IPv6” na stránce 25

## Testování spojení adres URL s webovým prolézacím modulem

Po zadání adres URL, které má prolézací modul prolézat, můžete otestovat konfiguraci pravidel prolézání.

Test můžete spustit klepnutím na tlačítko **Testovat** při zadávání prolézáných domén, předpon HTTP nebo adres IP, a vyberete-li stránku Testovat adresy URL, můžete otestovat schopnost prolézacího modulu připojovat se kromě adres URL, které jste zadali, také k počátečním adresám URL.

Výsledky testu ukazují, zda má prolézací modul přístup k adresám URL s použitím názvu uživatelského agenta určeného ve vlastnostech prolézacího modulu. Výsledky testu mohou také ukázat, že určitou adresu URL nelze prolézat kvůli pravidlům vyloučení (pokud například některý dokument nelze prolézat kvůli tomu, že jeho přípona odpovídá příponám vyloučeným z prolézaného prostoru).

Pokud byl určitý web již alespoň jednou prolezen, můžete na základě testu adres URL získat další informace. Prostřednictvím sestavy s výsledky testu můžete zjistit aktuální stavový kód HTTP (informující o tom, zda byla operace prolézání dané adresy URL úspěšná či nikoli), zobrazit údaje o tom, kdy byla daná adresa URL naposledy prolezena, kdy má být podle časového plánu prolézána znovu a zda uživatelský agent používá aktuální soubor robots.txt webového serveru.

## Nastavení intervalu opakovaného prolézání ve webovém prolézacím modulem

Chcete-li určit, jak často se má webový prolézací modul vracet k již prolezeným adresám URL, můžete nastavit příslušné volby ve vlastnostech webového prolézacího modulu.

Většina ostatních typů prolézacích modulů v systému podnikového vyhledávání se spouští podle plánů nastavených administrátorem. Webový prolézací modul naproti tomu po spuštění obvykle pracuje trvale. Chcete-li určit, jak často má tento modul navštěvovat již jednou zpracované adresy URL, můžete nastavit minimální a maximální interval opakovaného prolézání.

Použijete-li konzolu pro správu podnikového vyhledávání k vytvoření webového prolézacího modulu nebo k úpravě vlastností webového prolézacího modulu, můžete vybrat volbu, která slouží ke konfigurování rozšířených vlastností. Na stránce Rozšířené vlastnosti webového prolézacího modulu můžete zadat volby minimálního a maximálního intervalu opakovaného prolézání. Webový prolézací modul použije zadané hodnoty k výpočtu intervalů opakovaného prolézání dat.

Při prvním prolézání stránky prolézací modul použije datum a čas prolezení stránky a průměr ze zadané minimální a maximální hodnoty intervalu opakovaného prolézání k nastavení data příštího prolézání. Před tímto datem stránka nebude znovu prolézána. Doba opětovného prolézání stránky po tomto datu závisí na zátěži prolézacího modulu a poměru starých a nových adres URL v prolézaném prostoru.

Při každém dalším prolézání stránky prolézací modul kontroluje, zda se její obsah změnil. Pokud ano, bude příští interval opakovaného prolézání kratší než předcházející, nikdy však jeho délka neklesne pod zadaný minimální interval opakovaného prolézání. Pokud se obsah stránky nezměnil, bude příští interval opakovaného prolézání delší než předcházející, nikdy však jeho délka nepřesáhne zadaný maximální interval opakovaného prolézání.

## Volby pro návštěvy adres URL webovým prolézacím modulem

V případě potřeby je možné vynutit co nejdřívější návštěvu konkrétní adresy URL webovým prolézacím modulem.

Pokud potřebujete aktualizovat prolézáný prostor na základě informací získaných z určitých webových serverů, můžete aktivovat monitorování prolézacího modulu, vybrat volbu **Adresy URL, které mají být navštíveny nebo znovunavštíveny** a poté zadat adresy nebo vzory URL určující stránky, které chcete prolézt (poprvé nebo znovu).

Pokud například komunikační oddělení přidá na intranet webovou stránku nebo některou stránku reviduje, aby odrážela důležitou změnu zásad, můžete zadat adresu URL nové nebo změněné stránky. Pokud je prolézací modul spuštěný, zařadí určenou adresu URL pro prolézání do fronty při příští kontrole stránek čekajících na zpracování (obvyklá doba je každých 10 minut). Není-li spuštěn, zařadí zadanou stránku URL do fronty tak, aby mohla být prolezena při příštím spuštění prolézacího modulu.

Zkontrolujte, zda je v pravidlech prolézání obsaženo pravidlo, které prolézacímu modulu umožňuje navštívit určené adresy URL. Prolézací modul může navštívit určené adresy URL dříve než obvykle. Aby mohla být adresa URL vůbec prolezena, musí existovat pravidlo prolézání, které adrese URL umožní být prolézána.

Nově prolezená data budou k dispozici pro vyhledávání při příštím sestavení hlavního indexu.

## Způsob práce webového prolézacího modulu se stránkami méně závažných chyb

Webový prolézací modul lze konfigurovat pro zpracování přizpůsobených stránek, které administrátoři webových serverů vytvářejí namísto standardních chybových kódů vrácených jako odpovědi na požadavky na některé stránky.

Pokud server HTTP nemůže vrátit stránku požadovanou klientem, za normálních okolností vrací odezvu tvořenou záhlavím se stavovým kódem. Stavový kód informuje o povaze problému (například chyba 404 informuje o tom, že soubor nebyl nalezen). Administrátoři některých webových serverů vytvářejí speciální stránky objasňující problém podrobněji a konfiguruji servery HTTP tak, aby místo standardní odezvy vracely tyto přizpůsobené stránky. Přizpůsobené stránky se nazývají *stránky méně závažných chyb*.

Stránky s málo závažnými chybami mohou zkreslovat výsledky webového prolézacího modulu. Namísto hlavičky informující o existenci problému obdrží prolézací modul stránku méně závažné chyby a stavový kód 200, který oznamuje úspěšné stažení platné stránky HTML. Tato stažená stránka méně závažné chyby však nesouvisí s požadovanou adresou URL a její obsah je prakticky totožný vždy, když je vrácena místo požadované stránky. Tyto irelevantní a v podstatě duplicitní stránky zkreslují obsah indexu a výsledky vyhledávání.

Chcete-li tuto situaci vyřešit, můžete při konfigurování webového prolézacího modulu zadat volby pro práci se stránkami méně závažných chyb. Pro každý webový server vracející stránky méně závažných chyb potřebuje webový prolézací modul následující informace:

- Vzor adresy URL pro web, který používá stránky s málo závažnými chybami. Tento vzor adresy URL se skládá z protokolu (HTTP nebo HTTPS), názvu hostitele, čísla portu (jestliže není použit standardní port) a názvu cesty. Ve vzoru můžete použít hvězdičku (\*) jako zástupný znak odpovídající jednomu či více standardním znakům až do nejbližšího výskytu znaku, který není zástupným znakem. V zadaném vzoru se rozlišují velká a malá písmena.

- Vzor názvu pro text odpovídající značce <TITLE> v dokumentu HTML. Při zadávání tohoto vzoru můžete použít hvězdičku (\*) jako zástupný znak. V zadaném vzoru se rozlišují velká a malá písmena.
- Vzor obsahu pro text odpovídající obsahu dokumentu HTML. Nejedná se pouze o obsah značky <BODY>, pokud je značka <BODY> v dokumentu obsažena. Pojmeme obsah označujeme veškerá data následující v souboru za záhlavím HTTP. Při zadávání tohoto vzoru můžete použít hvězdičku (\*) jako zástupný znak. V zadaném vzoru se rozlišují velká a malá písmena.
- Celé číslo reprezentující stavový kód, který má být použit pro dokumenty odpovídající zadaným vzorům adresy URL, názvu a obsahu.

## Příklad

Následující konfigurace předává webovému prolézacímu modulu pokyn k porovnání všech platných stránek HTML (stavový kód 200) vrácených z webového serveru `http://www.mysite.com/hr/*` se zadanými vzory názvu a obsahu. Pokud obsah značky <TITLE> na stránce začíná textem "Sorry, the page" (Omlouváme se, stránka) a obsah dokumentu je libovolný (\*), prolézací modul pracuje se stránkou stejným způsobem, jako kdyby obsahovala stavový kód 404 (Stránka nebyla nalezena).

Tabulka 3. Příklad stránky méně závažné chyby

Vzor adresy URL	Vzor názvu	Vzor obsahu	Stavový kód HTTP
<code>http://www.mysite.com/hr/*</code>	<code>Sorry, the page*</code>	<code>*</code>	404

Pro tentýž webový server lze vytvořit více položek umožňujících zpracovávat různé stavové kódy. Pro každý stavový kód téhož webového serveru musí existovat vlastní položka v konfiguraci webového prolézacího modulu.

## Použití zástupných znaků

Vzory adresy URL, názvu a obsahu nejsou regulární výrazy. Znak hvězdička se shoduje s libovolnými znaky až do nejbližšího výskytu libovolného znaku, který není zástupným znakem. Příklad:

\*404 odpovídá hodnotě *libovolné znaky*404  
 Položka 404: \* odpovídá hodnotě 404: libovolné znaky  
 Položka `http://*.mysite.com/*` odpovídá hodnotě `http://libovolný hostitel.mysite.com/libovolný soubor`.  
 Položka \* odpovídá hodnotě *libovolné znaky*

## Vliv na výkon

Konfigurováním voleb pro práci se stránkami méně závažných chyb se prodlužuje doba zpracování prolézacího modulu, protože je nutné zkontrolovat všechny úspěšně prolezené stránky. Další strojový čas spotřebuje kontrola odpovídajících vzorů a zjišťování, zda má být vrácena stránka nebo náhradní stavový kód.

## Podpora prolézání zabezpečených webových serverů

Zadáním pověřovacích údajů na konzole pro správu podnikového vyhledávání můžete webovému prolézacímu modulu umožnit přístup k chráněnému obsahu, například k dokumentům, které nejsou dostupné bez zadání hesla.



Pokud webový server omezuje přístup k webům pomocí základního ověřování HTTP nebo pomocí ověřování založeného na formulářích HTML, můžete v konfiguraci webového prolézacího modulu zadat pověřovací údaje umožňující prolézání webových stránek chráněných heslem. Rovněž můžete zadat volby pro ruční konfigurování souborů cookie.

## Webové servery chráněné základním ověřováním HTTP

Pokud webový server používá k omezení přístupu k webům základní ověřování HTTP, můžete zadat pověřovací údaje pro ověřování, které webovému prolézacímu modulu umožní přístup na stránky chráněné heslem.

Při zjišťování, zda má uživatel (nebo klientská aplikace) oprávnění pro přístup k určitému webu, používá mnoho webových serverů schéma ověřování klientů nazývané základní ověřování HTTP, které slouží k prokazování identity uživatele. Toto ověřování obvykle probíhá interaktivně:

- Když si uživatelský agent HTTP (například webový prohlížeč) vyžádá stránku chráněnou základním ověřováním HTTP, webový server odpoví stavovým kódem 401, který oznamuje, že žadatel nemá oprávnění pro přístup k požadované stránce.
- Webový server rovněž vyzve žadatele k poskytnutí pověřovacích údajů, na jejichž základě lze ověřit, zda je uživateli povolen přístup k chráněnému obsahu.
- Webový prohlížeč zobrazí dialogové okno s výzvou k zadání jména uživatele, hesla a dalších informací tvořících pověření uživatele.
- Webový prohlížeč pověřovací údaje zakóduje a později je uvede při opakované žádosti o přístup k chráněné stránce.
- Pokud jsou pověřovací údaje platné, webový server odpoví návratovým kódem 200 a vrátí obsah požadované stránky.
- V dalších požadavcích na stránky ze stejného webového serveru jsou obvykle uvedeny stejné pověřovací údaje, čímž je autorizovanému uživateli umožněn přístup k dalšímu chráněnému obsahu bez zadávání dalších pověřovacích údajů.

Po prokázání identity uživatele si webový server obvykle vymění s uživatelským agentem HTTP prvky nazývané *položky cookie*, které umožňují zachovat informace o stavu přihlášení uživatele mezi dvěma požadavky HTTP.

Vzhledem k tomu, že webový prolézací modul nepracuje interaktivně, je nutné zadat pověřovací údaje potřebné pro přístup ke stránkám chráněným heslem ještě předtím, než prolézací modul začne s prolézáním. Při vytváření webových prolézacích modulů a úpravách prolézaného prostoru zadejte informace o jednotlivých zabezpečených webových serverech, které potřebujete prolézat.

Při zadávání těchto informací je nutné úzce spolupracovat s administrátory webů nebo webových serverů chráněných pomocí základního ověřování HTTP. Tito administrátoři vás musí informovat o bezpečnostních požadavcích platných pro webové servery, které chcete prolézat, včetně veškerých informací používaných k ověření totožnosti webového prolézacího modulu a zjištění, zda má tento modul oprávnění k prolézání stránek s omezeným přístupem.

Pokud bylo pro kolekci při jejím vytvoření povoleno zabezpečení, můžete při konfigurování prolézacího modulu zadat prvky zabezpečení pro řízení přístupu k dokumentům, například jméno uživatele, ID skupin nebo role uživatelů. Webový prolézací modul asociuje tyto prvky zabezpečení se všemi dokumenty, které prolézá ve stromu souborového systému s danou kořenovou adresou URL. Tyto prvky doplňují prvky zabezpečení na úrovni dokumentu, které konfigurujete pro celý prolézáný webový prostor.

Pořadí adres URL má svůj význam. Po přidání informací o webech chráněných heslem je třeba tyto informace seřadit podle toho, jak je má prolézací modul zpracovávat. V seznamu uvádějte nejprve konkrétnější adresy URL a teprve za nimi obecnější adresy URL. Při

vyhodnocování kandidátské adresy URL používá webový prolézací modul ověřovací data uvedená u první adresy URL v seznamu, která odpovídá kandidátské adrese URL.

## **Webové servery chráněné ověřováním založeným na formulářích**

Pokud webový server používá pro omezení přístupu k webům formuláře HTML, můžete zadat pověřovací údaje pro ověřování, které webovému prolézacímu modulu umožní přístup na stránky chráněné heslem.

Při zjišťování, zda má uživatel (nebo klientská aplikace) oprávnění pro přístup k určitému webu, používá mnoho webových serverů k prokazování identity uživatele formuláře HTML. Toto ověřování obvykle probíhá interaktivně:

- Když si uživatelský agent HTTP (například webový prohlížeč) vyžádá stránku chráněnou ověřováním založeným na formulářích, webový server zkontroluje, zda požadavek obsahuje položku cookie prokazující identitu uživatele.
- Pokud není položka cookie nalezena, webový server vyzve uživatele k zadání dat zabezpečení do formuláře. Když uživatel formulář odešle, webový server vrátí požadované položky cookie a povolí zpracování požadavku na stránku chráněnou heslem.
- Současně je povoleno zpracování dalších požadavků obsahujících požadované položky cookie. Autorizovanému uživateli je umožněn přístup k dalšímu chráněnému obsahu bez nutnosti vyplňovat formulář a zadávat pověřovací údaje s každým požadavkem.

Vzhledem k tomu, že webový prolézací modul nepracuje interaktivně, je nutné zadat pověřovací údaje potřebné pro přístup ke stránkám chráněným heslem ještě předtím, než prolézací modul začne s prolézáním. Při vytváření webových prolézacích modulů a úpravách prolézáního prostoru zadejte informace o jednotlivých zabezpečených webových serverech, které potřebujete prolézat.

Zadávaná pole odpovídají polím, která uživatel vyplňuje v interaktivním režimu na výzvu webového prohlížeče, a všem skrytým a statickým polím vyžadovaným pro úspěšné přihlášení.

Při zadávání těchto informací je nutné úzce spolupracovat s administrátory webů nebo webových serverů chráněných pomocí ověřování založeného na formulářích. Tito administrátoři vás musí informovat o bezpečnostních požadavcích platných pro webové servery, které chcete prolézat, včetně veškerých informací používaných k ověření totožnosti webového prolézacího modulu a zjištění, zda má tento modul oprávnění k prolézání stránek s omezeným přístupem.

Pořadí vzorů adres URL má svůj význam. Po přidání informací o webech chráněných heslem je třeba tyto informace seřadit podle toho, jak je má prolézací modul zpracovávat. Na začátek seznamu vložte konkrétnější vzory adres URL a na konec seznamu obecnější vzory adres URL. Při vyhodnocování kandidátské adresy URL používá webový prolézací modul data formuláře uvedená v prvním vzoru adresy URL v seznamu, který odpovídá kandidátské adrese URL.

## **Použití modulu plug-in k prolézání zabezpečených webů produktu WebSphere Portal**

Je-li v produktu WebSphere Application Server povoleno globální zabezpečení a chcete-li pomocí webového prolézacího modulu prolézat zabezpečené weby produktu WebSphere Portal, je třeba vytvořit modul plug-in prolézacího modulu, který obsluží požadavky ověřování založeného na použití formulářů. Výhody a nevýhody ověřování založeného na použití formulářů a ukázkový program, který můžete přizpůsobit pro svůj vlastní modul plug-in webového prolézacího modulu, najdete v tématu <http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani>.

Modul plug-in je vyžadován, pokud používáte webový prolézací modul k prolézání libovolných webů produktu WebSphere Portal, včetně webů produktů Workplace Web Content Management a Lotus Quickr.

## Webové servery obsluhované servery proxy

Pokud nemá webový prolézací modul povolen přímý přístup k síti, můžete jej konfigurovat pro přístup k prolézanému obsahu přes server proxy HTTP.

Nemá-li počítač, ve kterém má pracovat webový prolézací modul, přístup k síti TCP/IP nebo je-li přístup omezen na privilegované procesy, můžete webový prolézací modul konfigurovat pro použití serveru proxy HTTP. Server proxy HTTP je proces, který přijímá požadavky HTTP na určeném portu zadaného hostitele. Server proxy předává požadavky webovému serveru a odezvy webového serveru vrací klientovi, který si je vyžádal (webovému prolézacímu modulu). Server proxy může pracovat ve stejném počítači jako webový prolézací modul nebo v jiném počítači.

U prolézání bez použití serveru proxy je požadavek na adresu URL zasílán přímo hostiteli. U prolézání s použitím serveru proxy je požadavek zasílán serveru proxy.

Při vytváření webových prolézacích modulů a úpravách prolézaného prostoru zadejte informace o serverech proxy, které má webový prolézací modul použít při prolézání stránek v doméně serveru proxy. Dříve, než do prolézaného prostoru přidáte server proxy, zjistěte si názvy domén obsluhovaných tímto serverem proxy, hostitelský název nebo adresu IP serveru proxy a číslo portu, který server proxy používá.

Pokud server proxy vyžaduje ověřování, zjistěte si také jméno uživatele a heslo, které může prolézací modul používat pro přístup ke stránkám obsluhovaným serverem proxy. Webový prolézací modul podporuje pouze základní ověřování serveru proxy pomocí protokolu HTTP popsané v dokumentu RFC2616 (<http://rfc.net/rfc2616.html>). Ostatní typy ověřování včetně ověřování Windows NTLM (NT LAN Manager) podporovány nejsou.

Po přidání je třeba server proxy vybrat a umístit tak, aby prolézací modul zpracovával servery proxy v požadovaném pořadí. Na začátek seznamu vložte konkrétnější názvy domén a na konec seznamu obecnější názvy domén. Při vyhodnocování kandidátské adresy URL používá webový prolézací modul data serveru proxy uvedené u první domény v seznamu, která odpovídá kandidátské adrese URL. (Adresy URL, které neodpovídají žádnému pravidlu serveru proxy, považuje prolézací modul za přímo dostupné.)

## Správa souborů cookie

Správa souborů cookie probíhá obvykle automaticky, bez aktivní účasti administrátora podnikového vyhledávání. V případě potřeby však můžete položky souboru cookie pro relaci webového prolézání zadat ručně.

Položky cookie jsou skryté prvky, které webový server vrací agentovi uživatele v rámci hlavičky odpovědi HTTP. Mají význam pouze pro webový server, který je vygeneroval, a slouží k udržování stavu mezi požadavky HTTP. Během ověřování klienta může server například vrátit položku cookie, která mu později umožní zjistit, že ověřený uživatel je již přihlášen. Přítomnost souboru cookie umožňuje uživateli zadávat další požadavky na stránky na stejném webovém serveru, aniž by se musel znovu přihlašovat.

Webový prolézací modul uchovává položky cookie přijaté od webových serverů a používá je po celou dobu existence dané instance prolézacího modulu. Položky cookie se ukládají do souboru `cookies.ini`. Na konci každé relace prolézacího modulu je tento soubor přepsán. Při zastavení webového prolézacího modulu jsou všechny položky v souboru cookie, jejichž platnost dosud nevypršela, uloženy a při spuštění další relace jsou opět načteny.

Pokud zadáváte položky cookie ručně, uložte je do samostatného souboru a v okamžiku, kdy je potřebujete použít, je slučte s položkami v souboru `cookies.ini`. Prolézací modul neodstraňuje položky v souboru cookie, jejichž platnost dosud nevypršela, pokud by však nastal problém bránící zapsání celé kolekce položek cookie, při dodržení tohoto postupu nepřijdete o položky v souboru cookie, které jste zadali ručně. Sloučení vlastních položek cookie s položkami udržovanými automaticky prolézacím modulem musíte provést před spuštěním relace prolézacího modulu.

## Formát souborů cookie

Soubory cookie, které chcete sloučit se souborem `cookies.ini` podnikového vyhledávání, musí dodržovat stanovený formát.

- Každá položka v souboru cookie musí být uvedena na samostatném řádku. Použití prázdných řádků a komentářů je povoleno, nebudou však přeneseny do souboru `cookies.ini`.
- Každá položka v souboru cookie musí mít následující formát:

```
CookieN(délka_položky,délka_URL)text_položky,ověřovací_URL
```

### Cookie

Povinné klíčové slovo označující začátek položky cookie.

Klíčové slovo `Cookie` nesmí obsahovat mezery a musí k němu být připojena číslice 0, 1 nebo 2. Tato číslice označuje typ položky cookie: `version-0` (Netscape), `version-1` (RFC2109) nebo `version-2` (RFC2965). V položkách cookie typu RFC2965 nejsou podporovány seznamy portů.

### *délka\_položky*

Délka (počet znaků) připojeného textu položky cookie.

### *délka\_URL*

Délka (počet znaků) připojené ověřovací adresy URL.

### *text\_položky*

Obsah položky cookie, určený k odeslání webovému serveru, který položku vytvořil. Tento řetězec (reprezentující pravou stranu direktivy `Set-Cookie` v hlavičce odezvy HTTP) obsahuje dvojici názvu a hodnoty položky cookie a veškerý další obsah (například cestu, nastavení zabezpečení apod.), který má být odeslán spolu s položkou cookie. Za tímto řetězcem následuje oddělovací čárka (,).

### *ověřovací\_URL*

Adresa URL, odkud daná položka cookie pochází. Tato adresa URL určuje, kam má být položka cookie odeslána (například na základě názvu domény a názvu cesty). Ověřovací adresa URL musí odpovídat omezením pro položky cookie, která zavádí původní webový server v zájmu zabezpečení a ochrany soukromí.

Následující příklad je pro přehlednost uveden na dvou řádcích, položky v souboru cookie je však nutné uvádět vždy na jednom řádku:

```
Cookie0(53,40)ASPSESSIONIDQSQTACSD=SLNSIDFNLSIDNFLSINFLSINL;path=/  
https://www.ibm.com:443/help/solutions/
```

## Konfigurování souboru cookie pro webový prolézací modul

V případě potřeby můžete ručně zadat položky v souboru cookie pro relaci webového prolézacího modulu a sloučit je s položkami spravovanými webovým prolézacím modulem.

### Než začnete

Chcete-li ručně konfigurovat položky souboru cookie používaného webovým prolézacím modulem, musíte být přihlášen jako administrátor podnikového vyhledávání.

## Postup

Při ručním konfigurování položek v souboru cookie pro webový prolézací modul postupujte takto:

1. Z konzoly pro správu podnikového vyhledávání spusíte monitorování kolekce, pro kterou chcete nastavit soubor cookie, a zastavíte webový prolézací modul.
2. Přihlašte se k prolézacímu serveru jako administrátor podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition).
3. Přejděte do datového adresáře prolézacího modulu, který chcete konfigurovat, kde *ID\_relace\_prolézacího\_modulu* je ID přiřazené k relaci prolézacího modulu systémem podnikového vyhledávání. Příklad:  
`ES_NODE_ROOT/data/co1_56092.WEB_88534`
4. Otevřete soubor `cookies.ini` pro úpravy, připojte ručně zadané položky cookie k těm, které jsou již v souboru uvedeny, a poté soubor uložte a zavřete. Zkontrolujte, zda položky cookie, které zadáváte, nepřepíší existující položky.
5. Z konzoly pro správu podnikového vyhledávání znovu spusíte webový prolézací modul, který jste zastavili.

## Konfigurace globálního prolézaného webového prostoru

Pro webové prolézací moduly můžete konfigurovat globální prolézaný prostor, což vám poskytne lepší kontrolu nad odebíráním adres URL z indexu.

Každý webový prolézací modul je konfigurován spolu s prolézaným prostorem definujícím adresy URL, které mají nebo nemají být prolézány. Nalezené adresy URL obsažené v prolézaném prostoru jsou uchovány (v databázi) pro pozdější prolézání. Adresy URL, které nepatří do prolézaného prostoru, jsou odstraněny. Pokud prolézací modul zahájí prolézání s prázdnou databází, zůstává definice prolézaného prostoru konzistentní s databází po celou dobu práce prolézacího modulu.

V některých případech dochází k zastavení prolézacího modulu a k omezení jeho prolézaného prostoru (například novými pravidly zakazujícími prolézání některých stránek). Po novém spuštění prolézacího modulu je definice prolézaného prostoru nekonzistentní s databází. Databáze obsahuje adresy URL (prolezené i neprolezené), které nepatří do nového menšího prolézaného prostoru.

Pokud kolekce obsahuje pouze jeden webový prolézací modul, může tento modul obnovit konzistenci změnou stavových kódů HTTP pro popisované adresy URL na hodnotu 760 (která určuje, že daná adresa má být vyloučena) a odesláním žádosti o odebrání nově vyloučených stránek z indexu.

Je-li prolézaný prostor rozdělen mezi dva a více webových prolézacích modulů (například proto, aby bylo zajištěno častější prolézání některých důležitých stránek), udržuje každý z těchto modulů vlastní nezávislé databázové tabulky (na počátku prázdné) a prolézá jinou část webového prolézaného prostoru. Prolézaný prostor původního prolézacího modulu je poté omezen na zbytek prolézaného prostoru po odebrání částí pokrytých jinými prolézacími moduly. Pokusí-li se původní prolézací modul obnovit konzistenci odebráním přesunutých stránek z indexu, nastávají problémy. Vzhledem k tomu, že s přesunutými stránkami nyní pracují jiné prolézací moduly, měly by tyto stránky zůstat v indexu.

Konfigurováním globálního prolézaného prostoru na vyšší úrovni můžete určit adresy URL, které nemají být prolézány původním prolézacím modulem, avšak ani odebrány z indexu. Adresy URL, které již nejsou součástí žádného z prolézaných prostorů daného prolézacího modulu, budou i nadále označovány zjišťovacím procesem jako určené k vyloučení a při opakovaném prolézání budou odebrány z indexu.

K definování globálního prolézaného prostoru slouží konfigurační soubor s názvem `global.rules`, který se musí nacházet v konfiguračním adresáři prolézacího modulu (přítomnost souboru `global.rules` aktivuje funkci globálního prolézaného prostoru). Pokud tento soubor existuje, je načten během inicializace prolézacího modulu. Pokud tento soubor neexistuje, prolézací modul pracuje s jednovrstvým prolézaným prostorem a podle potřeby odebírá dokumenty z indexu, je-li třeba zajistit konzistenci definice prolézaného prostoru s databází.

V případě použití globálního prolézaného prostoru prolézací modul zahrnuje a vylučuje adresy URL jako obvykle, ale o odebrání adresy URL z indexu požádá jen v případě, že se daná adresa URL nenachází v žádném z prolézaných webových prostorů.

Soubor `global.rules` používá stejnou syntaxi jako místní soubor `crawl.rules`, obsahuje však pouze pravidla pro názvy domén. Toto omezení umožňuje rozdělit prolézaný prostor mezi prolézací moduly pouze na základě hostitelských názvů DNS, nikoli podle adres IP nebo vzorů předpon HTTP. Na adresy URL vyloučené z lokálního prolézaného prostoru pomocí pravidel pro předpony adres URL nebo pomocí pravidel pro adresy IP (dle definice v souboru `crawl.rules`) nemá globální prolézaný prostor vliv - takové adresy URL jsou i nadále vylučovány.

Globální prolézaný prostor se používá pouze jako prostředek zabráňující odebrat z indexu adresy URL, které jsou vyloučeny z prolézaného prostoru jednoho prolézacího modulu pomocí pravidla pro lokální domény. Platí následující pravidla v uvedeném pořadí:

1. Je-li adresa URL vyloučena z databáze prolézacího modulu pomocí pravidla pro lokální předpony nebo pomocí pravidla pro adresy, je jí přiřazen stavový kód 760 a poté je odebrána z indexu. Taková adresa URL již nebude prolézána.
2. Je-li adresa URL obsažená v databázi prolézacího modulu vyloučena pomocí pravidla pro lokální domény a neexistuje-li globální prolézaný prostor, je této adrese URL přiřazen stavový kód 760 a poté je odebrána z indexu. Taková adresa URL již nebude prolézána.
3. Je-li adresa URL obsažená v databázi prolézacího modulu vyloučena pomocí pravidla pro lokální domény, avšak zároveň je explicitně zahrnuta pomocí pravidla v globálním prolézaném prostoru, je této adrese URL přiřazen stavový kód 761. Prolézací modul již takovou adresu URL nebude prolézat, ale neodebere ji z indexu (předpokládá se, že je obsažena v lokálním prolézaném prostoru jiného prolézacího modulu).
4. Je-li adresa URL obsažená v databázi prolézacího modulu vyloučena pomocí pravidla pro lokální domény a není-li ani explicitně zahrnuta pomocí pravidla v globálním prolézaném prostoru, je této adrese URL přiřazen stavový kód 760 a poté je odebrána z indexu.

Vzhledem k tomu, že globální prolézaný prostor se používá pouze k zabránění výmazu adres URL, které již byly vyloučeny z lokálního prolézaného prostoru, platí zásada, že pokud se na kandidátskou adresu URL nevztahuje žádné pravidlo globálního prolézaného prostoru, je prolézání této adresy URL zakázáno.

Soubor `global.rules` musí existovat v adresáři `master_config` všech prolézacích modulů sdílejících globální prolézaný prostor. Všechny kopie souboru `global.rules` a jednotlivé lokální soubory `crawl.rules` je nutné pečlivě upravovat tak, aby byly trvale vzájemně konzistentní.

## Direktivy `no-follow` a `no-index`

Kvalitu vyhledávání lze zlepšit zadáním direktiv pro webový prolézací modul, které určují, zda má prolézací modul použít odkazy na stránce a zda mají být stránky indexovány.

Některé webové stránky obsahují direktivy no-follow a no-index informující roboty (například webové prolézací moduly) o tom, že nemají používat odkazy uvedené na těchto stránkách, že nemají obsah těchto stránek zahrnout do indexu, nebo že nemají provádět žádnou z těchto akcí.

Řízením těchto nastavení lze zvýšit kvalitu prolézání. Některé adresářové stránky mohou například obsahovat tisíce odkazů, ale žádný použitelný obsah. Tyto stránky je třeba prolézt a použít odkazy, které obsahují, k přechodu na další stránky, indexování těchto stránek jako takových však nepřinese žádný užitek.

V některých případech můžete rovněž požadovat, aby prolézací modul nepokračoval do hlubších vrstev hierarchie, avšak požadované listové stránky obsahují odkazy a neobsahují direktivy no-follow. Vzhledem k tomu, že některé z těchto stránek jsou generovány automaticky, nemají žádné vlastníky, kteří by do nich mohli vložit potřebné direktivy.

Chcete-li zadat pravidla pro prolézání takových stránek, vytvořte nebo upravte konfigurační soubor s názvem `followindex.rules`. Při definování pravidel v tomto souboru postupujte podle následujících pokynů:

- V konfigurovaných pravidlech musí být uvedeny předpony adres URL (webové servery tedy nelze identifikovat podle adres IP ani podle hostitelských názvů DNS).
- Předpony adres URL mohou obsahovat hvězdičky (\*) použité jako zástupné znaky. Tímto způsobem lze povolit nebo zakázat více webů s podobnými adresami URL.
- Použití pravidel závisí na jejich pořadí v seznamu (prolézací modul použije první pravidlo odpovídající kandidátské adrese URL).
- Pravidla, která explicitně povolují a zakazují použití odkazů nebo indexování stránek, mají přednost před jinými nastaveními včetně nastavení obsaženého v cílovém dokumentu.

## Potlačení direktiv no-follow a no-index na webových stránkách

V konfiguračním souboru můžete zadat pravidla určující, zda má webový prolézací modul používat odkazy na stránky obsahující direktivu no-follow a indexovat stránky obsahující direktivu no-index.

### Než začnete

Chcete-li nastavit použití direktiv no-follow a no-index webovým prolézacím modulem, musíte být přihlášení jako administrátor podnikového vyhledávání. Zadané direktivy mají přednost před direktivami obsaženými v prolézáných stránkách.

### Postup

Chcete-li přepsat direktivy no-follow a no-index, postupujte takto:

1. Z konzoly pro správu podnikového vyhledávání spusíte monitorování kolekce, pro kterou chcete konfigurovat pravidla, a zastavte webový prolézací modul.
2. Přihlašte se k prolézacímu serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
3. Přejděte do konfiguračního adresáře prolézacího modulu, který chcete konfigurovat, kde *ID\_relace\_prolézacího\_modulu* je ID přiřazené k relaci prolézacího modulu systémem podnikového vyhledávání. Příklad:  
`ES_NODE_ROOT/master_config/col_56092.WEB_88534`
4. Vytvořte nebo otevřete pro úpravy soubor s názvem `followindex.rules`.
5. Zadejte pravidla pro prolézací modul v následujícím formátu, kde *předpona\_URL* jsou první znaky webových serverů, jejichž prolézání nebo indexování chcete povolit nebo zakázat:

```

forbid follow předpona_URL
allow follow předpona_URL
forbid index předpona_URL
allow index předpona_URL

```

- Uložte soubor a zavřete jej.
- Z konzoly pro správu podnikového vyhledávání znovu spusťte webový prolézací modul, který jste zastavili.

## Konfigurování data používaného webovým prolézacím modulem pro prolézané dokumenty

V konfiguračním souboru můžete zadat volbu určující datum, které webový prolézací modul používá jako datum prolézaného dokumentu.

### Než začnete

Ve výchozím nastavení webový prolézací modul nastaví pole Datum v prolezených dokumentech na datum, kdy byl daný dokument prolezen. Pokud chcete, můžete webový prolézací modul konfigurovat tak, aby toto datum nastavil na datum a čas poslední úpravy vrácené webovým serverem.

Datum poslední úpravy nemusí být k dispozici pro všechny dokumenty. Jestliže prolézací modul konfiguruje tak, aby používal tuto hodnotu, a webový server pro daný dokument nevrátí datum poslední úpravy, nastaví prolézací modul do pole Datum datum a čas, kdy byl daný dokument prolezen.

*Tabulka 4. Způsob, jakým webový prolézací modul konfiguruje pole Datum pro prolezené dokumenty*

Konfigurace prolézacího modulu	Data vrácená webovým serverem	Pole metadat Datum pro prolezené dokumenty
Výchozí konfigurace	Webový server vrací pro prolezený dokument hodnotu data poslední úpravy	Datum a čas prolezení
	Webový server nevrací pro prolezený dokument hodnotu data poslední úpravy	Datum a čas prolezení
Konfigurace pro použití data poslední úpravy	Webový server vrací pro prolezený dokument hodnotu data poslední úpravy	Datum poslední úpravy vrácené webovým serverem
	Webový server nevrací pro prolezený dokument hodnotu data poslední úpravy	Datum a čas prolezení

Chcete-li zadat, které datum má webový prolézací modul používat pro prolezené dokumenty, musíte být přihlášení jako administrátor podnikového vyhledávání.

### Postup

Chcete-li konfigurovat webový prolézací modul tak, aby datem prolezeného dokumentu bylo datum a čas poslední úpravy vrácené webovým serverem, postupujte takto:

- Přihlaste se k prolézacímu serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.



2. Upravte soubor `crawl.properties` pro webový prolézací modul, který chcete konfigurovat. Parametr `ID_relace_prolézacího_modulu` reprezentuje ID, které bylo relaci prolézacího modulu přiřazeno systémem podnikového vyhledávání.

`ES_NODE_ROOT/master_config/ID_relace_prolézacího_modulu/crawl.properties`

Příklad:

```
/home/esadmin/master_config/col_00112.WEB_23344/crawl.properties  
C:\Program Files\IBM\es\esadmin\master_config\col_55667.WEB_78899\crawl.properties
```

3. Přidejte následující řádek a soubor uložte.  
`which_date=LastModified`
4. Restartujte webový prolézací modul z konzoly pro správu podnikového vyhledávání.

V případě, že webový server pro dokument prolezený tímto webovým prolézacím modulem vrátí datum poslední úpravy, použije prolézací modul vrácené datum a čas jako hodnotu pole Datum prolezeného dokumentu.

---

## Prolézací moduly Web Content Management

Chcete-li do kolekce podnikového vyhledávání zahrnout dokumenty produktu IBM Workplace Web Content Management, musíte konfigurovat prolézací modul Web Content Management.

### Konfigurace serveru WebSphere Portal

Pokud na server WebSphere Portal verze 6 nainstalujete produkt Workplace Web Content Management, můžete k prolézání webů spravovaných produktem Web Content Management použít prolézací modul Web Content Management. Volby pro prolézání těchto webů můžete nastavit odděleně od voleb, které zadáte pro portálové weby prolézané prolézacím modulem WebSphere Portal.

Před vytvořením prolézacího modulu Web Content Management musíte provést postup, který nastaví podnikové vyhledávání v portálu WebSphere. K nastavení prostředí podnikového vyhledávání spusíte skript (`wp6_install.sh` v systémech AIX, Linux a Solaris a `wp6_install.bat` v systémech Windows), který se instaluje na vyhledávací servery společně s produktem OmniFind Enterprise Edition.

### Formát adres URL

Prolézané adresy URL musíte zadávat v následujícím formátu:

```
protokol_http://hostitelský_název_portálu:číslo_portu/předpona_portálu  
/cesta_k_výchozímu_vyhledávacímu_servletu_WCM/searchseed?site=název_webu  
_WCM&lib=název_knihovny_WCM
```

Následující příklad ukazuje adresu URL webu ve výchozí instalační cestě produktu Workplace Web Content Management v prostředí portálu WebSphere:

```
http://portal.server.ibm.com:80/wps/wcmsearchseed/searchseed?site=SiteTest01  
&lib=Web+Content
```

Pokud název webu nebo knihovny obsahuje mezery, musíte je nahradit znakem plus (+). Název **Obsah webu** tedy například nahradíte textem **Obsah+webu**.

## Přehled konfigurace

Prolézací modul Web Content Management lze použít k prolézání libovolného počtu webů spravovaných produktem Web Content Manager. Při konfigurování prolézacího modulu zadáváte adresy URL webů, které chcete prolézat. Prolézací modul pak bude stahovat stránky propojené se zadaným webem.

Prolézané weby musí být dostupné pod stejným ID a heslem administrátora portálu WebSphere. Chcete-li prolézat weby s použitím jiných pověřovacích údajů, musíte k tomuto účelu konfigurovat samostatný prolézací modul Web Content Management.

Chcete-li vytvořit nebo změnit prolézací modul Web Content Management, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny dokumenty v prolézaném prostoru.
- Zadat adresy URL prolézaných webů a informace, které prolézacímu modulu umožní připojit se k těmto webům.

Pokud vytvoříte nebo upravíte prolézací modul, můžete otestovat, zda se tento modul dokáže úspěšně připojit k adresám URL, které mají být prolézány. Před spuštěním prolézacího modulu tak můžete zobrazit zprávy informující o tom, zda má prolézací modul přístup k dokumentům, které chcete prolézat.

- Nastavit zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení.

Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat.

**Důležité:** Chcete-li prohledávat zabezpečené stránky produktu Web Content Management, musíte vyhledávací požadavky zadávat pomocí vyhledávacího portletu pro podnikové vyhledávání v rámci portálu WebSphere. Vyhledávací požadavky odeslané z ukázkové vyhledávací aplikace ESSearchApplication nebudou mít k dispozici potřebné pověřovací údaje a nemohou tedy ověřit oprávnění uživatele pro přístup k dokumentům.

- Zadat informace umožňující prolézacímu modulu komunikaci se serverem proxy, pokud produkt Web Content Management při poskytování dokumentů využívá server proxy.
- Pokud k ochraně serveru WebSphere Portal a webů používáte jiný produkt (například IBMTivoli Access Manager WebSEAL nebo CA SiteMinder SSO Agent for PeopleSoft), zadejte údaje pověření pro jednotné přihlášení, které prolézacímu modulu umožní přístup k dokumentům na serveru.
- Zadat informace o souboru úložiště klíčů, aby prolézací modul mohl pro připojení k webům produktu Web Content Management použít protokol SSL (Secure Sockets Layer).
- Zadat jazyk a kódovou stránku prolézaných dokumentů.
- Nastavit volby pro prolézání a vyhledávání metadat v dokumentech produktu Web Content Management.
- Nastavit plán prolézání webů spravovaných produktem Web Content Management.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

#### **Související pojmy**

“Integrace s produktem WebSphere Portal” na stránce 323

#### **Související úlohy**

“Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 6” na stránce 330

---

## **Prolézací moduly WebSphere Portal**

Chcete-li do kolekce podnikového vyhledávání zahrnout stránky z webu IBM WebSphere Portal, musíte konfigurovat prolézací modul WebSphere Portal.

### **Konfigurace serveru WebSphere Portal**

Před vytvořením prolézacího modulu WebSphere Portal musíte spustit skript, který nastaví prostředí podnikového vyhledávání v portálu WebSphere. Různé verze portálu WebSphere vyžadují různé skripty. Tyto skripty jsou na vyhledávací servery instalovány při instalaci produktu OmniFind Enterprise Edition.

- U produktu WebSphere Portal verze 5.1.0 nebo vyšší se v systémech AIX, Linux a Solaris spouští skript `wp5_install.sh` a v systémech Windows skript `wp5_install.bat`.
- U produktu WebSphere Portal verze 6 se v systémech AIX, Linux a Solaris spouští skript `wp6_install.sh` a v systémech Windows skript `wp6_install.bat`.

#### **Tip:**

Podrobné příklady konfigurace zabezpečeného prolézacího modulu WebSphere Portal najdete ve scénáři pro střední organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

### **Přehled konfigurace**

Prolézací modul WebSphere Portal lze použít k prolézání jednoho webu WebSphere Portal. Při konfigurování prolézacího modulu zadáváte adresu URL webu portálu, který má být prolezen. Prolézací modul pak bude stahovat portlety a stránky propojené se zadanou adresou URL portálu. Chcete-li prolézat jiný portálový web, vytvořte nový prolézací modul.

Chcete-li vytvořit nebo změnit prolézací modul WebSphere Portal, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu stanovují způsob prolézání všech stránek na webu.
- Zadat adresu URL pro web portálu, který má být prolezen, a informace, které prolézacímu modulu umožní připojit se k příslušnému webu. Vzhledem k tomu, že tyto typy adres URL mohou být dlouhé a mohou obsahovat kódované znaky mimo sadu znaků ASCII, může být vhodné zkopírovat adresu URL ze serveru WebSphere Portal a vložit ji do konzoly pro správu podnikového vyhledávání.

Pokud vytvoříte nebo upravíte prolézací modul, můžete otestovat, zda se tento modul dokáže úspěšně připojit k adrese URL, která má být prolézána. Před spuštěním prolézacího modulu tak můžete zobrazit zprávy informující o tom, zda má prolézací modul přístup k dokumentům, které chcete prolézat.

- Nastavit zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit kontrolu přístupu založenou na uložených seznámech přístupových práv nebo prvcích zabezpečení.

Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat.

**Důležité:** Chcete-li prohledávat zabezpečené stránky portálu WebSphere, musíte vyhledávací požadavky zadávat pomocí vyhledávacího portletu pro podnikové vyhledávání v rámci portálu WebSphere. Vyhledávací požadavky odeslané z ukázkové vyhledávací aplikace ESSearchApplication nemají k dispozici potřebné pověřovací údaje a nemohou tedy ověřit oprávnění uživatele pro přístup k dokumentům.

- Zadat informace umožňující prolézacímu modulu komunikaci se serverem proxy, pokud portálový web WebSphere při poskytování stránek využívá server proxy.
- Pokud k ochraně serveru WebSphere Portal a webů používáte jiný produkt (například IBM Tivoli Access Manager WebSEAL nebo CA SiteMinder SSO Agent for PeopleSoft), zadejte údaje pověření pro jednotné přihlášení, které prolézacímu modulu umožní přístup k dokumentům na serveru.
- Zadat informace o souboru úložiště klíčů, aby prolézací modul mohl pro připojení k portálovému webu WebSphere použít protokol SSL (Secure Sockets Layer).
- Zadat jazyk a kódovou stránku prolézáných dokumentů.
- Nastavit volby pro prolézání a vyhledávání metadat v dokumentech portálu WebSphere.
- Nastavit plán prolézání dokumentů portálu WebSphere.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

#### **Související pojmy**

“Integrace s produktem WebSphere Portal” na stránce 323

#### **Související úlohy**

“Kopírování adresy URL pro prolézání ze systému WebSphere Portal”

“Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 6” na stránce 330

“Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 5.1” na stránce 325

## **Kopírování adresy URL pro prolézání ze systému WebSphere Portal**

Chcete-li snížit pravděpodobnost zadání nesprávné adresy URL, můžete při konfigurování prolézacího modulu WebSphere Portal nebo Web Content Management zkopírovat adresu URL webu, který chcete prolézat, a vložit ji do příslušného pole.

### **O této úloze**

Při vytvoření prolézacího modulu WebSphere Portal nebo Web Content Management určujete adresu URL webu umístěného na portálovém serveru WebSphere, který chcete prolézat.

Vzhledem k tomu, že adresy URL jsou dlouhé a obvykle obsahují kódované znaky mimo sadu znaků ASCII, může být vhodné zkopírovat adresu URL ze serveru WebSphere Portal a vložit ji do konzoly pro správu podnikového vyhledávání podle tohoto postupu.

### Postup

Chcete-li zadat adresu URL, kterou má prolézací modul prolézat, postupujte takto:

1. Pokud jste připraveni zadat prolézané adresy URL prostřednictvím konzoly pro správu podnikového vyhledávání, ověřte, že je server WebSphere Portal spuštěn, a poté se přihlašte k serveru WebSphere Portal jako administrátor.
2. Pokud používáte portál WebSphere verze 5.1, proveďte na serveru WebSphere Portal následující kroky:
  - a. V pravém horním rohu klepněte na položku **Správa**.
  - b. V navigační oblasti vlevo klepněte na volbu **Nastavení portálu** a poté klepněte na volbu **Správa hledání**.
  - c. Na stránce Správa kolekcí vyhledávání klepněte na volbu **PortalCollection** v oblasti Kolekce vyhledávání. Jsou-li k dispozici další kolekce, můžete vybrat jinou kolekci.
  - d. V části Zdroje obsahu v kolekci klepněte na volbu **Přidat zdroj obsahu**.
  - e. V rámečku **Typ zdroje prohledávání** klepněte na volbu **Portálové stránky**. Adresa URL webu je zobrazena v poli **Sbírat dokumenty propojené z této adresy URL**.
  - f. Zkopírujte adresu URL do schránky. Můžete například označit adresu URL, držet klávesu Ctrl stisknutou a stisknout klávesu Insert.
3. Pokud používáte portál WebSphere verze 6, proveďte na serveru WebSphere Portal následující kroky:
  - a. Klepněte na položku **Správa** v levém dolním rohu.
  - b. Klepněte na volbu **Správa vyhledávání** umístěnou vlevo.
  - c. Na stránce Správa vyhledávání klepněte na odkaz **Kolekce vyhledávání**.
  - d. Na stránce Správa vyhledávání klepněte v oblasti voleb vyhledávací služby na položku **Výchozí vyhledávací služba portálu**.
  - e. V tabulce Kolekce vyhledávání klepněte na kolekci **Obsah portálu**.
  - f. V tabulce Zdroje obsahu klepněte na ikonu Upravit zcela vpravo (ikona tužky) u pole **Zdroj obsahu portálu**.
  - g. Pro volbu **Typ zdroje obsahu** vyberte hodnotu **Portálové stránky**. Adresa URL webu je zobrazena v poli **Sbírat dokumenty propojené z této adresy URL**.
  - h. Zkopírujte adresu URL do schránky. Můžete například označit adresu URL, držet klávesu Ctrl stisknutou a stisknout klávesu Insert.
4. Vraťte se do konzole pro správu podnikového vyhledávání a vložte zkopírovanou adresu URL do pole adresy URL webu.

#### Související pojmy

“Prolézací moduly WebSphere Portal” na stránce 105

---

## Prolézací moduly souborového systému Windows

Chcete-li do kolekce podnikového vyhledávání zahrnout dokumenty uložené v souborových systémech Microsoft Windows, musíte konfigurovat prolézací modul Souborový systém Windows.

Prolézací modul Souborový systém Windows lze použít k prolézání libovolného počtu souborových systémů Windows. Při konfigurování prolézacího modulu vyberete lokální a vzdálené adresáře a podadresáře, které chcete prolézat.

Pokud instalujete prolézací server v systému AIX, Linux nebo Solaris, nemůžete tento server použít k prolézání zdrojů v souborových systémech Windows (v seznamu dostupných typů prolézacích modulů se neobjeví prolézací modul Souborový systém Windows).

**Tip:**

Podrobné příklady konfigurace zabezpečeného prolézacího modulu Souborový systém Windows najdete ve scénáři pro malé organizace v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

## **Prolézání sdílených síťových adresářů**

Prolézací modul Souborový systém Windows při prolézání dokumentů využívá oprávnění ke čtení nastavené pro administrátora podnikového vyhledávání. Administrátor je účet služeb produktu OmniFind Enterprise Edition.

Pro prolézané adresáře můžete nastavit jméno uživatele a heslo. Toto jméno uživatele a heslo však bude použito pouze pro připojení ke sdíleným síťovým adresářům. Prolézací modul prolézá soubory v souladu s oprávněními ke čtení, která jsou zadanému uživateli udělena pro sdílené síťové adresáře, nikoli pro lokální jednotky.

Připojení k síťovým adresářům zůstává aktivní, dokud nerestartujete službu OmniFind Enterprise Edition. Po navázání připojení je možné získat přístup k adresáři s nesprávným jménem uživatele a heslem. Přístup k tomuto připojení však mají pouze zjišťovací a prolézací relace Souborový systém Windows řízené systémem podnikového vyhledávání. Abyste předešli případným bezpečnostním rizikům, zkontroluje, zda jsou správně nastavena oprávnění pro účet administrátora podnikového vyhledávání (se kterým jsou spuštěny služby produktu OmniFind Enterprise Edition).

Chcete-li se v budoucnosti vyvarovat problémů s připojováním k síťovému adresáři, zadejte pro stejný síťový adresář stejné jméno uživatele a heslo. Zadáte-li nesprávné jméno uživatele a heslo a restartujete službu OmniFind Enterprise Edition, v prolézacím modulu Souborový systém Windows může dojít k selhání při pokusu o prolézání, protože se bude pokoušet o připojení k adresáři s nesprávnými pověřovacími údaji. Prolézání může uspět pouze za předpokladu, že síťové připojení naváže jiný prolézací modul Souborový systém Windows, který používá správné jméno uživatele a heslo.

## **Přehled konfigurace**

Chcete-li vytvořit nebo změnit prolézací modul Souborový systém Windows, přihlašte se ke konzole pro správu podnikového vyhledávání. Musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce kolekci, která vlastní prolézací modul.

Po vytvoření prolézacího modulu vám průvodce pomůže s následujícími úlohami:

- Určit vlastnosti, které řídí způsob, jakým prolézací modul pracuje a využívá systémové prostředky. Vlastnosti prolézacího modulu určují, jakým způsobem bude tento prolézací modul zpracovávat všechny podadresáře v prolézaném prostoru.
- Nastavit časový plán pro prolézání souborových systémů.
- Vybrat prolézané podadresáře.

Můžete určit, kolik úrovní podadresářů má prolézací modul prolézat. Při prolézání vzdálených souborových systémů zadáte také ID uživatele a heslo, které prolézacímu modulu umožní přístup k datům.

- Zadat volby umožňující prohledávání dokumentů v podadresářích. Z prolézaného prostoru můžete například vyloučit některé typy dokumentů nebo zadat ID uživatele a heslo umožňující prolézacímu modulu přístup k souborům v určitém podadresáři.
- Konfigurovat zabezpečení na úrovni dokumentu. Pokud bylo při vytváření kolekce povoleno zabezpečení, prolézací modul může přiřadit k dokumentům v indexu údaje o zabezpečení. Tato data umožňují vynutit u vyhledávacích aplikací kontrolu přístupu založenou na uložených seznamech přístupových práv nebo prvcích zabezpečení. Můžete rovněž vybrat volbu aktivující ověřování pověřovacích údajů uživatele při zadání dotazu uživatelem. V tomto případě systém neporovnává pověřovací údaje uživatele s indexovanými daty zabezpečení, ale s aktuálními seznamy přístupových práv spravovanými přímo původním zdrojem dat. Chcete-li vynutit zabezpečení na úrovni dokumentu, musíte ověřit, že jsou na prolézacím serveru správně konfigurovány informace o účtu uživatele a domény.

Chcete-li se dozvědět další informace o polích průvodce a způsobu, jak prolézacímu modulu poskytnout informace potřebné k prolézání dat, klepněte při vytváření prolézacího modulu na tlačítko **Nápověda**.

#### **Související pojmy**

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows” na stránce 271

“Bezpečné vyhledávání v důvěryhodných doménách systému Windows” na stránce 272

---

## **Konfigurování podpory pro aplikace modulu pro příjem dat**

Podnikové vyhledávání lze rozšířit vytvořením externího prolézacího modulu prostřednictvím rozhraní API modulu pro příjem dat. Vlastní aplikace modulu pro příjem dat mohou přidávat data do kolekce, odebírat data z kolekce nebo vydávat prolézacímu modulu pokyny k návštěvě či opětné návštěvě adres URL.

### **Než začnete**

Chcete-li konfigurovat aplikace modulu pro příjem dat, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

**Důležité:** Modul pro příjem dat nebude v budoucích verzích podporován. Při vývoji klientských aplikací pro podnikové vyhledávání používejte místo něj rozhraní API pro hledání a indexování (SI-API). Následující informace jsou určeny uživatelům, kteří dříve aplikace využívající modul pro příjem dat vytvořili.

### **O této úloze**




Klientská aplikace modulu pro příjem dat umožňuje prolézat typy zdrojů dat, které nelze prolézat výchozími prolézacími moduly pro podnikové vyhledávání. Než použijete aplikaci modulu pro příjem dat, je nutné konfigurovat pověření, která aplikaci umožní získat přístup ke kolekcím a aktualizovat je.

Při připojení klientské aplikace modulu pro příjem dat k modulu pro příjem dat musí modul klientské aplikace předat ID a heslo aplikace klienta a ID kolekce, která má být aktualizována. Tyto informace musí odpovídat informacím, které pro aplikaci konfiguruje prostřednictvím konzoly pro správu.

Modul pro příjem dat se spouští automaticky současně se systémem podnikového vyhledávání. Pokud jste změnili číslo portu po konfiguraci aplikace v konzole pro správu, je nutné modul pro příjem dat restartovat.

## Postup

Konfigurování aplikací modulu pro příjem dat:

1. Klepnutím na volbu **System** otevřete pohled System.
  2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
  3. Na stránce Modul pro příjem dat klepněte na volbu **Konfigurovat aplikace modulu pro příjem dat**.
  4. Na stránce Aplikace modulu pro příjem dat určete počet podprocesů, které může modul pro příjem dat vytvořit pro zpracování požadavků z klientských aplikací, a číslo portu, na kterém modul pro příjem dat přijímá požadavky. Určete také maximální počet dokumentů (pro jednu kolekci), které mohou být umístěny v dočasném úložišti, dokud je analyzátor nezačne analyzovat.
  5. Po klepnutí na volbu **Přidat aplikaci modulu pro příjem dat** můžete přidat informace o klientské aplikaci.
  6. Na stránce Přidat aplikaci modulu pro příjem dat zadejte ověřovací informace, které umožní klientským aplikacím modulu pro příjem dat získat přístup ke kolekcím podnikového vyhledávání. Identifikátory klientů modulu pro příjem dat musí být v rámci systému podnikového vyhledávání jedinečné.
  7. Vyberte kolekce, které může aplikace modulu pro příjem dat aktualizovat:
    - Pokud chcete, aby aplikace aktualizovala všechny kolekce, vyberte volbu **Všechny kolekce**.
    - Pokud chcete, aby aplikace aktualizovala pouze určené kolekce, vyberte volbu **Specifické kolekce**.  
Po výběru této volby se zobrazí seznam názvů kolekcí. U každé kolekce, kterou aplikace může aktualizovat, zaškrtněte políčko **Vybrat**.
  8. Klepněte na tlačítko **OK**.
  9. Pokud jste změnilí číslo portu modulu pro příjem dat nebo počet dokumentů, které mohou být uloženy v dočasném úložišti, restartujte modul pro příjem dat:
    - a. Chcete-li přejít do pohledu pro monitorování systému, klepněte na volbu  **Monitor**.
    - b. Na stránce Modul pro příjem dat klepněte na volbu  **Restartovat**.
- Související úlohy**  
“Monitorování modulu pro příjem dat” na stránce 301

---

## Vlastní moduly plug-in pro prolézací moduly

Při konfigurování vlastností pro prolézací moduly můžete určit třídu Java, kterou chcete použít k vynucení řízení přístupu na úrovni dokumentů. Pomocí třídy Java můžete také aktualizovat index přidáním, změnou nebo odebráním metadat a obsahu dokumentů. Vytvořením modulu plug-in můžete také rozšířit funkce prolézacího modulu při prolézání archivních souborů.

Modul plug-in obsahuje třídu Java, která je volána pro každý dokument zpracováváný prolézacím modulem. Tříde Java je předán identifikátor dokumentu (URI) z indexu podnikového vyhledávání, prvky zabezpečení, metadatum a obsah dokumentu. Třída může vrátit nové nebo změněné prvky zabezpečení, metadatum a obsah nebo odebrat prvky zabezpečení, metadatum a obsah.

Po prvním prolezení všech dokumentů v prolézaném prostoru je modul plug-in volán pouze pro nové nebo upravené dokumenty. Chcete-li změnit prvky zabezpečení, metadatum nebo



obsah dokumentů, které jsou obsaženy v indexu podnikového vyhledávání, ale nebyly aktualizovány v původním zdroji dat, spusťte úplné prolézání všech dokumentů v prolézaném prostoru a poté znovu sestavte hlavní index.

## **Moduly plug-in pro vynucení zabezpečení**

Zabezpečení na úrovni dokumentu je vynuceno prostřednictvím přiřazení jednoho nebo více prvků zabezpečení (řetězce oddělované čárkami) ke každému dokumentu, který prolézací modul zpracovává. Jako prvky zabezpečení se běžně používají identifikátory skupin.

Standardně je ke každému dokumentu přiřazen veřejný prvek, který dokument zpřístupní všem. Veřejný prvek lze nahradit hodnotou zadanou administrátorem nebo hodnotou, která byla extrahována z pole v prolézaném dokumentu.

Modul plug-in umožňuje používat k určování hodnoty prvků zabezpečení pro prolézané dokumenty vlastní obchodní pravidla. Prvky zabezpečení přiřazené k jednotlivým dokumentům se ukládají do indexu. Používají se k filtrování dokumentů, které odpovídají prvkům zabezpečení, a zaručují, že ve výsledcích vyhledávání budou vráceny pouze dokumenty, k jejichž prohlížení má uživatel oprávnění.

## **Moduly plug-in pro přidávání, úpravy a odebírání metadat**

Pro všechny prolézané dokumenty se vytvářejí metadata dokumentů, například datum poslední změny dokumentu. Modul plug-in prolézacího modulu vám umožňuje použít k určení hodnoty metadat, která mají být indexována pro jednotlivé dokumenty, vlastní obchodní pravidla.

Metadata jsou vytvářena jako dvojice název-hodnota. K prohledávání metadat mohou uživatelé použít dotaz pro volný text nebo dotaz s uvedením názvu pole metadat.

## **Moduly plug-in pro přidávání, úpravy a odebírání obsahu dokumentů**

Obsah dokumentů je tvořen těmi částmi dokumentu, které lze prohledávat a které se mohou stát součástí dynamického shrnutí dokumentu ve výsledcích vyhledávání. Modul plug-in prolézacího modulu vám umožňuje použít k určení obsahu, který má být indexován pro jednotlivé dokumenty, vlastní obchodní pravidla.

## **Moduly plug-in webového prolézacího modulu**

Prostřednictvím rozhraní API pro webový prolézací modul můžete řídit způsob prolézání dokumentů a způsob jejich přípravy na analýzu. Můžete například do hlavičky požadavků HTTP přidat pole, které bude použito, když prolézací modul odešle požadavek na dokument. Po prolezení dokumentu a před jeho analýzou a převodem na prvky můžete změnit obsah, prvky zabezpečení a metadata. Můžete rovněž zastavit odesílání dokumentu do modulu analýzy.

Výhody a nevýhody ověřování založeného na použití formulářů a ukázkový program, který můžete přizpůsobit pro svůj vlastní modul plug-in webového prolézacího modulu, najdete v tématu <http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani>.

## **Moduly plug-in pro archivní soubory**

Vytvořením modulu plug-in můžete rozšířit funkce prolézacích modulů a umožnit podporu prolézání jiných formátů archivních souborů než formátů ZIP a TAR. Můžete například vytvořit modul plug-in podporující prolézání dokumentů ve formátu LZH.

## Nechráněný režim

Při konfiguraci prolézacího modulu, který není webový, lze vybrat volbu spuštění modulu plug-in v nechráněném režimu. V tomto režimu je proces modulu plug-in spuštěn v rámci procesu prolézacího modulu, což může zvýšit výkon prolézacího modulu.

**Důležité:** Pokud modul plug-in narazí během spuštění v tomto režimu na problém, ze kterého se nelze zotavit, může být ukončen proces prolézacího modulu.

### Související pojmy



Moduly plug-in prolézacího modulu

---

## Podpora prolézání souborů archivu

Prolézací moduly podnikového vyhledávání mohou extrahovat soubory z archivních souborů (například ze souborů ZIP a TAR) a umožnit tak indexování a prohledávání jednotlivých souborů v těchto archivech.

### Podporované formáty archivních souborů

Podporovány jsou následující formáty archivních souborů:

*Tabulka 5. Formáty archivních souborů podporované prolézacími moduly podnikového vyhledávání*

Přípona souboru	Typ MIME	Typ dat	Poznámky
.zip, .ZIP	application/zip	zip	<ul style="list-style-type: none"><li>• Závisí na možnostech balíku java.util.zip</li><li>• Podporuje kompresi typu deflate (metoda 8):<ul style="list-style-type: none"><li>– Nepodporuje šifrování souborů</li><li>– Nepodporuje standard zip64</li></ul></li></ul>
.tar	application/tar	tar	Podporované formáty tar: <ul style="list-style-type: none"><li>• GNU tar 1.13</li><li>• POSIX 1003.1-1998 (ustar)</li><li>• POSIX 1003.1-2001 (pax)</li></ul>
.tar.gz, .tgz	nedefinován	tgz	Závisí na možnostech balíku java.util.zip

### Omezení a pokyny

Pro soubory extrahované z archivních souborů není k dispozici automatická detekce kódové stránky. Při extrakci těchto souborů použijte prolézací modul nastavení kódové stránky převzaté z konfigurace pro prostý text a neznámé typy dokumentů. Pokud nastavíte konfiguraci jazyka a kódové stránky prolézacího modulu pomocí konzoly pro správu podnikového vyhledávání, určujete tím kódovou stránku, kterou má prolézací modul použít pro dokumenty obsahující prostý text a pro dokumenty, jejichž kódovou stránku nelze detekovat automaticky.

Při zjišťování, které soubory obsažené v archivním souboru je třeba prolézt znovu, používá prolézací modul datum poslední úpravy jednotlivých souborů uvedené v datech záhlaví archivní položky. Při monitorování prolézacího modulu jsou do zobrazovaných statistik prolézaných dokumentů včetně statistiky vložených, aktualizovaných a odstraněných dokumentů zahrnuty informace o souborech extrahovaných z archivních souborů.

Chcete-li prolézacímu modulu umožnit prolézání archivních souborů dalších formátů, například souborů LZH, musíte vytvořit modul plug-in prolézacího modulu a konfigurovat

prolézací modul tak, aby tento modul plug-in používal.

## Migrace

Chcete-li umožnit prolézání archivních souborů prolézacím modulům existujícím před instalací produktu OmniFind Enterprise Edition verze 8.4, musíte upravit prolézací prostor prolézacího modulu. Z nastavení vyloučených typů MIME a vyloučených přípon souborů odeberte všechny odkazy na soubory typu .zip, .tar, .tgz a .gz.

**Důležité:** Pokud změníte toto nastavení pro prolézací modul Content Edition nebo DB2 Content Manager a chcete, aby se změny uplatnily, musíte provést nové prolezení všech dokumentů.

---

## Formáty identifikátorů URI v indexu podnikového vyhledávání

Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

Identifikátory URI a jejich vzory můžete zadávat při konfigurování kategorií, oborů a rychlých odkazů pro kolekci. Identifikátor URI je nutné zadat také při odebírání dokumentů z indexu a při zobrazení podrobných stavových informací o konkrétním identifikátoru.

Prohledáním kolekce můžete zjistit identifikátory URI a vzory URI pro dokument. Klepnutím na identifikátory URI ve výsledcích vyhledávání můžete načíst dokumenty, které vás zajímají. Identifikátor URI můžete z výsledků vyhledávání zkopírovat a použít jej v konzole pro správu podnikového vyhledávání. Zadáním vzoru identifikátoru URI můžete například automaticky asociovat dokumenty, které odpovídají danému vzoru identifikátoru URI, s kategorií podnikového vyhledávání.

### Soubory archivu

Formát identifikátoru URI pro dokumenty extrahované z archivního souboru (např. soubory ZIP či TAR) a následně prolezené:

```
původní_URI(?|&)ArchiveEntry=název_položky(&ArchiveEntry=název_položky)
```

#### Parametry

*původní\_URI*

Umístění archivního souboru ve zdroji dat.

*název\_položky*

Název archivní položky v archivním souboru, zakódovaný ve formě adresy URL.

#### Příklady

```
file:///d:/Archiv1.zip  
file:///d:/Archiv1.zip?ArchiveEntry=Složka1/PowerPoint.ppt  
file:///d:/Archiv1.zip?ArchiveEntry=Složka2/Text.txt
```

### Prolézací moduly Content Edition

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Content Edition v režimu přístupu k serveru je následující:

```
vbr://název_serveru/systémové_ID_úložiště/trvalé_ID_úložiště  
/ID_položky/ID_verze  
/typ_položky/?[Page=číslo_stránky&] vlastnosti_JNDI
```

Formát identifikátorů URI pro dokumenty prolézané prolézacím modulem Content Edition v režimu přímého přístupu je následující:

```
vbr://systémové_ID_úložiště/trvalé_ID_úložiště  
/ID_položky/ID_verze  
/typ_položky/[?Page=číslo_stránky]
```

### Parametry

Pro všechna pole je použito kódování adres URL.

*název\_serveru*

Název serveru WebSphere Information Integrator Content Edition.

*systémové\_ID\_úložiště*

Systémový identifikátor úložiště.

*trvalé\_ID\_úložiště*

Trvalý identifikátor úložiště.

*ID\_položky*

Identifikátor položky.

*ID\_verze*

Identifikátor verze. Pokud je identifikátor verze prázdný, označuje tato hodnota nejnovější verzi dokumentu.

*typ\_položky*

Typ položky (CONTENT - obsah nebo FOLDER - složka).

*číslo\_stránky*

Číslo stránky.

*vlastnosti\_JNDI*

Vlastnosti JNDI aplikačního klienta J2EE. Existují dva typy vlastností:

#### **java.naming.factory.initial**

Název třídy aplikačního serveru používaný k vytvoření manipulátoru EJB.

#### **java.naming.provider.url**

Adresa URL služby názvů aplikačního serveru používaného k zadání požadavku na manipulátor EJB.

### Příklady

Documentum:

```
vbr://vbrsrv.ibm.com/Documentum/c06b/094e827780000302//CONTENT/?  
java.naming.provider.url=iiop%3A%2F%2Fmyvbr.ibm.com%3A2809&  
java.naming.factory.initial=com.ibm.websphere.naming.WsnInitContextFactory
```

FileNet PanagonCS:

```
vbr://vbrsrv.ibm.com/PanagonCS/4a4c/003671066//CONTENT/?Page=1&  
java.naming.provider.url=iiop%3A%2F%2Fmyvbr.ibm.com%3A2809&  
java.naming.factory.initial=com.ibm.websphere.naming.WsnInitContextFactory
```

## Prolézací moduly DB2

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem DB2:

```
db2://název_databáze/název_tabulky  
/jedinečný_identifikátor_názvu_sloupce_1/jedinečný_identifikátor_hodnoty_1  
[/jedinečný_identifikátor_názvu_sloupce_2/jedinečný_identifikátor_hodnoty_2/...  
/jedinečný_identifikátor_názvu_sloupce_N/jedinečný_identifikátor_hodnoty_N]
```

### Parametry:

Pro všechna pole je použito kódování adres URL.

*název\_databáze*

Interní název nebo alias databáze.

*název\_tabulky*

Název cílové tabulky včetně názvu schématu.

*název\_sloupce\_jedinečného\_identifikátoru\_1*

Název prvního sloupce jedinečného identifikátoru v tabulce.

*hodnota\_jedinečného\_identifikátoru\_1*

Hodnota v prvním sloupci jedinečného identifikátoru.

*název\_sloupce\_jedinečného\_identifikátoru\_N*

Název *ntého* sloupce jedinečného identifikátoru v tabulce.

*hodnota\_jedinečného\_identifikátoru\_N*

Hodnota v *ntém* sloupci jedinečného identifikátoru v tabulce.

### Příklady

Lokální katalogizovaná databáze:

db2://LOCALDB/SCHEMA1.TABLE1/MODEL/ThinkPadA20

Vzdálená nekatalogizovaná databáze:

db2://server.spolecnost.com:50001/REMOTEDB/SCHEMA2.TABLE2/NAME/DAVID

## Prolézací moduly DB2 Content Manager

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem DB2 Content Manager:

cm://název\_serveru/název\_typu\_položky/PID

### Parametry

Pro parametr *PID* je použito kódování adres URL.

*název\_serveru*

Název serveru knihovny IBM DB2 Content Manager.

*název\_typu\_položky*

Název cílového typu položky.

*PID* Trvalý identifikátor DB2 Content Manager.

### Příklad

cm://cmsrvctg/ITEMTYPE1/92+3+ICM8+icm1sdb12+ITEMTYPE159+26+A1001001A  
03F27B94411D1831718+A03F27B+94411D183171+14+1018

## Prolézací moduly Domino Document Manager

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Domino Document Manager:

dominodoc://název\_serveru:číslo\_portu/ID\_repliky\_databáze/název\_databáze\_s\_cestou  
/univerzální\_ID\_pohledu/univerzální\_ID\_dokumentu  
/?AttNo=číslo\_přílohy&AttName=název\_souboru\_přílohy

### Parametry

Pro všechna pole je použito kódování adres URL.

*název\_serveru*

Název serveru Domino Document Manager.

*číslo\_portu*

Volitelné: Číslo portu pro server Domino Document Manager.

### *ID\_repliky\_databáze*

Identifikátor repliky databáze.

### *název\_databáze\_s\_cestou*

Cesta a název souboru databáze NSF dokumentu na cílovém serveru Domino Document Manager.

### *univerzální\_ID\_pohledu*

Univerzální identifikátor pohledu používaný pro prolézání dokumentů Domino Document Manager.

### *univerzální\_ID\_dokumentu*

Univerzální identifikátor dokumentu definovaný v prolézáném dokumentu.

### *číslo\_přílohy*

Volitelné: Pořadové číslo přílohy. Přílohy se číslují od nuly.

### *název\_souboru\_přílohy*

Volitelné: Původní název souboru přílohy.

## **Příklady**

Dokument Domino Document Manager:

```
dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
```

Příloha dokumentu:

```
dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226?AttNo=0&AttName=AttachedFile.doc
```

## **Prolézací moduly Exchange Server**

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Exchange Server:

```
exchange://cesta_OWA[?useSSL=true]
```

### **Parametry**

#### *cesta\_OWA*

Cesta OWA (Outlook Web Access) bez protokolu.

#### **useSSL=true**

Přidává se, pokud původní cesta OWA používá protokol HTTPS.

### **Příklady**

Tělo dokumentu:

```
exchange://exchangesvr.ibm.com/public/RootFolder1/Folder1/Document.EML
```

Příloha dokumentu:

```
exchange://exchangesvr.ibm.com/public/RootFolder1/Folder1/Document.EML/AttachedFile.doc
```

Podpora zabezpečení SSL:

```
exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/Document.EML?useSSL=true
```

## **Prolézací moduly Databáze JDBC**

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Databáze JDBC:

```
jdbc://URL_databáze/název_tabulky
/název_sloupce_jedinečného_identifikátoru_1/hodnota_jedinečného
_identifikátoru_1
/[název_sloupce_jedinečného_identifikátoru_2/hodnota_jedinečného
_identifikátoru_2
/.../název_sloupce_jedinečného_identifikátoru_N/hodnota_jedinečného
_identifikátoru_N]
```

### Parametry

Pro všechna pole je použito kódování adres URL.

*URL\_databáze*

Adresa URL databáze.

*název\_tabulky*

Název cílové tabulky včetně názvu schématu.

*název\_sloupce\_jedinečného\_identifikátoru\_1*

Název prvního sloupce jedinečného identifikátoru v tabulce.

*hodnota\_jedinečného\_identifikátoru\_1*

Hodnota v prvním sloupci jedinečného identifikátoru.

*název\_sloupce\_jedinečného\_identifikátoru\_N*

Název *n*tého sloupce jedinečného identifikátoru v tabulce.

*hodnota\_jedinečného\_identifikátoru\_N*

Hodnota v *n*tém sloupci jedinečného identifikátoru v tabulce.

### Příklady:

Databáze DB2:

```
jdbc:db2://host01.svl.ibm.com:50000/SAMPLE/DB2INST1.ORG/DEPTNUMB/51
```

Databáze Oracle:

```
jdbc:oracle:thin:@/host01.svl.ibm.com:1521:ora/SCOTT.EMP/EMPNO/7934
```

Databáze MS SQL Server 2000:

```
jdbc:microsoft:sqlserver://host01.svl.ibm.com:1433;
DatabaseName=Northwind/dbo.Region/RegionID/100
```

Databáze MS SQL Server 2005:

```
jdbc:sqlserver://host01.svl.ibm.com:1433;
DatabaseName=Northwind/dbo.Region/RegionID/100
```

## Prolézací moduly Notes

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Notes:

```
domino://název_serveru[:číslo_portu]/ID_repliky_databáze/název_databáze_s_cestou
/[univerzální_ID_pohledu]/univerzální_ID_dokumentu
[?AttNo=číslo_přílohy&AttName=název_souboru_přílohy]
```

### Parametry

Pro všechna pole je použito kódování adres URL.

*název\_serveru*

Název serveru Lotus Notes.

*číslo\_portu*

Číslo portu serveru Lotus Notes. Číslo portu je nepovinné.

*ID\_repliky\_databáze*

Identifikátor repliky databáze.

#### *název\_databáze\_s\_cestou*

Cesta a název souboru databáze NSF na cílovém serveru Lotus Notes.

#### *univerzální\_ID\_pohledu*

Univerzální identifikátor pohledu definovaný v cílové databázi. Toto ID se zadává jen tehdy, je-li vybrán dokument obsažený v pohledu nebo složce. Pokud neurčíte pohled nebo složku, kterou chcete prolézat (například tehdy, chcete-li prolézat všechny dokumenty v databázi), univerzální ID pohledu se nezadá.

#### *univerzální\_ID\_dokumentu*

Univerzální ID dokumentu definované v dokumentu zpracovávaném prolézacím modulem.

#### *číslo\_přílohy*

Pořadové číslo přílohy. Přílohy se číslijí od nuly. Číslo přílohy je nepovinné.

#### *název\_souboru\_přílohy*

Původní název souboru přílohy. Název souboru přílohy je nepovinný.

### **Příklady**

Dokument, který byl vybrán k prolézání v pohledu nebo složce:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf/  
8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
```

Dokument, který nebyl vybrán k prolézání v pohledu nebo složce:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf/  
0205F44FA3F45A9049256DB20042D226
```

Příloha dokumentu:

```
domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf/  
0205F44FA3F45A9049256DB20042D226?AttNo=0&AttName=AttachedFile.doc
```

### **Prolézací moduly QuickPlace**

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem QuickPlace:

```
quickplace://název_serveru:číslo_portu/ID_repliky_databáze/název_databáze_s_cestou  
/univerzální_ID_pohledu/univerzální_ID_dokumentu  
/?AttNo=číslo_přílohy&AttName=název_souboru_přílohy
```

#### **Parametry**

Pro všechna pole je použito kódování adres URL.

#### *název\_serveru*

Název serveru Lotus QuickPlace.

#### *číslo\_portu*

Volitelné: Číslo portu pro server QuickPlace.

#### *ID\_repliky\_databáze*

Identifikátor repliky databáze.

#### *název\_databáze\_s\_cestou*

Cesta a název souboru databáze NSF dokumentu na cílovém serveru QuickPlace.

#### *univerzální\_ID\_pohledu*

Univerzální identifikátor pohledu používaný pro prolézání dokumentů QuickPlace.

#### *univerzální\_ID\_dokumentu*

Univerzální identifikátor dokumentu definovaný v prolézaném dokumentu.



### *číslo\_přílohy*

Volitelné: Pořadové číslo přílohy. Přílohy se číslovají od nuly.

### *název\_souboru\_přílohy*

Volitelné: Původní název souboru přílohy.

## **Příklady**

Dokument:

```
quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace%5CPageLibrary4925704300021490.nsf/A7986FD2A9CD47090525670800167225/2B02B1DE3A82B2CE49257043001C2498
```

Příloha stránky:

```
quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace%5CPageLibrary4925704300021490.nsf/A7986FD2A9CD47090525670800167225/2B02B1DE3A82B2CE49257043001C2498?AttNo=0&AttName==QPCons3.ppt
```

## **Prolézací moduly Počáteční seznam**

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Počáteční seznam:

```
seedlist://URL_stránky?pageID=ID_stránky [&useSSL;=true]
```

### **Parametry**

Pro všechna pole je použito kódování adres URL.

#### *URL\_stránky*

Adresa URL dokumentu (jedinečná pro každý dokument).

#### *ID\_stránky*

Identifikátor objektu pro dokument.

#### **useSSL**

Pokud je použit protokol HTTPS, přidává se k identifikátoru URI parametr `&useSSL;=true`. V opačném případě je parametr `useSSL` vynechán.

### **Příklad**

Protokol HTTPS:

```
seedlist://quickrserver.ibm.com:10035/lotus/my poc?uri=dm:bec6090046f1cd52bc5cfc b06e9f4550&verb;=view&pageID;=N1FSZUR1MkJQNjZSMdZQMUMwM1FPNjZCQzY2SUw2SUhPNk1RQ0M2Uk80Nk9PNjVCRUM2UUs2TDFDMA==&useSSL;=true
```

## **Prolézací moduly Souborový systém UNIX**

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Souborový systém UNIX:

```
file:///název_adresáře/název_souboru
```

### **Parametry**

Pro všechna pole je použito kódování adres URL.

#### *název\_adresáře*

Absolutní cesta k adresáři.

#### *název\_souboru*

Název souboru.

### **Příklad**

```
file:///home/user/test.doc
```

## Prolézací moduly Web Content Management

Formát identifikátorů URI pro dokumenty produktu WebSphere Content Management zpracovávané prolézacím modulem Web Content Management:

wcm://URL\_stránky?pageID=ID\_stránky[&useSSL=true]

### Parametry

Pro všechna pole je použito kódování adres URL.

*URL\_stránky*

Adresa URL dokumentu (jedinečná pro každý dokument).

*ID\_stránky*

Identifikátor stránky.

### useSSL

Pokud je použit protokol HTTPS, přidává se k identifikátoru URI parametr useSSL=true. V opačném případě je parametr useSSL vynechán.

### Příklady

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
```

## Prolézací moduly WebSphere Portal: WebSphere Portal verze 5

Formát identifikátorů URI pro dokumenty produktu WebSphere Portal verze 5 zpracovávané prolézacím modulem WebSphere Portal:

wps://URL\_stránky?portletDefID=ID\_definice\_portletu&portletID=ID\_portletu  
&pageID=ID\_stránky[&useSSL=true]

### Parametry

Pro všechna pole je použito kódování adres URL.

*URL\_stránky*

Adresa URL dokumentu (jedinečná pro každý dokument).

*ID\_definice\_portletu*

Identifikátor definice portletu.

*ID\_portletu*

Identifikátor portletu..

*ID\_stránky*

Identifikátor stránky.

### useSSL

Pokud je použit protokol HTTPS, přidává se k identifikátoru URI parametr useSSL=true. V opačném případě je parametr useSSL vynechán.

### Příklady

Tělo dokumentu:

```
wps://wpserver.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
```

## Příklady

Podpora zabezpečení SSL:

```
wps://wpserver.ibm.com:9081/wps/myportal!/ut/p/kcxm1/04_Sj9SPykssy0x+
LKnPy1vM0Y_QjzKCN4g3cQbJgQi0-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8qHRj/delta
/base64xm1/L01DU1kvd0NrQUpORUeVnFBVR0VoQSEvN18wXzZPLzZFMF80RA!!
?portletDefID=7_0_A4&pageID=6_0_6J&useSSL=true
```

## Prolézací moduly WebSphere Portal: WebSphere Portal verze 6

Formát identifikátorů URI pro dokumenty produktu WebSphere Portal verze 6 zpracovávané prolézacím modulem WebSphere Portal:

```
wp6://URL_stránky?portletURL=URL_portletu?portletDefID=ID_definice_portletu
&pageID=ID_stránky[&useSSL=true]
```

### Parametry

Pro všechna pole je použito kódování adres URL.

*URL\_stránky*

Adresa URL dokumentu (jedinečná pro každý dokument).

*URL\_portletu*

Jedinečná adresa URL dokumentu.

*ID\_definice\_portletu*

Identifikátor definice portletu.

*ID\_stránky*

Identifikátor stránky.

**useSSL**

Pokud je použit protokol HTTPS, přidává se k identifikátoru URI parametr `useSSL=true`. V opačném případě je parametr `useSSL` vynechán.

## Příklady

Protokol HTTP:

```
wp6://wp6server.ibm.com:9081/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy
8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UESTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_
Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGh
gYWjobuhmaGDka0Bu5uHqFRpkaAAAEisaBQ!!
?portletUrl=/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59
gQwMLQ0P_IDMnAyNHA3f3UESTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdF
RUBTu-saA!!/d12/d0/Y2BiUZnBwMqsyabykYGBmS2tcDoDE40MebyRiU-woYGFoaG7oZu
hgZGjgbuZh6uXY4ihAYOMGQ7Z0CBDAwDcXPkM
&portletDefID=6_24LS10811G1F102A0G6HEJUR10
&pageID=3_24LS108110R6B02A0GGU94LN00
```

Protokol HTTPS:

```
wp6://wp6server.ibm.com:9444/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy
8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UESTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_
Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGh
gYWjobuhmaGDka0Bu5uHq5RtqaAAA50L41Q!!
?portletUrl=/wps/myportal!/ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59
gQwMLQ0P_IDMnAyNHA3f3UESTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdF
RUBTu-saA!!/d12/d0/Y2BiUZnBwMqsyabykYGBmS2tcDoDE40MebyRiU-woYGFoaG7oZu
hgZGjgbuZh6uXWZihAYOMGQ5Z31BDAAwAk73P2
&portletDefID=6_24LS10811G1F102A0G6HEJMU10
&pageID=3_24LS108110R6B02A0GGU94T410&useSSL=true
```

## Prolézací moduly Souborový systém Windows

Formáty identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Souborový systém Windows:

file:///název\_adresáře/název\_souboru  
file:///název\_síťové\_složky/název\_adresáře/název\_souboru

### **Parametry**

Pro všechna pole je použito kódování adres URL.

*název\_adresáře*

Absolutní cesta k adresáři.

*název\_souboru*

Název souboru.

*název\_síťové\_složky*

Pouze u dokumentů umístěných na vzdálených serverech: název sdílené složky v síti Windows.

### **Příklady**

Lokální souborový systém:

file:///d:/adresář/test.doc

Síťový souborový systém:

file:///server.ibm.com/adresář/soubor.doc

---

## Administrace analyzátorů

Chcete-li zlepšit dostupnost dokumentů, můžete zadat volby určující způsob analýzy a kategorizace dokumentů a metadat před jejich přidáním do indexu podnikového vyhledávání.

Pro analýzu obsahu dokumentů a optimalizaci možností načítání informací můžete nastavit následující volby:

### **Konfigurování voleb pro analýzu dokumentů v čínštině, japonštině a korejštině**

Pro analýzu dokumentů napsaných v čínštině, japonštině a korejštině můžete nastavit volby použití n-gramové segmentace. Z prázdného prostoru v čínských a japonských dokumentech můžete rovněž odebrat znaky nového řádku.

### **Povolení nativního vyhledávání XML**

Pokud kolekce obsahuje dokumenty XML, můžete povolit jejich prohledávání pomocí nativní syntaxe dotazů XML, například pomocí fragmentů XPath a XML. Nativní vyhledávání XML umožňuje uživatelům zadávat dotazy založené na vztazích mezi různými prvky XML.

### **Konfigurování kategorií**

Dokumenty s podobnými identifikátory URI nebo dokumenty obsahující určitá slova lze seskupit do kategorií. Když uživatelé prohledávají kolekci, mohou omezit výsledky vyhledávání pouze na dokumenty patřící do určitých kategorií.

### **Konfigurování vyhledávacích polí**

Prvky obsažené v dokumentech XML lze namapovat na vyhledávací pole indexu. Podobně lze na vyhledávací pole namapovat také prvky metadat v dokumentech HTML. Vytvořením vyhledávacích polí v indexu podnikového vyhledávání umožníte uživatelům zadávat dotazy na konkrétní části dokumentů XML a HTML a zvýšíte tak přesnost výsledků vyhledávání.





### **Konfigurování voleb zpracování textu**

Pokud byly do systému podnikového vyhledávání přidány vlastní stroje pro analýzu textu, můžete určit, který z nich má být použit pro jednotlivé kolekce. Po přiřazení analytického stroje ke kolekci můžete zadat volby mapování obsahu, podle nichž bude prováděna lingvistická analýza a vytvářeny anotace dokumentů. Můžete také určit způsob, jakým mají být výsledky analýzy mapovány na index podnikového vyhledávání nebo na databázové tabulky JDBC.

### **Mapování polí na třídy upřednostnění**

Můžete určit, že dokumenty s poli odpovídajícími výrazům dotazu mají mít ve výsledcích vyhledávání vyšší hodnocení než jiné dokumenty, které odpovídají výrazům dotazu. Při mapování polí na třídy upřednostnění můžete určit, která pole s obsahem a pole metadat mají být upřednostňována. Můžete rovněž konfigurovat skóre, která bude každá třída upřednostnění používat pro ohodnocení důležitosti dokumentů.

### **Související pojmy**

-  Lingvistická podpora pro sémantické vyhledávání
-  Analýza textu integrovaná v podnikovém vyhledávání
-  Základní koncepce používané při zpracování analýzy textu
-  Aplikace sémantického vyhledávání



Výrazy dotazu sémantického vyhledávání

“Podpora jazyků a kódových stránek” na stránce 157

“Detekce formátu dokumentu” na stránce 143

---

## Práce s kategoriemi

Kategorie umožňují seskupit dokumenty se společnými charakteristikami a prohledávat nebo načítat pouze dokumenty, které splňují kritéria platná pro členy dané skupiny.

Pokud přiřadíte dokumenty ke kategoriím a vaše vyhledávací aplikace tuto funkci podporuje, mohou uživatelé zadat název kategorie a prohledat pouze příslušnou podmnožinu kolekce. Při prohledávání celé kolekce mohou uživatelé zpřesnit výsledky vyhledávání a procházet pouze dokumenty, které patří do stejné kategorie jako některý z výsledných dokumentů.

Při konfigurování kategorie zadáváte pravidla, která prolézací modul informují o tom, jak má přiřazovat dokumenty ke kategoriím. Dokumenty můžete seskupit podle toho, zda sdílejí určité části vzoru URI, nebo podle toho, zda obsahují určitá data (například dokumenty obsahující, nebo neobsahující určitá slova a slovní spojení).

K vytváření a administraci kategorií se používá konzola pro správu podnikového vyhledávání:

- Při vytvoření kolekce zvolíte typ kategorizace. Můžete zvolit práci bez kategorií nebo použití kategorií založených na pravidlech.
- Při konfigurování pravidel analýzy pro kolekci můžete v případě nutnosti změnit typ kategorizace. Jestliže změníte typ kategorizace po zpracování dokumentů prolézacím modulem a sestavení indexu, bude kvalita vyhledávání snížena, dokud neprovedete nové prolezení všech dokumentů a nové sestavení hlavního indexu.
- Pokud se rozhodnete používat kategorie založené na pravidlech, můžete pomocí konzoly pro správu spravovat strom kategorií, kategorie a pravidla kategorií. Jestliže po zpracování dokumentů prolézacím modulem a sestavení indexu změníte kategorie nebo pravidla kategorií, bude kvalita vyhledávání snížena, dokud neprovedete nové prolezení všech dokumentů a nové sestavení hlavního indexu.

## Kategorie založené na pravidlech

Přiřazení dokumentů ke kategoriím v kolekci podnikového vyhledávání lze určovat konfigurováním pravidel.

Pravidla kategorií lze vytvořit pro kolekce vytvořené v prostředí podnikového vyhledávání a pro kategorie založené na pravidlech, které importujete z kolekci produktu IBM WebSphere Portal. Chcete-li konfigurovat pravidla pro kategorizaci dokumentů, musíte při vytvoření kolekce nebo při zadávání voleb analýzy pro kolekci určit, že chcete používat kategorie založené na pravidlech.

Analýzátor používá zadaná pravidla pro přiřazení dokumentů k jedné nebo více kategoriím:

- Pokud dokument splňuje podmínky alespoň jednoho pravidla kategorie, analyzátor jej přiřadí k této kategorii.
- Pokud dokument splňuje podmínky alespoň jednoho pravidla u více kategorií, analyzátor jej přiřadí ke všem těmto kategoriím.
- Pokud dokument nespĺňuje podmínky žádného z pravidel všech kategorií, analyzátor jej nepřijadí k žádné kategorii. Uživatelé mohou tento dokument vyhledat a načíst při prohledávání kolekce, nemohou jej však načíst, vyberou-li některou kategorii.

Při administraci stromu kategorií (taxonomie) kolekce určujete, na které místo v hierarchii kategorií chcete novou kategorii přidat. Strom kategorií slouží také k výběru kategorie, kterou

chcete upravit, a k následnému přidání pravidel pro kategorizaci dokumentů, odstranění pravidel nebo změně obsahu jednotlivých pravidel.

Při konfigurování pravidla pro kategorizaci dokumentů určujete, zda má podnikové vyhledávání rozhodovat o přiřazení dokumentu ke kategoriím podle jeho identifikátoru URI nebo podle obsahu.

## Pravidla vzorů URI

Pravidla URI se vztahují na identifikátory URI dokumentů. Zadáte-li část identifikátoru URI (vzor), podmínce pravidla vyhoví všechny dokumenty, v jejichž identifikátoru URI se zadaný vzor vyskytuje.

Zadáte-li například text pravidla `/hr/`, první z následujících identifikátorů URI bude podmínku pravidla splňovat a druhý nikoli:

```
file:///corporate/hr/medicalform.doc  
http://company.com/human_resources/medicalform.htm
```

Vzhledem k tomu, že všechny identifikátory URI jsou zpracovávány jako vzory, systém ignoruje hvězdičky, které zadáte jako zástupné znaky na začátku nebo konci vzoru. Vzory `*/hr/*` a `/hr/` například vrátí stejnou sadu identifikátorů URI.

V pravidlech vzorů URI se nerozlišují velká a malá písmena. Pokud identifikátor URI obsahuje mezery, musí vzor URI respektovat pravidla podnikového vyhledávání pro kódování identifikátorů URI. V následujícím příkladu jsou uvedeny správné a nesprávné způsoby zadání identifikátoru URI pro cestu v souborovém systému Windows:

Nesprávný identifikátor URI:

```
file:///c:/program files/
```

Správný identifikátor URI: `file:///c:/program+files/`

## Pravidla obsahu dokumentu

Pravidla obsahu dokumentu zapisujete ve stejném formátu jako dotaz. Dokument splňuje podmínku pravidla, pokud je platný z hlediska zadaného dotazu. Při konfigurování pravidla zadáváte slova a fráze, které se v dokumentu musí nebo nesmí vyskytovat, a vybíráte jazyk, podle kterého bude pravidlo převádět slova na základní tvary.

Následující pravidlo například určuje, že podmínku splní dokumenty, které obsahují buď slovo `lz` nebo frázi `lidské zdroje`:

```
lz "lidské zdroje"
```

Pravidlo v následujícím příkladu určuje, že podmínku splní dokumenty, jejichž název obsahuje slovo `"zdraví"`, nikoli však slovní spojení `"zaměstnanecké výhody"`:

```
+title:zdraví -title:"zaměstnanecké výhody"
```

Pravidla obsahu procházejí stejnou jazykovou normalizací jako dotazy rozhraní SI-API (Search and Index API). Syntaxe pravidel obsahu však podporuje jen podmnožinu operací dostupných v syntaxi dotazů SI-API. V pravidlech obsahu jsou povoleny pouze následující operátory dotazů:

- + Znaménko plus vložte před výrazy, které se v dokumentu musí vyskytovat.
- Znaménko minus vložte před výrazy, které se v dokumentu vyskytovat nesmí.

" " Uvedením dvou nebo více slov v uvozovkách určíte, že v dokumentu se musí vyskytnout celá fráze přesně tak, jak je napsána.

*název\_pole:*

Uvedením názvu pole před výrazem nebo frází určíte, že se příslušný výraz nebo fráze musí (nebo nesmí) vyskytovat v určeném poli dokumentu.

Podporována jsou všechna pole obsahu a metadat konfigurovaná jako pole kolekce s možností prohledávání. Následující typy polí a klíčová slova pro pole SI-API nejsou podporována:

- site:
- url:
- link:
- docID:
- samegroupas:
- parametrická pole
- prvky zabezpečení
- atributy (například \$source, \$language, \$doctype a další)

## Stromy kategorií

Strom kategorií vám umožňuje zobrazit všechny kategorie založené na pravidlech v kolekci. Slouží také k vytváření a odstraňování kategorií a k úpravám pravidel přiřazujících dokumenty ke kategoriím.

Strom kategorií, nazývaný také taxonomie, má hierarchické uspořádání. Strom začíná kořenovou kategorií a všechny další kategorie pocházejí z kořenové kategorie. Můžete vnořit libovolný počet kategorií a podkategorií, a poskytnout tak uživatelům různé možnosti procházení a načítání dokumentů.

Pokud například dokument splňuje pravidla v několika kategoriích, je asociován se všemi těmito kategoriemi. Když uživatel vyhledává kategorii nebo při práci s výsledky vyhledávání prochází dokumenty, které do kategorie náleží, skutečnost, že dokument patří do více kategorií, zvyšuje pravděpodobnost, že uživatel dokument najde.

Při administraci stromu kategorií můžete určit, které dokumenty patří do jedné nebo více kategorií, zařazením nových kategorií pod stávající kategorie. Při vytvoření kategorie určujete, zda má být vytvořena na kořenové úrovni nebo jako podkategorie jiné kategorie. Pomocí stromu kategorií můžete také odstranit kategorie z kolekce a změnit pravidla přiřazování dokumentů ke kategoriím. V rámci úprav kategorie můžete kategorii přejmenovat, přidat nebo odstranit pravidla kategorizace a změnit obsah jednotlivých pravidel.

Při správě stromu kategorií použijte jako pomůcku následující popis chování stromu při vyhledávání a procházení:

- Pokud uživatel prohledává kategorii na vyšší úrovni, jsou dokumenty odpovídající zadaným kritériím vyhledávány v této kategorii a ve všech jejích podkategoriích. V případě prohledávání kategorie, která nemá žádné další podkategorie, je prohledávána pouze zadaná kategorie.
- Pokud uživatel prochází výsledky vyhledávání a vybere volbu procházení dokumentů, které patří do konkrétní kategorie, zobrazí se pouze dokumenty v dané kategorii. Ve výsledcích vyhledávání se zobrazí rovněž názvy podkategorií, takže uživatelé mohou procházet mezi kategoriemi a zobrazovat podmnožiny dokumentů jednu po druhé.



## Výběr typu kategorizace

Při výběru typu kategorizace určujete přístup, který chcete použít k asociování dokumentů s kategoriemi v kolekci.

### Než začnete

Chcete-li změnit typ kategorizace, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, kterou měníte.

### O této úloze

Typ kategorizace je určen při vytváření kolekce. V případě potřeby lze způsob kategorizace dokumentů pro kolekci změnit. Můžete použít kategorie založené na pravidlech, které konfiguruje přímo pro kolekci, nebo nepoužít žádné kategorie.

**Důležité:** Pokud změníte typ kategorizace po prolézání dat a vytvoření indexu pro kolekci, bude index nekonzistentní. Chcete-li zajistit přesnost výsledků vyhledávání, proveďte nové prolezení dokumentů v kolekci a poté znovu sestavte hlavní index.

### Postup

Výběr typu kategorizace:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Vybrat typ kategorizace**.
2. Na stránce Vybrat typ kategorizace vyberte jednu z následujících voleb:

**Žádný** Tuto volbu vyberte, nechcete-li kategorizovat dokumenty v dané kolekci.

#### Založený na pravidlech

Tuto volbu vyberte, chcete-li dokumenty kategorizovat na základě pravidel konfigurovaných specificky pro tuto kolekci.

## Konfigurování kategorií

Pro kolekci lze vytvořit libovolný počet kategorií a každá kategorie může obsahovat libovolný počet pravidel. Pravidla určují, které dokumenty budou automaticky asociovány s danou kategorií.

### Než začnete

Chcete-li konfigurovat kategorie, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci s danými kategoriemi.

Jako typ kategorizace je nutné vybrat volbu použití kategorií založených na pravidlech.

Po klepnutí na volbu **Nápověda** při vytváření nebo úpravách kategorie můžete zobrazit příklady zadávání pravidel pro asociování dokumentů s kategoriemi.

### O této úloze

Pokud vyhledávací aplikace umožňují podporu kategorií, uživatelé mohou po zadání názvu kategorie prohledávat pouze určitou část dané kolekce. Mohou také vybrat kategorie ve výsledcích vyhledávání a procházet pouze dokumenty, které patří do vybrané kategorie.

**Důležité:** Pokud po prolézání dat a vytvoření indexu pro kolekci změníte kategorie nebo pravidla kategorií, přestane být index konzistentní. Chcete-li zajistit přesnost výsledků vyhledávání, proveďte nové prolezení dokumentů v kolekci a sestavte znovu hlavní index.

## Postup

Konfigurování kategorie:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat strom kategorií**.
2. Na stránce Strom kategorií vyberte ve stromu umístění, do kterého chcete přidat kategorii, a klepněte na volbu **Vytvořit kategorii**.

Pokud vyberete kořenovou položku, bude nová kategorie vytvořena na úrovni kořene. Pokud vyberete název kategorie, bude nová kategorie vnořena pod vybranou kategorií ve stromu kategorií.

Otevře se průvodce umožňující zadat pravidla pro asociování dokumentů s novou kategorií:

- a. Na stránce Vytvořit kategorii zadejte popisný název pro kategorii a klepněte na tlačítko **Další**.
- b. Na stránce Vytvořit pravidla kategorií klepněte na tlačítko **Přidat pravidlo**.
- c. Na stránce Vytvořit pravidlo kategorií zadejte do pole **Název pravidla** jedinečný název pro pravidlo. Tento název musí být jedinečný pro všechny kategorie v kolekci.
- d. Zadejte pravidlo, které chcete použít pro asociování dokumentů s danou kategorií, a potom klepněte na tlačítko **OK**.

- Chcete-li k určení, zda dokument patří do dané kategorie, použít identifikátor URI dokumentu, klepněte na volbu **Vzor identifikátoru URI** a zadejte vzor identifikátoru URI.

Jestliže zadaný text v identifikátoru existuje, pak je dokument asociován s danou kategorií.

Příklad: file:///c:/program+files/finance

- Chcete-li k určení, zda dokument patří do dané kategorie, použít dotaz na prohledávatelný obsah, klepněte na volbu **Obsah dokumentu**, vyberte jazyk dokumentů a zadejte slova a slovní spojení, která se musí nebo nesmí vyskytovat v obsahu dokumentu.

Pravidla se zapisují ve stejném formátu jako dotazy, povoleny jsou však pouze dotazovací operátory zahrnutí (+), vyloučení (-), slovního spojení (" ") a názvu pole (*název\_pole*:). Segmentace typu N-gram není s pravidly obsahu podporována.

Pokud dokument obsahuje nebo neobsahuje zadaná slova, pak je asociován s kategorií.

Příklad: +finance -účetnictví +title:"fiskální rok"

- e. Klepněte na tlačítko **Dokončit**.

Nová kategorie se nyní zobrazí na stránce Strom kategorií společně s ostatními kategoriemi, které patří do dané kolekce.

### Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 113

---

## Práce s vyhledávacími poli XML

Chcete-li uživatelům umožnit prohledávání určitých částí dokumentů XML, namapujte prvky XML na vyhledávací pole.

K mapování prvků XML na vyhledávací pole slouží konzola pro správu podnikového vyhledávání.

Obvykle jsou indexována všechna textová data v dokumentu XML. Namapujete-li prvky XML na vyhledávací pole, můžete využít informace o struktuře dokumentů XML k zajištění

podpory konkrétnějších dotazů. Po namapování názvu prvku XML na název pole můžete například povolit prohledávání dat v daném prvku XML podle názvu pole a vracení těchto dat ve výsledcích vyhledávání. Mapování můžete zpřesnit zadáním atributů prvku XML (a hodnot těchto atributů) jako kritérií pro výběr prvků, které se mohou stát vyhledávacími poli.

## Vyhledávací pole XML

Vyhledávací pole XML umožňují uživatelům zadávat dotazy na konkrétní části dokumentů XML.

Používání dokumentů XML se stále rozšiřuje, protože mohou obsahovat jak částečně strukturovaný, tak i nestrukturovaný text. Formát XML má zapouzdřenou strukturu a využívá kontext explicitně definovaný pomocí prvků XML obklopujících vlastní text. Jméno autora může být například uvedeno takto:

```
<author>Jan Novák</author>
```

V tomto kontextu text `Jan Novák` identifikuje autora dokumentu XML.

V rámci podnikového vyhledávání lze prvky XML přiřazovat (mapovat) k názvům vyhledávacích polí. Při konfigurování voleb analýzy pro kolekci určujete, které prvky XML mají být namapovány na určité názvy vyhledávacích polí. Mapováním prvků XML na vyhledávací pole umožňujete uživatelům vyhledávat hodnoty těchto prvků zadáním názvů polí v dotazech. Dotazy, které prohledávají konkrétní pole, mohou vracet přesnější výsledky než volné textové dotazy prohledávající celý obsah dokumentu.

Obsahuje-li například vaše kolekce dokumenty XML a určíte-li, že prvky `<title>` a `<author>` mají být v indexu označeny jako vyhledávací pole, uživatelé budou moci zadávat dotazy přímo na tyto prvky. Dotaz na text `author:Novák` nalezne dokumenty XML, u nichž je v prvcích namapovaných na pole `author` zadána hodnota `Novák`.

Jiný příklad: Prvek XML s názvem `<summary>` může obsahovat informace, které je vhodné zobrazit ve výsledcích vyhledávání. Namapujete-li prvek `<summary>` na vyhledávací pole a určíte-li, že hodnota tohoto prvku se má zobrazit ve výsledcích vyhledávání, bude obsah prvku součástí výsledného dokumentu.

Namapujete-li prvek XML na název pole, bude text obsažený v tomto prvku možné prohledávat pod názvem pole, který jste zadali. Pokud však prvek XML obsahuje atributy, hodnoty těchto atributů nebudou indexovány a nebude v nich možné přímo vyhledávat. Chcete-li zadávat dotazy na hodnoty atributů, musíte v konzole pro správu podnikového vyhledávání změnit nastavení voleb analýzy a povolit nativní vyhledávání XML. Po povolení nativního vyhledávání XML je v indexu dostupná struktura dokumentu XML a je možné zadávat dotazy na obsah dokumentu pomocí omezujících podmínek dotazů XPath. Příklad:

```
@xpath::'//název_prvku[@název_atributu="hodnota_atributu"]
```

### Související pojmy



Výraz dotazu sémantického vyhledávání

### Související úlohy

“Povolení podpory nativního vyhledávání XML” na stránce 142

## Mapování prvků XML na vyhledávací pole

Při mapování prvků XML na vyhledávací pole můžete určit, které prvky XML mohou uživatelé vyhledávat zadáním názvu pole v dotazu.

**Než začnete**

Chcete-li mapovat prvky XML na vyhledávací pole, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, do níž dokumenty XML náleží.

## Omezení

Pro každou kolekci může existovat více mapování polí XML, avšak pouze jedno mapování kořenového prvku XML. Kořenový prvek dokumentu XML zajistí správné použití mapování.

## O této úloze

Při vytváření mapování polí XML či přidávání, změnách nebo odstraňování polí v existujícím mapování polí XML se změna projeví až po restartování modulu analýzy. Nová a změněná mapování se týkají pouze nových dat, která jsou analyzována po restartování modulu analýzy, a na data, která jsou již v indexu obsažena, nemají žádný vliv. Chcete-li aktualizovat dokumenty, které jsou již v indexu obsaženy, musíte dokumenty prolézt a indexovat znovu.

V této úloze je na následujícím vzorovém dokumentu XML demonstrováno, jakým způsobem lze mapovat osobní záznamy a umožnit uživatelům přímé dotazy na určité prvky.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<personnel>
  <personnelrecord>
    <phone>5555</phone>
    <email>jannovak@praha.spolecnost.cz</email>
    <jobroles>manažer, architekt
      <jobrole>Skupina pro řízení vývoje technologií vyhledávání</jobrole>
      <jobrole>Technologie vyhledávání pro architekturu</jobrole>
    </jobroles>
    <location>Praha</location>
    <section id="expertise">
      <text>Lingvistika</text>
    </section>
  </personnelrecord>
</personnel>
```

## Postup

Mapování prvků XML z tohoto příkladu na vyhledávací pole:

1. Upravte kolekci, vyberte stránku Analýza XML klepněte na volbu **Mapovat prvky XML na pole**.
2. Na stránce Mapování polí XML klepněte na volbu **Vytvořit mapování XML**. Otevře se stránka Vytvořit mapování polí XML.
3. Do pole **Název kořenového prvku XML** zadejte název kořenového prvku: **personnel**.  
Ujistěte se, že zadaný název přesně odpovídá kořenovému prvku v dokumentech XML, které chcete vyhledat. Při analýze a indexování dokumentů XML podnikové vyhledávání vybere mapování podle názvu kořenového prvku.
4. Do pole **Název mapování XML** zadejte název pro danou sadu pravidel mapování polí XML.  
Po vytvoření sady pravidel mapování XML bude tento název zobrazen na stránce Mapování polí XML a po výběru tohoto názvu můžete přidávat, odstraňovat nebo měnit pravidla mapování.
5. Mapování prvku XML **jobrole** na vyhledávací pole s názvem **jobrole**:
  - a. Do pole **Název pole** zadejte hodnotu **jobrole**.

**Tip:** Klepnutím na tlačítko **Nápověda** zobrazíte informace o znacích ASCII a metaznacích, které nejsou v názvech polí podporovány.

- b. Do pole **Název prvku XML** zadejte hodnotu **jobrole**.
  - c. Chcete-li uživatelům umožnit zadávání dotazů na pole **jobrole**, zaškrtněte políčko **Vyhledávání s použitím polí**.
  - d. Chcete-li určit, že shoda má nastat jen tehdy, jestliže podmínka vyhledávání odpovídá celé hodnotě v poli **jobrole** (tj. toto pole neobsahuje žádná další slova), zaškrtněte políčko **Úplná shoda**.
  - e. Chcete-li uživatelům umožnit seřazení výsledků vyhledávání podle hodnot v poli **jobrole**, zaškrtněte políčko **Lze řadit**.  
Pole se zobrazí jako pole s možností řazení pouze tehdy, podporuje-li tuto funkci vyhledávací aplikace. Ukázková vyhledávací aplikace pro podnikové vyhledávání neobsahuje v seznamu polí, která lze vybrat pro řazení výsledků vyhledávání, pole mapovaná z prvků XML.
  - f. Chcete-li uživatelům umožnit zobrazení hodnot pole **jobrole** ve výsledcích vyhledávání, zaškrtněte políčko **Výsledky vyhledávání**.
6. Mapování prvku XML **jobroles** na stejné vyhledávací pole:
- a. Klepnutím na volbu **Přidat pole** přidejte do seznamu pravidel mapování polí prázdný řádek.
  - b. Do pole **Název pole** zadejte hodnotu **jobrole**.
  - c. Do pole **Název prvku XML** zadejte hodnotu **jobroles**.
- Tip:** Názvy prvků XML nemusí odpovídat názvům vyhledávacích polí. Na jedno vyhledávací pole lze mapovat více prvků XML.
- d. Chcete-li uživatelům umožnit zadávání dotazů na pole **jobrole**, určit, že vyhledávací požadavky musí zcela odpovídat hodnotě v tomto poli, používat toto pole k seřazení výsledků vyhledávání a zobrazovat je ve výsledcích vyhledávání, zaškrtněte příslušná políčka.
7. Mapování prvku XML **section** s atributem ID obsahujícím hodnotu **expertise** na vyhledávací pole s názvem **expertise**:
- a. Klepnutím na volbu **Přidat pole** přidejte do seznamu pravidel mapování polí prázdný řádek.
  - b. Do pole **Název pole** zadejte hodnotu **expertise**.
  - c. Do pole **Název prvku XML** zadejte hodnotu **section**.
  - d. Do pole **Název atributu XML** zadejte hodnotu **id**.
  - e. Do pole **Hodnota atributu XML** zadejte hodnotu **expertise**.
  - f. Chcete-li uživatelům umožnit zadávání dotazů na pole **expertise**, zaškrtněte políčko **Vyhledávání s použitím polí**.
  - g. Chcete-li určit, že shoda má nastat jen tehdy, jestliže podmínka vyhledávání odpovídá celé hodnotě v poli **expertise** (tj. toto pole neobsahuje žádná další slova), zaškrtněte políčko **Úplná shoda**.
  - h. Chcete-li uživatelům umožnit seřazení výsledků vyhledávání podle hodnot v poli **expertise**, zaškrtněte políčko **Lze řadit**, pokud vyhledávací aplikace tuto funkci podporuje.
  - i. Chcete-li uživatelům umožnit zobrazení hodnot pole **expertise** ve výsledcích vyhledávání, zaškrtněte políčko **Výsledky vyhledávání**.

### Příklady:

Chcete-li vyhledat všechny osoby v organizaci, které pracují na produktech pro vyhledávání, zadejte následující dotaz:

jobrole:vyhledávání

Chcete-li vyhledat všechny osoby v organizaci, které mají kvalifikaci v oboru lingvistika, zadejte následující dotaz:

expertise:lingvistika

---

## Práce s vyhledávacími poli HTML

Chcete-li uživatelům umožnit prohledávání určitých sekcí metadat v dokumentech HTML, namapujte prvky metadat HTML na vyhledávací pole.

K mapování prvků metadat HTML na vyhledávací pole slouží konzola pro správu podnikového vyhledávání.

Mapováním prvků metadat HTML na vyhledávací pole umožňujete uživatelům používat k prohledávání dokumentů HTML přesnější dotazy.

## Vyhledávací pole HTML

Vyhledávací pole HTML umožňují uživatelům zadávat dotazy na atributy dokumentů HTML.

Prvky metadat v dokumentech HTML se podobají atributům dokumentu v tom smyslu, že poskytují informace o dokumentu, o způsobu jeho formátování a o tom, jaký způsob webového přístupu k danému dokumentu je povolen. Příklad:

```
<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="Toto téma popisuje produkt společnosti IBM." />
<meta name="format" content="XHTML" />
```

V rámci podnikového vyhledávání lze názvy prvků metadat HTML přiřazovat (mapovat) k názvům vyhledávacích polí. Při konfigurování voleb analýzy pro kolekci určujete, které prvky metadat HTML mají být namapovány na určité názvy vyhledávacích polí. Mapováním prvků metadat HTML na vyhledávací pole umožňujete uživatelům prohledávat dokumenty s těmito prvky zadáním názvů vyhledávacích polí v dotazech. Dotazy, které prohledávají konkrétní pole, mohou vracet přesnější výsledky než volné textové dotazy prohledávající celý obsah dokumentu.

Obsahuje-li například vaše kolekce dokumenty HTML a určíte-li, že prvky metadat `copyright` a `abstract` mají být indexovány jako vyhledávací pole, uživatelé budou moci zadávat dotazy na tyto konkrétní prvky. Operace vyhledání řetězce `copyright:IBM` nalezne dokumenty HTML, v nichž prvek metadat `copyright` obsahuje hodnotu `IBM`.

Při mapování prvků metadat HTML na vyhledávací pole určujete, zda chcete namapovat všechny prvky metadat HTML, pouze prvky patřící k sadě prvků metadat Dublin Core nebo pouze prvky metadat HTML, které sami zadáte. Popis prvků v sadě prvků metadat Dublin Core naleznete na webu Dublin Core Metadata Initiative:

<http://dublincore.org/documents/dcmi-terms/#H2>

Pokud se rozhodnete vytvořit mapování pro všechny prvky metadat HTML nebo pro všechny prvky metadat Dublin Core, bude výchozí název vyhledávacího pole odpovídat názvu prvku metadat. Výchozí název vyhledávacího pole můžete přepsat a zadat jiné volby vyhledávání pro určité prvky přidáním těchto prvků do seznamu prvků, které chcete konfigurovat individuálně.

## Mapování prvků metadat HTML na vyhledávací pole

Při mapování prvků metadat HTML na vyhledávací pole můžete určit, které prvky metadat HTML mohou uživatelé vyhledávat zadáním názvu pole v dotazu.

### Než začnete

Chcete-li mapovat prvky metadat HTML na vyhledávací pole, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, do níž dokumenty HTML náleží.

### O této úloze

Při vytváření mapování polí HTML či přidávání, změnách nebo odstraňování polí v existujícím mapování polí HTML se změny projeví až po restartování modulu analýzy. Nová a změněná mapování se týkají pouze nových dat, která jsou analyzována po restartování modulu analýzy, a na data, která jsou již v indexu obsažena, nemají žádný vliv. Chcete-li aktualizovat dokumenty, které jsou již v indexu obsaženy, musíte dokumenty prolézt a indexovat znovu.

### Postup

Mapování prvků metadat HTML na vyhledávací pole:

1. Upravte kolekci, vyberte stránku Analýza XML klepněte na volbu **Mapovat metadata HTML na pole**.
2. Na stránce Mapování polí HTML určete, které prvky metadat HTML chcete namapovat na vyhledávací pole:
  - Pokud mapujete všechny prvky HTML nebo všechny prvky, které patří do sady prvků Dublin Core, určete, zda chcete uživatelům umožnit prohledávání polí podle názvu pole a zda se mají pole zobrazit ve výsledcích vyhledávání. Můžete také určit, zda musí vyhledávací požadavek uživatele přesně odpovídat celé hodnotě prvku metadat (tj. v poli se smí vyskytovat pouze slova odpovídající vyhledávacímu požadavku) a zda může uživatel seřadit výsledky vyhledávání podle daného pole.

Jako názvy polí systém automaticky používá názvy prvků metadat. Pokud chcete pro některé prvky přepsat výchozí názvy polí nebo použít jiné volby vyhledávání, můžete namapovat název prvku na jedinečný název pole a poté zadat volby vyhledávání, které chcete pro daný prvek použít.
  - Mapování prvků metadat specifických na vyhledávací pole:
    - a. Klepnutím na volbu **Přidat pole** přidejte do seznamu pravidel mapování polí prázdný řádek.
    - b. Zadejte název, který chcete přiřadit mapovanému prvku HTML. Uživatelé mohou tento název pole zadat při dotazování na dokumenty HTML v této kolekci.

**Tip:** Klepnutím na tlačítko **Nápověda** zobrazíte informace o znacích ASCII a metaznacích, které nejsou v názvech polí podporovány.
    - c. Zadejte název prvku metadat, který chcete namapovat.
    - d. Chcete-li uživatelům umožnit zadávání dotazů na toto pole, zaškrtněte políčko **Vyhledávání s použitím polí**.

**Tip:** Pokud toto políčko není zaškrtnuto, pole nelze prohledávat pomocí dotazů využívajících pole ani pomocí dotazů na volný text. Takové nastavení může být užitečné například tehdy, jestliže jste vybrali volbu zahrnutí všech prvků metadat HTML nebo všech prvků metadat Dublin Core, ale chcete zabránit prohledávání některých polí.

- e. Chcete-li uživatelům umožnit prohledávání tohoto pole a určit, že dokument splní podmínku dotazu jen tehdy, jestliže tato podmínka odpovídá celé hodnotě v poli (tj. pole neobsahuje žádná další slova), zaškrtněte políčko **Úplná shoda**.
- f. Chcete-li uživatelům umožnit abecední seřazení výsledků vyhledávání (řetězců) podle tohoto pole, zaškrtněte políčko **Lze řadit**.  
Pole se zobrazí jako pole s možností řazení pouze tehdy, podporuje-li tuto funkci vyhledávací aplikace. Ukázková vyhledávací aplikace pro podnikové vyhledávání neobsahuje v seznamu polí, která lze vybrat pro řazení výsledků vyhledávání, pole mapovaná z prvků metadat jazyka HTML.
- g. Pokud je datový typ tohoto pole DECIMAL, DOUBLE, INTEGER, SHORT, TIME nebo TIMESTAMP a chcete uživatelům umožnit při prohledávání tohoto pole zadávat parametrické dotazy nebo seřadit výsledky numericky podle hodnot tohoto pole, zaškrtněte políčko **Parametrické vyhledávání**.
- h. Chcete-li uživatelům umožnit zobrazení tohoto pole ve výsledcích vyhledávání, zaškrtněte políčko **Výsledky vyhledávání**.

### Příklad:

Uživatelé mohou zadáním dotazu na název mapovaných polí hledat dokumenty HTML s konkrétními metadaty. Pokud jste například namapovali prvek metadat HTML `description` na vyhledávací pole `abstract`, uživatelé mohou při hledání dokumentů HTML, které se zabývají počítači Thinkpad, zadat dotaz podobný následujícímu:

```
abstract:thinkpad
```

---

## Vlastní zpracování textu

Kvalitu a přesnost výsledků vyhledávání lze zvýšit integrací vlastních algoritmů pro zpracování textu do kolekce podnikového vyhledávání.

Produkt OmniFind Enterprise Edition podporuje architekturu IBM UIMA (Unstructured Information Management Architecture), která tvoří rámec pro vytváření, zjišťování, sestavování a implementaci funkcí analýzy textu. Vývojáři aplikací vytvoří a otestují algoritmy analýzy pro obsah, který má být prohledáván, a poté vytvoří archiv stroje pro zpracování (soubor `.pear`) zahrnující všechny prostředky potřebné k použití archivu pro podnikové vyhledávání. Chcete-li prohledávat kolekce pomocí vlastních algoritmů analýzy, je nutné tento archiv (obsahující stroj pro analýzu textu) přidat do systému podnikového vyhledávání.

Logická komponenta pro analýzu obsažená ve stroji pro analýzu textu se nazývá *anotační modul*. Každý anotační modul vykonává přesně definované úlohy lingvistické analýzy. Stroj pro zpracování textu může obsahovat libovolné množství anotačních modulů nebo může být utvořen sloučením několika strojů pro analýzu textu, z nichž každý obsahuje vlastní anotační moduly.

Informace generované anotačními moduly označujeme jako *výsledky analýzy*. Výsledky analýzy, které odpovídají hledaným informacím, jsou zapisovány do datové struktury nazývané *obecná struktura analýzy*.

Při konfigurování voleb zpracování textu pro kolekci je třeba provést následující úlohy:

- Vyberte stroj pro analýzu textu, který chcete používat k anotaci dokumentů v kolekci.
- Pokud daná kolekce obsahuje dokumenty XML s informačně relevantními značkami a chcete-li tyto značky použít při vlastní analýze textu, můžete soubory mapování asociovat s kolekci a mapovat výstup mapování XML na obecnou strukturu analýzy.



Můžete například namapovat prvky <addressee> a <customer> na anotace osoby v obecné struktuře analýzy. K těmto anotacím pak budou mít přístup vaše vlastní moduly pro anotaci, které mohou zjišťovat další informace (například pohlaví osoby). Anotace osoby můžete rovněž namapovat na index podnikového vyhledávání, takže uživatelé budou moci hledat osoby bez znalosti původních prvků XML.







Chcete-li uživatelům umožnit zadávat v dotazech původní prvky XML, nemusíte definovat žádné mapování XML. Namísto toho můžete konfigurovat volby analýzy a povolit v kolekci nativní vyhledávání XML.

- Namapujte obecnou strukturu analýzy na index podnikového vyhledávání, který umožňuje prohledávat anotované dokumenty pomocí sémantického vyhledávání.

V závislosti na entitách a vzájemných vztazích zjištěných anotačními moduly mohou uživatelé vyhledávat pojmy, které se vyskytují ve stejné větě (například konkrétní osoba a libovolné jméno konkurenta), nebo klíčová slova a koncepce (například jméno Aleš a telefonní číslo).

- Namapujte obecnou strukturu analýzy na relační databázi. Můžete mapovat data na tabulky databáze IBM DB2 nebo Oracle. Tento typ mapování umožňuje použití výsledků analýzy v databázových aplikacích, například při dolování dat. Umožňuje také používat dotazy SQL k prohledávání dat mimo systém podnikového vyhledávání.

### Související pojmy

-  [Integrace vlastní analýzy textu](#)
-  [Základní koncepce používané při zpracování analýzy textu](#)
-  [Workflow pro integraci vlastní analýzy](#)
-  [Algoritmy analýzy textu](#)
-  [Aplikace sémantického vyhledávání](#)
-  [Výraz dotazu sémantického vyhledávání](#)

## Přidávání strojů pro analýzu textu do systému

Pokud vytvoříte vlastní stroj pro analýzu textu, lze jej použít pro podnikové vyhledávání až po přidání do systému. Kolekce mohou stroj používat k analýze a anotacím dokumentů a k dosažení přesnějších výsledků vyhledávání.

### Než začnete

Chcete-li do systému přidávat stroje pro analýzu textu, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### O této úloze


Vývojáři aplikací mohou vytvořit archiv stroje pro zpracování (soubor PEAR), který je v souladu s rámcem UIMA pro analýzu textu. Archiv obsahuje všechny prostředky potřebné k prohledávání kolekci podnikového vyhledávání. Chcete-li prohledávat kolekce pomocí vlastních algoritmů analýzy, je nutné tento archiv (obsahující stroj pro analýzu textu) přidat do systému podnikového vyhledávání.

Po přidání stroje pro analýzu textu do systému můžete změnit jeho zobrazovaný název a vybrat volbu zobrazení zdroje XML. Prostřednictvím zdroje XML můžete zobrazit informace vytvářené tímto strojem.





Pokud je stroj pro analýzu textu asociován s kolekci, nelze jej odebrat ze systému.

## Postup

Přidání vlastního stroje pro analýzu textu do systému podnikového vyhledávání:

1. Klepnutím na volbu **Systém** otevřete pohled Systém.
2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
3. Na stránce Analýza klepněte na volbu **Konfigurovat stroje pro analýzu textu**.
4. Na stránce Stroje pro analýzu textu klepněte na volbu **Přidat stroj pro analýzu textu**.
5. Na stránce Přidat stroj pro analýzu textu zadejte popisný název pro nový stroj. Systém použije tento zobrazovaný název k identifikaci stroje pro analýzu textu v rámci celé konzoly pro správu.
6. Určete umístění souboru PEAR. Je-li soubor menší než 8 MB, může být umístěn v lokálním počítači a lze jej vyhledat procházením adresářů. Je-li soubor větší než 8 MB, musí se nacházet na indexovém serveru a je nutné zadat úplnou cestu k němu.
7. Klepněte na tlačítko **OK**. Stroj pro analýzu textu bude uveden na stránce Stroje pro analýzu textu.

### Související pojmy

-  Workflow pro integraci vlastní analýzy
-  Integrace vlastní analýzy textu
-  Základní koncepce používané při zpracování analýzy textu
-  Značky jazyka XML při analýze a vyhledávání

### Související úlohy

-  Vytváření prvků XML do souboru mapování obecné struktury analýzy

## Asociování stroje pro analýzu textu s kolekcí

Pokud jsou se systémem podnikového vyhledávání asociovány vlastní stroje pro analýzu textu, můžete jeden z nich vybrat pro použití s kolekcí. Uživatelé potom mohou při prohledávání kolekce zadávat sémantické dotazy a získat tak kvalitnější a přesnější výsledky vyhledávání.

### Než začnete

Chcete-li asociovat stroj pro analýzu textu s kolekcí, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

Pokud je stroj pro analýzu textu již asociován s danou kolekcí, při asociování jiného stroje budou provedeny následující akce:

- Pokud jste vybrali volbu **K dispozici není žádná vlastní analýza**, dojde k vynulování všech mapování, která jste pro danou kolekci definovali dříve. Kolekce začne používat výchozí systémové hodnoty.
- Pokud vyberete název jiného vlastního stroje pro analýzu textu, budou všechna dříve definovaná mapování analýzy textu pro danou kolekci zachována. Pokud například změníte hodnotu stroj\_1 na hodnotu stroj\_2, zdědí stroj stroj\_2 soubory mapování XML, které jste konfigurovali pro stroj stroj\_1.

## Postup

Asociování stroje pro analýzu textu s kolekcí:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat volby zpracování textu**.
2. Klepněte na volbu **Vybrat stroj pro analýzu textu**. Pokud do systému podnikového vyhledávání nebyly přidány žádné vlastní stroje pro analýzu textu nebo pokud kolekce používá výchozí algoritmy analýzy, je název stroje **Výchozí**.
3. Na stránce Vybrat stroj pro analýzu textu pro tuto kolekci vyberte název stroje, který chcete použít s danou kolekcí. Pokud nejsou k dispozici žádné stroje pro analýzu textu nebo pokud jste vybrali volbu **K dispozici není žádná vlastní analýza**, modul pro analýzu použije při anotaci dokumentů a jejich přípravě pro index výchozí pravidla analýzy textu.

#### Související pojmy



Workflow pro integraci vlastní analýzy



Integrace vlastní analýzy textu



Základní koncepce používané při zpracování analýzy textu

## Mapování prvků XML na obecnou strukturu analýzy

Pokud daná kolekce obsahuje dokumenty XML s informačně relevantními značkami a pokud chcete tyto značky použít, abyste uživatelům umožnili prohledávat index podnikového vyhledávání nebo tabulky relační databáze s použitím sémantického vyhledávání, můžete namapovat prvky XML na obecnou strukturu analýzy.

#### Než začnete

Chcete-li mapovat prvky XML na obecnou strukturu analýzy, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

#### Omezení

Maximální velikost souboru mapování je 8 MB.

#### O této úloze

Chcete-li procesům vlastní analýzy textu umožnit přístup ke konkrétním prvkům v dokumentech XML nebo namapovat některé prvky XML na obecné typy používané při sémantickém vyhledávání, můžete vytvořit vlastní soubory mapování. Soubory mapování musí dodržovat specifikaci architektury UIMA pro analýzu textu.

Přidáním souborů mapování do kolekce, která používá vlastní stroj pro analýzu textu, umožníte mapování prvků XML ve zdrojových dokumentech na anotace v obecné struktuře analýzy. Tyto anotace pak může používat vlastní stroj pro analýzu textu. Obecnou strukturu analýzy lze namapovat na index a povolit uživatelům zadávat dotazy na anotace při vyhledávání v kolekci pomocí sémantického vyhledávání.

Můžete například namapovat prvky adresáta a zákazníka na anotace osoby v obecné struktuře analýzy. K těmto anotacím pak budou mít přístup vaše vlastní moduly pro anotaci, které mohou zjišťovat další informace (například pohlaví osoby). Anotace osoby můžete rovněž namapovat na index podnikového vyhledávání, takže uživatelé budou moci hledat osoby bez znalosti původních prvků XML.

Chcete-li uživatelům umožnit zadávat v dotazech původní prvky XML, nemusíte konfigurovat žádné soubory mapování. Namísto toho můžete konfigurovat volby analýzy a povolit v kolekci nativní vyhledávání XML.

### Postup

Mapování prvků XML na obecnou strukturu analýzy:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat volby zpracování textu**.
2. V oblasti **Mapovat prvky XML na obecnou strukturu analýzy** klepněte na volbu **Přidat mapování**.
3. Na stránce Mapovat prvky XML na obecnou strukturu analýzy zadejte popisný zobrazovaný název pro soubor mapování.
4. Určete umístění souboru. Pokud je soubor mapování uložen v lokálním systému, můžete cestu k němu vyhledat procházením adresářů. Pokud je soubor mapování umístěn na indexovém serveru, musíte zadat úplnou cestu sami.
5. Klepněte na tlačítko **OK**. Nový soubor mapování bude přidán na stránku Volby zpracování textu.

#### Související pojmy



Workflow pro integraci vlastní analýzy



Integrace vlastní analýzy textu



Základní koncepce používané při zpracování analýzy textu



Značky jazyka XML při analýze a vyhledávání

#### Související úlohy



Vytváření prvků XML do souboru mapování obecné struktury analýzy

## Mapování obecné struktury analýzy na index

V případě potřeby můžete určit, které výsledky textové analýzy z obecné struktury analýzy mají být namapovány na index a zpřístupněny uživatelům, kteří zadávají dotazy na kolekce s využitím sémantického vyhledávání.

### Než začnete

Chcete-li mapovat obecnou strukturu analýzy na index, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Omezení

Maximální velikost souboru mapování je 8 MB.

### O této úloze

Mapováním obecné struktury analýzy na index podnikového vyhledávání umožníte uživatelům zadávat sémanticky přesné dotazy a zvýšit tak kvalitu výsledků vyhledávání.





V závislosti na entitách a vzájemných vztazích určených anotátory mohou uživatelé vyhledávat pojmy, které se vyskytují ve stejné větě (například konkrétní osoba a libovolné jméno konkurenta), nebo klíčová slova a koncepce (například jméno Aleš a telefonní číslo).

## Postup

Chcete-li mapovat obecnou strukturu analýzy na index, postupujte takto:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat volby zpracování textu**.
2. V oblasti **Mapovat obecnou strukturu analýzy na index** klepněte na volbu **Vybrat soubor mapování**.
3. Na stránce Vybrat soubor mapování pro tuto kolekci vyberte mapování, které chcete používat pro index podnikového vyhledávání:
  - Chcete-li v indexu podnikového vyhledávání používat výchozí pravidla mapování, vyberte volbu **Výchozí**.
  - Chcete-li mapovat vlastní obecnou strukturu analýzy na index, určete umístění souboru mapování. Pokud je soubor uložen v lokálním systému, můžete jej vyhledat. Pokud je soubor umístěn na indexovém serveru, zadejte úplnou cestu k souboru.
4. Klepněte na tlačítko **OK**. Určený soubor mapování se zobrazí na stránce Volby zpracování textu.

### Související pojmy

-  Workflow pro integraci vlastní analýzy
-  Integrace vlastní analýzy textu
-  Základní koncepce používané při zpracování analýzy textu
-  Mapování indexů pro výsledky vlastní analýzy

### Související úlohy

-  Vytvoření obecné struktury analýzy do souboru mapování indexu

## Mapování obecné struktury analýzy na relační databázi

V případě potřeby můžete určit, které výsledky textové analýzy z obecné struktury analýzy mají být namapovány na relační databázi, aby je bylo možné používat v databázových aplikacích.

### Než začnete

Chcete-li mapovat obecnou strukturu analýzy na relační databázi, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Omezení

Maximální velikost souboru mapování je 8 MB.

### O této úloze

Mapováním obecné struktury analýzy na relační databázové tabulky umožníte použití dat v databázových aplikacích. Uživatelé například mohou prostřednictvím dotazů SQL zadávaných mimo rámec podnikového vyhledávání prohledávat anotace přidané při analýze textu. Informace můžete rovněž použít pro další zpracování textu, například v aplikacích dolování dat.

Obecnou strukturu analýzy lze mapovat na relační databázi vždy pouze pomocí jednoho souboru mapování. Pokud odešlete nový soubor mapování, bude aktuální soubor mapování přepsán.

## Postup

Chcete-li namapovat obecnou strukturu analýzy na relační databázi, postupujte takto:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat volby zpracování textu**.
2. V oblasti **Mapovat obecnou strukturu analýzy na relační databázi** klepněte na volbu **Přidat mapování**.
3. Na stránce Mapovat obecnou strukturu analýzy na relační databázi zadejte popisný zobrazovaný název souboru mapování, který chcete použít pro mapování informací na relační databázi.
4. Určete umístění souboru mapování. Pokud je soubor uložen v lokálním systému, můžete jej vyhledat. Pokud je soubor umístěn na indexovém serveru, zadejte úplnou cestu k souboru.
5. Klepněte na tlačítko **OK**. Zobrazovaný název souboru mapování je uveden na stránce Volby zpracování textu.

### Související pojmy



Workflow pro integraci vlastní analýzy



Integrace vlastní analýzy textu



Základní koncepce používané při zpracování analýzy textu



Mapování databázi pro vybrané výsledky analýzy

### Související úlohy



Vytvoření obecné struktury analýzy do souboru mapování databázi

---

## Konfigurování podprocesů pro službu analyzátoru

Máte-li k dispozici dostatek prostředků paměti, můžete zvýšit počet podprocesů, které jsou k dispozici pro modul analýzy při analýze dokumentů.

### Než začnete

V případě většího počtu kolekcí může být vhodné zvýšit počet podprocesů analyzátoru. Ověřte, že je v systému dostatek paměti pro podporu dalších podprocesů. Analyzátor s jedním podprocesem vyžaduje 200 MB paměti. Pro každý další podproces je vyžadováno dalších 50 paměti.

Chcete-li konfigurovat počet podprocesů spouštěných pro modul analýzy, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

## Postup

Chcete-li konfigurovat počet podprocesů modulu analýzy, postupujte takto:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat volby analýzy**.
2. Zadejte maximální počet podprocesů modulu analýzy, které mají být spuštěny při spuštění modulu analýzy, a klepněte na tlačítko **OK**.
3. Restartujte modul analýzy.

---

## Povolení rozšířené analýzy složených výrazů

Povolením použití rozšířené analýzy složených výrazů pro analyzátor lze zvýšit kvalitu vyhledávání. V případě použití rozšířené analýzy jsou složené výrazy rozkládány, aby mohla být každá část zpracována jako samostatný výraz.

### Než začnete

Chcete-li zadat volby pro analýzu složených výrazů, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

V některých jazycích jsou výrazy o více slovech slučovány do jednoho slova bez mezer (*složené výrazy*). Rozšířená analýza a rozklad složených výrazů jsou užitečné pro prohledávání jazyků, jako je němčina, a zásadní pro prohledávání jazyků, jako je korejština.

Povolíte-li pro složené výrazy rozšířenou analýzu, budou uživatelé moci vyhledávat výrazy, aniž by bylo nutné používat zástupné znaky, aby byly vyhledány složené tvary výrazů dotazu. Při hledání výrazu **Organ** (varhany) mohou být například vráceny dokumenty obsahující výraz **Organspender** (dárce varhan), nikoli však dokumenty obsahující výraz **Organisation** (organizace). Na rozdíl od dotazu s použitím zástupného znaku **Organ\***, pro který může být vrácen libovolný řetězec obsahující posloupnost znaků **Organ**, je při tomto vyhledávání dosaženo shody pouze pro lingvisticky zcela korektní podřadky v rámci rozsáhlejšího složeného výrazu.

Dalším příkladem je rozložení složeniny **Mustermann** na dva tokeny (muster a mann), které jsou v indexu uloženy odděleně. Při zadání dotazu se zástupným znakem **Musterma\*** nemůže proces identifikovat řetězec **Musterma** jako předponu rozloženého slova. Výsledkem je, že dokument s termínem **Mustermann** nebude nalezen. Chcete-li, aby uživatelé mohli zadávat dotazy se zástupnými znaky pro složené výrazy, nepovolujte rozšířenou analýzu složených výrazů.

Výrazy slovníků definované uživatelem, například synonyma a slova pro upřednostnění, jsou také používána pro části složených výrazů použitých jako jednotlivá slova v dotazu.

### Postup

Chcete-li povolit rozšířenou analýzu složených výrazů, postupujte takto:

1. Upravte kolekci, vyberte volbu Analýza a klepněte na volbu **Konfigurovat volby analýzy**.
2. Zaškrtněte políčko **Povolit rozšířenou analýzu pro složené výrazy** a klepněte na tlačítko **OK**.

#### Související pojmy

“Zástupné znaky v dotazech” na stránce 174



Lingvistická podpora pro sémantické vyhledávání



Analýza textu integrovaná v podnikovém vyhledávání

---

## Povolení podpory nativního vyhledávání XML

Pokud kolekce obsahuje dokumenty XML, můžete uživatelům umožnit použití značek XML při vyhledávání dokumentů, povolíte-li pro tuto kolekci nativní vyhledávání XML.

### Než začnete

Chcete-li povolit podporu prohledávání dokumentů XML pomocí nativního vyhledávání XML, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

Nativní vyhledávání XML například XPath nebo fragmenty XML, může poskytovat přesnější výsledky vyhledávání díky využití značek XML dokumentů v dotazu. Uživatelé mohou určit, že výraz uvedený v dotazu se musí vyskytnout v určitém prvku nebo atributu XML.





Faktury ve formátu XML od prodejce počítačové techniky mohou například obsahovat položky <objednavka> obsahující prvky <spolecnost> a <typ\_pocitace>. Chcete-li načíst faktury obsahující objednávky na notebooky IBM, může vyhledávání klíčových slov IBM a notebook vrátit dokumenty, které se zabývají notebooky Dell a stolními počítači IBM. Pomocí vyhledávání XML můžete určit, že slovo IBM se musí objevit v prvku <spolecnost>, slovo notebook v prvku <typ\_pocitace> a oba tyto prvky musí být obsaženy ve stejném prvku <objednavka>. Tímto způsobem načtete pouze faktury, které se zabývají výslovně notebooky IBM.

### Postup

Chcete-li uživatelům povolit prohledávání kolekce s použitím nativního vyhledávání XML, postupujte takto:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat volby analýzy**.
2. Zaškrtněte políčko **Povolit uživatelům vyhledávání v dokumentech ve formátu XML s použitím nativního vyhledávání XML**.
3. Klepněte na tlačítko **OK**.

#### Související pojmy

-  Lingvistická podpora pro sémantické vyhledávání
  -  Analýza textu integrovaná v podnikovém vyhledávání
  -  Aplikace sémantického vyhledávání
  -  Výraz dotazu sémantického vyhledávání
- “Vyhledávací pole XML” na stránce 129



---

## Detekce formátu dokumentu

Podnikové vyhledávání využívá k určování typů dokumentů a analyzátorů, které mají být použity pro jednotlivé dokumenty, výchozí mapování přípon adres URL a typů MIME.

Úpravou konfiguračního souboru `parserTypes.cfg` můžete přepsat a rozšířit výchozí mapování přípon adres URL a typů MIME na typy analyzátorů. Pravidla pro mapování přípon souborů a typů MIME na typy analyzátorů jsou definována v souboru `parserTypes.cfg`. Můžete například vytvořit mapování pro příponu názvů souborů, např. `.content`, a určit, že dokumenty tohoto typu mají být analyzovány pomocí analyzátoru HTML.

Různé formáty dokumentů používají různé interní reprezentace. Systém podnikového vyhledávání analyzuje dokumenty pomocí interních filtrů i filtrů jiných dodavatelů a mnoho dokumentů je analyzováno pomocí služeb analýzy specializovaných na konkrétní formát.

Detekce formátu dokumentů a přiřazení analyzátoru probíhá takto:

1. Detekční algoritmus zkontroluje příponu adresy URL zpracovávaného dokumentu.
2. Systém zkontroluje typ MIME dokumentu, který je součástí metadat poskytnutých prolézacím modulem.
3. Systém se pokusí každému z dokumentů přiřadit odpovídající typ analyzátoru. Dokumentům typu HTML, text (TXT) a XML systém přiřadí typ analyzátoru, který je specifický pro daný formát dokumentu.

U ostatních formátů dokumentů používá systém analyzátor Stellent. Technologii filtrování dokumentů Stellent nyní vlastní společnost Oracle. Odkazy na technologii Stellent v tomto dokumentu jsou ekvivalenty odkazů na technologii Outside In Content Access společnosti Oracle.

Analyzátor Stellent podporuje několik set formátů dokumentů, pro účely podnikového vyhledávání je však povolena jen část filtrů dokumentů. Konfigurační soubory ovšem můžete sami upravit a povolit analýzu dalších typů dokumentů pomocí analyzátoru Stellent.

**Důležité:** Přidané filtry dokumentů, které nepřísluší k dílčí sadě filtrů dokumentů s podporou podnikového vyhledávání, nebyly testovány a nejsou podporovány.

4. Pokud není k dispozici soubor `parserTypes.cfg`, je k určení typu dokumentu a odpovídajícího analyzátoru použito výchozí mapování. Při zjišťování typu dokumentu postupuje systém takto:
  - a. Porovná příponu adresy URL s uživatelem definovanými pravidly pro přípony v souboru `parserTypes.cfg`.
  - b. Porovná typ MIME s uživatelem definovanými pravidly pro typy MIME v souboru `parserTypes.cfg`.
  - c. Porovná příponu adresy URL s výchozími pravidly podnikového vyhledávání.
  - d. Porovná typ MIME s výchozími pravidly podnikového vyhledávání pro typy MIME.
5. Dokument, jehož formát se systému nepodaří identifikovat, je odmítnut. Přitom se může zobrazit chybová zpráva informující o tom, že daný typ dokumentu není podporován.

Chybová zpráva se může zobrazit také tehdy, je-li danému typu dokumentu přiřazen analyzátor Stellent a nerozpozná-li tento analyzátor formát dokumentu. Možné příčiny chyb:

- Dokument je poškozen.
- Dokument je uložen ve formátu, který analyzátor Stellent nepodporuje. Chcete-li tento problém vyřešit, musíte přidat odmítané formáty souborů do souboru `stellentTypes.cfg`.

Dále je třeba aktualizovat soubor `parserTypes.cfg` a určit, že typ MIME nebo přípona odmítnutých formátů dokumentů mají být přiřazeny k analyzátoru Stellent.

---

## Výchozí podporované typy dokumentů

Při detekci formátu dokumentu jsou vyhodnocovány jen některé typy dokumentů.

Následující formáty dokumentů jsou nativními typy, které jsou automaticky detekovány a analyzovány vestavěnými službami analyzátoru kolekce.

HTML  
prostý text  
XML

Ve výchozím nastavení jsou následující formáty dokumentů analyzovány analyzátozem Stellent:

Adobe PDF (Portable Document Format)  
Lotus 1-2-3  
Lotus Freelance Graphics  
Lotus Word Pro  
Just System Ichitaro  
Microsoft Excel (verze až do 2007)  
Microsoft PowerPoint (verze až do 2007)  
Microsoft Visio  
Microsoft Word (verze až do 2007)  
RTF (Rich Text Format)  
StarOffice/OpenOffice Calc  
StarOffice/OpenOffice Impress  
StarOffice/OpenOffice Draw  
StarOffice/OpenOffice Writer

Formáty souborů Office Open XML v sadě Microsoft Office a formáty OpenDocument v sadě OpenOffice se zpracovávají bez nutnosti provádět změny v konfiguračních souborech.

Chcete-li analyzovat další typy dokumentů, musíte do konfiguračních souborů (`parserTypes.cfg` a `stellenttypes.cfg`) doplnit pravidla pro mapování požadovaných typů dokumentů na službu analyzátoru kolekce nebo na filtr Stellent.

**Omezení:** Zpracování obousměrného textu v souborech PDF, aby odpovídal logickému pořadí čtení, je mimo možnosti technologie prohlížeče Stellent. Analyzátor Stellent negarantuje pořadí textu, které extrahuje ze souborů PDF. Při použití obousměrného textu v souborech PDF pořadí, ve kterém je text analyzován, pravděpodobně neodpovídá logickému pořadí čtení textu. Toto omezení způsobuje problémy při zpracování souborů PDF, které jsou psány v jazycích Středního východu, jako je například hebrejščina a arabština, které jsou psány většinou zprava doleva (obousměrně).

---

## Typy dokumentů přiřazené k analyzátorům kolekcí a k analyzátorům Stellent

Chcete-li zajistit přesnou a efektivní analýzu dokumentů v prolézáném prostoru, můžete vytvořit konfigurační soubory určující, které typy dokumentů mají být analyzovány modulem analýzy kolekce a které pomocí filtrů dokumentů Stellent.

Většinu formátů dokumentů v kolekci podnikového vyhledávání analyzují vestavěné analytické moduly HTML a XML. Určité typy dokumentů obvykle nejsou analyzovány (například dokumenty ve formátu Postscript) a jiné typy dokumentů zpracovávají analytické

funkce Stellent (jedná se například o dokumenty typu Microsoft Word, Microsoft Excel, Microsoft PowerPoint, Lotus Freelance, Lotus 123, PDF, RT a Ichitaro).

V důsledku nejednoznačnosti metadat může být do analyzátoru Stellent omylem odeslán prostý text nebo dokumenty ve formátu HTML. Tato data jsou zasílána zpět ke zpracování některým z vestavěných analyzátorů, což může ovlivnit výkon systému. U jiných dokumentů nemusí být možné detekovat typ dokumentu - takové dokumenty jsou pak přeskočeny. Chcete-li takovým situacím předejít, můžete vytvořit konfigurační soubory určující, kde a jakým způsobem mají být analyzovány různé typy dokumentů.

Přiřazení typů dokumentů k analyzátorům kolekcí a k analyzátorům Stellent probíhá v následujících krocích:

1. Konfigurování typů dokumentů pro analyzátor kolekce. V tomto kroku je třeba vytvořit konfigurační soubor, který mapuje různé typy dokumentů na analyzátor využívaný kolekcí. Pro každou kolekci lze vytvořit jeden takový konfigurační soubor.
2. Konfigurování typů dokumentů pro analyzátor Stellent. V tomto kroku je třeba vytvořit konfigurační soubor, který mapuje různé typy dokumentů na filtry dokumentů Stellent využívané kolekcí. Pro každou kolekci lze vytvořit jeden takový konfigurační soubor.
3. Zastavení a nové spuštění analyzátoru. Chcete-li, aby změny vstoupily v platnost, použijte konzolu správy podnikového vyhledávání k monitorování kolekce, pro niž jste nastavili typy dokumentů, a poté zastavte a znovu spusíte analyzátor.

---

## Přiřazení typů dokumentů k analyzátoru kolekce

Chcete-li k analyzátoru kolekce přiřadit konkrétní typy dokumentů, musíte vytvořit konfigurační soubor `parserTypes.cfg`. Konzola pro správu podnikového vyhledávání neposkytuje žádnou podporu této úlohy.

### Než začnete

Chcete-li dokončit tuto úlohu, musíte se přihlásit jako administrátor podnikového vyhledávání.

### O této úloze

Pokud konfigurační soubor neexistuje, analyzátor kolekce bude používat výchozí pravidla služby analýzy. Pokud konfigurační soubor existuje, určují pravidla v tomto souboru následující nastavení:

- Které přípony URL a které typy MIME jsou namapovány na jednotlivé typy analyzátorů.
- Jakým způsobem mají být analyzovány soubory, jejichž typ je neznámý vzhledem k neúplným metadatům.

Formát souboru `parserTypes.cfg` je tvořen posloupností řádků, z nichž každý obsahuje jedno z následujících pravidel:

#### **EXTENSION** *přípona analyzátor*

Zadaný analyzátor bude zpracovávat všechny dokumenty, jejichž adresa URL končí zadanou příponou. Příponu uvádějte bez tečky na začátku. Při porovnávání se nerozlišují velká a malá písmena.

#### **CONTENTTYPE** *typ/podtyp analyzátor*

Zadaný analyzátor bude zpracovávat všechny dokumenty, jejichž typ obsahu odpovídá zadanému typu nebo dílčímu typu. Použijeme-li například typ obsahu dokumentu `t/s`, dojde ke shodě, pokud se hodnota `t` rovná typu a hodnota `s` se rovná podtypu nebo pokud je místo podtypu použit zástupný znak (hvězdička, `*`).

### UNKNOWN *analyzátor*

Zadaný analyzátor bude zpracovávat všechny dokumenty, u nichž není známa přípona a typ obsahu (tj. dokumenty, které nebyly zpřístupněny prolézacím modulem).

### DEFAULT *analyzátor*

Zadaným analyzátozem budou zpracovávány všechny dokumenty, na které se nevztahuje žádné jiné pravidlo.

Ve všech případech musí mít parametr *analyzátor* hodnotu `text`, `html`, `xml`, `stellent` nebo `none`, kde `none` znamená, že dokumenty tohoto typu nemají být analyzovány.

Pokud dokumentu odpovídá více pravidel než jedno, má přednost konkrétnější pravidlo bez ohledu na to, v jakém pořadí jsou pravidla uvedena:

- Pravidlo `EXTENSION` je konkrétnější než pravidlo `CONTENTTYPE`.
- Pravidlo `CONTENTTYPE` s uvedením podtypu je konkrétnější než pravidlo se zástupným znakem. Pravidlo pro typ obsahu `application/postscript` tak má například přednost před pravidlem pro typ `application/*`.
- Pro stejnou příponu nebo typ obsahu by neměla být definována dvě různá pravidla. V takovém případě závisí priorita pravidel na konkrétní implementaci.

### Postup

Chcete-li přiřadit typy dokumentů k analyzátoru kolekce, postupujte takto:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
2. Vytvořte konfigurační soubor následujícím příkazem, kde *ID\_kolekce* identifikuje kolekci, kterou chcete konfigurovat:  
`ES_NODE_ROOT/master_config/ID_kolekce.parserdriver/parserTypes.cfg`
3. Otevřete soubor v textovém editoru, zadejte pravidla služby analyzátoru a poté soubor uložte a zavřete.
4. Chcete-li, aby změny vstoupily v platnost, spusťte monitorování analyzátoru kolekce pomocí konzoly pro správu podnikového vyhledávání a poté analyzátor zastavte a znovu spusťte.

### Příklad

V tomto příkladu bude vestavěný analyzátor HTML zpracovávat všechny dokumenty s příponami `TXT`, `HTM` a `HTML`, všechny dokumenty s typem obsahu, který začíná předponou `text/`, a všechny dokumenty s neznámou příponou i typem obsahu. Vestavěný analyzátor XML bude zpracovávat všechny dokumenty s příponou `xml` a všechny dokumenty s typem obsahu `text/xml`. Všechny ostatní dokumenty včetně dokumentů s typem obsahu, který začíná předponou `application/`, budou odeslány do analyzátoru `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
```

### Související úlohy

“Změna pravidel náhrady pro některé značky HTML” na stránce 148

“Přiřazení typů dokumentů k analyzátoru Stellent” na stránce 150

### Související odkazy

“Výchozí pravidla služby analyzátoru kolekce”

## Výchozí pravidla služby analyzátoru kolekce

Nevytvoříte-li pro kolekci konfigurační soubor pro mapování typů souborů a typů obsahu na analyzátor, budou pro analýzu dokumentů použita výchozí pravidla.

Výchozí pravidla používaná analyzátozem kolekce jsou následující:

```
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 jtddx 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
```

### Související úlohy

“Přiřazení typů dokumentů k analyzátoru kolekce” na stránce 145

## Analýza neznámých typů dokumentů

Je-li typ dokumentu neznámý (například pokud dokument nezahrnuje příponu souboru nebo nemá přiřazen typ MIME), můžete konfigurovat pravidla zabraňující analyzátoru zrušení tohoto dokumentu.

### O této úloze

Pokud analyzátor nerozpozná formát souboru, pokusí se analyzovat dokument pomocí výchozího analyzátoru HTML. Pokud obsah není ve formátu HTML, analyzátor dokument zamítne. Záznam všech zamítnutých dokumentů je zapsán do následujícího umístění na indexovém serveru:

ES\_NODE\_ROOT/data/ID\_kolekce/dropped\_doc\_logs/dropped\_docs\_pd\_date.log

### Postup

Chcete-li se této situaci vyhnout, konfiguruje analyzátor tak, aby pro neznámé typy dokumentů používal analyzátor ASCII:

1. Přihlaste se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
2. Vytvořte nebo upravte následující konfigurační soubor, kde *ID\_kolekce* identifikuje kolekci, kterou chcete konfigurovat:  
ES\_NODE\_ROOT/master\_config/ID\_kolekce.parserdriver/parserTypes.cfg
3. Přidejte do souboru následující pravidlo:  
UNKNOWN text
4. Chcete-li, aby změny vstoupily v platnost, spusťte monitorování analyzátoru kolekce pomocí konzoly pro správu podnikového vyhledávání a poté analyzátor zastavte a znovu spusťte.

---

## Změna pravidel náhrady pro některé značky HTML

Můžete změnit pravidla náhrady značek HTML, která analyzátor používá při analýze dokumentů HTML. Konzola pro správu podnikového vyhledávání neposkytuje žádnou podporu této úlohy.

### Než začnete

Chcete-li dokončit tuto úlohu, musíte se přihlásit jako administrátor podnikového vyhledávání.

### O této úloze

Analyzátor HTML nahrazuje informace o značkách (značky HTML) jinými znaky, které modelují význam těchto značek tak, jak s nimi pracuje webový prohlížeč, za účelem poskytování textových informací pro index podnikového vyhledávání. Například výsledkem značky odstavce (<p>) je oddělovač odstavců ve struktuře CAS (Common Analysis Structure) a v indexu. Pravidla náhrady ovlivňují, které části textových informací se zobrazí ve stejném odstavci.

Chcete-li blíže spojit konkrétní značky HTML se standardem HTML, můžete aktualizovat konfigurační soubor analyzátoru kolekce a změnit některá pravidla náhrady.

**Tip:** Informace o dalších postupech řízení způsobu zpracování značek HTML v dokumentech HTML naleznete v tématu <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011251>.

### Postup

Chcete-li změnit pravidla náhrady značek HTML, postupujte takto:

1. Přihlaste se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.

- Upravte soubor `ES_NODE_ROOT/master_config/ID_kolekce.parserdriver/collection.properties`, kde *ID\_kolekce* identifikuje kolekci, kterou chcete konfigurovat.
- Nastavte následující parametry na hodnotu `true`:  
`treve.tokenizer.newHtmlTagReplacement=true`

Pravidla náhrady pro následující značky HTML se změní na níže uvedené hodnoty. Veškerá ostatní pravidla náhrady značek HTML nadále používají výchozí pravidla.

```
dfn    EMPTY
div    NEW LINE
q      BLANK
span   EMPTY
```

- Chcete-li, aby změny vstoupily v platnost, spusíte monitorování analyzátoru kolekce pomocí konzoly pro správu podnikového vyhledávání a poté analyzátor zastavíte a znovu spusíte.

#### Související úlohy

“Přiřazení typů dokumentů k analyzátoru kolekce” na stránce 145

#### Související odkazy

“Výchozí pravidla náhrady kódu HTML”

## Výchozí pravidla náhrady kódu HTML

Analyzátor HTML interpretuje a nahrazuje značky v dokumentech HTML, aby vytvořil textové informace pro vyhledávací index podnikového vyhledávání.

Analyzátor HTML pro podnikové vyhledávání používá při analýze značek HTML následující substituční pravidla. První sloupec obsahuje název značky HTML. Druhý sloupec obsahuje hodnotu použitou pro náhradu.

```
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

#### Související úlohy

“Změna pravidel náhrady pro některé značky HTML” na stránce 148

---

## Přiřazení typů dokumentů k analyzátoru Stellent

Chcete-li určit, které typy dokumentů mají být analyzovány pomocí dokumentových filtrů Stellent, musíte vytvořit konfigurační soubor `stellentypes.cfg`. Konzola pro správu podnikového vyhledávání neposkytuje žádnou podporu této úlohy.

#### Oracle Outside In Technology:

Technologie filtrování dokumentů Stellent nyní vlastní společnost Oracle. Odkazy na technologii Stellent v tomto dokumentu jsou ekvivalenty odkazů na technologii Outside In Content Access společnosti Oracle.

#### Omezení

Produkt OmniFind Enterprise Edition podporuje filtry Stellent pro následující typy dokumentů:

- Adobe PDF (Portable Document Format)
- Lotus 1-2-3
- Lotus Freelance Graphics
- Lotus Word Pro
- Ichitaro
- Microsoft Excel
- Microsoft PowerPoint
- Microsoft Visio
- Microsoft Word
- RTF (Rich Text Format)

Chcete-li do indexu podnikového vyhledávání zahrnout další typy dokumentů a formáty těchto dokumentů jsou podporovány filtry Stellent, můžete konfigurovat pravidla analýzy v konfiguračním souboru `stellentypes.cfg`.

**Důležité:** Přidané formáty dokumentů, které se neshodují s typy dokumentů uvedenými v předchozím seznamu, nebyly testovány pro podnikové vyhledávání a nejsou podporovány.

Úplný seznam formátů dokumentů Stellent naleznete v části *Outside In Technology: Supported File Formats* v příručce [http://www.oracle.com/technology/products/content-management/oit/ds\\_oitFiles.pdf](http://www.oracle.com/technology/products/content-management/oit/ds_oitFiles.pdf).

#### Než začnete



Chcete-li dokončit tuto úlohu, musíte se přihlásit jako administrátor podnikového vyhledávání.

### O této úloze

Konfigurační soubor `stellenttypes.cfg` určuje následující nastavení:

- Přijmout pravidla pro typy souborů, které mají být analyzovány pomocí analyzátoru Stellent. Typ souboru odpovídá jednomu z typů souborů rozpoznávaných knihovnou Stellent.
- Nativní pravidla pro typy souborů, které mají být odeslány zpět analyzátoru kolekce ke zpracování pomocí některého z vestavěných analyzátorů. Tuto akci je nutné nastavit pro případ, že analyzátor kolekce odešle dokument do analyzátoru Stellent omylem, v důsledku nesprávných metadat.
- Zamítnout pravidla pro typy souborů, které mají být odmítnuty, protože nejsou podporovány v podnikovém vyhledávání.

Pokud byl konfigurační soubor zadán, ale neexistuje, analyzátor se nepodaří spustit. Pokud nebyl pro vlastnost `OutsideInSupportedTypes` v souboru `stellent.properties` zadán žádný konfigurační soubor, budou použita výchozí pravidla analýzy pro analyzátor Stellent.

V konfiguračním souboru jsou uvedeny typy dokumentů a způsob, jakým mají být zpracovány. Formát souboru je tvořen posloupností řádků, z nichž každý obsahuje jedno pravidlo, které se shoduje s jedním z následujících formátů:

```
accept DEFAULT
accept ALL typ_dokumentu
accept typ_stellent typ_dokumentu
native DEFAULT
native typ_stellent typ_dokumentu
reject typ_stellent
```

kde:

*typ\_dokumentu*

Hodnota, která má být použita pro prvek dotazu `doctype`. Dokumenty lze vyhledávat podle jejich typů. Chce-li uživatel prohledávat například dokumenty PDF, může zadat hodnotu `$doctype::pdf`.

*typ\_stellent*

Jedna z hodnot typu filtru v knihovně Stellent, jako například hodnota `FI_123R1`.

### DEFAULT

Znamená, že seznam přijímaných nebo nativních typů (v závislosti na typu pravidla) zahrnuje všechna výchozí pravidla. Tato volba vám umožňuje rozšířit výchozí konfiguraci namísto jejího nahrazení.

**All** Znamená, že mají být přijaty všechny typy s daným prvkem `doctype`, které nejsou v seznamu explicitně uvedeny.

Pravidla v konfiguračním souboru jsou zpracovávána takto:

1. Existuje-li pro typ *typ\_stellent* pravidlo `reject`, nebude tento dokument přijat.
2. Pokud pro typ *typ\_stellent* existuje nativní pravidlo (včetně výchozích pravidel analýzy, jestliže bylo zadáno pravidlo `native DEFAULT`), je dokument zaslán zpět vestavěnému analyzátoru spolu s hodnotou prvku *typ\_dokumentu* uvedenou v použitém pravidle. Prvek *typ\_dokumentu* musí mít hodnotu `txt` (prostý text), `htm` (formát HTML) nebo `xml` (formát XML).
3. Existuje-li pravidlo `accept` pro typ *typ\_stellent* (včetně výchozího seznamu, byla-li zadána volba `accept DEFAULT`), bude tento dokument přijat.

4. Nejsou-li splněny předchozí podmínky a bylo-li zadáno pravidlo `accept ALL`, bude tento dokument přijat.
5. V opačném případě je dokument odmítnut a nebude analyzován.

Jestliže je typ dokumentu přijat, bude použita hodnota `typ_dokumentu` uvedená v použitém pravidle. Tato hodnota je zaslána zpět analyzátoru kolekce spolu s analyzovaným obsahem.

## Postup

Chcete-li přiřadit typy dokumentů k analyzátoru Stellent, postupujte takto:

1. Přihlaste se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
2. Otevřete pro úpravy soubor `ES_NODE_ROOT/master_config/ID_kolekce.stellent/stellent.properties`, kde `ID_kolekce` identifikuje kolekci, kterou chcete konfigurovat.
3. Jako hodnotu vlastnosti `OutsideInSupportedTypes` zadejte absolutní cestu k vytvářenému konfiguračnímu souboru.

Můžete například vytvořit následující konfigurační soubor pro jednotlivou kolekci a uložit jej společně s dalšími specifickými soubory kolekci:

```
ES_NODE_ROOT/master_config/collection_ID.stellent/stellenttypes.cfg
```

Jiným příkladem může být vytvoření následujícího konfiguračního souboru určujícího stejné nastavení pro všechny kolekce. Tento soubor uložíte společně s dalšími soubory na úrovni systému. Pokud použijete tento přístup, nezapomeňte uvést příslušnou cestu v souboru `stellent.properties` pro každou kolekci, jak je uvedeno v kroku 2.

```
ES_INSTALL_ROOT/default_config/stellent/stellenttypes.cfg
```

4. Pomocí textového editoru vytvořte konfigurační soubor a zadejte pravidla analýzy Stellent. Soubor uložte a zavřete.
5. Chcete-li, aby změny vstoupily v platnost, spusťte monitorování analyzátoru kolekce pomocí konzoly pro správu podnikového vyhledávání a poté analyzátor zastavte a znovu spusťte.

## Příklady

Při použití následujícího konfiguračního souboru relace Stellent přijímá kromě podporovaných typů dokumentů ve výchozím seznamu také dokumenty ve formátu Microsoft Visio.

```
accept DEFAULT
accept FI_VISI03 visio
accept FI_VISI04 visio
accept FI_VISI05 visio
accept FI_VISI06 visio
```

Při použití následujícího konfiguračního souboru budou akceptovány dokumenty ve formátu Postscript a bude je možné prohledávat jako dokumenty typu `ps`, soubory formátu `X pixmap` (XPM) budou zasilány zpět vestavěnému textovému analyzátoru, dokumenty ve formátu obrázků `PNG` budou odmítnuty a všechny ostatní typy souborů budou přijaty a bude je možné prohledávat jako dokumenty typu `other`.

```
accept DEFAULT
accept FI_POSTSCRIPT ps
native FI_XPIXMAP txt
reject FI_PNG
accept ALL other
```

### Související úlohy

“Přiřazení typů dokumentů k analyzátoru kolekce” na stránce 145

### Související odkazy

## Výchozí pravidla analýzy pro analyzátoři Stellent

Nevytvoříte-li konfigurační soubor pro mapování typů souborů na filtry dokumentů Stellent, budou pro analýzu dokumentů použita výchozí pravidla.

Analyzátor Stellent přijímá a analyzuje následující typy dokumentů:

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

Analyzátor Stellent vrací analyzátoru kolekce následující typy dokumentů ke zpracování pomocí některého z vestavěných analyzátorů:

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_CYRILLICKOI8	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_KOI8R	htm
NATIVE FI_HTML_CYRILLIC_MAC	htm
NATIVE FI_HTML_CYRILLIC_WINDOWS	htm
NATIVE FI_HTML_CYRILLIC1251	htm
NATIVE FI_HTML_CYRILLICKOI8	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_JAPANESEEUC	htm
NATIVE FI_HTML_JAPANESEJIS	htm
NATIVE FI_HTML_JAPANESESJIS	htm
NATIVE FI_HTML_KOREAN_JOHAB	htm
NATIVE FI_HTML_KOREAN_WINDOWS	htm
NATIVE FI_HTML_KOREANHANGUL	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_HTMLUNICODE	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_CHINESEEUC	xml
NATIVE FI_WML_CHINESEGB	xml
NATIVE FI_WML_CYRILLIC1251	xml
NATIVE FI_WML_CYRILLICKOI8	xml
NATIVE FI_WML_JAPANESEEUC	xml
NATIVE FI_WML_JAPANESEJIS	xml
NATIVE FI_WML_JAPANESESJIS	xml
NATIVE FI_WML_KOREANHANGUL	xml
NATIVE FI_WML_LATIN2	xml
NATIVE FI_XHTML	htm

NATIVE FI_XL2KHTML	htm
NATIVE FI_XML	xml
NATIVE FI_XML_DOCTYPE_HTML	htm

### **Související úlohy**

“Přiřazení typů dokumentů k analyzátoru Stellent” na stránce 150

---

## Podpora jazyků a kódových stránek

Lingvistické zpracování u podnikového vyhledávání probíhá v analyzátoru jinak než na vyhledávacích serverech.

Pro účely lingvistického zpracování analyzátor nerozlišuje mezi jazyky a lokalitami. Pokud však uživatel prohledává kolekci obsahující dokumenty ve více jazycích, vyhledávací servery umožňují omezit výsledky vyhledávání pouze na konkrétní jazyk nebo lokalitu.

Pokud je například v metadatech anglického dokumentu uvedena lokalita `en_US`, je dokument indexován jednak jako dokument v angličtině (`en`) a jednak jako dokument používající anglickou lokalitu USA (`en_US`). Tento typ indexování umožňuje správně reprezentovat informace ve formátu specifickém pro lokalitu, například čísla, kalendářní data a časové údaje. Když uživatel prohledává kolekci, bude daný dokument nalezen bez ohledu na to, zda uživatel hledá dokumenty s označením `en` nebo `en_US`.

Pokud je dokument indexován pouze kódem jazyka, např. `en`, je k indexování použit pouze kód jazyka a nikoli lokalita. Vyhledávají-li uživatelé například dokumenty s označením `en_US`, daný dokument nebude nalezen.

System podnikového vyhledávání poskytuje jazykovou podporu pro následující jazyky a dvouznakové kódy jazyka zdokumentované ve standardu ISO 639:

### Jazyky s jednoduchým textem:

- en=angličtina
- sq=albánština
- az=ázerbájdžánština - latinka
- bg=bulharština
- be=běloruština
- ca=katalánština
- hr=chorvatština
- cs=čeština
- da=dánština
- nl=holandština
- et=estonština
- fi=finština
- fr=francouzština
- de=němčina
- el=řečtina
- hu=maďarština
- is=islandština
- id=indonéština
- in=indonéština
- it=italština
- kk=kazaština
- lv=lotyština
- lt=litevština
- lo=laoština
- mk=makedonština
- ms=malajština
- mt=maltština
- no=norština
- nb=norština (Bokmal)
- pl=polština
- pt=portugalština
- ro=rumunština
- ru=ruština
- sr=srbština (azbuka)
- sh=srbština (latinka)

sk=slovenština  
sl=slovinština  
es=španělština  
sv=švédština  
tr=turečtina  
uk=ukrajinština  
cy=velština

#### **Ideografické jazyky:**

Pro zjednodušenou a tradiční čínštinu se používají rozšířené kódy jazyků namísto dvouznakových kódů.

zh-CN=čínština (zjednodušená)  
zh-TW=čínština (tradiční)  
ja=japonština  
ko=korejština

#### **Jazyky se složitým textem:**

ar=arabština  
as=assaměština  
bn=bengálština  
gu=gujaratština  
iw=hebrejština  
he=hebrejština  
hi=hindština  
kn=kannadština  
ml=malabarština  
mr=maráthština  
or=urijština  
pa=pandžábština  
ta=tamilština  
te=telugština  
th=thajština  
ur=urdština  
vi=vietnamština

System podnikového vyhledávání nabízí automatickou detekci mnoha z těchto jazyků a kódových stránek používaných v dokumentech ve formátu prostého textu. Pokud chcete při konfigurování prolézacího modulu sami určit použitý jazyk nebo kódovou stránku, můžete automatickou detekci jazyka a kódové stránky zakázat.

---

## **Automatická detekce jazyka**

System podnikového vyhledávání může zpracovávat dokumenty prakticky v jakémkoli jazyce.

Pokud je dokument psán v některém z následujících jazyků, systém je schopen detekovat jazyk automaticky. Pokud víte, v jakém jazyce jsou dokumenty napsány, můžete určit použitý jazyk při konfigurování prolézacího modulu a nepoužívat automatickou detekci jazyka systémem.

Angličtina  
Arabština  
Bulharština  
Čeština  
Čínština, tradiční  
Čínština, zjednodušená  
Dánština  
Finština  
Francouzština, Kanada  
Francouzština, národní  
Hebrejština  
Holandština  
Islandština  
Italština



Japonština  
Korejština  
Maďarština  
Němčina, národní  
Němčina, Švýcarsko  
Norština, Bokmal  
Polština  
Portugalština, Brazílie  
Portugalština, národní  
Rumunština  
Ruština  
Řečtina  
Španělština  
Švédština  
Thajština  
Turečtina

---

## Automatická detekce kódové stránky

System podnikového vyhledávání podporuje mnoho různých kódových stránek dokumentů.

U textových souborů může systém automaticky detekovat následující kódové stránky. U dokumentů v jiných formátech systém používá k detekci kódové stránky metadata obsažená v dokumentu, například prvky metadat HTML. Pokud víte, v jaké kódové stránce jsou dokumenty uloženy, můžete určit použitou kódovou stránku při konfigurování prolézacího modulu a nepoužívat automatickou detekci kódové stránky systémem.

### Formáty kódování Unicode:

UTF-8  
UTF-16BE  
UTF-16LE

### Formáty vícebajtového kódování:

Shift-JIS  
ISO-2022-CN  
ISO-2022-JP  
ISO-2022-KR  
GB18030  
EUC-JP  
EUC-KR

### Formáty jednobajtového kódování:

ISO-8859-1: angličtina, dánština, francouzština, italština, holandská, němčina, norština, portugalština, španělština, švédština  
ISO-8859-2: čeština, maďarština, polština, rumunština  
ISO-8859-5: ruština  
ISO-8859-6: arabština  
ISO-8859-7: řečtina  
ISO-8859-8: hebrejština, hebrejština ve vizuálním pořadí  
ISO-8859-9: turečtina  
Windows-1250: čeština, maďarština, polština, rumunština  
Windows-1251: ruština  
Windows-1252: angličtina, dánština, francouzština, italština, holandská, němčina, norština, portugalština, španělština, švédština  
Windows-1253: řečtina  
Windows-1254: turečtina  
Windows-1255: hebrejština  
Windows-1256: arabština  
KOI8-R: ruština

Detekce znakové sady je nepřesná operace. Proces detekce kódové stránky se pokusí identifikovat znakovou sadu, která nejlépe odpovídá charakteristikám bajtových dat, využívá přitom zčásti statistických metod, u nichž nelze zaručit správnost výsledků.

Aby bylo možné dosáhnout co nejvyšší přesnosti, měla by být vstupní data převážně v jednom jazyce. Je také nutné mít k dispozici nejméně několik set bajtů prostého textu.

Pokud není detekované kódování podporováno, systém použije pro danou kolekci výchozí kódovou stránku.

---

## Lingvistická analýza dokumentů v čínštině, japonštině a korejštině

Chcete-li zlepšit možnosti načítání dokumentů psaných v čínštině, japonštině a korejštině, můžete zadat volby lingvistické analýzy.

U čínských, japonských a korejských dokumentů můžete určit, zda má analyzátor používat k lexikální analýze metodu n-gramové segmentace. U čínských a japonských dokumentů lze analyzátor rovněž konfigurovat tak, aby z prázdného prostoru odebíral znaky nového řádku.

### N-gramová segmentace

Při vytváření kolekce volíte typ lexikální analýzy, který chcete použít při zpracování dokumentů psaných v jazycích, jež pro oddělování slov nepoužívají mezerové znaky.

Segmentace na základě mezer ve formátu Unicode používá jako oddělovač mezi slovy mezeru. N-gramová segmentace považuje překrývající se posloupnosti libovolného počtu znaků za jediné slovo. U jazyků, jako je čínština, japonština nebo korejština, které nepoužívají jako oddělovače mezery, může n-gramová segmentace vrátit lepší výsledky vyhledávání než segmentace na základě mezer ve formátu Unicode.

Metodu segmentace, kterou chcete použít k analýze dokumentů, nastavujete při vytvoření kolekce. Po vytvoření kolekce můžete nastavené volby analýzy zobrazit a prohlédnout si je, nemůžete je však změnit.

Informace o konfiguraci podpory pro úplnou n-gramovou analýzu a převod na prvky v kolekcích podnikového vyhledávání a zpracování znaků v kolekcích konfigurovaných pro úplnou n-gramovou podporu najdete v dokumentu <http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088>.

### Odebrání mezerových znaků z textu

Analyzátor lze nakonfigurovat tak, aby z textu odebíral mezerové znaky.

#### Než začnete

Chcete-li dokončit tuto úlohu, musíte se přihlásit jako administrátor podnikového vyhledávání.

#### O této úloze

Povolíte-li pro kolekci tuto volbu, analyzátor bude odebírat posloupnosti mezerových znaků oddělující dvě písmena. Mezerové znaky může být vhodné odebrat například z dokumentů psaných v jazyku, který nepoužívá mezery k oddělování slov, jako je čínština nebo japonština.

Když konfiguruje analyzátor pro odebírání mezerových znaků, můžete určit, zda chcete odebírat pouze mezerové znaky, které se vyskytují mezi znaky z dvoubajtové znakové sady (DBCS), nebo všechny mezerové znaky bez ohledu na kontext. Druhou možností můžete použít například tehdy, pracujete-li s japonským dokumentem obsahujícím anglický text a chcete-li mezerové znaky odebrat také z anglického textu.

Analyzátor odebírá následující znaky:

- tabelátor (0x09),
- LF - nový řádek (0x0A),
- CR - návrat vozíku (0x0D).

Změna se projeví po zastavení a novém spuštění analyzátoru. Chcete-li změnu použít v dokumentech, které již byly uloženy do indexu, proveďte nové prolezení dokumentů a poté znovu sestavte hlavní index.

### Postup

Chcete-li z textu odebrat mezerové znaky, postupujte takto:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
2. Otevřete následující soubor pro úpravy v textovém editoru. *ID\_kolekce* je ID kolekce zadané (nebo přiřazené systémem) při vytvoření kolekce:  
`ES_NODE_ROOT/master_config/ID_kolekce.parserdriver/collection.properties`
3. Určete požadovaný způsob odebrání mezerových znaků:
  - Chcete-li odebírat mezerové znaky mezi znaky DBCS, nastavte vlastnost `removeCjNewlineChars` na hodnotu `true`:  
`removeCjNewlineChars=true`
  - Chcete-li odebírat všechny mezerové znaky, ať se v dokumentu vyskytují kdekoli, nastavte vlastnost `removeCjNewlineChars` na hodnotu `true` a vlastnost `removeCjNewlineCharsMode` na hodnotu `all`:

```
removeCjNewlineChars=true
removeCjNewlineCharsMode=all
```



---

## Administrace indexů

Aby měli uživatelé zajištěn trvalý přístup k nejaktuálnějším informacím, podnikové vyhledávání vytváří pro každou kolekci index a udržuje jej pravidelnou aktualizací jeho obsahu.

Aby bylo možné prohledávat data shromážděna prolézacími moduly, je nutné vytvořit indexy. Při prvotním vytvoření kolekce podnikové vyhledávání vytvoří index pro všechna data, která byla na začátku prolezena. V průběhu zpracování nových a změněných zdrojů dat vytvářejí prolézací moduly aktualizace (*rozdílový index*) pro nový obsah. Aktualizace je po určité době nutné sloučit se základním indexem. Tento proces sloučení se nazývá sestavení *hlavního indexu*. Při každém sestavení rozdílového indexu nebo hlavního indexu je nový obsah zkopírován na vyhledávací servery a zpřístupněn pro vyhledávání.

Prolézací moduly shromažďují data trvale nebo na základě pravidelného časového plánu. Častou aktualizací indexů umožníte uživatelům vyhledávat v nejaktuálnějších datech. Průběžně aktualizovaný index je nutné po určité době znovu sestavit. Zvětšující se index spotřebovává stále více systémových prostředků. V zájmu udržení optimálního výkonu sestavujte hlavní index v pravidelných intervalech.

Četnost sestavování hlavního indexu závisí na následujících faktorech:

- systémové prostředky (místo v souborovém systému, rychlost procesoru a velikost paměti),
- počet dokumentů, které je třeba prolézat a vracet se k nim,
- typ prolézáných dat,
- četnost změn pravidel kategorií (změny se projeví až po sestavení hlavního indexu),
- četnost, s jakou je prolézací modul spouštěn ručně a nikoli v naplánovanou dobu,

U kolekcí obsahujících několik miliónů dokumentů, sestavovaných zejména z webových dokumentů, by měl být hlavní index sestaven přibližně jednou denně a aktualizace indexu by měla probíhat každou hodinu nebo každé dvě hodiny.

Při udržování indexu ve stále aktuálním stavu umožňujícím vyhledávání se používají následující postupy:

- nastavení plánu sestavování indexu,
- změna plánu indexu,
- povolení a zákaz plánu indexu,
- konfigurování souběžného sestavování indexů.

Pokud potřebujete zadat volby ovlivňující chování indexu z pohledu uživatele, můžete provést také následující kroky:

- konfigurování podpory pro použití zástupných znaků v dotazech,
- konfigurování oborů omezujících rozsah dokumentů, které smí uživatelé prohledávat,
- sbalení dokumentů pocházejících ze stejného zdroje ve výsledcích vyhledávání,
- odebrání identifikátorů URI z indexu.

### Související úlohy

“Monitorování aktivity indexu pro kolekci” na stránce 298

“Monitorování fronty indexů podnikového vyhledávání” na stránce 298

---

## Plánování sestavení indexu

Pro sestavení hlavního indexu a aktualizaci indexu novým obsahem můžete určit časové plány.

### Než začnete

Chcete-li plánovat sestavení indexu, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

Chcete-li uživatelům vždy zajistit přístup k nejnovějším informacím z prohledávaných zdrojů, naplánujte pravidelné sestavování indexu. Při sestavování hlavního indexu je celý index sestaven znovu. Procesy indexování načtou všechna data shromážděná prolézacími moduly a analyzovaná modulem analýzy. Při sestavování rozdílového indexu jsou pro prohledávání zpřístupněny informace získané prolézáním od posledního sestavení hlavního indexu.

Při výchozím nastavení je volba plánování sestavení indexu vybrána. Tato volba určuje, že proces plánovače naplánuje úlohy sestavování hlavního indexu a rozdílových indexů při spuštění systému podnikového vyhledávání. Pokaždé, když budete chtít zabránit naplánovanému spuštění sestavování indexu, můžete zrušit zaškrtnutí políčka **Povolit při spuštění systému**. Plán můžete chtít zakázat například při odstraňování problémů.

Z důvodů šetření prostředků a zvýšení výkonu systém automaticky kontroluje, zda došlo ke změnám, které je třeba promítnout do indexu. Pokud není v indexu třeba provádět žádné změny, naplánovaný požadavek na sestavení indexu je zrušen.

### Postup

Naplánování sestavování indexu:

1. Upravte kolekci, vyberte stránku Index a klepněte na volbu **Plánovat sestavení indexu**.
2. Chcete-li určit, jak často má být index aktualizován s použitím nového obsahu, zadejte na stránce Časový plán sestavení indexu v oblasti **Zadat plán sestavení rozdílového indexu** následující volby:
  - a. V oblasti **Čas spuštění** zadejte do polí **Rok**, **Měsíc**, **Den**, **Hodina** a **Minuta** datum a čas sestavení prvního rozdílového indexu.
  - b. V oblasti **Interval aktualizace** určete pomocí polí pro **dny**, **hodiny** a **minuty** frekvenci sestavování rozdílových indexů.

Obvykle je vhodné sestavovat rozdílové indexy často, například každou hodinu nebo každé dvě hodiny. V závislosti na tom, jak často se mění zdrojový obsah, můžete zadat delší nebo kratší interval. Můžete nastavit například hodinový interval (0 dnů a 1 hodina) nebo dvanáctihodinový interval (0 dnů a 12 hodin).
3. Chcete-li určit, jak často má být index zcela sestaven znovu, určete v oblasti **Zadat plán sestavení hlavního indexu** následující volby:
  - a. V oblasti **Čas spuštění** zadejte do polí **Rok**, **Měsíc**, **Den**, **Hodina** a **Minuta** datum a čas prvního sestavení hlavního indexu.
  - b. V oblasti **Interval aktualizace** určete pomocí polí pro **dny**, **hodiny** a **minuty** frekvenci sestavování hlavního indexu.

Obvykle je vhodné sestavovat hlavní index pravidelně, například každých 24 hodin. V závislosti na tom, jak často se mění zdrojový obsah, můžete zadat delší nebo kratší interval. Můžete nastavit například 12hodinový interval (0 dnů a 12 hodina) nebo interval dvou a půl dne (2 dny a 12 hodin).

4. Klepněte na tlačítko **OK**.

## Změna časového plánu indexu

Plán sestavování indexu můžete v případě potřeby změnit.

### Než začnete

Chcete-li změnit časový plán indexu, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo být administrátorem dané kolekce.

### Postup

Chcete-li změnit časový plán indexu, postupujte takto:

1. Upravte kolekci, vyberte stránku Index a změňte příslušné hodnoty v polích **Měsíc**, **Den**, **Rok** a **Hodina**. Určete, jak často má být index aktualizován a jak často má být sestavován hlavní index.
2. Klepněte na tlačítko **Použít**.

## Povolení a zákaz časových plánů indexování

Časové plány sestavení indexu lze povolit a zakázat.

### Než začnete




Chcete-li povolit nebo zakázat plán indexu, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo být administrátorem dané kolekce.

### O této úloze

Plán indexu můžete zakázat, pokud potřebujete zabránit spuštění naplánovaného sestavení indexu. Zakázáním plánu můžete například zabránit sestavení indexu v naplánované datum a čas, abyste mohli vyřešit vzniklý problém.

Plán můžete povolit nebo zakázat během úprav kolekce a během monitorování kolekce.

### Postup

1. Chcete-li povolit nebo zakázat plán indexu v rámci úprav kolekce, postupujte takto:
  - a. Upravte kolekci, kterou chcete změnit.
  - b. Na stránce Index zaškrtněte políčko **Povolit při spuštění systému**, chcete-li povolit časový plán aktualizace indexu, nebo zaškrtnutí tohoto políčka zrušte, chcete-li plán zakázat.
  - c. Výběrem nebo zrušením výběru políčka **Povolit při spuštění systému** povolte nebo zakažte plán sestavování hlavního indexu.
  - d. Klepněte na tlačítko **Použít**.
2. Chcete-li povolit nebo zakázat plán indexu v rámci monitorování kolekce, postupujte takto:
  - a. Monitorujte kolekci, kterou chcete změnit.
  - b. Je-li naplánováno sestavení indexu a nechcete-li, aby toto sestavení proběhlo v naplánované datum a čas, klepněte na stránce Index na ikonu  **Zakázat časový plán**. Index nebude sestaven, dokud plán nepovolíte nebo nespustíte proces sestavení indexu klepnutím na ikonu  **Spustit**.
  - c. Je-li index naplánován, ale časový plán pro sestavení je zakázán, klepněte na volbu  **Povolit časový plán**.

Index bude zařazen do fronty pro sestavení k datu a času zadanému v jeho časovém plánu.

## Konfigurování souběžných sestavení indexu

Používání prostředků indexace můžete řídit na základě určení počtu kolekcí, pro které mohou být souběžně zpracovávány požadavky na sestavení indexu. Máte-li dostatek systémových prostředků, můžete zvýšit kvalitu vyhledávání povolením aktualizace indexu souběžně se sestavením hlavního indexu.

### Než začnete

Chcete-li pro systém určit volby sestavování indexu, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### O této úloze

System podnikového vyhledávání může sestavovat několik indexů současně díky sdílení prostředků mezi kolekcemi, což umožňuje paralelní zpracování požadavků na sestavování indexu pro více kolekcí. Tím, že procesy budete sdílet, můžete zajistit, že sestavování rozsáhlého indexu nebude blokovat dostupnost ostatních indexů, které čekají ve frontě na sestavení.

Jakmile je požadováno nebo naplánováno sestavení indexu, je index zařazen do fronty indexů a čeká na zpracování. Protože má každá kolekce svůj vlastní index, může fronta indexů obsahovat několik požadavků na sestavení indexu z různých kolekcí současně. Při konfiguraci voleb indexování pro systém můžete určit, pro kolik kolekcí mohou být současně se sdílením prostředků indexování zpracovávány požadavky.


Můžete rovněž určit, že požadavky na aktualizaci indexu mají být zpracovávány souběžně se sestavováním hlavního indexu kolekce. Pokud tuto volbu povolíte, budou na vyhledávacích serverech aktualizovány údaje týkající se nejnovějších dokumentů (prostřednictvím rozdílového indexu), zatímco bude zpracováno sestavení hlavního indexu, které probíhá pomaleji. Sestavování indexu je však proces náročný na prostředky. Při sestavování indexu je využíváno velké množství systémové paměti a prostoru na disku. Pokud povolíte tuto volbu a máte nedostatek místa na disku nebo paměti, může dojít ke snížení celkového výkonu systému.

Zvýšíte-li počet souběžných sestavení indexu, nedojde k automatickému spuštění požadavků na sestavení indexů, které již čekají ve frontě. Změna ovlivní další požadavky na sestavení indexu, které budou zařazeny do fronty po změně této hodnoty.

Snížíte-li počet souběžných sestavení indexu, nedojde k automatickému zastavení zpracovávaných požadavků na sestavení indexů. Změna se projeví po dokončení aktuálních sestavení indexů a uvolnění prostoru pro spuštění požadavků čekajících ve frontě.

### Postup

Chcete-li určit volby pro sestavování indexu pro systém, postupujte takto:

1. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
3. Na stránce **Index** klepněte na volbu **Konfigurovat volby vytváření indexu**.
4. Na stránce **Volby indexování** na úrovni systému zadejte počet kolekcí, pro které mohou být sdíleny systémové prostředky souběžně se zpracováním požadavků na sestavení indexu.



Počet kolekcí, které sdílejí prostředky indexování, nesmí být vyšší než počet kolekcí v systému podnikového vyhledávání. Pokud máte například pět kolekcí, je nutné zadat počet menší nebo roven pěti.

5. Máte-li k dispozici dostatek systémových prostředků pro podporu více souběžných operací sestavování indexu pro jednotlivé kolekce, můžete vybrat volbu, která povolí souběžné sestavování rozdílových indexů a hlavního indexu.

**Omezení:** Pokud vyberete tuto volbu v době, kdy probíhá sestavení indexu pro některou kolekci, sestavení indexu se nemusí podařit. Před povolením souběžného sestavování rozdílových indexů a hlavního indexu zakažte všechna naplánovaná sestavení indexů a počkejte, než se aktuálně spuštěná sestavení dokončí (nebo je případně předčasně ukončete). Po zapnutí této volby naplánovaná sestavení indexů opět povolte.

---

## Sestavení indexů pouze při detekci změn

U časového plánu sestavení indexu dochází k sestavení pouze při detekci změn, které je nutné použít na index. Pomocí příkazu **startIndexBuild** můžete spustit sestavení indexu a vybrat, zda má systém zjišťovat změny.

### Než začnete

Chcete-li ručně zadat, že k sestavení indexu má dojít pouze při detekci změn, musíte se přihlásit jako administrátor podnikového vyhledávání.

### Omezení

Systém nezjišťuje změny, které by mohly ovlivnit index, dokud nebyl hlavní index sestaven alespoň dvakrát. Tato implementace používá změny (které nelze zjistit, dokud nebyl hlavní index sestaven alespoň dvakrát), jako jsou například změny definic oboru nebo pravidel sbalení výsledků vyhledávání.

### O této úloze

Sestavení indexu spotřebovává systémové prostředky, jako například výkon procesoru, paměť a šířku pásma I/O. Z důvodů šetření prostředků a zvýšení výkonu může systém podnikového vyhledávání detekovat, zda došlo ke změnám, které je nutné použít na index, a sestavit index pouze v případě nutnosti použití změn. Systém může detekovat následující události úpravy indexu:

#### Pouze sestavení hlavního indexu

- Odebrání identifikátorů URI podle vzoru, což odebere dokumenty z indexu.
- Odebrání prolézacího modulu ze systému, což odebere dokumenty z indexu.

#### Sestavení hlavního a rozdílového indexu

- Úprava definic oboru nebo pravidel sbalení výsledků vyhledávání.
- Přidání nebo odebrání dokumentů přímo prostřednictvím nastavení prolézacího modulu nebo rozhraní Push API.
- Úprava faktorů upřednostnění k ovlivnění statických skóre ohodnocení důležitosti dokumentů v indexu.
- Úprava pravidel použití zástupných znaků v dotazech.

Detekce změn se vztahuje na všechny požadavky časového plánu sestavení indexu. Sestavení indexu je prováděno pouze v případě existence nevyřízených změn, které je nutné použít na index. Pokud není detekována žádná změna, dojde ke zrušení požadavku na sestavení indexu.

Chcete-li vynutit sestavení indexu bez ohledu na změny, můžete je spustit pomocí konzoly pro správu podnikového vyhledávání. Dále můžete pomocí příkazu **startIndexBuild** ručně spustit sestavení indexu a zadat, zda má systém před zpracováním požadavku zjišťovat změny.

### Postup

Chcete-li spustit sestavení indexu pomocí příkazového řádku a požadovat sestavení pouze při zjištění změn, postupujte takto:

1. Přihlaste se jako administrátor podnikového vyhledávání. V konfiguraci s více servery se přihlaste k indexovému serveru.
2. Zadejte následující příkaz:

```
esadmin controller startIndexBuild -volby
```

Volby:

**-cid** *ID\_kolekce*

ID kolekce, pro kterou chcete spustit sestavení indexu.

**-buildType** *typ\_sestavení*

Určuje typ indexu k sestavení. Platnými hodnotami jsou main (hlavní) a delat (rozdílový).

**-detectChanges**

Volitelné. Způsobuje spuštění testů zjišťování změn pro požadavek sestavení indexu. Sestavení indexu je prováděno pouze v případě zjištění změn, které je nutné použít na index.

V následujícím příkladu dojde ke zpracování požadavku na spuštění sestavení hlavního indexu kolekce col\_1, pouze pokud jsou zjištěny změny, které je nutné použít na index:

```
esadmin controller startIndexBuild -cid col_1 -buildType main -detectChanges
```

---

## Zastavení sestavení indexu

Příkaz **stopIndex** lze použít k zastavení sestavení hlavního nebo rozdílového indexu, aby nebylo potřeba používat konzolu pro správu podnikového vyhledávání.

### Postup

Postup při zastavení sestavení indexu pomocí příkazu místo použití konzoly pro správu:

1. Přihlaste se jako administrátor podnikového vyhledávání. V instalaci s více servery se přihlaste k indexovému serveru.
2. Zadejte následující příkaz:

```
esadmin stopIndex -cid ID_kolekce -buildType typkde:
```

**-cid** *ID\_kolekce*

Určuje ID kolekce, která vlastní index.

*typ* Určuje typ zastavovaného sestavení indexu. Platné hodnoty jsou main (hlavní) a delat (rozdílový).

### Příklad

```
esadmin stopIndex -cid col1 -buildType delta
```

---

## Volby ovlivňující prohledatelné zobrazení indexu

Po vytvoření indexu pro dokumenty můžete zadat volby určující, jakým způsobem mohou uživatelé dokumenty vyhledávat a zobrazovat ve výsledcích vyhledávání.

Pokud potřebujete zadat volby ovlivňující chování indexu z pohledu uživatele, můžete provést následující kroky:

- Konfigurování podpory zástupných znaků v dotazovacích výrazech. Podporu dotazů se zástupnými znaky můžete zabudovat do indexu nebo použít volby expanze dotazovacích výrazů během zpracování dotazů.
- Konfigurování oborů omezujících rozsah dokumentů, které smí uživatelé prohledávat. Když pak uživatelé prohledávají kolekci, nehledají v celém indexu, ale pouze v dokumentech patřících do daného oboru.
- Sbalení dokumentů pocházejících ze stejného zdroje ve výsledcích vyhledávání. Dokumenty se shodným identifikátorem URI nebo vzorem URI lze v indexu seskupit a zobrazit ve výsledcích vyhledávání pouze první výsledné dokumenty (uživatelé mohou sbalené výsledné dokumenty zobrazit výběrem příslušných voleb).
- Odebrání některých identifikátorů URI z indexu. V případě potřeby můžete uživatelům dočasně zabránit v prohledávání některých dokumentů v indexu.

Pro některé typy prolézacích modulů a pro kolekce, které nepovolují zabezpečení, se používá zjišťování duplicitních dokumentů, aby se zabránilo zobrazení stejných nebo skoro stejných dokumentů ve výsledcích vyhledávání.

## Volby vytváření indexu pro prohledávání dokumentů

Při konfigurování voleb pro prohledávání prolezených dat nebo při mapování prvků XML a prvků metadat HTML na vyhledávací pole můžete určit způsob prohledávání dokumentů a jejich zobrazování ve výsledcích vyhledávání.

Volby vyhledávání, které určíte, jsou uloženy spolu s dokumenty v indexu. Umožňují vám omezit cíle dotazů uživatelů a položky, které se uživatelům zobrazí ve výsledcích vyhledávání.

### Volby prolézacího modulu:

Při konfigurování prolézacího modulu pro prolézání zdrojů dat, které obsahují pole, můžete určit následující volby týkající se určování, zda lze pole prohledávat, jakým způsobem je lze prohledávat a zda je lze vracet ve výsledcích vyhledávání:

- Volné vyhledávání textu
- Vyhledávání s použitím polí
- Úplná shoda
- Lze řadit
- Parametrické vyhledávání
- Výsledky vyhledávání
- Obsah dokumentu

### Volby mapování polí XML a HTML:

Při konfigurování analyzátoru a určování, že mají být prvky XML a prvky metadat HTML mapovány na prohledávací pole v indexu, můžete nastavit následující volby:

- Vyhledávání s použitím polí
- Úplná shoda
- Lze řadit
- Výsledky vyhledávání

Pokud konfiguruje volby vyhledávání pro konkrétní prvky metadat HTML (tj. nikoli pro všechny prvky nebo pro prvky obsažené v sadě prvků metadat Dublin Core), můžete mimo jiné povolit prohledávání polí, která obsahují číselné hodnoty, pomocí parametrických dotazů.

## Volné vyhledávání textu

Index podnikového vyhledávání je realizován jako plnotextový index s obsahem pocházejícím z různých zdrojů dat. Obsah lze prohledávat zadáním jednoduchého dotazu v přirozeném jazyce. Vyhledávací procesy procházejí pole a obsah dokumentů a hledají dokumenty relevantní pro daný dotaz.

Chcete-li povolit prohledávání polí pomocí dotazů na volný text, zaškrtněte při konfigurování prolézacího modulu políčko **Prohledávání volného textu**. Chcete-li prohledávat pole s názvy, klíčovými slovy a popisy, zaškrtněte také políčko **Vyhledávání s použitím polí**.

### Příklad 1:

Volné vyhledávání textu může mít například podobu následujícího jednoduchého dotazu:

řetěz na kolo

Chcete-li určit, která slova se v dokumentu musí nebo naopak nesmí objevit, můžete použít speciální způsoby zápisu. Chcete-li například stanovit, že za vyhovující mají být považovány pouze dokumenty obsahující určité slovo, uveďte před tímto slovem znaménko plus (+). Chcete-li z výsledků vyhledávání vyloučit dokumenty obsahující určité slovo, vložte před toto slovo znaménko mínus (-). Chcete-li vyhledat určitou frázi přesně v zadané podobě, uzavřete dvě nebo více slov do uvozovek (").

### Příklad 2:

U následujícího volného textového dotazu nastane shoda jen tehdy, obsahuje-li dokument frázi **science fiction** přesně tak, jak je zapsána, a neobsahuje-li slovo **robot**:

+ "science fiction" -robot

## Vyhledávání s použitím polí

Vyhledávání s použitím polí vám umožňuje omezit dotaz pouze na konkrétní datová pole a pole metadat v dokumentu. Můžete například určit, že zadaná slova se musí vyskytnout v názvu dokumentu.

Chcete-li povolit prohledávání polí podle jejich názvů, zaškrtněte políčko **Vyhledávání s použitím polí** při konfigurování prolézacího modulu nebo při konfigurování voleb mapování polí pro prvky XML a HTML. Chcete-li prohledávat pole s názvy, klíčovými slovy a popisy, zaškrtněte také políčko **Prohledávání volného textu**.

### Příklad:

Chcete-li v prostředí podnikového vyhledávání vyhledávat s použitím polí, uveďte v dotazu název pole a slovo nebo frázi, které se v tomto poli musí vyskytovat.

Následující dotaz hledá dokumenty obsahující slovo **ibm** a frázi podnikové vyhledávání v poli názvu:

title:ibm title:"podnikové vyhledávání"

## Úplná shoda

Vyhledávání úplných shod může zkvalitnit výsledky vyhledávání tím, že vám umožní zadávat přesně cílené dotazy. V případě vyhledávání úplných shod můžete zadávat dotazy na pole a prvky XML a načítat pouze dokumenty, v nichž podmínkám dotazu vyhovuje celá hodnota pole nebo celá hodnota prvku XML. Pokud hodnota v poli nebo prvku obsahuje méně obsahu či více obsahu, nedochází ke shodě.

**Tip:** Při hledání úplných shod systém převede všechna písmena v podmínkách dotazu na malá, z řetězců dotazů odebere přebytečné mezery a provede porovnání se vzorem se zástupnými znaky. Neprobíhá však rozklad na hesla ani vyhledávání synonym a nejsou odebrána zakázaná slova. K úplné shodě prvku XML je třeba, aby název prvku bez dalších vnořených prvků a celá hodnota prvku přesně odpovídaly podmínkám dotazu.

Úplné shody můžete v polích vyhledávat pomocí syntaxe dotazů podnikového vyhledávání (SI-API) nebo syntaxe dotazů XMLFrag2 (dotazy XPath nejsou podporovány). Požadavek na vyhledání úplné shody lze zadat vložení znaku rovná se (=) před podmínky dotazu.

Chcete-li povolit vyhledávání úplných shod v polích, zaškrtněte políčko **Úplná shoda** při konfigurování prolézacího modulu nebo při konfigurování voleb mapování polí pro prvky XML a HTML. Povolíte-li uživatelům při konfigurování voleb analýzy pro kolekci prohledávat dokumenty XML pomocí nativního vyhledávání XML, lze úplné shody s podmínkami vyhledávání hledat ve všech prvcích XML.

### Příklad 1:

Pole s názvem `barva` obsahuje hodnotu `tmavě modrá`.

- Následující dotaz na úplnou shodu toto pole nalezne, protože dotaz neobsahuje žádné další podmínky:  
`barva:="tmavě modrá"`
- Následující dotaz na úplnou shodu pole nenalezne, protože kromě slov `tmavě modrá` obsahuje ještě slovo `sukně`:  
`barva:="tmavě modrá sukně"`
- Následující dotaz na úplnou shodu pole nenalezne, protože pole `barva` obsahuje navíc slovo `tmavě`:  
`barva:="modrá"`

### Příklad 2:

Bez použití úplné shody může následující dotaz XMLFrag2 vrátit dokumenty obsahující položku `<diagnóza>intraduktální karcinom typu komedokarcinom</diagnóza>` nebo jiné pojmy, které podmínkám dotazu přesně neodpovídají:

```
@xmlf2::'<diagnóza>intraduktální karcinom</diagnóza>'
```

Při použití úplné shody bude zaručeno, že následující dotaz XMLFrag2 vrátí pouze dokumenty, u nichž celý obsah hodnoty prvku XML splňuje podmínky dotazu:

```
@xmlf2::'=<diagnóza>intraduktální karcinom</diagnóza>'
```

## Lze řadit

Pokud zdroj dat obsahuje pole nebo pokud uživatelé prohledávají dokumenty XML či HTML, může být vhodné povolit řazení výsledků podle hodnot v určitém poli. V ukázkové aplikaci pro podnikové vyhledávání jsou uvedeny názvy všech polí, pro něž byla nakonfigurována

možnost řazení. Uživatelé se mohou rozhodnout seřadit výsledky abecedně (podle řazení řetězců) podle některého z uvedených polí a nikoli podle relevance nebo data dokumentu. Dále mohou uživatelé určit, zda mají být dokumenty seřazeny vzestupně nebo sestupně.

Výsledné dokumenty, které neobsahují pole pro řazení, jsou uvedeny na konci výsledků vyhledávání. Výsledné dokumenty, které obsahují pole pro řazení, ale byly indexovány před tím, než bylo pole konfigurováno pro řazení, jsou také uvedeny na konci výsledků vyhledávání.

Chcete-li uživatelům umožnit abecední řazení výsledků vyhledávání podle hodnot určitého pole, zaškrtněte políčko **Lze řadit** při konfigurování prolézacího modulu nebo při konfigurování voleb mapování polí pro prvky XML a HTML. Pokud pole obsahuje číselné hodnoty, můžete zaškrtnutím políčka **Parametrické vyhledávání** určit, že hodnoty pole lze použít k numerickému seřazení výsledků vyhledávání.

## Parametrické vyhledávání

Parametrické vyhledávání je speciální typ vyhledávání s použitím polí, při němž můžete zadávat porovnávací a vyhodnocovací dotazy na číselná a kalendářní pole a metadata. Můžete například vyhledat dokumenty určité velikosti nebo dokumenty vytvořené po určitém datu. Rovněž můžete vyhledávat dokumenty s atributy, které jsou větší či menší než zadaná hodnota nebo se jí rovnají.

Chcete-li řadit výsledky numericky podle hodnoty pole, musíte u pole povolit parametrické vyhledávání.

Chcete-li prohledávat pole pomocí parametrických dotazů nebo řadit výsledky numericky, zaškrtněte políčko **Parametrické vyhledávání** při konfigurování prolézacího modulu nebo voleb mapování polí pro konkrétní prvky metadat HTML.

### Příklad 1:

Následující dotaz vyhledává položky, které stojí přesně 50 korun (nebo jiných měnových jednotek indexovaných v poli *cena*):

```
#cena:=50
```

### Příklad 2:

Následující dotaz vyhledává dokumenty s velikostí souboru větší než 1024, ale menší nebo rovnou 2048:

```
#filesize::>1024<=2048
```

## Výsledky vyhledávání

Můžete se rozhodnout, že některá pole chcete prohledávat, ale nechcete je zobrazit ve výsledcích vyhledávání. Můžete také chtít zobrazit ve výsledcích vyhledávání určité pole i přesto, že jste na něj nezadali dotaz. Může být například potřebné zadat dotaz na finanční data, abyste obdrželi smysluplnou sestavu, ve výsledcích obsahujících jména zaměstnanců však nemusí být vhodné zobrazit mzdy zaměstnanců.

Chcete-li umožnit zobrazení pole ve výsledcích vyhledávání, zaškrtněte políčko **Výsledky vyhledávání** při konfigurování prolézacího modulu nebo voleb mapování polí pro prvky XML a HTML.

## Obsah dokumentu

U některých typů dokumentů, jako jsou například webové dokumenty, je za obsah považován celý dokument. U jiných typů dokumentů, například u dokumentů obsahujících pole, můžete určit, která pole nesou užitečný obsah a která metadata.

Chcete-li určit, že v daném poli je uložen obsah dokumentu, zaškrtněte při konfigurování prolézacího modulu políčko **Obsah dokumentu**. Jsou-li zaškrtnuta políčka **Obsah dokumentu** i **Prohledávání volného textu**, bude hodnota pole použita pro zjišťování duplicitních dokumentů a stane se součástí oblasti dynamického shrnutí dokumentu ve výsledcích vyhledávání.

### Související pojmy

 [Syntaxe dotazů](#)

## Detekce duplicitních dokumentů

Detekce duplicitních dokumentů je technika používaná k zabránění tomu, aby výsledky vyhledávání obsahovaly více dokumentů se stejným nebo skoro stejným obsahem.

Je-li ve výsledcích vyhledávání uvedeno více kopií stejných (nebo skoro stejných) dokumentů, může dojít ke snížení kvality vyhledávání. K analýze duplicitních dokumentů dochází pouze v případě splnění obou následujících podmínek:

- Kolekce používá model ohodnocení důležitosti založený na odkazech. Tento model se vztahuje na prolézací moduly, které prolézají webové stránky, jako je například webový prolézací modul nebo prolézací modul WebSphere Portal.
- Zabezpečení kolekce je vypnuto.

V průběhu globální analýzy proces indexování detekuje duplikáty pomocí procházení obsahu jednotlivých dokumentů. Mají-li dva dokumenty stejný obsah, jsou považovány za duplikáty.

Chcete-li, aby byla při analýze detekce duplikátů brána v úvahu i metadata dokumentu, zaškrtněte při konfiguraci prolézacích modulů pro kolekci políčko **Obsah dokumentu** a zadejte volby prolézání metadat. V tomto případě prolézací modul prolézá pole metadat jako obsah dokumentů a zahrnuje metadata do analýzy obsahu duplicitních dokumentů. K podobné analýze dochází při konfiguraci voleb analýzy dokumentů HTML a XML a při zaškrtnutí políčka **Obsah dokumentu**.

Zadáte-li, že pole nebo pole metadat tvoří obsah dokumentu, je obsah těchto polí přidan do dynamického souhrnu dokumentu ve výsledcích vyhledávání, což může ovlivnit, zda dojde k zobrazení dokumentu ve výsledcích vyhledávání. Při povolení detekce téměř duplicitních dokumentů ve vyhledávací aplikaci (vlastnost `NearDuplicateDetection` v metodě `setProperty` je nastavena na hodnotu `ANO`), dojde k potlačení dokumentů s podobnými názvy a souhrny v zobrazení výsledků vyhledávání. Klepnutím na odkaz lze téměř duplicitní potlačené dokumenty zobrazit.

Ve skupině duplicitních dokumentů je jeden dokument hlavní a ostatní jsou duplikáty. Všechny dokumenty ve skupině duplikátů mají stejnou kanonickou reprezentaci obsahu. V průběhu indexování dojde k indexování obsahu (tokenů) hlavního dokumentu. U duplicitních dokumentů jsou indexovány pouze tokeny metadat. Při odstranění hlavního dokumentu z indexu se hlavním dokumentem stane jeden z duplikátů. Při prohledávání kolekce dojde k vrácení pouze hlavního dokumentu.

### Související pojmy

“Analýza duplicitních dokumentů a zabezpečení kolekce” na stránce 247

## Zástupné znaky v dotazech

V případě potřeby můžete uživateli povolit použití zástupných znaků v dotazovacích výrazech a hledání slov, která odpovídají určitému vzoru.

Dotazovací výraz se zástupným znakem je výraz obsahující znak \* (hvězdička). Když uživatel odešle dotaz obsahující zástupný znak, budou do výsledků vyhledávání zahrnuty všechny dokumenty v indexu, které odpovídají dotazovacímu výrazu, spolu se všemi dokumenty v indexu, které odpovídají vzoru reprezentovanému zástupným znakem. Koncový zástupný znak ve výrazu dotazu `sea*` může odpovídat například slovům `search`, `season` a `seals`.

Při nastavování voleb zástupných znaků pro index určete, zda chcete uživateli umožnit použití zástupných znaků v dotazech a pokud ano, jakým způsobem má být tato podpora poskytována:

- Pro vyhledávání slov shodujících se se vzorem zástupných znaků v dokumentu můžete povolit všechny části dokumentu nebo můžete hledání shody se vzorem omezit na pole.
- Podporu dotazů obsahujících zástupné znaky můžete povolit pro všechna pole nebo můžete hledání shody se vzorem omezit na určená pole.
- Výskyt zástupného znaku můžete omezit na poslední pozici ve výrazu dotazu (koncový zástupný znak). Můžete rovněž povolit výskyt zástupného znaku kdekoli ve výrazu dotazu. (Zástupný znak se nemůže vyskytovat v názvu pole.)
- V závislosti na povolených pozicích pro zástupné znaky můžete určit způsob expanze výrazů dotazu (výrazy dotazu obsahující zástupné znaky jsou expandovány na všechny výrazy v indexu, které jim odpovídají). Veškeré možné expanze výrazů mohou být uloženy v indexu nebo mohou vyhledávací procesy expandovat výrazy během zpracování dotazu.

Změny nastavení pro práci se zástupnými znaky se projeví po příštím sestavení hlavního indexu.

### Expanze indexu

Chcete-li zahrnout expanzi výrazů do indexu, určete, kolik počátečních znaků slova se musí shodovat se vzorem zástupných znaků ve výrazu dotazu, aby došlo ke shodě. Výsledky budou vráceny pouze pro výrazy dotazu obsahující alespoň tento počet znaků (kromě znaku \*). Zadáte-li například hodnotu 4, musí výraz dotazu obsahovat alespoň čtyři znaky, aby došlo ke shodě.

Zadáte-li hodnotu 4, bude se slovo `technologie` shodovat s výrazem dotazu `tech*` a výrazem dotazu `techno*`, nikoli však s výrazem dotazu `te*`.

Během sestavování rozdílového nebo hlavního indexu jsou spolu s původními výrazy indexovány všechny potenciální rozšířené tvary jednotlivých výrazů v dokumentu. Výhodou tohoto přístupu je skutečnost, že při zpracování dotazu již expanze výrazů nezabere žádný čas navíc. Zároveň však roste velikost indexu, takže je nutné zajistit dostatek systémových prostředků pro rozsáhlý index.

Tento přístup je výhodný zejména u relativně malých kolekcí a tam, kde jsou prostor a doba potřebná k sestavení indexu méně důležité než rychlost odezvy na dotazy. Příkladem je prohledávání katalogu nebo adresáře zaměstnanců.

Tento přístup je k dispozici pouze v případě, že povolíte podporu pro koncové zástupné znaky. Pokud povolíte podporu pro zástupné znaky, které se mohou ve výrazu dotazu vyskytovat kdekoli, nebudete moci vybrat volbu zahrnutí expanzí výrazu do dotazu.



## Expanze dotazů

V případě expanze dotazů a použití pravidel pro porovnávání se vzorem, když uživatel odešle dotaz obsahující zástupné znaky, je třeba určit, kolik variant vyhledávacího výrazu tvoří shodu. Zadáte-li například hodnotu 50, bude za shodné s výrazem uvedeným v dotazu považováno až 50 variant tohoto výrazu.

Tento příklad ilustruje dotazovací výraz **tech\***, který se shoduje se slovy **technický**, **technika**, **technologie**, apod., celkem až s 50 různými slovy, která začínají znaky **tech**.

Expanze dotazů má sice jen malý vliv na velikost indexu, může však zpomalit zpracování dotazů. Vyhledávací procesy musí projít všechny existující varianty vyhledávacího výrazu se zástupnými znaky až od limitu, který určíte v nastavení zástupných znaků.

Tento přístup je výhodný zejména u relativně velkých kolekcí a v případech, kdy je třeba minimalizovat nároky na prostor a dobu potřebnou k sestavení indexu. Příkladem je vyhledávání v úložištích e-mailů, kde se index musí přizpůsobovat rychlým změnám v dokumentech, zatímco doba odezvy na dotazy je méně důležitá.

Tento přístup je k dispozici bez ohledu na to, zda povolíte podporu pro koncové zástupné znaky nebo zda povolíte podporu pro zástupné znaky, které se mohou vyskytovat kdekoli ve výrazu dotazu.

## Podpora pro zástupné znaky v dotazech

Sada expanzí pro výrazu dotazu se zástupným znakem obsahuje všechny výrazy indexu, které lze získat nahrazením zástupného znaku libovolnou posloupností znaků. Sada je určena následujícím způsobem:

- Pokud kolekce podporuje zástupné znaky, které se ve výrazu dotazu mohou vyskytovat kdekoli, je každý výraz obsahující hvězdičku interpretován jako výraz se zástupným znakem.
- Sada obsahuje nejvýše maximální počet expanzí povolený administrátorem podnikového vyhledávání. Pokud index obsahuje více expanzí, než určuje tato hodnota, budou tyto expanze ignorovány. (Ve výsledcích vyhledávání je uvedeno, zda byly některé expanze zástupných znaků ignorovány.)
- Je-li podpora zástupných znaků omezena na sadu polí, bude sada obsahovat pouze výrazy, které se vyskytují v jednom z určených polí. Výraz se musí vyskytovat pouze v jednom poli nejméně v jednom dokumentu v indexu.
- Je-li výraz dotazu výrazem pole, musí být zástupný znak uveden po specifikačtoru pole (příklad: `nazevpole:*sphere`). Název pole nesmí obsahovat dvojtečku (:).
- Je-li podpora zástupných znaků omezena na sadu polí, musí být název pole ve výrazu dotazu se zástupným znakem jedním z polí určených v konzole pro správu podnikového vyhledávání. V opačném případě nebudou pro příslušný výraz nalezeny žádné expanze.
- Zástupné znaky jsou podporovány pouze pro výrazy ve formátu prostého textu, nikoli pro názvy prvků XML, názvy atributů či hodnoty atributů. Výraz sestávající pouze ze zástupného znaku není podporován.

## Vliv zástupných znaků na index

Podpora zástupných znaků založená na expanzi indexu zvětšuje index a prodlužuje dobu potřebnou k sestavení indexu.

Při expanzi indexu jsou kromě pojmu samotného indexovány také všechny předpony tohoto pojmu. Pro pojem **podpora** jsou například indexovány následující řetězce:

p po pod podp podpo podpor podpora

Počet pojmů uložených v indexu roste o násobek daný průměrnou délkou slova. Komprimací lze velikost indexu snížit, ale nijak významně. Doba potřebná k sestavení indexu roste s průměrnou délkou slova.

Index anglických dokumentů se zvětšuje přibližně čtyřikrát, protože délka průměrných anglických slov se pohybuje od pěti do šesti znaků. Index s n-gramovými tokeny se zvětšuje přibližně dvakrát, protože každý n-gram obsahuje dva znaky.

Expanze indexu pro podporu zástupných znaků se doporučuje použít v následujících situacích:

- Kolekce je dostatečně malá, takže nárůst spotřeby prostoru a času způsobený expanzí indexu nepovede k problémům s výkonem.
- Pro splnění požadavků uživatele (nebo podniku) jsou do výsledků vyhledávání zahrnuty všechny existující expanze zástupných znaků.

Vliv na index můžete zmírnit zadáním minimální délky předpony při expanzi zástupných znaků. Je-li například minimální délka předpony nastavena na hodnotu 3, nebudou pro slovo **podpora** indexovány předpony **p** a **po** a velikost indexu pro angličtinu klesne ze čtyřnásobné na trojnásobnou.

Při použití podpory zástupných znaků metodou expanze dotazů nejsou do indexu zapisovány žádné předpony. Expanze pojmů se provádí při zadání dotazu a velikost indexu se zvětšuje pouze o malou datovou strukturu potřebnou pro podporu tohoto typu expanze. Index s podporou expanze dotazů je obvykle o 10 až 20 % větší než index bez podpory zástupných znaků a doba sestavení indexu se prodlužuje o méně než 10 %. Konfigurace maximálního počtu expanzí nemá na velikost indexu ani na dobu jeho sestavování žádný vliv.

## Konfigurování voleb pro použití zástupných znaků v dotazech

Pokud konfiguruje volby indexování pro kolekci podnikového vyhledávání, můžete zadat, zda chcete uživatelům povolit používání zástupných znaků ve výrazech dotazů.

### Než začnete

Chcete-li konfigurovat volby zástupných znaků, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, do které daný index náleží.

### O této úloze

Určíte-li volby zástupných znaků, projeví se změny po příštím sestavení hlavního indexu.

### Postup

Konfigurování podpory pro použití zástupných znaků v dotazech:

1. Upravte kolekci, vyberte stránku Index a klepněte na volbu **Konfigurovat volby pro zástupné znaky**.
2. Na stránce Volby pro zástupné znaky zaškrtněte políčko **Podpora zástupných znaků v dotazech**.
3. Volitelné: Můžete určit, zda chcete v dotazech vyhledávajících volný text podporovat zástupné znaky. Dotaz na volný text **tech\***, který neprohledává pojmenované pole, například vrátí rozbalené výsledky (například výrazy **technologie** nebo **technika**) pouze v případě, že je toto políčko zaškrtnuto.
4. Určete, která pole podporují zástupné znaky:

- Chcete-li určit, že v dotazech prohledávajících pole nelze používat zástupné znaky, vyberte volbu **Žádná pole**.
  - Chcete-li povolit podporu dotazů obsahujících zástupné znaky pro všechna pole v dokumentu, vyberte volbu **Všechna pole**.
  - Chcete-li omezit podporu pro zástupné znaky na některá pole, vyberte volbu **Specifická pole** a zadejte názvy polí. Rozbalené výsledky budou vráceny pouze pro určená pole. Dotaz `author:john*` například vrátí rozbalené výsledky pouze v případě, že pole `author` podporuje zástupné znaky.
5. Určete, zda se musí zástupný znak vyskytovat na poslední pozici výrazu dotazu (koncový zástupný znak), nebo zda jeho použití není omezeno a může se ve výrazu dotazu vyskytovat kdekoli.
- Vyberete-li umístění a typ zástupného znaku, musíte rovněž určit, zda chcete povolit podporu pro zástupné znaky. Podrobné informace zobrazíte klepnutím na tlačítko **Nápověda** v konzole pro správu.

## Obory

Konfigurované obory využijete v případě, že potřebujete uživatelům poskytnout omezený pohled na kolekci.

Obor je skupina souvisejících identifikátorů URI v indexu. Konfigurováním oboru omezíte množinu dokumentů, kterou uživatelé v kolekci vidí. Když pak uživatelé prohledávají kolekci, nehledají v celém indexu, ale pouze v dokumentech z daného oboru. Chcete-li tuto funkci použít, vaše vyhledávací aplikace musí podporovat obory vyhledávání.

Při vytvoření oboru zadáte rozsah identifikátorů URI v indexu, který mohou uživatelé prohledávat. Omezením dokumentů, v nichž mohou uživatelé vyhledávat, se zajistí, že dokumenty ve výsledcích vyhledávání se budou vztahovat přímo k vyhledávaným informacím.

Můžete například vytvořit jeden obor, který obsahuje identifikátory URI pro oddělení technické podpory a jiný obor zahrnující identifikátory URI pro personální oddělení. Pokud vyhledávací aplikace podporuje obory, budou uživatelé v oddělení technické podpory načítat dokumenty z oboru pro technickou podporu, zatímco uživatelé v personálním oddělení budou načítat dokumenty z oboru pro personální oddělení.

Můžete vytvořit libovolný počet oborů, i když vytvoření příliš mnoha oborů může ovlivnit výkon. Konfigurujte obory tak, aby většina vyhledávacích požadavků musela filtrovat maximálně jeden nebo dva obory. Vzhledem k tomu, že obory mohou obsahovat celé identifikátory URI nebo vzory identifikátorů URI, může stejný dokument patřit k více oborům.

Pokud konfiguruje obory, mohou se změny v některých případech projevit až po dvojitým sestavení hlavního indexu. Pokud konfiguruje obory před prvním sestavením hlavního indexu kolekce, uživatelé budou moci prohledávat kolekci, ale nebudou moci ve výsledcích vyhledávání zobrazit údaje o oboru. Opakovaným sestavením hlavního indexu zajistíte, že výsledky vyhledávání budou odpovídat rozsahu identifikátorů URI v daném oboru.

Konfigurujete-li obory po sestavení hlavního indexu, začnou změny platit při nejbližším dalším sestavení hlavního indexu.

## Konfigurování oborů

Při konfigurování oboru pro kolekci podnikového vyhledávání můžete určit identifikátory URI nebo vzory identifikátorů URI pro rozsah dokumentů v indexu, v nichž uživatelé mohou vyhledávat.

## Než začnete

Chcete-li konfigurovat obory, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, do níž obory náležejí.

## O této úloze

Pokud vyhledávací aplikace podporuje použití oborů, uživatelé mohou při prohledávání kolekce vyhledávat pouze dokumenty odpovídající identifikátorům URI, které definují hranice oboru.

Pokud konfiguruje obory, mohou se změny v některých případech projevit až po dvojitým sestavení hlavního indexu. Pokud konfiguruje obory před prvním sestavením indexu, uživatelé budou moci prohledávat kolekci, ale nebudou moci ve výsledcích vyhledávání zobrazit údaje o oboru. Opakovaným sestavením hlavního indexu zajistíte, že výsledky vyhledávání budou odpovídat rozsahu identifikátorů URI v daném oboru.

Konfiguruje-li obory po sestavení hlavního indexu, začnou změny platit při nejbližším dalším sestavení hlavního indexu.

## Postup

Konfigurování oboru:

1. Upravte kolekci, vyberte stránku Index a klepněte na volbu **Konfigurovat obory**.
2. Na stránce Obory klepněte na volbu **Vytvořit obor**.
3. Zadejte název oboru, identifikátory URI a vzory identifikátorů URI, které definují hranice daného oboru. Můžete rovněž zadat identifikátory URI a vzory identifikátorů URI, které chcete z oboru vyloučit.
4. Klepněte na tlačítko **OK**.

Nový obor se zobrazí na stránce Obory společně s ostatními obory, které náležejí do dané kolekce.

### Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 113

## Sbalené identifikátory URI

V prostředí podnikového vyhledávání lze výsledky vyhledávání uspořádat tak, že dokumenty ze zdrojů se stejnou předponou identifikátoru URI budou ve výsledcích vyhledávání sbaleny.

Jsou-li výsledky sbaleny, první výsledek se obvykle zobrazuje zcela vlevo. Výsledky s nižším výsledným ohodnocením jsou seskupeny a odsazeny na řádku pod prvním výsledkem.

Chcete-li sbalit jako jednu skupinu dokumenty s různými předponami identifikátorů URI, můžete předpony příslušných identifikátorů URI přiřadit k názvu skupiny, kterou vytvoříte. Používáte-li například tři různé servery pro správu finančních dat, můžete ve výsledcích vyhledávání seskupit dokumenty ze všech tří serverů a výsledky s nižším ohodnocením sbalit pod prvním výsledným dokumentem.

Vyhledávací aplikace mohou používat předponu identifikátoru URI nebo název skupiny ke sbalení dokumentů ve výsledcích vyhledávání. V ukázkové aplikaci pro podnikové vyhledávání se zobrazují první dva výsledné dokumenty. Jsou-li vráceny více než dva výsledné dokumenty se stejnou předponou identifikátoru URI (nebo dokumenty patřící do stejné skupiny předpon identifikátorů URI), můžete sbalené výsledky zobrazit výběrem příslušné volby.

Uživatelé mohou na základě syntaxe dotazů podnikového vyhledávání (`samegroupas:předpona_identifikátoru_URI`) prohledávat všechny dokumenty, které jsou obsaženy ve stejné skupině jako předpona identifikátoru URI uvedená v dotazu.

## Uspořádání předpon identifikátorů URI a názvů skupin

Při konfigurování pravidel pro sbalení výsledků vyhledávání pomocí konzoly pro správu zadáváte předpony identifikátorů URI dokumentů, které chcete sbalit, a můžete také přiřadit předpony identifikátorů URI k názvu skupiny.

Pořadí, v němž předpony identifikátorů URI konfiguruje, má svůj význam. Indexový server používá pořadí předpon identifikátorů URI k výpočtu hodnot jednotlivých identifikátorů URI v kolekci. Pro každý identifikátor URI jsou provedeny následující kroky:

1. Indexový server prohledává předpony identifikátorů URI uvedené v pravidlech pro sbalení výsledků vyhledávání postupně.
2. Nalezne-li indexový server v indexu první předponu identifikátoru URI, která se shoduje s předponou dokumentu, přidá název skupiny (nebo předponu identifikátoru URI, pokud pravidlo neurčuje název skupiny) k výrazům vyhledávaným v dokumentech.

Pokud nelze webový dokument přiřadit k předponě identifikátoru URI, použije indexový server jako předponu identifikátoru URI název hostitele odpovídající dané adrese URL. Pokud nelze k předponě identifikátoru URI přiřadit dokument NNTP, indexový server použije jako předponu identifikátoru URI první ID zprávy v hodnotě referenční hlavičky.

Po přidání předpony identifikátoru URI do seznamu předpon, které mají být ve výsledcích vyhledávání sbaleny, musíte tuto předponu identifikátoru URI umístit podle toho, v jakém pořadí ji má indexový server použít, a případně ji přiřadit jako zvláštní vyhledávací výraz k dokumentům v indexu:

- Pokud přidáte předponu identifikátoru URI a nepřijedíte ji k názvu skupiny, můžete tuto předponu identifikátoru URI individuálně vybrat a přesunout v seznamu nahoru nebo dolů.
- Pokud přidanou předponu identifikátoru URI přiřadíte k názvu skupiny, bude se spolu s ní přesouvat v seznamu nahoru nebo dolů celá skupina předpon identifikátorů URI. Na pořadí předpon identifikátorů URI v rámci skupiny nezáleží - výběrem jedné předpony identifikátoru URI automaticky vyberete celou skupinu.

## Sbalení identifikátorů URI ve výsledcích vyhledávání

Můžete určit volby pro seskupení a sbalení výsledných dokumentů ze zdrojů, které mají stejnou předponu identifikátoru URI. Můžete rovněž vytvořit název skupiny, který umožňuje společné sbalení výsledných dokumentů s různými předponami identifikátorů URI.

### Než začnete

Chcete-li zadat volby pro sbalení výsledků vyhledávání, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze



Změny voleb pro sbalení výsledků vyhledávání se projeví až po příštím sestavení hlavního indexu.

### Postup

Určení voleb pro sbalení výsledků vyhledávání:

1. Upravte kolekci, vyberte stránku Index a klepněte na volbu **Sbalit výsledky vyhledávání**.

2. Na stránce Sbalit výsledky vyhledávání klepněte na volbu **Přidat předponu identifikátoru URI**.
3. Na stránce Přidat předponu identifikátoru URI pro sbalení výsledků zadejte předponu identifikátoru URI pro dokumenty, které chcete ve výsledcích vyhledávání sbalit. Příklad:  

```
http://finance/ROI/
http://server1.com/finance/
db2://LOCALDB/SCHEMA1.TABLE1/
exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/
```
4. Můžete zadat také popisný název skupiny, kterou chcete asociovat s danou předponou identifikátoru URI. Chcete-li sbalit výsledné dokumenty z více zdrojů jako jednu skupinu, zadejte při přidávání jednotlivých předpon identifikátoru URI stejný název skupiny.
5. Klepněte na tlačítko **OK**.
6. Na stránce Sbalit výsledky vyhledávání přesuňte nové pravidlo na místo odpovídající pořadí, ve kterém je má indexový server použít:
  - Pokud jste přidali předponu identifikátoru URI a neasociovali ji s názvem skupiny, bude nová předpona identifikátoru URI zobrazena na konci seznamu. Pomocí kláves se šipkami ji přesuňte na požadované místo.
  - Pokud jste novou předponu identifikátoru URI asociovali s názvem skupiny, bude nová předpona identifikátoru URI zobrazena na konci sady předpon identifikátoru URI, které náleží ke stejné skupině. Pomocí kláves se šipkami přesuňte celou skupinu předpon identifikátoru URI na požadované místo.
7. Chcete-li změnit předponu identifikátoru URI nebo název skupiny, vyberte předponu identifikátoru URI a klepněte na volbu  **Upravit**.
8. Chcete-li vzor identifikátoru URI odebrat ze seznamu, vyberte příslušný vzor identifikátoru URI a klepněte na volbu  **Odebrat**.

## Odebrání identifikátorů URI z indexu

Chcete-li uživatelům zabránit ve vyhledávání dokumentů v kolekci, můžete odebrat identifikátory URI pro tyto dokumenty z indexu.

### Než začnete

Chcete-li odebrat identifikátory URI z indexu, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

Zadáte-li úplný identifikátor URI, uživatelům se tento identifikátor URI přestane zobrazovat ve výsledcích vyhledávání. Pokud uživatel odešle stejný dotaz a výsledné dokumenty pro tento dotaz budou uloženy ve vyhledávací mezipaměti, potom dočasně uložená výsledná stránka pro odebraný identifikátor URI bude vrácena ve výsledcích vyhledávání. Vyhledávací mezipaměť bude aktualizována a identifikátor URI bude z indexu odebrán až při příštím sestavení hlavního nebo rozdílového indexu.

Zadáte-li vzor identifikátoru URI pro odebrání více identifikátorů URI, uživatelům se budou identifikátory URI, které odpovídají zadanému vzoru, ve výsledcích vyhledávání nadále zobrazovat, a to až do příštího sestavení hlavního indexu.

Odeberete-li identifikátor URI z indexu, neodeberete jej z prolézaného prostoru. Při příštím prolézání dokumentu prolézacím modulem bude identifikátor URI vložen do indexu a bude opět dostupný pro hledání. Chcete-li identifikátor URI odebrat z prolézaného prostoru, je třeba aktualizovat pravidla prolézání tak, aby vylučovala daný dokument, a potom prolézací modul zastavit a restartovat.

## Postup

Odebrání identifikátorů URI pro specifické dokumenty z indexu:

1. Upravte kolekci, vyberte stránku Index a klepněte na volbu **Odebrat identifikátory URI z indexu**.
2. Na stránce Odebrat identifikátory URI z indexu zadejte identifikátory URI (nebo vzory identifikátorů URI), které chcete odebrat z indexu.

Příklad:

```
http://domain.org/hr/*  
db2://knowledgeManagement/ROI*  
cm://enterprise/finance*
```

### Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 113





---

## Administrace vyhledávacího serveru

Volby, které lze zadat pro vyhledávací servery, zahrnují použití prostoru mezipaměti pro vrácení výsledků vyhledávání, určování maximální délky shrnutí dokumentů ve výsledcích vyhledávání, zvyšování kvality vyhledávání na základě asociování vlastních slovníků a vrácení předdefinovaných identifikátorů URI ve výsledcích vyhledávání vždy, když se v dotazu objeví určité výrazy.

Když uživatel odešle dotaz, vyhledávací servery použijí index k rychlému vyhledání odpovídajících dokumentů. K načítání metadat odpovídajících dokumentů využívají vyhledávací servery datový sklad podnikového vyhledávání obsahující analyzovaná data rozložená na prvky. Metadata mohou mimo jiné obsahovat identifikátor URI dokumentu, jeho název, popis, datum, typ dat apod.

Při konfigurování vyhledávacích serverů pro kolekci zadáte volby ovlivňující způsob zpracování dotazů včetně voleb, které mohou ovlivnit výkon dotazu:

### **Konfigurování mezipaměti vyhledávání**

Chcete-li optimalizovat výkon dotazů, můžete určit, že mají být výsledky vyhledávání (odpovědi na dotazy) ukládány do mezipaměti, a můžete nastavit velikost prostoru, který bude výsledkům vyhledávání v mezipaměti přidělen.

### **Konfigurování maximální délky zobrazených shrnutí dokumentů**

Pro většinu výsledných dokumentů se zobrazují shrnutí obsahu dokumentů, podle nichž se mohou uživatelé rozhodnout, který dokument načtou. Velikost prostoru ve výsledcích vyhledávání, který bude použit k zobrazení těchto souhrnných informací, můžete určit podle potřeby.

### **Zadání jiného výchozího jazyka**

Výchozí jazyk pro prohledávání dokumentů se nastavuje při vytvoření kolekce, v případě potřeby však můžete zadat jiný jazyk.

### **Asociování vlastních slovníků**

Pokud vývojáři aplikace vytvořili vlastní slovníky synonym, zakázaných slov nebo slov pro upřednostnění, můžete určit, které slovníky mají být použity při prohledávání kolekce uživateli.

### **Konfigurování rychlých odkazů**

V případě potřeby lze předem určit identifikátory URI, které mají být vráceny pro určitá klíčová slova a fráze. Když uživatel zadá v dotazu některé z těchto klíčových slov či frází, bude ve výsledcích vyhledávání vrácen předdefinovaný identifikátor URI. Identifikátory URI definované rychlými odkazy jsou vráceny spolu s identifikátory URI, které vyhledávací servery vrátí po prohledání indexu.

### **Související pojmy**

“Ohodnocení důležitosti dokumentů” na stránce 195

“Vlastní slovníky slov pro upřednostnění” na stránce 198

---

## Mezipaměti pro vyhledávání

Pokud je zatížení vyhledávacích serverů poměrně vysoké, lze jejich výkon zvýšit ukládáním výsledků vyhledávání do mezipaměti.

Při zpracování výsledků vyhledávání vyhledávací servery nejprve zkontrolují, zda již mezipaměť neobsahuje výsledky pro stejný dotaz. Pokud naleznou odpověď na příslušný

dotaz, mohou uživatelé rychle vrátit výsledky vyhledávání. Pokud vyhledávací servery nenaleznou odpověď na příslušný dotaz, prohledají index.

Po zaplnění mezipaměti vyhledávání jsou odstraňovány nejstarší výsledky vyhledávání a výsledky odpovídající méně často zadávaným dotazům, aby uvolnily místo novým výsledkům vyhledávání.

Z konzoly pro správu podnikového vyhledávání můžete povolit ukládání výsledků vyhledávání do mezipaměti a také zadat kapacitu mezipaměti (počet výsledků dotazů, které lze do mezipaměti současně uložit).

Změny nastavení mezipaměti pro vyhledávání vejdou v platnost až po restartování vyhledávacích serverů.

---

## Konfigurování mezipaměti vyhledávání

Pro kolekci lze povolit nebo zakázat mezipaměť vyhledávání. Lze také určit volby pro řízení velikosti mezipaměti vyhledávání.

### Než začnete

Chcete-li konfigurovat mezipaměť vyhledávání pro kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Postup

Konfigurování mezipaměti vyhledávání:

1. Upravte kolekci, vyberte stránku Vyhledat a klepněte na volbu **Konfigurovat volby vyhledávacího serveru**.
2. Na stránce Volby vyhledávacího serveru zaškrtněte políčko **Použít mezipaměť vyhledávání**.
3. Do pole **Maximální počet položek v mezipaměti** zadejte maximální počet odpovědí na dotaz, které lze uložit do mezipaměti vyhledávání.
4. Klepněte na tlačítko **OK**.
5. Chcete-li, aby se změny projevily, použijte funkci monitorování vyhledávacích serverů a restartujte procesy serveru.

---

## Vlastní slovníky synonym

Chcete-li zlepšit kvalitu výsledků vyhledávání, můžete uživatelům povolit vyhledávat při prohledávání kolekce synonyma výrazů uvedených v dotazu.

Vytvoříte-li slovník synonym, přidáte jej do systému podnikového vyhledávání a přiřadíte jej ke kolekci, mohou uživatelé při prohledávání kolekce vyhledávat dokumenty obsahující synonyma výrazů uvedených v dotazu. Díky tomuto rozšíření dotazů uživatelé snáze naleznou všechny dokumenty, které potřebují, a nikoli jen dokumenty přesně odpovídající výrazům zadaným v dotazu. Tím, že při vytvoření slovníku synonym definujete slova, která mají být považována za vzájemná synonyma, pomůžete uživatelům nalézt potřebné dokumenty, aniž by museli zadávat všechny varianty dotazovacího výrazu.

Vaše organizace může například používat zkratky označující oddělení, vybavení apod., nebo se v dokumentech ve vašich kolekcích mohou vyskytovat odborné výrazy z vašeho oboru. Vytvořením slovníku synonym můžete zajistit, že dotazy obsahující zkratku (např. ACL) vrátí dokumenty, v nichž je uveden význam této zkratky (např. ACL, seznam přístupových práv, řízení přístupu apod.).

Jazyk dotazů podnikového vyhledávání umožňuje uživatelům vyhledávat synonyma připojením operátoru vlnovka před dotazovací výraz. Dotaz ~WAS může například vrátit dokumenty, které se zabývají aplikačním serverem WebSphere. Vývojáři aplikací mohou zpřístupnit podporu synonym také prostřednictvím vlastností dotazu - tento přístup nevyžaduje použití speciální syntaxe.

Slovníky synonym obsahují varianty slov a mají následující charakteristické vlastnosti:

- Slova nejsou rozlišena podle jazyků, je však možné zadávat slova v různých jazycích. Pro každou kolekci existuje pouze jeden slovník synonym.
- Slova nejsou při vyhledávání skloňována a časována. Ve slovníku synonym je nutné uvést všechny gramatické tvary daného slova. U podstatného jména je tedy třeba uvést všechny tvary jednotného i množného čísla (např. seznam, seznamu, seznamem, seznamy, seznamů, seznamům, seznamech).

Většina výrazů, které přidáte do slovníku synonym, jsou přesné sémantické ekvivalenty, tzn. je-li výraz A synonymem výrazu B, je také výraz B synonymem výrazu A. Vždy, když je v dotazu použit výraz A, lze jej nahradit výrazem B a naopak.

Můžete však přidat také výrazy odpovídající různým použitím téhož pojmu včetně jeho obecnějších a konkrétnějších variant. Můžete například vytvořit jednu skupinu synonym obsahující výrazy **budova** a **dům** a jinou skupinu obsahující výrazy **banka**, **peněžní ústav** a **spořitelna**.

Čím volnější je vztah mezi jednotlivými výrazy, tím rozsáhlejší budou výsledky vyhledávání, některé výsledky vyhledávání však nemusí být z hlediska dotazu zajímavé. Vyhledávací a indexovací rozhraní API nabízí metody, které uživatelům umožňují vybrat při odeslání vyhledávacího požadavku vhodná synonyma, a metody umožňující uživateli zjistit, jaká synonyma byla odvozena z jednotlivých výrazů v dotazu.

Chcete-li vytvořit slovník synonym, je třeba, aby nejprve odborník na danou problematiku případně ve spolupráci s vývojářem aplikací vytvořil seznam synonym ve formátu XML. Soubor XML musí být převeden na binární soubor (.dic) pomocí nástroje podnikového vyhledávání **essyndictbuilder**.

Administrátor podnikového vyhledávání odešle binární soubor do systému a přiřadí mu zobrazovaný název. Administrátoři kolekcí mohou vybrat slovník synonym, který chtějí použít pro prohledávání dokumentů v kolekci, při konfigurování voleb vyhledávacího serveru pro danou kolekci.

**Omezení:** Vlastní slovník synonym po přidání do systému již nelze upravit. Chcete-li změnit synonyma dostupná v kolekci, musíte postupovat takto:

1. Aktualizujte zdrojový soubor XML.
2. Převedte zdrojový soubor XML na nový slovníkový soubor.
3. Odeberte starý slovník synonym z kolekce, která jej využívá.
4. Odstraňte starý slovník synonym ze systému.
5. Přidejte do systému nový slovník synonym.
6. Asociujte nový slovník synonym s kolekcemi, které jej mají používat.

Můžete vytvořit skript, který obsahuje tyto kroky, a pak jej použít k opakované implementaci slovníku v systému podnikového vyhledávání.

#### Související pojmy



Podpora synonym ve vyhledávacích aplikacích

#### Související úlohy

 Vytvoření souboru XML se synonymy

 Vytvoření slovníku synonym

“Opakovaná implementace vlastních slovníků” na stránce 189

## Přidávání slovníků synonym do systému

Pokud pro vyhledávání dokumentů v kolekci vytvoříte vlastní slovníky synonym, je nutné asociovat je se systémem podnikového vyhledávání. Později můžete zvolit, který ze slovníků synonym chcete použít pro prohledávání kolekce.

### Než začnete


Chcete-li přidávat vlastní slovníky synonym pro použití spolu s dotazy podnikového vyhledávání, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### Omezení

Maximální velikost slovníku synonym je 8 MB.

### Postup

Asociování slovníků synonym se systémem podnikového vyhledávání:

1. Klepnutím na volbu **Systém** otevřete pohled Systém.
2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
3. Na stránce Vyhledávání klepněte na volbu **Konfigurovat slovníky synonym**.
4. Na stránce Konfigurovat slovníky synonym klepněte na volbu **Přidat slovník synonym**.
5. Na stránce Přidat slovník synonym zadejte jedinečný zobrazovaný název pro slovník synonym a případně zadejte popis.
6. Určete umístění souboru .dic. Pokud je soubor uložen v lokálním systému, můžete jej vyhledat. Pokud je soubor umístěn na indexovém serveru, zadejte úplnou cestu k souboru.
7. Klepněte na tlačítko **OK**. Vlastní slovník synonym bude přidán do systému podnikového vyhledávání a bude k dispozici pro prohledávání kolekci.

## Asociování slovníku synonym s kolekcí

Pokud jsou se systémem podnikového vyhledávání asociovány slovníky synonym, můžete jeden z nich vybrat pro vyhledávání v kolekci. Pokud výraz v dotazu odpovídá výrazu ve slovníku, budou ve výsledcích vyhledávání vráceny také dokumenty, které obsahují synonyma daného výrazu.

### Než začnete

Chcete-li vybrat slovník synonym pro kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Postup

Asociování slovníku synonym s kolekcí:

1. Upravte kolekci, vyberte stránku Vyhledat a klepněte na volbu **Konfigurovat volby vyhledávacího serveru**.
2. V poli **Název slovníku synonym** na stránce Volby vyhledávacího serveru vyberte slovník synonym, který chcete použít při dotazech na tuto kolekci.

Seznam dostupných slovníků synonym obsahuje všechny slovníky synonym, které byly přidány do systému podnikového vyhledávání.

3. Klepněte na tlačítko **OK**.

---

## Vlastní slovníky zakázaných slov

Chcete-li zlepšit kvalitu výsledků vyhledávání, můžete určit, že určitá slova mají být z výrazů dotazu během jeho zpracování automaticky odebrána.

Slovník zakázaných slov obsahuje výrazy specifické pro příslušný podnik, které jsou často používány, a tudíž nejsou jako výrazy dotazu užitečné. Vyloučením těchto slov z dotazů můžete zaručit, že uživatelé nebudou zaplaveni výslednými dokumenty, které jsou relevantní pouze okrajově (vráceny budou pouze dokumenty odpovídající ostatním výrazům v dotazu). Během zpracování dotazu vyhledávací servery odstraňují zakázaná slova z dotazů. Mezi odebíraná slova patří zakázaná ve vlastním slovníku a zakázaná slova, která jsou předdefinována pro podnikové vyhledávání (například běžné předložky a členy).

U podnikového vyhledávání standardně probíhá rozpoznávání zakázaných slov závislých na jazyce. V rámci tohoto procesu jsou z dotazu odebrána často se vyskytující běžná slova jako *a* či *the*. Vlastní slovník zakázaných slov je třeba definovat pouze pro zakázaná slova specifická pro konkrétní podnik nebo doménu.

Při zpracování dotazu jsou zakázaná slova odebrána ještě před vygenerováním návrhů na opravy pravopisu. Pokud dotaz obsahuje pouze zakázaná slova, nejsou při zpracování dotazu odebrána žádná zakázaná slova. Aby bylo zaručeno, že budou vráceny výsledky vyhledávání, je u dotazů s podmínkami obsahujícími pouze zakázaná slova odebrání zakázaných slov vypnuto. Je-li například slovo *auto* definováno jako zakázané slovo a vyhledáte-li pouze slovo *auto*, budou vráceny výsledky vyhledávání, které odpovídají slovu *auto*. Pokud vyhledáte výraz *auto volvo*, budou výsledky vyhledávání obsahovat pouze dokumenty odpovídající slovu *volvo*.

Chcete-li vytvořit slovník zakázaných slov, je třeba, aby nejprve odborník na danou problematiku případně ve spolupráci s vývojářem aplikací vytvořil seznam zakázaných slov ve formátu XML. Soubor XML musí být převeden na binární soubor (.dic) pomocí nástroje podnikového vyhledávání **esstopworddictbuilder**.

Administrátor podnikového vyhledávání odešle binární soubor do systému a přiřadí mu zobrazovaný název. Administrátoři kolekcí mohou vybrat slovník zakázaných slov, který chtějí použít pro prohledávání dokumentů v kolekci, při konfigurování voleb vyhledávacího serveru pro danou kolekci.

**Omezení:** Vlastní slovník zakázaných slov po přidání do systému již nelze upravit. Chcete-li upravit zakázaná slova, která jsou k dispozici pro zpracování dotazu, musíte provést následující operace:

1. Aktualizujte zdrojový soubor XML.
2. Převeďte zdrojový soubor XML na nový slovníkový soubor.
3. Odeberte starý slovník zakázaných slov z kolekce, která jej využívá.
4. Odstraňte starý slovník zakázaných slov ze systému.
5. Přidejte do systému nový slovník zakázaných slov.
6. Asociujte nový slovník zakázaných slov s kolekcemi, které jej mají používat.

Můžete vytvořit skript, který obsahuje tyto kroky, a pak jej použít k opakované implementaci slovníku v systému podnikového vyhledávání.

### Související pojmy

 Vlastní slovníky zakázaných slov

### Související úlohy

 Vytvoření souboru XML se zakázanými slovy

 Vytvoření slovníku zakázaných slov

“Opakovaná implementace vlastních slovníků” na stránce 189

## Přidávání slovníků zakázaných slov do systému

Pokud vytvoříte vlastní slovníky zakázaných slov pro odebrání slov z dotazů, je nutné přidat je do systému podnikového vyhledávání. Později můžete zvolit, který ze slovníků zakázaných slov chcete použít pro prohledávání kolekce.

### Než začnete


Chcete-li do systému přidávat vlastní slovníky zakázaných slov, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### Omezení

Maximální velikost slovníku zakázaných slov je 8 MB.

### Postup

Asociování vlastních zakázaných slov se systémem podnikového vyhledávání:

1. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
3. Na stránce **Vyhledávání** klepněte na volbu **Konfigurovat slovníky zakázaných slov**.
4. Na stránce **Konfigurovat slovníky zakázaných slov** klepněte na volbu **Přidat slovník zakázaných slov**.
5. Na stránce **Přidat slovník zakázaných slov** zadejte jedinečný zobrazovaný název pro slovník.
6. Určete umístění souboru .dic. Pokud je soubor uložen v lokálním systému, můžete jej vyhledat. Pokud je soubor umístěn na indexovém serveru, zadejte úplnou cestu k souboru.
7. Klepněte na tlačítko **OK**. Vlastní slovník zakázaných slov bude přidán do systému podnikového vyhledávání a bude k dispozici pro prohledávání kolekci.

## Asociování slovníku zakázaných slov s kolekcí

Pokud jsou se systémem podnikového vyhledávání asociovány slovníky zakázaných slov, můžete jeden z nich vybrat pro vyhledávání v kolekci. Pokud výraz v dotazu odpovídá výrazu ve slovníku, bude tento výraz z dotazu před zpracováním odebrán.

### Než začnete

Chcete-li vybrat slovník zakázaných slov pro kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Postup

Asociování slovníku zakázaných slov s kolekcí:

1. Upravte kolekci, vyberte stránku **Vyhledat** a klepněte na volbu **Konfigurovat volby vyhledávacího serveru**.

2. V poli **Název slovníku zakázaných slov** na stránce Volby vyhledávacího serveru vyberte slovník zakázaných slov, který chcete použít při dotazech na tuto kolekci.  
Seznam dostupných slovníků obsahuje všechny slovníky zakázaných slov, které byly přidány do systému podnikového vyhledávání.
3. Klepněte na tlačítko **OK**.

---

## Opakovaná implementace vlastních slovníků

Konzolu pro správu podnikového vyhledávání nelze použít k provádění změn slovníku synonym, ignorovaných slov nebo slov pro upřednostnění. Je však možné zahrnout kroky do skriptu a poté skript použít k opakované implementaci slovníku.

### Než začnete

Chcete-li opakovaně implementovat vlastní slovník, musíte se přihlásit jako administrátor podnikového vyhledávání.

### Postup

Postup při opakované implementaci vlastního slovníku pro podnikové vyhledávání:

1. Na indexovém serveru otevřete jeden z následujících souborů a vyhledejte název slovníku, který chcete opakovaně implementovat:
  - Chcete-li vyhledat název slovníku synonym, otevřete soubor `ES_NODE_ROOT/master_config/SynonymConfiguration.xml`.
  - Chcete-li vyhledat název slovníku ignorovaných slov, otevřete soubor `ES_NODE_ROOT/master_config/StopWordDictionaryConfiguration.xml`
  - Chcete-li vyhledat název slovníku slov pro upřednostnění, otevřete soubor `ES_NODE_ROOT/master_config/BoostingWordDictionaryConfiguration.xml`.

V následujícím příkladu je uveden soubor `SynonymConfiguration.xml` obsahující dva načtené slovníky synonym s názvy `hello` a `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>
```

Při načítání systém přiřazuje slovníkům jedinečné názvy souborů. V předchozím příkladu jsou to názvy `synonym_hello1.dic` a `synonym_goodbye2.dic`. Úplné cesty k těmto souborům na indexovém serveru vypadají následovně:

- Cesta ke slovníku synonym: `ES_NODE_ROOT/data/custom_dictionary/synonym_*.dic`.
- Cesta ke slovníku ignorovaných slov: `ES_NODE_ROOT/data/custom_dictionary/stopwordDictionary_*.dic`.
- Cesta ke slovníku slov pro upřednostnění: `ES_NODE_ROOT/data/custom_dictionary/boostingwordDictionary_*.dic`.

V operačních systémech AIX, Linux a Solaris můžete použít příkaz **ls** k vypsání seznamu dostupných slovníků. Příklad:

```
% ls -l $ES_NODE_ROOT/data/custom_dictionary/synonym_*.dic
```

```
-rw-rw-r-- 1 esuser users 9 Jan 25 15:11 /home/esuser/node/data/custom_dictionary/synonym_hello1.dic  
-rw-rw-r-- 1 esuser users 9 Jan 25 15:11 /home/esuser/node/data/custom_dictionary/synonym_goodbye2.dic
```

2. Po identifikaci slovníku, který chcete aktualizovat, zastavte systém podnikového vyhledávání.
3. Přepište slovník, který chcete aktualizovat, novým souborem slovníku. Například přepište soubor `synonym_hello1.dic` nebo `synonym_goodbye2.dic`.
4. Je-li systém podnikového vyhledávání spuštěn na dvou nebo čtyřech serverech, přepište soubor slovníku ručně i na vyhledávacích serverech. Umístění souborů slovníku má stejnou cestu jako na indexovém serveru (`ES_NODE_ROOT/data/custom_dictionary/`).
5. Restartujte systém podnikového vyhledávání.

#### **Související pojmy**

“Vlastní slovníky synonym” na stránce 184

“Vlastní slovníky zakázaných slov” na stránce 187

“Vlastní slovníky slov pro upřednostnění” na stránce 198

---

## **Dynamické shrnutí**

Dynamické shrnutí je technologie určující, které fráze ve výsledném dokumentu nejlépe reprezentují hledané tématické okruhy.

U podnikového vyhledávání se dynamické shrnutí pokouší zachytit v dokumentu věty obsahující co nejrozmanitější výskyty hledaných výrazů. Ve výsledcích vyhledávání se zobrazí několik vybraných vět nebo částí vět. Hledané výrazy jsou ve výsledcích vyhledávání zvýrazněny pomocí formátování HTML.

Při konfigurování voleb vyhledávacího serveru pro kolekci můžete zadat maximální délku zobrazených shrnutí dokumentů ve výsledcích vyhledávání. Vzhledem k tomu, že souhrny obsahují znaky definující zvýraznění, bude velikost vyrovnávací paměti vrácené vyhledávací aplikaci vyšší než zadaná maximální hodnota. Délka zobrazeného textu však zadanou maximální hodnotu nepřekročí, ačkoli souhrn může být i kratší (v závislosti na souhrnných datech extrahovaných ze zdrojového dokumentu).

## **Přizpůsobení shrnutí dokumentů v konzole pro správu**

Množství informací zobrazovaných v souhrnech dokumentů můžete přizpůsobit určením voleb pro vyhledávací server v rámci konzoly pro správu podnikového vyhledávání.

### **Než začnete**

Chcete-li řídit délku zobrazovaných souhrnů pro kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### **O této úloze**

Hodnota určená pro maximální zobrazovanou délku souhrnů dokumentů je použita v kombinaci s hodnotou určenou pro počet vět, které může každý souhrn obsahovat. Přednost má hodnota, která vede ke kratšímu shrnutí dokumentu.

Pokud například zadáte mezní hodnotu čtyři věty, bude shrnutí dokumentu obsahovat pouze čtyři věty i přesto, že délka textu pro zobrazení povoluje více znaků, než je celkový počet



znaků v těchto větách. Jiný příklad: Zadání mezní hodnoty 10 vět v kombinaci s mezní hodnotou 500 znaků pro délku textu pro zobrazení může vést k tomu, že shrnutí dokumentu může obsahovat méně než 10 vět.

### Postup

Konfigurování zobrazované délky pro souhrny dokumentů:

1. Upravte kolekci, vyberte stránku Vyhledat a klepněte na volbu **Konfigurovat volby vyhledávacího serveru**.
2. Na stránce Volby vyhledávacího serveru určete maximální zobrazovanou délku pro souhrny dokumentů. Když uživatelé zobrazí výsledky vyhledávání, souhrny dokumentů nepřesáhnou určenou hodnotu.
3. Určete počet vět, které může každý souhrn dokumentu obsahovat (souhrny mohou obsahovat nejvýše deset vět).
4. Klepněte na tlačítko **OK**.
5. Chcete-li, aby se změny projevily, použijte funkci monitorování vyhledávacích serverů a restartujte procesy serveru.

## Přizpůsobení souhrnů dokumentů provedením úprav vlastností

Součástí každého výsledného dokumentu dotazu podnikového vyhledávání je souhrn. Množství informací obsažených v jednotlivých souhrnech můžete podle potřeby přizpůsobit provedením úprav souboru vlastností.

### O této úloze

Popis výsledků vyhledávání můžete přizpůsobit změnou hodnot následujících vlastností v souboru `ES_NODE_ROOT/master_config/ID_kolekce.runtime.node1/runtime-generic.properties`:

#### **MinWordsPerSentence**

Minimální počet slov v každé větě v souhrnu. Kratší věty jsou v souhrnu zahrnuty v případě, že není k dispozici dostatek vět s více slovy, než kolik určuje hodnota `MinWordsPerSentence`. Výchozí hodnota je 4.

#### **MaxWordsPerSentence**

Maximální počet slov v každé větě, která bude uvedena v souhrnu. Pokud věta obsahuje více slov, než kolik určuje tento limit, zobrazí se v souhrnu pouze část věty (ta část, která obsahuje výrazy dotazu, až po hodnotu `MaxWordsPerSentence`). Zbytek věty zobrazen nebude. Výchozí hodnota je 20.

Věty jsou pro souhrny dokumentů vybírány na základě vlastního interního algoritmu, který určuje relevantnost všech vět, které obsahují hledané výrazy. Výběr podle relevance se provede před odfiltrováním některých vět na základě délky.

#### **NumberOfReturnedSentences**

Počet vět, které tvoří popis dokumentu. Výchozí hodnota je 5.

#### **MaxSentencesPerDocument**

Maximální počet vět v dokumentu, které budou považovány za kandidáty při vytváření popisu. Výchozí hodnota je 1000.

### Postup

Chcete-li přizpůsobit souhrny dokumentů ve výsledcích vyhledávání, postupujte takto:

1. Přihlaste se k vyhledávacím serverům jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.

2. Otevřete následující soubor pro úpravy v textovém editoru. *ID\_kol* je ID kolekce zadané (nebo přiřazené systémem) při vytvoření kolekce:

```
ES_NODE_ROOT/master_config/ID_kol.runtime.node1/runtime-generic.properties
```

**Tip:** Informace o mapování mezi názvy kolekcí a jejich ID naleznete v souboru `ES_NODE_ROOT/master_config/collections.ini`.

3. Změňte vlastnosti, které chcete přizpůsobit, poté soubor uložte a zavřete jej.
4. Zastavte a znovu spusťte vyhledávací servery, aby se změny projevíly.

---

## Práce s rychlými odkazy

Rychlé odkazy jsou dokumenty vracené ve výsledcích vyhledávání vždy, když uživatel odešle dotaz obsahující určitá slova a fráze.

Ke konfigurování rychlých odkazů pro kolekci slouží konzola správy podnikového vyhledávání.

### Rychlé odkazy

Rychlé odkazy vám umožňují poskytnout uživatelům odkazy na dokumenty, které předem označíte jako relevantní pro určité dotazovací výrazy.

Rychlý odkaz je identifikátor URI zahrnovaný podnikovým vyhledáváním automaticky do výsledků vyhledávání, pokud dotaz obsahuje určitá slova nebo fráze. Identifikátory URI rychlých odkazů jsou obvykle uváděny na začátku seznamu výsledků, takže je zaručeno, že uživatelé uvidí dokumenty, které jste předem označili jako významné pro daný dotaz.

Rychlé odkazy jsou vráceny ještě s jinými výsledky vyhledávání. Vyhledávací procesy vyhledávají v indexu dokumenty, které odpovídají podmínkám dotazu, a vrací identifikátory URI těchto dokumentů spolu s identifikátory URI rychlých odkazů.

Při konfigurování rychlého odkazu můžete zadat popisný název a souhrnné informace o identifikátoru URI, které uživatelům pomohou rychle určit, zda chtějí daný dokument načíst.

Pro identifikátor URI `http://www.ibm.com/education/us/` můžete například použít název **Vzdělávací služby společnosti IBM v USA** a zadat text shrnutí **Řešení, produkty a prostředky pro odborníky, pedagogy a studenty v USA**.

Chcete-li v kolekci podnikového vyhledávání používat rychlé odkazy, musí být ve vyhledávací aplikaci k dispozici volba pro zobrazení rychlých odkazů. V některých vyhledávacích aplikacích mohou mít uživatelé možnost povolit či zakázat vracení rychlých odkazů při prohledávání kolekce.

### Konfigurování rychlých odkazů

Při vytváření rychlého odkazu pro kolekci podnikového vyhledávání asociujete identifikátor URI dokumentu s klíčovými slovy, která spouštějí jeho zahrnutí do výsledků vyhledávání.

#### Než začnete

Chcete-li konfigurovat rychlé odkazy, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro kolekci, do které daný rychlý odkaz náleží.

#### O této úloze

Po klepnutí na volbu **Nápověda** při vytváření nebo úpravách rychlého odkazu můžete zobrazit příklady zadávání klíčových slov a identifikátorů URI pro rychlé odkazy.

Změny se projeví, aniž by bylo nutné restartovat vyhledávací servery.

### **Postup**

Konfigurování rychlého odkazu:

1. Upravte kolekci, vyberte stránku Vyhledat a klepněte na volbu **Konfigurovat rychlé odkazy**.
2. Na stránce Rychlé odkazy klepněte na volbu **Vytvořit rychlý odkaz**.
3. Zadejte klíčová slova a fráze, které způsobí, že daný rychlý odkaz bude vrácen ve výsledcích vyhledávání, identifikátor URI dokumentu, který jste předem určili jako související s tímto dotazem, a další volby pro daný rychlý odkaz.

Na každý řádek můžete zadat jedno klíčové slovo, několik klíčových slov nebo jednu frázi (dvě či více slov v uvozovkách). Klíčová slova oddělujte mezerami (pro oddělování nelze použít čárku). Nový řádek vytvoříte stisknutím klávesy Enter.

4. Klepněte na tlačítko **OK**.

Nový rychlý odkaz bude uveden na stránce Rychlé odkazy spolu s dalšími rychlými odkazy, které náleží do dané kolekce.

#### **Související odkazy**

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 113



---

## Ohodnocení důležitosti dokumentů

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

Vyhledávací servery podporují rozšířenou syntaxi dotazů a využívají několik různých technik k tomu, aby dokázaly poskytnout co nejhodnotnější výsledky vyhledávání, například hodnocení na základě textu a statické třídy důležitosti. Výchozí chování při řazení podle důležitosti lze rozšířit konfigurováním voleb ovlivňujících důležitost dokumentů ve výsledcích vyhledávání:

- Můžete vytvářet vlastní slovníky slov pro upřednostnění, které ovlivní způsob ohodnocení důležitosti dokumentů obsahujících určená slova pro upřednostnění ve výsledcích vyhledávání.
- Můžete ovlivnit skóre dokumentů shodujících se s určeným vzorem identifikátoru URI.
- Můžete ovlivnit skóre dokumentů obsahujících pole mapovaná na třídy upřednostnění.

### Související pojmy

“Hodnocení důležitosti dokumentů na základě tříd upřednostnění” na stránce 202

“Vlastní slovníky slov pro upřednostnění” na stránce 198

---

## Hodnocení na základě textu

Podnikové vyhledávání dynamicky vypočítává hodnotící skóre pro každý dokument odpovídající výrazům v dotazu.

Textové skóre dokumentu reprezentuje důležitost výrazů dotazu v dokumentu. Při určování textového skóre všech dokumentů vyhovujících podmínkám dotazu pracuje podnikové vyhledávání s mnoha faktory, k nimž patří například:

- Zda výrazy odlišují dokument od jiných dokumentů. Například pokud se výraz dotazu zobrazí v jednom dokumentu, ale ne v jiných dokumentech, znamená to, že tyto výrazy jsou důležité pro daný dokument a odlišují jej od jiných dokumentů. Dotazovací výrazy, které se vyskytují ve většině dokumentů, přispívají k celkovému hodnocení dokumentu méně než výrazy obsažené pouze v užší skupině dokumentů.
- Počet výskytů výrazů dotazu v dokumentu. Skóre výrazu je vypočítáno na základě každého výskytu výrazu v dokumentu. Čím vyšší je počet výskytů výrazu dotazu v dokumentu, tím vyšší je skóre dokumentu. Například při hledání výrazu `thinkpad` je dokument, který tento výraz obsahuje vícekrát, ohodnocen výše než jiné dokumenty s nižším počtem výskytů dotazu.
- V dokumentech HTML jsou zvaženy i atributy výrazu dotazu (tedy kontext výrazu). Skóre výrazu je vypočítáno na základě atributů daného výrazu (například umístění, tučné písmo, kurzíva, kotva atd.) Obecně výrazy, které se vyskytují v názvu dokumentu, mají vyšší skóre než výrazy v odstavci. Zvýrazněné výrazy (například tučné písmo) mají vyšší skóre než výrazy v prostém textu. Důležitost atributů můžete konfigurovat v nastaveních analyzátoru.
- Váha výrazů dotazu. Chcete-li přizpůsobit důležitost výrazu v dokumentu, můžete pro výrazy konfigurovat hodnoty pro upřednostnění. V tomto případě přiřadíte slovník slov pro upřednostnění (který obsahuje výrazy a jejich hodnoty pro upřednostnění) ke kolekci. Tento slovník je používán při vyhledávání a hodnoty pro upřednostnění výrazů ve slovníku přispívají ke skóre dokumentu. Čím vyšší je hodnota pro upřednostnění, tím vyšší je příspěvek výrazu ke skóre dokumentu.
- Blížkost výrazů dotazu v dokumentu. Pokud se výrazy dotazu zobrazí v dokumentu blízko sebe, je jejich lexikální blízkost použita k výpočtu textového skóre. Například

předpokládejme existenci dvou dokumentů. V jednom dokumentu je uvedeno Národní divadlo (slova národní a divadlo jsou blízko sebe). V druhém dokumentu je uvedeno divadlo Viola na Národní třídě (slova národní a divadlo nejsou blízko sebe). Hledáte-li frázi Národní divadlo, způsobí blízkost těchto slov v prvním dokumentu, že bude lépe ohodnocen než druhý dokument.

- Dalšími faktory zahrnovanými do hodnocení dokumentu je délka dokumentu a bohatost použití slovní zásoby (například počet jedinečných slov).

---

## Statické řazení

K některým dokumentům lze přiřadit statický faktor řazení, který zvyšuje důležitost těchto dokumentů uváděnou ve výsledcích vyhledávání.

Při vytváření kolekce je nutné zadat volbu **Důležitost dokumentů**. Vybraný typ důležitosti dokumentů závisí na tom, zda je k dokumentům v kolekci přidružen statický faktor ohodnocení důležitosti. Při prohledávání kolekce používající statické ohodnocení důležitosti má statický faktor důležitosti vliv na způsob ohodnocení dokumentů vrácených ve výsledcích vyhledávání.

Pro webový obsah je statický faktor důležitosti založen na odkazech. Důležitost dokumentu ve výsledcích vyhledávání může být zvyšována počtem odkazů na tento dokument z jiných dokumentů a původem těchto odkazů.

U dokumentů zahrnujících kalendářní pole či metadata je statický faktor důležitosti založen na datu dokumentu. Pole data dokumentu poskytované prolézacím modulem může obsahovat datum poslední úpravy dokumentu nebo datum posledního prolezení dokumentu v závislosti na nastavení konfigurace prolézacího modulu.

Datum dokumentu může zvýšit jeho relevanci. Novější články v diskusních skupinách NNTP mohou mít například vyšší důležitost než články starší. Pokud zdroj dat zahrnuje více kalendářních hodnot, můžete určit, podle které z nich má být určována důležitost dokumentů při konfiguraci prolézacího modulu.

Pokud v kolekci použijete statické řazení, ujistěte se, že ve stejné kolekci nekombinujete zdroje dat s různými typy řazení. Chcete-li jako faktor statického řazení použít například odkazy na dokument, ujistěte se, že daná kolekce obsahuje pouze webové dokumenty. Ohodnocení důležitosti dokumentu je méně přesné při použití kombinace zdrojů s různými modely ohodnocení důležitosti ve stejné kolekci a pořadí výsledků vyhledávání se může lišit od očekávaného pořadí.

Dále byste měli zajistit, že dokumenty v kolekci obsahují pole a hodnoty, které umožňují použití statického ohodnocení důležitosti. Představte si například kolekci, která je konfigurována k používání statického ohodnocení důležitosti založeného na datech dokumentů, a prolézací modul v kolekci, která je konfigurována k používání konkrétních polí jako dat dokumentů. Pokud dokument neobsahuje dané pole, nemusí být důležitost dokumentu správně ohodnocena a pořadí výsledků vyhledávání se může lišit od očekávaného pořadí.

### Důsledky ohodnocení důležitosti založeného na odkazech

Statické ohodnocení důležitosti spolu s faktory, jako je například přiřazení skóre za účelem zvýšení důležitosti vzorů identifikátoru URI, přispívají ke statickému skóre dokumentu a ovlivňují důležitost dokumentu. Model ohodnocení důležitosti založeného na odkazech je typicky používán pro webové kolekce, protože tento model počítá statické ohodnocení důležitosti dokumentu v závislosti na počtu odkazů na dokument. Dokument, na který odkazuje vyšší počet jiných dokumentů, je ohodnocen jako důležitější.

Z tohoto důvodu může dojít při konfiguraci daného modelu pro jinou než webovou nebo smíšenou kolekci (kolekci obsahující webové a jiné než webové dokumenty) ke snížení kvality vyhledávání, protože jiné než webové dokumenty neobsahují žádnou koncepci odkazování.

Při povolení ohodnocení důležitosti založeného na odkazech je povolena také detekce duplicitních dokumentů. Duplicitní dokumenty mají stejné statické ohodnocení důležitosti jako hlavní dokument. Pokud nejsou faktory upřednostnění vzoru identifikátoru URI konfigurovány v žádných dokumentech v duplicitní skupině, mají všechny duplicitní dokumenty stejné statické skóre.

---

## Obnovení výchozích hodnot pro statické ohodnocení důležitosti dokumentů

Pokud při vytvoření kolekce konfiguruje statické ohodnocení důležitosti dokumentů, můžete původní hodnoty vlastností obnovit úpravou souborů `runtime.properties` pro danou kolekci.

### Než začnete

Chcete-li obnovit výchozí hodnoty ohodnocení důležitosti dokumentů v kolekci, musíte být přihlášen jako administrátor podnikového vyhledávání.

### O této úloze

Pro obnovení výchozích hodnot ohodnocení důležitosti dokumentů v kolekci je nutné aktualizovat soubory `runtime.properties` dané kolekce a všech vyhledávacích serverů v systému podnikového vyhledávání. V konfiguraci s více servery je soubor `runtime.properties` umístěn na indexovém serveru v adresáři `ES_NODE_ROOT/master_config/ID_kolekce.runtime.ID_uzlu`, kde `ID_kolekce` je ID kolekce a `ID_uzlu` je ID vyhledávacích serverů.

Chcete-li například aktualizovat kolekci `col1` v systému podnikového vyhledávání s více servery, proveďte aktualizaci souborů `runtime.properties` pro tuto kolekci a pro oba vyhledávací servery (`node3` a `node4`):

```
ES_NODE_ROOT/master_config/col1.runtime.node3/runtime.properties
ES_NODE_ROOT/master_config/col1.runtime.node4/runtime.properties
```

### Postup

Postup při obnovení výchozích hodnot ohodnocení důležitosti dokumentů v kolekci:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání.
2. Zjistěte ID kolekce, v níž chcete obnovit výchozí hodnoty ohodnocení důležitosti. ID kolekce je uvedeno v souboru `ES_NODE_ROOT/master_config/collections.ini`. Obsah tohoto souboru můžete seřadit, aby se vám lépe procházel. V následujícím příkladu je použito ID kolekce `col1`:

```
% sort $ES_NODE_ROOT/master_config/collections.ini | more
collection1.configfile=col1_config.ini
collection1.datadir=/home/esearch/node/data/col1
collection1.description=
collection1.displayname=Collection1
collection1.flags=0
collection1.id=col1
collection1.sectiontype=collection
collection1.type=1
...
```

3. Otevřete pro úpravy soubor `runtime.properties` kolekce, kterou chcete obnovit, a proveďte v něm následující změny:
  - a. Odstraňte následující vlastnosti:
 

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```
  - b. Pokud soubor `runtime.properties` obsahuje parametr `trevi.sourcetype=1`, což znamená, že důležitost dokumentů je určována podle odkazů, otevřete soubor `ES_INSTALL_ROOT/default_config/runtime.1/runtime.properties`, zkopírujte z něj následující výchozí vlastnosti a vložte je do souboru `runtime.properties`:
 

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```
  - c. Pokud soubor `runtime.properties` obsahuje parametr `trevi.sourcetype=2`, což znamená, že důležitost dokumentů je určována podle data, otevřete soubor `ES_INSTALL_ROOT/default_config/runtime.2/runtime.properties`, zkopírujte z něj následující výchozí vlastnosti a vložte je do souboru `runtime.properties`:
 

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```
  - d. Pokud soubor `runtime.properties` obsahuje parametr `trevi.sourcetype=3`, což znamená, že k určení důležitosti dokumentů v kolekci není použit statický faktor důležitosti, otevřete soubor `ES_INSTALL_ROOT/default_config/runtime.0/runtime.properties`, zkopírujte z něj následující výchozí vlastnosti a vložte je do souboru `runtime.properties`:
 

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```
4. V konfiguraci s více servery opakujte krok 3 a aktualizujte soubor `runtime.properties` stejné kolekce na druhém vyhledávacím serveru.
5. V konzole pro správu zobrazte stránku Vyhledat a restartujte proces vyhledávání v této kolekci.

Tyto kroky podle potřeby opakujte se všemi kolekcemi, v nichž chcete obnovit výchozí hodnoty ohodnocení důležitosti dokumentů.

---

## Vlastní slovníky slov pro upřednostnění

Chcete-li zvýšit kvalitu výsledků vyhledávání, můžete ovlivnit ohodnocení důležitosti dokumentů ve výsledcích vyhledávání vytvořením vlastního slovníku slov pro upřednostnění.

Pokud dotaz obsahuje slovo uvedené ve slovníku slov pro upřednostnění, bude důležitost dokumentů obsahujících příslušné slovo zvýšena nebo snížena na základě faktoru upřednostnění konfigurovaného pro příslušné slovo ve slovníku.

Pomocí slovníku obsahujícího slova pro upřednostnění můžete zajistit, aby byly konkrétní dokumenty vráceny v případě, že uživatel zadá konkrétní termíny dotazu. Například předpokládejme, že existuje kolekce obsahující dokumenty o autech. Pro tyto dokumenty



může být vhodné použít některá klíčová slova související s modely aut, názvem výrobce atd. Chcete-li ovlivnit ohodnocení důležitosti výsledků vyhledávání, můžete klíčovým slovům (model, výrobce atd.) ve slovníku slov pro upřednostnění přiřadit důležitost přidělením hodnoty pro upřednostnění. Pokud uživatelé vyhledávají kolekce a zadávají dotazy zahrnující libovolné z těchto klíčových slov, budou dokumenty o autech ohodnoceny ve výsledcích vyhledávání lépe než jiné dokumenty.

Faktory upřednostnění leží v rozsahu -10 až 10. Během zpracování dotazu vyhledávací servery zvyšují důležitost dokumentů obsahujících slova s kladnými faktory upřednostnění a snižují důležitost dokumentů obsahujících slova se zápornými faktory upřednostnění.

Dokument vyhovující výrazům dotazu s vysokými faktory upřednostnění bude například ohodnocen výše než v případě, že by faktor upřednostnění nebyl použit. (Faktor upřednostnění je pouze jedním z faktorů ovlivňujících skóre dokumentu.)

Při vytváření slovníku můžete stejný faktor upřednostnění přiřadit k libovolnému počtu slov. Slovník může obsahovat jednoslovné i víceslovné výrazy. Pro víceslovné výrazy je shoda zjišťována jako pro frázi.

Pokud je slovo s váhou určenou hodnotou upřednostnění zadáno v dotazu s operátorem OR (příklad: `this | that`), bude pro výrazy dotazu vypočítán vážený průměr. Výsledné agregované skóre je použito pro všechny výskyty operandů dotazu OR. Pro různé operandy dotazu OR nejsou počítána různá skóre.

Upřednostňování na základě slovníků slov pro upřednostnění není podporováno pro výrazy dotazů s použitím polí. Při analýze výrazů dotazu je pro výpočet skóre dokumentu použit pouze text dotazu, nikoli název pole. Chcete-li použít faktory upřednostnění pro výrazy dotazu vyskytující se v polích, můžete mapovat názvy polí na třídy upřednostnění.

Chcete-li vytvořit slovník slov pro upřednostnění, je třeba, aby nejprve odborník na danou problematiku případně ve spolupráci s vývojářem aplikací vytvořil seznam slov pro upřednostnění ve formátu XML. Soubor XML musí být převeden na binární soubor (.dic) pomocí nástroje podnikového vyhledávání **esboosttermdictbuilder**.

Administrátor podnikového vyhledávání odešle binární soubor do systému a přiřadí mu zobrazovaný název. Administrátoři kolekcí mohou vybrat slovník slov pro upřednostnění, který chtějí použít pro prohledávání dokumentů v kolekci, při konfigurování voleb vyhledávacího serveru pro danou kolekci.

**Omezení:** Vlastní slovník slov pro upřednostnění po přidání do systému již nelze upravit. Chcete-li upravit slova pro upřednostnění, která jsou k dispozici pro zpracování dotazu, musíte provést následující operace:

1. Aktualizujte zdrojový soubor XML.
2. Převedte zdrojový soubor XML na nový slovníkový soubor.
3. Odeberte starý slovník slov pro upřednostnění z kolekce, která jej využívá.
4. Odstraňte starý slovník slov pro upřednostnění ze systému.
5. Přidejte do systému nový slovník slov pro upřednostnění.
6. Asociujte nový slovník slov pro upřednostnění s kolekcemi, které jej mají používat.

Můžete vytvořit skript, který obsahuje tyto kroky, a pak jej použít k opakované implementaci slovníku v systému podnikového vyhledávání.


#### **Související pojmy**

“Ohodnocení důležitosti dokumentů” na stránce 195

 Vlastní slovníky slov pro upřednostnění

### Související úlohy

 Vytvoření souboru XML se slovy pro upřednostnění

 Vytvoření slovníku slov pro upřednostnění

“Opakovaná implementace vlastních slovníků” na stránce 189

## Přidávání slovníků slov pro upřednostnění do systému

Pokud vytvoříte vlastní slovníky slov pro upřednostnění, je nutné asociovat je se systémem podnikového vyhledávání. Později můžete zvolit, který ze slovníků slov pro upřednostnění chcete použít pro prohledávání kolekce.

### Než začnete


Chcete-li do systému přidávat vlastní slovníky slov pro upřednostnění, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### Omezení

Maximální velikost slovníku slov pro upřednostnění je 8 MB.

### Postup

Asociování vlastních slov pro upřednostnění se systémem podnikového vyhledávání:

1. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
3. Na stránce **Vyhledávání** klepněte na volbu **Konfigurovat slovníky slov pro upřednostnění**.
4. Na stránce **Konfigurovat slovníky slov pro upřednostnění** klepněte na volbu **Přidat slovník slov pro upřednostnění**.
5. Na stránce **Přidat slovník slov pro upřednostnění** zadejte jedinečný zobrazovaný název pro slovník a případně zadejte popis.
6. Určete umístění souboru **.dic**. Pokud je soubor uložen v lokálním systému, můžete jej vyhledat. Pokud je soubor umístěn na indexovém serveru, zadejte úplnou cestu k souboru.
7. Klepněte na tlačítko **OK**. Vlastní slovník slov pro upřednostnění bude přidán do systému podnikového vyhledávání a bude k dispozici pro prohledávání kolekcí.

## Asociování slovníku slov pro upřednostnění s kolekcí

Pokud jsou se systémem podnikového vyhledávání asociovány slovníky slov pro upřednostnění, můžete jeden z nich vybrat pro vyhledávání v kolekci. Pokud výraz dotazu odpovídá výrazu ve slovníku, bude důležitost dokumentů obsahujících příslušný výraz zvýšena nebo snížena na základě faktoru upřednostnění přiřazeného k výrazu ve slovníku.

### Než začnete

Chcete-li vybrat slovník slov pro upřednostnění pro kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Postup

Asociování slovníku slov pro upřednostnění s kolekcí:

1. Upravte kolekci, vyberte stránku Vyhledat a klepněte na volbu **Konfigurovat volby vyhledávacího serveru**.
2. V poli **Název slovníku slov pro upřednostnění** na stránce Volby vyhledávacího serveru vyberte slovník slov pro upřednostnění, který chcete použít při dotazech na tuto kolekci. Seznam dostupných slovníků obsahuje všechny slovníky slov pro upřednostnění, které byly přidány do systému podnikového vyhledávání.
3. Klepněte na tlačítko **OK**.

---

## Hodnocení důležitosti dokumentů na základě vzorů identifikátoru URI

Přiřazením faktorů upřednostnění ke vzorům identifikátorů URI můžete zvýšit nebo snížit důležitost dokumentů.

Každému dokumentu je při přidávání do indexu přiřazeno výchozí skóre statické třídy důležitosti. Výchozí skóre se může lišit podle toho, zda byly pro kolekci povoleny statické třídy důležitosti, a, pokud ano, podle typu statické třídy důležitosti (podle data dokumentu nebo v případě webových dokumentů podle počtu jiných dokumentů, které na příslušný dokument odkazují).

Přiřazením faktorů upřednostnění ke vzorům identifikátorů URI můžete ovlivnit relativní důležitost dokumentu. Faktor upřednostnění se používá spolu s výchozími skóre statických třída důležitosti a dalšími faktory k určení konečného statického skóre pro dokument.

Pořadí, v němž vzory identifikátorů URI konfiguruje, je důležité. Indexový server při výpočtu hodnoty pro každý dokument v kolekci vyhodnocuje vzory identifikátorů URI v pořadí, v jakém jsou uvedeny. Pro každý identifikátor URI jsou provedeny následující kroky:

1. Indexový server postupně prohledává vzory identifikátorů URI.
2. Když indexový server nalezne první vzor identifikátoru URI odpovídající dokumentu v indexu, použije pro příslušný dokument faktor upřednostnění, který je konfigurován pro příslušný vzor identifikátoru URI.
3. Pokud dokument neodpovídá žádnému vzoru identifikátoru URI, je použito výchozí skóre statické třídy důležitosti.

Po konfigurování faktoru upřednostnění pro vzor identifikátoru URI je třeba vzor identifikátoru URI umístit tak, aby indexový server prohledával vzory v požadovaném pořadí.

## Ovlivňování skóre dokumentů shodujících se se vzory identifikátorů URI

Důležitost dokumentů odpovídajících vzoru identifikátoru URI můžete zvýšit nebo snížit použitím faktoru upřednostnění pro výchozí skóre ohodnocení důležitosti.

### Než začnete

Chcete-li ovlivnit důležitost dokumentů odpovídajících vzoru identifikátoru URI, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.



### O této úloze

Konfigurovaný faktor upřednostnění je použit v kombinaci s výchozím skóre statické třídy důležitosti pro výpočet nového statického skóre pro všechny dokumenty odpovídající určenému vzoru identifikátoru URI.

Faktory upřednostnění zvyšují pouze statická skóre a jsou pouze jedním z činitelů ovlivňujících výpočet určující konečné ohodnocení důležitosti dokumentu. Dokument, na který vede více odkazů (což má za následek vyšší počáteční skóre), bude vždy řazen výše než dokument, na který žádný odkaz nevede.

### Postup

Chcete-li ovlivnit skóre dokumentů odpovídajících vzoru identifikátoru URI, postupujte takto:

1. Upravte kolekci, vyberte stránku Index a klepněte na volbu **Ovlivnit skóre podle shody se vzory identifikátoru URI**.
2. Na stránce Ovlivnit skóre podle shody se vzory identifikátoru URI klepněte na volbu **Přidat vzor identifikátoru URI**.
3. Zadejte vzor identifikátoru URI pro dokumenty, jejichž důležitost ve výsledcích vyhledávání chcete zvýšit nebo snížit. Příklad:  
`http://domain.org/hr/*`  
`db2://*ROI*`  
`*/afs/*`
4. Zadejte pro faktor upřednostnění hodnotu v rozsahu -10 až 10. Konečné statické skóre pro všechny dokumenty odpovídající vzoru identifikátoru URI bude vypočteno na základě tohoto faktoru ohodnocení.
5. Klepněte na tlačítko **OK**.
6. Na stránce Ovlivnit skóre podle shody se vzory identifikátoru URI přesuňte nový vzor identifikátoru URI na místo odpovídající pořadí, ve kterém jej má indexový server procházet.  
Indexový server počítá skóre statické třídy důležitosti v pořadí, ve kterém jsou identifikátory URI uvedeny. Nejlepších výsledků dosáhnete, uvedete-li nejprve konkrétnější identifikátory URI. V následujícím příkladu podadresář `/forms` odpovídá vzoru identifikátoru URI `http://www.ibm.com/hr/*`. Chcete-li zaručit správný výpočet skóre pro dokumenty v podadresáři `/forms`, uveďte vzor identifikátoru URI pro podadresář `/forms` jako první:  
`http://www.ibm.com/hr/forms/* 8`  
`http://www.ibm.com/hr/* -2`
7. Chcete-li změnit vzor identifikátoru URI nebo faktor upřednostnění, vyberte vzor identifikátoru URI a klepněte na volbu  **Upravit**.
8. Chcete-li vzor identifikátoru URI odebrat ze seznamu, vyberte příslušný vzor identifikátoru URI a klepněte na volbu  **Odebrat**.
9. Chcete-li použít faktory upřednostnění pro dokumenty, které byly dříve indexovány, znovu sestavte hlavní index.

---

## Hodnocení důležitosti dokumentů na základě tříd upřednostnění

Mapováním polí na třídy upřednostnění můžete ovlivnit řazení dokumentů ve výsledcích vyhledávání.

Při analýze dokumentů analyzátor přiřazuje prvkům dokumentů *třídy upřednostnění* podle polí, k nimž prvky náleží. Tyto třídy upřednostnění jsou zahrnuty do indexu a používány během vyhodnocování dotazu pro výpočet skóre, která ovlivňují ohodnocení důležitosti výsledných dokumentů.

Chcete-li ovlivnit způsob výpočtu skóre, můžete pro třídy upřednostnění konfigurovat číselné faktory upřednostnění. Pokud výraz dotazu odpovídá prvku v poli mapovaném na třídu upřednostnění, tento výskyt prvku ovlivní celkové skóre dokumentu. Toto skóre je vypočteno na základě použití faktoru upřednostnění konfigurovaného ve třídě upřednostnění.

Můžete například zvýšit skóre polí s názvem. Pokud se výraz dotazu vyskytne v názvu, tento výskyt významně přispěje ke skóre dokumentu a pomůže dosáhnout vyššího ohodnocení důležitosti dokumentu ve výsledcích vyhledávání.

Chcete-li ovlivnit ohodnocení důležitosti dokumentů, můžete prostřednictvím konzoly pro správu podnikového vyhledávání určit faktory upřednostnění pro třídy upřednostnění a mapovat pole na třídy upřednostnění. Pro podnikové vyhledávání je předkonfigurováno šestnáct tříd upřednostnění. Osm ze tříd upřednostnění je určeno pro použití s poli s obsahem a zbývajících osm tříd upřednostnění je určeno pro pole metadat. Můžete upravit skóre asociovaná s výchozími třídami upřednostnění a asociovat různá nebo další pole s třídami upřednostnění.

Změníte-li mapování polí, musíte dokumenty prolézat a analyzovat znovu, aby se změny projevíly pro dokumenty indexované dříve. Změníte-li faktory určené pro třídu upřednostnění, monitorujte vyhledávací servery a restartujte vyhledávací procesy, aby se změny projevíly.

## Detekce duplicitních dokumentů a shrnutí dokumentů

Při mapování pole na třídu upřednostnění je nutné určit, zda má být toto pole použito pro zjišťování duplicitních dokumentů a zda lze obsah pole vložit do shrnutí dokumentů ve výsledcích vyhledávání.

- Používá-li se pole k detekci duplicitních dokumentů, je považováno za pole s obsahem; vybírat lze pouze třídy upřednostnění určené pro pole s obsahem. Obsah polí tohoto typu lze použít v dynamických souhrnech dokumentů ve výsledcích vyhledávání.
- Pokud se pole nepoužívá k detekci duplicitních dokumentů, je považováno za pole metadat; vybírat lze pouze třídy upřednostnění určené pro pole metadat. V takovém případě budou dva dokumenty, které se shodují s výjimkou určeného pole, považovány za vzájemné duplikáty a pole nebude použito v dynamických souhrnech dokumentů.

## Vysoké a nízké hodnoty návratnosti

Při vyhodnocování dotazu proces vyhledávání odhaduje počet výsledných dokumentů, které budou vráceny. Prahové hodnoty určují, zda je dotaz považován za dotaz s nízkou hodnotou návratnosti, s vysokou hodnotou návratnosti nebo s hodnotou návratnosti spadající do intervalu mezi těmito dvěma hodnotami:

### Nízká hodnota návratnosti

Pokud je odhadovaný počet výsledných dokumentů menší než dolní prahová hodnota, je dotaz považován za dotaz s nízkou návratností.

### Vysoká hodnota návratnosti

Pokud je odhadovaný počet výsledných dokumentů větší než horní prahová hodnota, je dotaz považován za dotaz s vysokou návratností.

### Směšená hodnota návratnosti

Pokud odhadovaný počet dokumentů leží mezi oběma prahovými hodnotami, je hodnota návratnosti dotazu kombinací obou prahových hodnot.

Každá třída upřednostnění určuje faktory upřednostnění, které jsou během zpracování dotazů asociovaný s dotazy s nízkou návratností a s dotazy s vysokou návratností. Nízký faktor upřednostnění ovlivňuje relativní důležitost dotazů s nízkou návratností a vysoký faktor upřednostnění ovlivňuje relativní důležitost dotazů s vysokou návratností. Kombinace obou faktorů upřednostnění ovlivňuje relativní důležitost dotazů se smíšenou hodnotou návratnosti.

Hodnoty faktorů upřednostnění řídí relativní důležitost každého výskytu výrazu dotazu v dokumentu. Každý výskyt výrazu dotazu v dokumentu je započítán podle příslušného faktoru upřednostnění.

Při konfigurování tříd upřednostnění pro kolekci můžete upravit výchozí faktory upřednostnění. Můžete například určit faktory upřednostnění tak, abyste zaručili, že výrazy dotazu, které se vyskytnou v polích názvů, budou započítány pětinasobně ve srovnání s výrazy dotazu, které se vyskytnou v běžném textu.

#### Související pojmy

“Ohodnocení důležitosti dokumentů” na stránce 195

## Mapování polí na třídy upřednostnění

Mapováním názvů polí na třídy upřednostnění můžete ovlivnit relativní důležitost polí.

### Než začnete

Chcete-li mapovat pole na třídy upřednostnění, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

System používá faktor upřednostnění k ovlivnění ohodnocení důležitosti dokumentů obsahující výrazy dotazu v polích mapovaných na třídy upřednostnění.

System podnikového vyhledávání rezervuje některá mapování pro interní pole a běžný text bez dalších definičních charakteristik. Na třídy upřednostnění využívané vyhrazenými poli lze mapovat jiná pole, vyhrazená pole však nelze upravovat ani odstraňovat.

### Postup

Chcete-li mapovat pole na třídy upřednostnění, postupujte takto:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Mapovat pole na třídy upřednostnění**.
2. Na stránce Mapovat pole na třídy upřednostnění klepněte na volbu **Přidat pole**.
3. Na stránce Přidat pole do třídy upřednostnění zadejte název pole, který chcete mapovat na třídu upřednostnění.



Můžete zadat název pole existujícího v prolézaném zdroji nebo v externím zdroji, název pole mapovaného z prvku XML, název pole mapovaného z prvku metadat HTML, nebo jeden z předdefinovaných názvů polí.

4. Určete, zda je pole použito pro detekci duplicitních dokumentů. Zaškrtnete-li toto políčko, bude seznam tříd upřednostnění k dispozici obsahovat třídy vztahující se k polím s obsahem. Je-li dokument s tímto polem vrácen v rámci výsledků vyhledávání, bude obsah tohoto pole zobrazen v oblasti pro shrnutí dokumentu.

Zrušíte-li zaškrtnutí tohoto políčka, bude seznam tříd upřednostnění k dispozici obsahovat třídy vztahující se k polím metadat. Obsah pole se nezobrazí v oblasti shrnutí dokumentu ve výsledcích vyhledávání.

5. Vyberte třídu upřednostnění a klepněte na tlačítko **OK**.

Přidané pole se zobrazí na stránce Mapovat pole na třídy upřednostnění. Můžete vybrat volbu úprav třídy upřednostnění a konfigurovat různé faktory upřednostnění pro určování skóre dokumentů obsahujících příslušné pole.

6. Chcete-li změnit nastavení použití pole pro detekci duplicitních dokumentů nebo mapovat pole na jinou třídu upřednostnění, klepněte na volbu  **Upravit**. (Pole vyhrazená pro použití podnikovým vyhledáváním nelze upravovat.)
7. Chcete-li pole odebrat ze třídy upřednostnění, klepněte na volbu  **Odebrat**. (Pole vyhrazená pro použití podnikovým vyhledáváním nelze odebrat.)

8. Chcete-li, aby se projevíly změny pro dokumenty indexované dříve, musíte provést nové prolezení a indexování těchto dokumentů.

#### **Související pojmy**

“Ohodnocení důležitosti dokumentů” na stránce 195

## **Konfigurování faktorů upřednostnění pro třídy upřednostnění**

Faktory upřednostnění konfigurované pro třídy upřednostnění reprezentují váš odhad míry relevance konkrétních polí ve výsledných dokumentech k dotazu. Třídy upřednostnění s vysokými faktory upřednostnění mohou zvýšit důležitost výsledných dokumentů obsahujících pole mapovaná na příslušnou třídu upřednostnění.

#### **Než začnete**


Chcete-li konfigurovat faktory upřednostnění pro třídy upřednostnění, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

#### **O této úloze**

System pomocí faktorů upřednostnění konfigurovaných pro třídu upřednostnění, výchozího skóre statické třídy důležitosti a dalších faktorů vypočítá nové skóre pro výsledné dokumenty obsahující pole mapovaná na třídu upřednostnění.

#### **Postup**

Chcete-li konfigurovat faktor upřednostnění pro třídu upřednostnění, postupujte takto:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Mapovat pole na třídy upřednostnění**.
2. Na stránce Mapovat pole na třídy upřednostnění klepněte na volbu **Upravit třídy upřednostnění**.
3. Na stránce Třídy upřednostnění vyhledejte třídu upřednostnění, kterou chcete změnit, a klepněte na volbu  **Upravit**.
4. Na stránce Upravit třídu upřednostnění zadejte nové hodnoty pro vysoký a nízký faktor upřednostnění. Pro oba faktory můžete zadat stejnou hodnotu.
5. Klepněte na tlačítko **OK**.
6. Chcete-li, aby se změny projevíly, použijte funkci monitorování vyhledávacích serverů a vyberte ikony pro zastavení a opětné spuštění vyhledávacích procesů. Když uživatelé odesílají dotazy, je relativní důležitost výsledných dokumentů obsahujících pole mapovaná na příslušnou třídu upřednostnění určena na základě nových faktorů upřednostnění.

#### **Související pojmy**

“Ohodnocení důležitosti dokumentů” na stránce 195

## **Výchozí hodnoty tříd upřednostnění**

Podnikové vyhledávání poskytuje 16 tříd upřednostnění, jejichž prostřednictvím můžete ovlivňovat ohodnocení důležitosti dokumentů ve výsledcích vyhledávání.

Chcete-li vypočítat skóre pro pole a text bez jakékoli další definované charakteristiky, jsou pro použití podnikovým vyhledáváním vyhrazena následující pole:

```
es_special_field.regular_text  
es_special_field.default_field  
es_special_field.default_metadata_field
```

Na třídy upřednostnění využívané vyhrazenými poli lze mapovat jiná pole, vyhrazená pole však nelze upravovat ani odstraňovat.

Pro všechna ostatní pole můžete upravovat faktory upřednostnění, podle kterých systém počítá ohodnocení důležitosti dokumentu. Můžete rovněž mapovat libovolný počet polí na kteroukoli třídu upřednostnění včetně tříd upřednostnění používaných vyhrazenými poli.

V následující tabulce jsou uvedeny názvy tříd upřednostnění, výchozí faktory upřednostnění pro dotazy s nízkou hodnotou návratnosti, výchozí faktory upřednostnění pro dotazy s vysokou hodnotou návratnosti a názvy předdefinovaných polí mapovaných na třídy upřednostnění ve výchozí konfiguraci.

Výchozí faktory upřednostnění se liší v závislosti na metodě určování statických tříd důležitosti vybrané pro kolekci při jejím vytváření. Mezi volby platí možnost bez statických tříd důležitosti, ohodnocení důležitosti určené počtem odkazů na dokument (pro webové zdroje) nebo ohodnocení důležitosti na základě data dokumentu.

Tabulka 6. Výchozí hodnoty tříd upřednostnění

Výchozí nízké a vysoké faktory upřednostnění				
Název třídy upřednostnění	Bez statických tříd důležitosti	Odkazy dokumentů	Datum dokumentu	Předdefinovaná mapování polí
Třída obsahu A	Nízký: 4 Vysoký: 2	Nízký: 6 Vysoký: 1	Nízký: 4 Vysoký: 2	es_special_field.regular_text
Třída obsahu B	Nízký: 5 Vysoký: 4	Nízký: 7 Vysoký: 3	Nízký: 5 Vysoký: 4	es_special_field.html_emphasized_text  Obsahuje následující prvky jazyka HTML: b, big, caption, dfn, em, h4, h5, h6, strong
Třída obsahu C	Nízký: 7 Vysoký: 4	Nízký: 9 Vysoký: 3	Nízký: 7 Vysoký: 4	es_special_field.html_headers  Obsahuje následující prvky jazyka HTML: h1, h2, h3
Třída obsahu D	Nízký: 2 Vysoký: 5	Nízký: 1 Vysoký: 5	Nízký: 2 Vysoký: 5	title
Třída obsahu E	Nízký: 1 Vysoký: 1	Nízký: 5 Vysoký: 10	Nízký: 1 Vysoký: 1	es_special_field.anchor
Třída obsahu F	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	es_special_field.anchor_same_dir
Třída obsahu G	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	es_special_field.anchor_same_host
Třída obsahu H	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	es_special_field.default_field
Třída metadat A	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	es_special_field.default_metadata_field
Třída metadat B	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	
Třída metadat C	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	
Třída metadat D	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	
Třída metadat E	Nízký: 1 Vysoký: 1	Nízký: 5 Vysoký: 1	Nízký: 1 Vysoký: 1	keywords
Třída metadat F	Nízký: 1 Vysoký: 1	Nízký: 3 Vysoký: 1	Nízký: 1 Vysoký: 1	es_special_field.urlhost



Tabulka 6. Výchozí hodnoty tříd upřednostnění (pokračování)

Výchozí nízké a vysoké faktory upřednostnění				
Název třídy upřednostnění	Bez statických tříd důležitosti	Odkazy dokumentů	Datum dokumentu	Předdefinovaná mapování polí
Třída metadat G	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	es_special_field.urlpath
Třída metadat H	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	popis

#### Související pojmy

“Ohodnocení důležitosti dokumentů” na stránce 195



---

## Vyhledávací aplikace podnikového vyhledávání

Vyhledávací aplikace vám umožňují prohledávat kolekce a externí zdroje v systému podnikového vyhledávání. Můžete vytvořit libovolný počet vyhledávacích aplikací a každá z těchto vyhledávacích aplikací může prohledávat libovolný počet kolekcí a externích zdrojů.

### Ukázková vyhledávací aplikace

Tato aplikace demonstruje mnoho funkcí vyhledávání a načítání, které jsou v prostředí podnikového vyhledávání k dispozici. Současně je ukázková vyhledávací aplikace funkčním příkladem demonstrujícím možnosti použití rozhraní IBM SI-API (Search and Index API) k sestavování vlastních interaktivních vyhledávacích aplikací přizpůsobených požadavkům a záměrům vašeho podniku.

Pokud nezměníte vlastnosti obsažené ve výchozím konfiguračním souboru, umožní vám ukázková vyhledávací aplikace prohledávat všechny aktivní kolekce a externí zdroje v systému podnikového vyhledávání. Prostřednictvím ukázkové vyhledávací aplikace můžete testovat nové kolekce a externí zdroje před jejich zpřístupněním pro uživatele.

Ukázková vyhledávací aplikace je automaticky asociována se všemi kolekcemi a externími zdroji. V provozním prostředí určují administrátoři podnikového vyhledávání, kterým vyhledávacím aplikacím má být povoleno prohledávání jednotlivých kolekcí.

### Vlastní vyhledávací aplikace

Vyhledávací aplikace, které vytvoříte, můžete spouštět jako samostatné webové aplikace v prostředí aplikačního serveru IBM WebSphere nebo jako portlety v prostředí IBM WebSphere Portal. Rozhraní SI-API vám umožňuje navrhovat vyhledávací aplikace, které podobně jako ukázková vyhledávací aplikace pracují bez dalších úprav v obou prostředích.

Při přizpůsobování vyhledávacích aplikací vám může pomoci nástroj Přizpůsobení vyhledávacích aplikací. Tato aplikace umožňuje vybírat volby v grafickém rozhraní a průběžně sledovat účinky provedených změn. Při uložení změn dojde k aktualizaci konfiguračního souboru vyhledávací aplikace.

#### Tip:

Podrobné příklady použití modulu Přizpůsobení vyhledávacích aplikací a instalace vyhledávací aplikace jako portletu portálu WebSphere Portal najdete v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

#### Související pojmy

“Volby vytváření indexu pro prohledávání dokumentů” na stránce 169

“Identifikátory vyhledávací aplikace” na stránce 248

“Zabezpečení na úrovni dokumentu” na stránce 248



Vyhledávací a indexové rozhraní API - přehled



Syntaxe dotazů



Nastavení vlastností dotazu

---

## Asociace vyhledávacích aplikací s kolekcemi

Než použijete novou vyhledávací aplikaci, je nutné asociovat ji s kolekcemi, které tato aplikace může prohledávat.

### Než začnete

Chcete-li asociovat vyhledávací aplikace s kolekcemi, které mohou prohledávat, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### Postup

Asociování vyhledávací aplikace s jednou či více kolekcemi:

1. Klepněte na panelu nástrojů konzoly pro správu na volbu **Zabezpečení**.
2. Na stránce Vyhledávací aplikace klepněte na volbu **Konfigurovat vyhledávací aplikace**.
3. Na stránce Konfigurovat vyhledávací aplikace klepněte na volbu **Přidat vyhledávací aplikaci**.
4. Zadejte název vyhledávací aplikace.
5. Vyberte kolekce, které může aplikace prohledávat:
  - Chcete-li, aby vyhledávací aplikace přistupovala ke všem kolekcím přidáním do systému, klepněte na volbu **Všechny kolekce a externí zdroje**.
  - Pokud chcete, aby vyhledávací aplikace měla přístup pouze k zadaným kolekcím, vyberte volbu **Specifické kolekce a externí zdroje**.

Po výběru této volby se zobrazí seznam názvů kolekcí a názvů externích zdrojů. U každé kolekce, kterou aplikace může prohledávat, zaškrtněte políčko **Vybrat**.
6. Klepněte na tlačítko **OK**.

---

## Funkce ukázkové vyhledávací aplikace

Ukázková vyhledávací aplikace pro podnikové vyhledávání demonstruje většinu vyhledávacích funkcí, které můžete používat ve vlastních vyhledávacích aplikacích.

Pomocí ukázkové vyhledávací aplikace můžete současně prohledávat všechny kolekce a externí zdroje. Pokud neupravíte výchozí vlastnosti aplikace, můžete ji používat k prohledávání všech kolekcí a externích zdrojů v systému podnikového vyhledávání.

### Dotazovací funkce

Tyto funkce vám poskytují následující možnosti:

- Zadávání jednoduchých dotazů pro volné prohledávání textu.
- Zadávání složitějších dotazů poskytujících přesnější výsledky vyhledávání. Můžete například prohledávat pouze určitá pole nebo prvky XML nebo na základě syntaxe dotazů vyhledávat dokumenty, které obsahují konkrétní slova a fráze nebo je naopak neobsahují.
- Určování kolekcí a externích zdrojů, které chcete prohledávat.
- Prohledávání konkrétních typů zdrojů nebo všech typů zdrojů.
- Hledání konkrétních typů dokumentů. Můžete například prohledávat pouze dokumenty Microsoft Word nebo dokumenty PDF (Portable Document Format).
- Určení jazyka, ve kterém jsou dotazovací výrazy zapsány. Můžete rovněž určit jazyky dokumentů, které chcete prohledávat.
- Výběr konkrétních podmnožin kolekce. Vyhledávací aplikace může například omezit zobrazení na předdefinovaný rozsah dokumentů (obor). Také můžete zadat dotaz, který prohledá pouze dokumenty patřící do určité kategorie.

- Rozšiřování dotazů na synonyma výrazů uvedených v dotazu. Pokud je ke kolekci přiřazen slovník synonym, budou ve výsledcích vyhledávání vráceny dokumenty obsahující synonyma výrazů uvedených v dotazu.

## Funkce výsledků vyhledávání

Tyto funkce vám poskytují následující možnosti:

- Zobrazení výsledků vyhledávání, které odpovídají vašemu dotazu.
- Určení počtu výsledných dokumentů, které se zobrazí na každé stránce, a procházení výslednou sadou směrem vpřed a vzad.
- Skrytí a zobrazení podrobností o výsledných dokumentech. Můžete například zobrazit stručné popisy dokumentů nebo podrobnosti, například názvy polí v každém výsledném dokumentu.
- Sbalení dokumentů pocházejících ze stejného zdroje. Vrací-li například jeden zdroj 100 dokumentů, ve výsledné sadě se zobrazí skupina tvořená prvními dvěma dokumenty, které nejlépe odpovídají zadanému dotazu. Zbylých 98 dokumentů lze zobrazit výběrem volby zobrazení dalších dokumentů ze stejného zdroje.
- Řazení dokumentů podle relevance, data dokumentu nebo podle hodnot v určitém poli. Při řazení podle data a podle pole můžete určit, zda chcete výsledky seřadit vzestupně nebo sestupně.
- Zobrazování navrhovaných oprav pravopisných chyb, pokud jsou v dotazovacím řetězci nalezena slova, která byla pravděpodobně chybně zapsána.
- Zobrazení informací o kategoriích, do nichž výsledný dokument patří (pokud kolekce používá kategorie), a omezení procházení na dokumenty patřící do určité kategorie.
- Zadání dalších výrazů dotazu, které mají být vyhledávány v rámci výsledků vyhledávání.

## Funkce načítání dokumentů

Tyto funkce vám poskytují následující možnosti:

- Načítání dokumentů klepnutím na identifikátor URI příslušného dokumentu - dokument se otevře ve webovém prohlížeči. Pokud je prolézací modul Notes nebo Domino Document Manager konfigurován pro použití protokolu DIOP, lze dokumenty zpracované tímto prolézacím modulem zobrazit v klientské aplikaci prohlížeče Lotus Notes namísto webového prohlížeče.  
Jde-li pro prolézací modul konfigurováno zabezpečení na úrovni dokumentu, budou moci načítat dokumenty pouze uživatelé s oprávněním pro přístup k zabezpečenému obsahu.
- Načtení dokumentů klepnutím na rychlé odkazy. Rychlé odkazy přiřazují identifikátory URI ke klíčovým slovům. Pokud dotaz obsahuje zadaná klíčová slova, budou asociované identifikátory URI (které jste předem označili jako vysoce významné pro tato klíčová slova) zobrazeny na začátku výsledků vyhledávání.

---

## Vlastnosti vyhledávací aplikace

Úpravou konfiguračního souboru vyhledávací aplikace můžete určit volby pro příslušné prostředí, změnit vzhled aplikace a řídit volby, které mají uživatelé k dispozici po spuštění vyhledávací aplikace.

Vlastnosti můžete upravit také pomocí nástroje Přizpůsobení vyhledávacích aplikací. Pokud vybíráte volby v tomto nástroji, vidíte ihned důsledky provedených změn. Jakmile budete spokojeni s volbami, které jste zadali pro prohledávání kolekcí a zobrazení výsledků vyhledávání, můžete tyto volby uložit a aktualizovat tak konfigurační soubor vyhledávací aplikace.

**Důležité:** Spustíte-li vyhledávací aplikaci jako portlet v produktu WebSphere Portal, nemůžete používat Přizpůsobení vyhledávacích aplikací k interaktivnímu provádění změn vyhledávací aplikace. Je nutné upravit vlastnosti a konfigurovat instanci portletu pomocí rozhraní správy produktu WebSphere Portal.

Konfigurační soubor ukázkové vyhledávací aplikace pro podnikové vyhledávání má název `config.properties`. Toto téma se zabývá vlastnostmi obsaženými v tomto souboru a popisuje výchozí vlastnosti. Pokud vytvoříte konfigurační soubory pro vlastní vyhledávací aplikaci, mohou se vlastnosti obsažené v těchto souborech a hodnoty těchto vlastností, které zadáte, lišit od popisovaných.

## Parametry prostředí

Můžete určit volby, které řídí činnost vyhledávací aplikace.

### **applicationName**

Určuje název platné vyhledávací aplikace. Výchozí hodnota je `Default`.

Chcete-li používat jinou než výchozí vyhledávací aplikaci, změňte výchozí hodnotu.

**Tip:** Je-li použit název aplikace `Default`, můžete pomocí ukázkové vyhledávací aplikace prohledávat všechny kolekce a externí zdroje jediným dotazem.

### **timeout**

Určuje dobu čekání na odezvu vyhledávacího serveru v sekundách, jejímž uplynutím vyprší časový limit vyhledávacího požadavku. Zadaná hodnota musí být celé číslo (např. 60, nikoli 60,5 ani "šedesát"). Pokud hodnotu časového limitu nezadáte, bude použita výchozí hodnota 30 sekund.

### **hostname**

Určuje úplný název hostitele webového serveru konfigurovaného pro podporu příslušné instance serveru WebSphere Application Server. Výchozí hodnota je `localhost`.

Chcete-li zajistit správnou práci aplikace, změňte výchozí hodnotu na úplný název hostitele, pro jehož použití je konfigurován produkt WebSphere Application Server. Pokud je například název hostitele lokálního počítače `Mujpocitac` a název hostitele webového serveru je `www.ibm.com`, zadejte `www.ibm.com`.

### **protocol**

Určuje protokol pro komunikaci s webovým serverem: `http` nebo `https`. Ponecháte-li tento parametr prázdný, bude použita výchozí hodnota `http`.

### **port**

Určuje číslo portu webového serveru konfigurovaného pro podporu použité instance produktu WebSphere Application Server. Výchozí hodnota je `80`, což je u protokolu HTTP typické. Typický port používaný u protokolu HTTPS je `443`.

### **trustStore**

Pokud používáte protokol HTTPS, zadejte úplnou kvalifikovanou cestu k souboru úložiště klíčů (databázový soubor s veřejnými klíči). Tyto informace, nazývané také *důvěryhodné úložiště*, umožňují použití protokolu SSL (Secure Sockets Layer) k důvěryhodné komunikaci. Chcete-li zadat cestu v systému Windows, musíte všechny znaky zpětných lomítek zdvojit. Příklad: `x:\\Application Server\\webserver.key`

### **trustPassword**

Pokud používáte protokol HTTPS, zadejte heslo k zadanému souboru úložiště klíčů.

### **username**

Vyhledávací aplikace tuto hodnotu automaticky nastaví na jméno uživatele, které uživatel zadal při přihlášení k vyhledávací aplikaci. Jiné jméno uživatele zde uveďte

jen tehdy, chcete-li změnit výchozí chování při ověřování uživatelů. Toto pole je používáno pouze v případě, že jste pro produkt WebSphere Application Server povolili globální zabezpečení.

#### **password**

Vyhledávací aplikace tuto hodnotu automaticky nastaví na heslo, které zadal uživatel při přihlášení k vyhledávací aplikaci. Heslo zde uveďte jen tehdy, jestliže jste zadali jméno uživatele. Toto pole je používáno pouze v případě, že jste pro produkt WebSphere Application Server povolili globální zabezpečení.

#### **ssoCookieName**

Určuje název souboru cookie, který obsahuje řetězec tokenu jednotného přihlášení (SSO). Výchozí hodnota je `LtpaToken`.

#### **proxyHost**

Určuje úplný hostitelský název serveru proxy, pokud je server proxy vyžadován pro přístup k vyhledávacímu serveru.

#### **proxyPort**

Určuje port odpovídající zadanému hostitelskému serveru proxy.

#### **proxyUser**

Určuje jméno uživatele, které má být použito pro přihlášení k serveru proxy, pokud server proxy vyžaduje základní ověřování.

#### **proxyPassword**

Určuje heslo odpovídající zadanému jménu uživatele pro server proxy.

#### **filter**

Určuje třídu, která má být použita pro načítání dokumentů uvedených ve výsledcích vyhledávání. Výchozí třída je `com.ibm.es.api.filters.SetDocumentURIFilterFetch`. Tuto hodnotu změňte pouze v případě, že máte k dispozici vlastní třídu, kterou chcete použít k načítání dokumentů místo výchozí třídy.

#### **logging.level**

Určuje množství zaznamenávaných podrobností:

**OFF** Nejsou zaznamenávány žádné zprávy.

#### **SEVERE**

Jsou protokolovány zprávy indikující závažné selhání. Tato hodnota je výchozí.

**INFO** Zaznamenávány jsou informační zprávy.

**FINE** Jsou protokolovány zprávy trasování s nízkou úrovní podrobností. (Tato volba odpovídá úrovni protokolování FINE ve třídě `java.util.logging.Level`.)

**ALL** Zaznamenávány jsou všechny zprávy.






## **Ikony typu zdroje**

Můžete přizpůsobit obrázky reprezentující typ zdroje dat, ke kterému výsledný dokument náleží. V souboru `config.properties` jsou předdefinovány následující ikony typu zdroje identifikující prolézací moduly a externí zdroje, které jsou podporovány v případě, že je instalován produkt OmniFind Enterprise Edition.



#### **documentSource.vbr.icon**

Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Content Edition. Výchozí ikona je `/images/sourceVBR.gif`.

-  **documentSource.db2.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem DB2. Výchozí ikona je /images/sourceDB2.gif.
-  **documentSource.cm.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem DB2 Content Manager. Výchozí ikona je /images/sourceCM.gif.
-  **documentSource.dominodoc.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Domino Document Manager. Výchozí ikona je /images/sourceDominoDoc.gif.
-  **documentSource.exchange.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Exchange Server. Výchozí ikona je /images/sourceExchange.gif.
-  **documentSource.database.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Databáze JDBC. Výchozí ikona je /images/sourceJDBC.gif.
-  **documentSource.nntp.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem NNTP. Výchozí ikona je /images/sourceNNTP.gif.
-  **documentSource.notes.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Notes. Výchozí ikona je /images/sourceNotes.gif.
-  **documentSource.quickplace.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem QuickPlace. Výchozí ikona je /images/sourceWorkplace.gif.
-  **documentSource.seedlist.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Počáteční seznam. Výchozí ikona je /images/sourceSeedlist.gif.
-  **documentSource.unixfs.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Souborový systém UNIX. Výchozí ikona je /images/sourceUnixFS.gif.
-  **documentSource.web.icon**  
Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen webovým prolézacím modulem. Výchozí ikona je /images/sourceWeb.gif.



**documentSource.wcm.icon**

Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Web Content Management. Výchozí ikona je `/images/sourceWorkplace.gif`.

**documentSource.wps.icon**

Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem WebSphere Portal. Výchozí ikona je `/images/sourceWPS.gif`.

**documentSource.winfs.icon**

Určuje název souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Souborový systém Windows. Výchozí ikona je `/images/sourceWindowsFS.gif`.

**documentSource.ldap.icon**

Určuje název souboru obrázku (včetně cesty), který indikuje, že dokument náleží k externímu zdroji vytvořenému pro server LDAP. Výchozí ikona je `/images/sourceLDAP.gif`.

**documentSource.jdbc.icon**

Určuje název souboru obrázku (včetně cesty), který indikuje, že dokument náleží k externímu zdroji vytvořenému pro databázovou tabulku JDBC (Java Database Connectivity). Výchozí ikona je `/images/sourceJDBC.gif`.

## Ikony prohlížeče klienta

Výsledné dokumenty mohou být zobrazeny ve webovém prohlížeči. Dokumenty prolezené prolézacími moduly Notes nebo Domino Document Manager konfigurovanými pro použití protokolu DIIOP lze také zobrazit v aplikaci prohlížeče klienta Lotus Notes.

Chcete-li umožnit zobrazení dokumentů v aplikaci prohlížeče klienta, zastavte následující vlastnost na hodnotu true (zapnuto):

```
clientViewer.show=true
```

V případě potřeby můžete přizpůsobit obrázky reprezentující aplikaci prohlížeče klienta. V následujícím příkladu indikuje ikona aplikace Lotus Notes, že dokument lze zobrazit v aplikaci prohlížeče:

```
client.notes.icon=/images/notes.gif
client.dominodec.icon=/images/notes.gif
```

Ve výsledcích vyhledávání jsou ikona a odkaz na aplikaci prohlížeče klienta zobrazeny následujícím způsobem:

**Prohlížeč klienta**

## Pole dokumentů

Pro typy zdrojů dat obsahujících pole můžete určovat, která pole budou ve výsledných dokumentech zobrazována.

**fields.URI prefix=***seznam\_názevů\_polí\_oddělených\_mezerou*

Význam znaku dvojtečka (:) v předponě identifikátoru URI musíte změnit zadáním

znaku obrácené lomítka (\) před tento znak. Chcete-li v seznamu názvů polí pokračovat na dalším řádku, ukončete předchozí řádek znakem obrácené lomítka (\).  
Příklad:

```
fields.db2\://=databasename tablename  
fields.domino\://=databasetitle filename creator  
fields.dominodoc\://=librarydbtitle documentdbtitle filename author  
fields.exchange\://=from creator  
fields.file\://=directory filename  
fields.https\://=documentID  
fields.http\://=documentID  
fields.jdbc\://=databasename tablename  
fields.news\://=group from  
fields.quickplace\://=placetitle roomtitle creator  
fields.seedlist\://=author  
fields.vbr\://=itemname repositorytype revisionuser  
fields.wcm\://=author owner modifier  
fields.web\://=  
fields.wp6\://=  
fields.wps\://=
```

## Ikony polí

Pro typy zdrojů dat a dokumenty obsahující pole můžete přizpůsobit obrázky reprezentující pole. Všechna pole nad souhrnem dokumentu obsahují identifikační obrázky. V souboru `config.properties` jsou předdefinovány následující ikony polí.



### **field.icon.databasetitle**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje název dokumentu. Výchozí ikona je `/images/notesdb.gif`.



### **field.icon.databasename**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje název databáze, do níž příslušný dokument náleží. Výchozí ikona je `/images/db2.gif`.



### **field.icon.tablename**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje název tabulky, do níž příslušný dokument náleží. Výchozí ikona je `/images/table.gif`.



### **field.icon.directory**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje název adresáře, do něž příslušný dokument náleží. Výchozí ikona je `/images/closedFolder.gif`.



### **field.icon.filename**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje název souboru s dokumentem. Výchozí ikona je `/images/document.gif`.

### **field.icon.documentID**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje identifikátor dokumentu. Tento prázdný obrázek můžete použít ve webových dokumentech například k určení obrázku pro adresu URL, aniž by byl tento obrázek viditelný pro uživatele. Výchozí ikona je `/images/dot.gif`.



### **field.icon.group**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje identifikátor dokumentu. Tento prázdný obrázek můžete použít ve webových dokumentech například k určení obrázku pro adresu URL, aniž by byl tento obrázek viditelný pro uživatele. Výchozí ikona je `/images/document.gif`.

**field.icon.from**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje informace o odesilatelci dokumentu. Výchozí ikona je `/images/author.gif`.

**field.icon.creator**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje informace o tvůrci dokumentu. Výchozí ikona je `/images/author.gif`.

**field.icon.author**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje informace o autorovi dokumentu. Výchozí ikona je `/images/author.gif`.

**field.icon.revisionuser**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje informace o korektorovi dokumentu. Výchozí ikona je `/images/author.gif`.

**field.icon.owner**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje informace o vlastníkovi dokumentu. Výchozí ikona je `/images/author.gif`.

**field.icon.modifier**

Určuje název souboru obrázku (včetně cesty), který indikuje, že pole obsahuje informace o uživateli, který dokument upravil. Výchozí ikona je `/images/author.gif`.

## Výchozí ikona pole

Můžete určit obrázek, který má být použit v případě, že pro pole zobrazovaná ve výsledcích vyhledávání nejsou konfigurovány žádné ikony polí. Výchozí ikona pole je předdefinována v souboru `config.properties`.

**field.defaultIcon**

Určuje název souboru obrázku (včetně cesty), který je výchozí ikonou pro pole ve výsledcích vyhledávání. Výchozí ikona je `/images/database.gif`.

## Pole s datem

Můžete určit, která pole jsou poli s datem. Pole s názvy zadanými v této části jsou ve výsledcích vyhledávání formátována jako data určující datum. Formát data odpovídá nastavením lokality ve webovém prohlížeči.

**date.fields=seznam\_názevů\_polí\_oddělených\_čárkami**

Chcete-li v seznamu názvů polí pokračovat na dalším řádku, ukončete předchozí řádek znakem obrácené lomítka (`\`).

Příklad:

```
date.fields=modifieddate createddate
```

## Názvy dokumentů

Můžete určit alternativní názvy pro dokumenty nahrazením textu názvu smyslupnějšími daty (tedy *pročistit* názvy). Místo zobrazování názvů dokumentů s nadpisem Snímek 1, který je zcela neinformativní, můžete například určit, že název Snímek 1 má být ve výsledcích vyhledávání potlačen. Pro identifikaci výsledného dokumentu může být místo názvu použito pole s vyšší informační hodnotou, například název souboru.

Můžete rovněž určit alternativní názvy pro dokumenty odebráním slov bez informační hodnoty z názvů dokumentů (tedy *oříznout* názvy). Pokud například značný počet výsledných dokumentů začíná textem Microsoft Word -, můžete zvýšit přehlednost výsledků vyhledávání potlačením opakujícího se textu na začátku.

**titles.clean=seznam\_názvů\_oddělených\_čárkami**

**titles.truncatePrefix=seznam\_předpon\_oddělených\_čárkami**

Seznamy s čárkami jako oddělovači mohou obsahovat mezery a další znaky s výjimkou čárek. Chcete-li v seznamu pokračovat na dalším řádku, ukončete předchozí řádek znakem obrácené lomítka (\).

Příklad:

```
titles.clean=Snímek 1, Vzhled 1, bez názvu, \
Dokument bez názvu, Presentace aplikace PowerPoint, \
(stránka bez názvu)
```

```
titles.truncatePrefix=Microsoft Word -, Microsoft Powerpoint -
```

## Výchozí hodnoty pro uživatelské předvolby

Můžete určit výchozí hodnoty pro stránku Předvolby ve vyhledávací aplikaci. Pokud uživatel změní předvolby, projeví se nové hodnoty pouze pro aktuální relaci uživatele. V souboru `config.properties` jsou předdefinovány následující předvolby.

**preferences.resultsRange=10**

Určuje, že na každé stránce výsledků vyhledávání může být uvedeno 10 výsledných dokumentů.

**preferences.siteCollapsing=Yes**

Určuje, že identifikátory URI ze stejného zdroje mají být ve výsledcích vyhledávání sbalené. Sbalení webu je k dispozici pouze v případě, že jsou výsledky seřazeny podle relevance. U zdrojů dat webu a NNTP se identifikátory URI, které odpovídají identifikátoru URI kořenového webu (například `www.ibm.com`), sbalují automaticky. Pro jiné typy zdrojů dat a pro weby s hlubší strukturou adresářů (například `www.ibm.com/hr`) musí být pravidla pro sbalení webu konfigurována v konzole správy podnikového vyhledávání.

**preferences.spellCorrections=Yes**

Určuje, že mají být zobrazovány navrhované opravy pravopisu, pokud uživatel odešle dotaz obsahující slovo, které pravděpodobně obsahuje chybu. Pověšimněte si, že před vygenerováním navrhovaných oprav jsou vždy odebrána ignorovaná slova.

**preferences.extendedHighlighting=No**

Určuje, že výrazy dotazu nebudou ve speciálních polích (například v názvu dokumentu) mimo pole souhrnu dokumentu zvýrazňovány.

## Výchozí kolekce a externí zdroje

Můžete určit, které kolekce a externí zdroje jsou předem vybrány na stránkách Předvolby a Rozšířené vyhledávání. Uživatelé mohou upravovat výchozí sadu a prohledávat méně kolekcí a externích zdrojů, než je povoleno při výchozím nastavení. Omezíte-li sadu kolekcí a externích zdrojů na tomto místě, uživatelé budou moci při úpravách předvoleb nebo voleb rozšířeného vyhledávání vybrat kteroukoli kolekci nebo externí zdroj dostupný pro vyhledávací aplikaci.

**preferences.defaultCollections=\***

**preferences.defaultCollections=seznam\_identifikátorů\_kolekcí\_oddělených\_mezerou**

Chcete-li povolit prohledávání všech kolekcí a externích zdrojů, zadejte znak

hvězdička (\*). (Kolekce a externí zdroje musí být asociovány s vyhledávací aplikací prostřednictvím konzoly pro správu podnikového vyhledávání.) Jde o výchozí nastavení v souboru `config.properties`.

Chcete-li omezit obor prohledávání pro případ, že uživatelé neupraví příslušné předvolby nebo volby rozšířeného vyhledávání, určete identifikátory kolekcí a externích zdrojů, jejichž prohledávání chcete uživatelům při výchozím nastavení povolit.

Příklad:

```
preferences.defaultCollections=*  
preferences.defaultCollections=ID_kolekce_1 ID_kolekce_2
```

## Další informace ve výsledcích vyhledávání

Můžete přizpůsobit množství informací zobrazovaných ve výsledcích vyhledávání a určovat, zda uživatelé budou moci filtrovat výsledky vyhledávání. Následující nastavení jsou výchozími nastaveními v souboru `config.properties`.

### **refreshButton.show=false**

Určuje, zda se má na stránce základního vyhledávání zobrazovat tlačítko **Aktualizovat**. U rozšířeného vyhledávání je tlačítko **Aktualizovat** k dispozici vždy. Pokud tuto volbu nastavíte na hodnotu `true` (zapnuto), budou uživatelé moci aktualizovat seznam kolekcí a externích zdrojů dostupných pro vyhledávání.

Pokud používáte nástroj Přizpůsobení vyhledávacích aplikací, tlačítko **Aktualizovat** nebudete potřebovat.

Pokud nástroj Přizpůsobení vyhledávacích aplikací nepoužíváte, může být vhodné zobrazit tlačítko **Aktualizovat** při testování změn v konfiguračním souboru. Po uložení změn můžete klepnout na tlačítko **Aktualizovat** a zjistit, jak změny ovlivní vyhledávací aplikaci. Pokud není zobrazeno tlačítko **Aktualizovat**, projeví se změny až po restartování aplikace `ESSearchServer` na aplikačním serveru `WebSphere`.

Pokud pro prohledávání nejsou k dispozici žádné kolekce ani externí zdroje (pokud například byl zadán nesprávný název hostitele, nebyly spuštěny vyhledávací servery nebo nebyla spuštěna aplikace `ESSearchServer` na serveru `WebSphere Application Server`), bude tlačítko **Aktualizovat** zobrazeno automaticky, aby vám usnadnilo odstraňování problému.

### **builtQueryString.show=false**

Řídí zobrazování plně expandované syntaxe dotazu v oblasti před seznam výsledných dokumentů. Chcete-li zobrazovat příslušný zpracovaný dotaz, nastavte tuto volbu na hodnotu `true` (zapnuto).

### **extraQueryData.show=false**

Řídí zobrazení dalších informací o dotazu. Tuto volbu nastavte na hodnotu `true` (zapnuto), chcete-li zobrazit informace o omezeních seznamu ACL, názvech prohledávaných kolekcí a externích zdrojů a o jazyku dotazu.

### **refineResults.show=true**

Řídí, zda mohou uživatelé zpřesňovat výsledky vyhledávání zadáváním dalších výrazů dotazu. Nastavíte-li tuto volbu na hodnotu `true` (zapnuto), zobrazí se ve spodní části stránky s výsledky vyhledávání pole dotazu s označením **Vyhledat ve výsledcích**.

### **sorting.show=true**

Určuje, zda se má zobrazit volba pro seřazení výsledků vyhledávání. Nastavením této volby na hodnotu `false` (vypnuto) potlačíte zobrazení voleb **Řadit podle a Pořadí řazení**, které slouží k seřazení výsledků vyhledávání.

**sourceTypeFilter.show=true**

Určuje, zda se má ve výsledcích vyhledávání zobrazit volba pro filtrování výsledků podle typu zdroje. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li uživatelům umožnit filtrování výsledků podle typu zdroje.

Chcete-li zabránit uživatelům ve filtrování výsledků podle typu dokumentu, odstraňte vybrané nebo všechny položky typů dokumentů v konfiguračním souboru (`documentType.popis=typy_dokumentů`).

**filter.showOnTwoLines=true**

Řídí, zda budou volby filtrování výsledků podle typu zdroje a filtrování výsledků podle typu souboru zobrazovány ve výsledcích vyhledávání na jednom nebo dvou řádcích. Při zobrazování výsledků vyhledávání mohou uživatelé vybrat typ zdroje a typ souboru a zobrazit pouze výsledné dokumenty odpovídající vybraným filtrům.

Chcete-li maximalizovat množství prostoru, který je k dispozici pro zobrazování výsledků vyhledávání, nastavte tuto vlastnost na hodnotu false (vypnuto). Chcete-li zvýšit přehlednost použití filtrů, zejména v případě, že filtry k dispozici přesahují jeden řádek, můžete nastavit tuto vlastnost na hodnotu true (zapnuto), aby byl každý filtr zobrazen na samostatném řádku.

**clientViewer.show=true**

Určuje, zda má být k zobrazení dokumentu s výsledky použita aplikace prohlížeče klienta Lotus Notes. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li, aby se dokumenty Domino zobrazovaly v aplikaci prohlížeče.

**showDetails.show=true**

Řídí zobrazení odkazů Zobrazit podrobnosti a Skrýt podrobnosti ve výsledcích vyhledávání. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li uživatelům umožnit zobrazení dalších podrobností o dokumentech s výsledky.

**showDetailsImage.show=true**

Řídí zobrazení podrobných informací o dokumentech s výsledky v okně. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li uživatelům umožnit zobrazení dalších podrobností o dokumentech s výsledky přesunutím ukazatele na identifikátor URI dokumentu.

**numberSearchResultsReturned.show=true**

Určuje, zda se má zobrazit celkový počet výsledků vyhledávání. Tuto volbu nastavte na hodnotu false (vypnuto), pokud nechcete, aby uživatelé viděli, kolik dokumentů bylo vráceno ve výsledcích vyhledávání.

**showMessage.error=true**

Řídí zobrazení chybových zpráv. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li, aby se v horní části okna vyhledávací aplikace zobrazovaly chybové zprávy.

**showMessage.warning=true**

Řídí zobrazení varovných zpráv. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li, aby se v horní části okna vyhledávací aplikace zobrazovaly varovné zprávy.

**showMessage.info=true**

Řídí zobrazení informativních zpráv. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li, aby se v horní části okna vyhledávací aplikace zobrazovaly informativní zprávy.

**showMessage.success=true**

Řídí zobrazení zpráv o úspěšném provedení operací. Tuto volbu nastavte na hodnotu false (vypnuto), nechcete-li, aby se v horní části stránky vyhledávací aplikace zobrazovaly zprávy oznamující úspěšné dokončení akce.

## Vlastní titulek a logo

Můžete přizpůsobit obrázky zobrazované v oblasti titulku v horní části vyhledávací aplikace. Můžete například nahradit výchozí obrázky pro produkt OmniFind Enterprise Edition obrázky odpovídající značce vašeho podniku. Nechcete-li zobrazovat titulek, můžete jeden nebo oba tyto řádky označit jako komentář. Vlastnost `banner.icon` identifikuje obrázek, který je zobrazován v levé části oblasti titulku. Vlastnost `banner2.icon` identifikuje obrázek, který je zobrazován v pravé části oblasti titulku.

```
banner.icon=/images/WS_II_OFEdition.gif  
banner2.icon=/images/WS_II_mosaic.gif
```

## Vlastní obrázek na pozadí

Můžete přizpůsobit obrázky zobrazované na pozadí stránek ve vyhledávací aplikaci. Můžete například nahradit výchozí obrázky pro podnikové vyhledávání obrázky vycházející z vaší firemní značky. Nechcete-li na stránce zobrazovat obrázek na pozadí, můžete jeden nebo více těchto řádků označit jako komentář.

```
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
```

## Odkazy

Vlastnosti v oblasti Links (Odkazy) souboru `config.properties` umožňují zobrazovat názvy stránek vyhledávací aplikace jako odkazy na každé stránce místo jejich zobrazování na panelu nástrojů a na stránkách s kartami. Zobrazování odkazů je užitečné v případě, že vyhledávací aplikaci spouštíte jako portlet a chcete minimalizovat množství místa použitého pro zobrazení vyhledávací aplikace na stránce portálu.

Pokud vám více vyhovuje navigace vyhledávací aplikace prostřednictvím výběru voleb na panelu nástrojů a na stránkách s kartami, označte tyto řádky jako komentář.

## Karty vyhledávání

Vlastnosti v oblasti Search tabs (Karty vyhledávání) souboru `config.properties` určují názvy stránek JSP (Java Server Pages) používaných pro stránky s kartami v pohledu Vyhledávání (Základní vyhledávání, Rozšířené vyhledávání a Strom kategorií) vyhledávací aplikace. Neupravujte tyto stránky, nemáte-li zkušenosti s programováním v jazyku Java a se stránkami JSP.

Mezi příklady možného přizpůsobení této oblasti patří následující:

- Nasměrování vyhledávací aplikace na vlastní stránky JSP s jiným vzhledem stránek s kartami.
- Označení položek stránky Strom kategorií jako komentáře. Pokud například nekonfigurujete kategorie pro kolekce, není třeba zobrazovat stránku Strom kategorií ve vyhledávací aplikaci.
- Zkopírování položek odpovídajících stránkám s kartami do oblasti Toolbars (Panely nástrojů) souboru `config.properties` a označení těchto řádků jako komentáře. Můžete například chtít zobrazit pouze panel nástrojů, ale nezobrazovat žádné stránky s kartami.

## Panely nástrojů

Vlastnosti v oblasti Toolbars (Panely nástrojů) souboru `config.properties` určují názvy stránek JSP (Java Server Pages) používaných pro panel nástrojů ve vyhledávací aplikaci. Neupravujte tyto stránky, nemáte-li zkušenosti s programováním v jazyku Java a se stránkami JSP.

Mezi příklady možného přizpůsobení této oblasti patří následující:

- Nasměrování vyhledávací aplikace na vlastní stránky JSP s jiným vzhledem panelu nástrojů.
- Označení položek odpovídajících panelu nástrojů pro položky, které nechcete zobrazovat, jako komentáře. Můžete se například rozhodnout, že na panelu nástrojů nebude zobrazen odkaz na stránku Informace.
- Přesunutí funkce pro zobrazení stránky Rozšířené vyhledávání z oblasti karty souboru `config.properties`, aby byla tato volba k dispozici pouze na panelu nástrojů.

## Smysluplné popisky typů dokumentů

Přehlednost použití filtru typů dokumentů můžete zvýšit mapováním skutečných názvů typů dokumentů na stručnější výrazy s vyšší informační hodnotou. Typy dokumentů, které jsou k dispozici pro vyhledávací aplikace, jsou definovány pomocí třídy `AvailableDocumentTypes` rozhraní SI-API (vyhledávací a indexové rozhraní API). Typy dokumentů, které jsou k dispozici, jsou pro pohodlnější přístup uvedeny i na konci souboru `config.properties`.

**documentType.popisek=seznam\_typů\_dokumentů\_oddělených\_čárkami**

Určuje název zobrazovaný na řádku filtru typů dokumentů ve výsledcích vyhledávání a seznam typů dokumentů, které mají být zobrazovány, když uživatel vybere filtr.

Můžete například zadat popisek `html` a mapovat přípony souborů a typy MIME pro různé webové dokumenty na tento název. Když uživatel klepnutím na volbu **html** filtruje výsledky vyhledávání, zobrazí se pouze dokumenty s určenými příponami a typy MIME.

V souboru `config.properties` jsou předdefinována následující mapování typů dokumentů:

```
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
```

Pokud se hodnota zadaná pro popisek typů dokumentů shoduje s názvem vlastnosti v souboru `application.properties`, zobrazí se hodnota této vlastnosti v souboru `application.properties`, a nikoli hodnota zadaná zde. Zadáte-li například pro popisek filtru typů souborů hodnotu `documentType.unixfs`, zobrazí se jako název filtru typů souborů hodnota vlastnosti `unixfs` v souboru `application.properties` (**Souborový systém UNIX**).

## Vlastní filtry

Zobrazení výsledných dokumentů můžete filtrovat pomocí vlastních dotazů, které zadáte.

**filterCustom.popisek=výraz\_dotazu**

Určuje název zobrazovaný na řádku vlastního filtru ve výsledcích vyhledávání a dotaz, který zpřesňuje výsledky vyhledávání, když uživatel vybere filtr. (Při



zobrazování výsledků vyhledávání mohou uživatelé vybrat vlastní filtr a zobrazit pouze výsledné dokumenty odpovídající předdefinovanému dotazu.)

V následujícím příkladu jsou výsledky vyhledávání filtrovány a zobrazeny jsou pouze dokumenty, které náležejí do databáze personálního oddělení (hr):

```
filterCustom.HR_database_only=databasename::hr
```

Když uživatel klepnutím na volbu **HR\_database\_only** filtruje výsledky vyhledávání, je zpracován dotaz `databasename::hr`. Ve výsledcích vyhledávání jsou zobrazeny pouze dokumenty z databáze hr.

Pokud se hodnota zadaná pro popis vlastního filtru shoduje s názvem vlastnosti v souboru `application.properties`, zobrazí se hodnota této vlastnosti v souboru `application.properties`, a nikoli hodnota zadaná zde. Zadáte-li například pro popis vlastního filtru hodnotu `filterCustom.hostData=`, zobrazí se jako název vlastního filtru hodnota vlastnosti `hostData` v souboru `application.properties` (**Nastavení serveru**).

V souboru `config.properties` je několik vlastních filtrů označeno jako komentář a uvedeno jako příklady.

## Detekce duplikátů

Při přidávání dokumentů do indexu podnikového vyhledávání je prováděna analýza k odebrání duplikátů, aby se uživatelům ve výsledcích vyhledávání nezobrazovaly dokumenty vícekrát. Chcete-li dále filtrovat výsledky vyhledávání, můžete zadat volbu k potlačení dokumentů, které jsou téměř identické, a zabránit jejich zobrazení ve výsledcích vyhledávání.

### **preferences.nearDuplicateDetection=No**

Tato volba určuje, že téměř duplicitní dokumenty nejsou v průběhu zpracování dotazů filtrovány.

Zadáte-li volbu **Ano**, budou dokumenty s podobnými názvy a souhrny při zobrazení ve výsledcích vyhledávání potlačeny. Zobrazí se zpráva informující uživatele, že některé dokumenty byly vynechány, protože jsou podobné jiným dokumentům v sadě výsledků. Klepnutím na odkaz mohou uživatelé zakázat potlačení a zobrazit všechny dokumenty v sadě výsledků.

K potlačení téměř duplicitních dokumentů musí být v objektu `Dotaz` rozhraní `SI-API` (Search and Index API) pro vyhledávací aplikaci zadána metoda `setProperty` s řetězcem `NearDuplicateDetection` nastaveným na hodnotu `Ano` (například `query.setProperty("NearDuplicateDetection", "Yes");`).

## Analýza nejlepších výsledků (sloupcové grafy pro pole metadat)

Můžete zadat volby pro zobrazení sloupcových grafů, které reprezentují analýzu nejlepších výsledků. Při výchozím nastavení je analyzováno 500 nejlepších výsledků. Každý graf odpovídá jednomu poli metadat a každý sloupec v grafu odpovídá jedné hodnotě pole. Délka sloupce indikuje četnost výskytu hodnoty pole. Čím delší je sloupec, tím větší je počet výskytů hodnoty pole.

Při konfiguraci prolézaného prostoru pro prolézací modul můžete zadat volby pro pole metadat. Chcete-li konfigurovat graf nejlepších výsledků pole metadat, je nutné zadat, že se jedná o pole s možností prohledávání a že je lze vrátit ve výsledcích vyhledávání.

Vlastnosti, které konfiguruje pro sloupcový graf, mají následující formát, kde *číslo* je číslem, které jednoznačně identifikuje sloupcový graf ve vyhledávací aplikaci, *volba* je volbou sloupcového grafu a *hodnota* je hodnotou následující volby:

```
topResultsChartsčíslo.volba=hodnota
```

**topResultsCharts***číslo.titleKey=klíč\_aplikace*

Určuje název grafu, kde *číslo* jednoznačně určuje graf ve vyhledávací aplikaci a *klíč\_aplikace* je popisec pro název grafu. Tento popisec může být klíčem v souboru `application.properties` pro vyhledávací aplikaci nebo hodnotou zadanou zde. V následujícím příkladu popisec pro název grafu je určen hodnotou klíče `topResults.mostRecentDocuments` v souboru `application.properties:topResultsCharts3.titleKey=topResults.mostRecentDocuments`.

Další příklad ukazuje, že název grafu je hodnota zadaná zde:  
`topResultsCharts3.titleKey=Organizations`

**topResultsCharts***číslo.enable=true*

Určuje, zda má být tento sloupcový graf zobrazen při zobrazení výsledků vyhledávání. Zadáte-li hodnotu `false`, graf zobrazen nebude.

**topResultsCharts***číslo.fieldName=název\_pole*

Určuje název pole metadat, jehož hodnoty budou pro tento graf analyzovány. Například `databasetitle`. Je nutné zadat jiný název pole pro každý přidávaný graf (stejně pole nelze použít pro více grafů).

**topResultsCharts***číslo.maxValues.collapsed=číslo*

Určuje počet sbalených položek, které se mají v grafu zobrazit. Například při zadání hodnoty 5 se zobrazí pět hodnot v tomto poli, k jejichž výskytu dochází nejčastěji.

**topResultsCharts***číslo.maxValues.expanded=číslo*

Určuje počet položek, které se mají v grafu zobrazit při úplném rozbalení zobrazení grafu. Například při zadání hodnoty 10 se při rozbalení grafu zobrazí deset různých hodnot polí.

**topResultsCharts***číslo.fieldValueSeparator=znak*

Určuje znak, který odděluje hodnoty v analyzovaném poli. Pokud pole například obsahuje více hodnot oddělených středníkem (například `agent;prodejce;zprostředkovatel`), můžete tuto vlastnost použít k určení středníku (;) jako oddělovače hodnot pole, aby byla každá hodnota přidána do sloupcového grafu jako samostatná položka. Pokud tato volba není vybrána, bude celá hodnota pole do grafu přidána jako jedna položka.

**topResultsCharts***číslo.canUserChangeFieldName=true*

Určuje, zda uživatel může vybrat jiné pole při zobrazení výsledků vyhledávání a dále zobrazit nejlepší výsledky pro dané pole. Zadáte-li hodnotu `false`, uživatel nemůže vybrat jiné pole k analýze při zobrazení výsledků vyhledávání.

Zadáte-li hodnotu `true`, zobrazí se seznam všech polí, která se nachází v původních nejlepších 500 výsledcích, spolu s výsledky pro dané pole. Pokud uživatel vybere pole z tohoto seznamu, změní se popisec grafu na **Graf dynamických polí** a výsledky sloupcového grafu pro vybrané pole budou zobrazeny do doby, než uživatel vybere jiné pole nebo zavře prohlížeč. Při dalším spuštění vyhledávací aplikace se zobrazí výsledky původního pole.

**topResultsCharts***číslo.width=číslo*

Určuje šířku zobrazení sloupcového grafu v pixelech. Například hodnota 300.

**topResultsCharts***číslo.barheight=číslo*


Určuje výšku jednotlivých sloupců grafu v pixelech. Například hodnota 10.

**topResultsCharts***číslo.color=#kód\_barvy*

Určuje hexadecimální kód základní barvy sloupce v grafu. Výchozí hodnotou je modrá (#0309C0).

**topResultsCharts***číslo.color.gradient=#kód\_barvy*

Určuje barvu gradientu sloupce v grafu. Výchozí hodnotou je tyrkysová (#00FFFF).

Při přesunu výsledků pro konkrétní hodnotu pole z nižší na vyšší frekvenci se změní odstín barvy sloupce z hodnoty zadané ve volbě **COLOR** na hodnotu zadanou ve volbě **color.gradient**. Příklad: 

#### **topResultsCharts***číslo.sortKey=frequency*

Určuje způsob řazení položek ve sloupcovém grafu. Podporované hodnoty:

**none** Položky ve sloupcovém grafu nejsou řazeny.

**label** Položky ve sloupcovém grafu jsou řazeny podle hodnoty pole. Pokud konfigurujete vlastní popisky s vlastností **topResultsCharts.číslo.listOfLabels.prefixKey=název\_pole**, budou položky ve sloupcovém grafu řazeny i podle názvu popisku.

#### **frequency**

Položky ve sloupcovém grafu jsou řazeny podle počtu výsledků vrácených pro hodnotu pole.

#### **topResultsCharts***číslo.sortOrder=descending*

Určuje pořadí řazení položek, které jsou řazeny podle popisku nebo frekvence. Podporované hodnoty:

#### **ascending**

Položky řazené podle popisku jsou uvedeny v abecedním pořadí od a do z. U položek řazených podle frekvence se hodnoty polí s častějším výskytem zobrazí v seznamu níže než hodnoty s méně častým výskytem.

#### **descending**

Položky řazené podle popisku jsou uvedeny v obráceném abecedním pořadí od z do a. U položek řazených podle frekvence se hodnoty polí s častějším výskytem zobrazí v seznamu výše než hodnoty s méně častým výskytem.

#### **topResultsCharts***číslo.listOfLabels.prefixKey=název\_pole*

Volitelné. Umožňuje zadat informace, které se vždy zobrazí pro konkrétní hodnoty pole, kde *název\_pole* určuje pole, jehož hodnota je analyzována pro daný graf. Popisky zadané pro tuto vlastnost se vždy zobrazí ve sloupcovém grafu, a to i v případě, že se hodnota pole ve výsledcích vyhledávání nevyskytuje.

Pro pole můžete konfigurovat libovolný počet popisků. Pro každý popisek zadejte dvě vlastnosti v následujícím formátu:

*název\_pole***číslo.value=hodnota**

*název\_pole***číslo.displayValue=zobrazená\_hodnota**

kde:

*název\_pole*

Jedná se o název pole, pro které konfigurujete popisky.

*číslo*

Jedná se o číslo, které jednoznačně identifikuje popisek.

*hodnota*

Jedná se o hodnotu, kterou chcete zobrazit ve výsledcích analýzy.

*zobrazená\_hodnota*

Jedná se o popisek zobrazený ve sloupcovém grafu. Tento popisek může být klíčem v souboru **application.properties** pro vyhledávací aplikaci nebo hodnotou zadanou zde.

Příklad:

```
topResultsCharts1.listOfLabels.prefixKey=databasetitle
```

```
databasetitle1.value=JK Enterprises Články a přednášky  
databasetitle1.displayValue=Články a přednášky
```

```
databasetitle2.value=JK Enterprises Prázdné formuláře  
databasetitle2.displayValue=Prázdné formuláře
```

```
databasetitle3.value=JK Enterprises Brožury a pokyny  
databasetitle3.displayValue=Brožury a pokyny
```

## Analýza nejlepších výsledků (vlastní kód HTML)

Volby zobrazení grafů nejlepších výsledků můžete zadat rozšířením rozhraní API `com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` pomocí vlastní třídy jazyka Java. Při použití tohoto přístupu k zobrazení výsledků můžete pomocí kódu HTML formátovat návratovou hodnotu libovolných výsledků vyhledávání, a ne pouze dotazů prohledávajících pole metadat. Výsledky mohou být vráceny pro dokumenty v kolekcích podnikového vyhledávání nebo z vyhledávání externích úložišť a webových stránek.

Pro vyhledávací aplikaci můžete konfigurovat libovolný počet grafů. K přiřazení názvů ke grafům použijte následující vlastnost:

**topResultsCharts***číslo*.titleKey=*klíč\_aplikace*

Určuje název grafu, kde *číslo* jednoznačně určuje graf ve vyhledávací aplikaci a *klíč\_aplikace* je popis pro název grafu. Tento popis může být klíčem v souboru `application.properties` pro vyhledávací aplikaci nebo hodnotou zadanou zde. V následujícím příkladu popis pro název grafu je určen hodnotou klíče `topResults.mostRecentDocuments` v souboru `application.properties:topResultsCharts3.titleKey=topResults.mostRecentDocuments`.

Každá položka v grafu odpovídá jedné hodnotě výsledků vyhledávání a skládá se ze sady vlastností s následujícím formátem, kde *číslo* je číslem, které jednoznačně identifikuje daný graf, *volba* je volbou grafu a *hodnota* je hodnotou následující volby:

`topResultsCharts`*číslo*.volba=*hodnota*

**topResultsCharts***číslo*.enable=true

Určuje, zda má být tento graf zobrazen při zobrazení výsledků vyhledávání. Zadáte-li hodnotu `false`, graf zobrazen nebude.

**topResultsCharts***číslo*.maxValues.collapsed=*číslo*

Určuje počet sbalených položek, které se mají v grafu zobrazit. Například při zadání hodnoty `5` se zobrazí pět hodnot odpovídajících dané podmínce, k jejichž výskytu dochází nejčastěji.

**topResultsCharts***číslo*.width=*číslo*

Určuje šířku zobrazení grafu v pixelech. Například hodnota `400`.

**topResultsCharts***číslo*.dynamicChartClass=*vlastní\_třída*

Určuje název vlastní třídy jazyka Java, která rozšiřuje rozhraní API `com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` a definuje způsob zobrazení výstupu v grafu. Příklad:

```
topResultsCharts.3.dynamicChartClass=com.ibm.es.searchui.charts.servlet.  
DynamicMostRecentDocuments
```

Ukázková třída jazyka Java `DogearSearchResults` používá následující vlastnosti zahrnuté ve výchozím souboru `config.properties` ukázkové vyhledávací aplikace pro účely příkladu. Další informace o parametrech dotazů, které je vhodné zahrnout do vlastní vyhledávací aplikace, naleznete v dokumentaci rozhraní Dogear API.

Tyto vlastnosti nelze nastavit pomocí Přizpůsobení vyhledávacích aplikací:

**topResultsCharts***číslo.xml.fileName=soubor\_stylů*

Určuje cestu a název listu stylu XSL, který slouží k formátování zobrazení nejlepších výsledků v grafu. Například soubor /styles/dogear.xml.

**topResultsCharts***číslo.url=adresa\_url*

Určuje adresu URL, která má být vyhledávána.

**topResultsCharts***číslo.url.parameters=ps=číslo*

Určuje velikost stránky. Například volba **ps=3** omezuje velikost stránky na velikost 3.

Jedná se o řetězec parametru definovaný v rozhraní Dogear REST API.

## Analyza nejlepších výsledků (maximální počet výsledků)

**topResult.resultSize=číslo**

Výchozí a maximální hodnota počtu výsledků, které mají být vráceny z analýzy nejlepších výsledků, je 500. Tuto hodnotu můžete snížit, ale nelze ji zvýšit. Zadat nižší hodnotu může být vhodné, například pokud dochází k problémům s vypršením platnosti požadavků v průběhu analýzy nejlepších výsledků. Tuto vlastnost nelze nastavit pomocí Přizpůsobení vyhledávacích aplikací.

### Související pojmy



Nastavení vlastností dotazu



Třídy Java pro zobrazení nejlepších výsledků

## Úpravy vlastností ukázkové vyhledávací aplikace

Ukázkovou vyhledávací aplikaci pro podnikové vyhledávání lze použít k prohledávání všech aktivních kolekcí a externích zdrojů v systému. Úpravou konfiguračního souboru můžete určit volby prostředí webového serveru, použít jinou výchozí vyhledávací aplikaci nebo řídit volby zobrazené při spouštění vyhledávací aplikace.

### O této úloze

Instalační program implementuje ukázkovou vyhledávací aplikaci pro podnikové vyhledávání do prostředí IBM WebSphere Application Server na vyhledávací servery podnikového vyhledávání. Tuto vyhledávací aplikaci lze konfigurovat úpravou konfiguračního souboru `config.properties`, který se instaluje spolu s aplikací.

Chcete-li, aby se změny projevíly, musíte zastavit a znovu spustit aplikaci `ESSearchServer` na serveru `WebSphere Application Server`.

### Postup

Chcete-li upravit vlastnosti ukázkové vyhledávací aplikace, postupujte takto:

1. Přihlašte se k vyhledávacímu serveru jako administrátor podnikového vyhledávání.
2. Otevřete soubor `config.properties` pro úpravy ve standardním textovém editoru.

Soubor `config.properties` je nainstalován v následujícím adresáři, kde `ES_INSTALL_ROOT` je instalační adresář produktu `OmniFind Enterprise Edition` na vyhledávacím serveru:

```
ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/  
ESSearchApplication.war/WEB-INF/config.properties
```

3. Upravte vlastnosti a určete informace o prostředí webového serveru a předvolby vyhledávání. Poté soubor uložte a zavřete. Znak křížku (`#`) v souboru označuje řádek s komentářem.

4. Zastavte aplikaci ESSearchServer a znovu ji spusťte.

#### **Systém AIX, Linux nebo Solaris**

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

#### **Systém Windows**

```
stopServer ESSearchServer  
startServer ESSearchServer
```

Tyto skripty jsou umístěny v adresáři KOŘEN\_INSTALLACE\_WAS/AppServer/bin:

- Pro produkt WebSphere Application Server verze 5 je výchozí instalační cesta v systémech AIX /usr/WebSphere, v systémech Linux nebo Solaris /opt/WebSphere a v systémech Windows C:\Program Files\WebSphere.
- Pro produkt WebSphere Application Server verze 6 je výchozí instalační cesta v systémech AIX /usr/IBM/WebSphere, v systémech Linux nebo Solaris /opt/IBM/WebSphere a v systémech Windows C:\Program Files\IBM\WebSphere.

#### **Související pojmy**

 Nastavení vlastností dotazu

 Třídy Java pro zobrazení nejlepších výsledků

#### **Související úlohy**

“Konfigurování vyhledávacích serverů tak, aby přijímaly pouze zabezpečené požadavky (SSL)” na stránce 235

## **Přizpůsobení podnikových aplikací**

Nástroj Přizpůsobení vyhledávacích aplikací je tvořen grafickým rozhraním, které vám umožní přizpůsobit vyhledávací aplikace pro podnikové vyhledávání nebo vaše vlastní vyhledávací aplikace.

### **Omezení**

Nástroj Přizpůsobení vyhledávacích aplikací je k dispozici jako samostatná aplikace. Nástroj Přizpůsobení vyhledávacích aplikací nelze v rámci produktu WebSphere Portal k přizpůsobení vyhledávacích aplikací spuštěných jako portlety. Chcete-li přizpůsobit vyhledávací aplikace spuštěné jako portlety, je třeba upravit vlastnosti a konfiguraci instance portletu v rámci rozhraní správy produktu WebSphere Portal.

### **O této úloze**

Nástroj Přizpůsobení vyhledávacích aplikací vám umožňuje zobrazit změny, které chcete provést, ve vizuální podobě a upravit vyhledávací aplikaci bez přímých zásahů do konfiguračního souboru. Můžete například změnit obrázky v pruhu nápisu a na pozadí, změnit rozložení vyhledávacího rozhraní a zadat volby pro práci s výsledky vyhledávání.

Když vybíráte volby v nástroji Přizpůsobení vyhledávacích aplikací, zobrazují se ihned důsledky vybraných voleb. Při uložení změn dojde k aktualizaci konfiguračního souboru vyhledávací aplikace.

Chcete-li, aby se změny projevily, musíte zastavit a znovu spustit aplikaci ESSearchServer na serveru WebSphere Application Server.

### **Tip:**

Podrobné příklady použití nástroje Přizpůsobení vyhledávacích aplikací naleznete v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

## Postup

Chcete-li přizpůsobit vyhledávací aplikaci, postupujte takto:

1. Chcete-li přizpůsobit ukázkovou aplikaci pro podnikové vyhledávání, zadejte ve webovém prohlížeči adresu URL nástroje Přizpůsobení vyhledávacích aplikací. Příklad:  
`http://SearchServer.com/ESSearchApplication/palette.do`  
*SearchServer.com* je název hostitele vyhledávacího serveru.

Není-li váš webový server konfigurován pro použití portu 80, je nutné zadat také správné číslo portu. Příklad:

`http://SearchServer.com:9080/ESSearchApplication/palette.do`

**Tip:** Pokud jste přihlášení jako administrátor podnikového vyhledávání, můžete nástroj Přizpůsobení vyhledávacích aplikací otevřít také výběrem volby **Přizpůsobení vyhledávání** na konzole pro správu podnikového vyhledávání.

2. Chcete-li přizpůsobit vlastní vyhledávací aplikaci, zadejte adresu URL nástroje Přizpůsobení vyhledávacích aplikací a připojte název konfiguračního souboru vaší vyhledávací aplikace. Příklad:

`http://SearchServer.com/ESSearchApplication/palette.do?configFile=/WEB-INF/myConfig.properties`

Pokud zadaný soubor neexistuje, zobrazí se hodnoty načtené ze souboru `config.properties` ukázkové vyhledávací aplikace.

**Tip:** Konfigurační soubor, který chcete použít pro vyhledávací aplikaci, můžete určit také klepnutím na tlačítko **Načíst** po spuštění nástroje Přizpůsobení vyhledávacích aplikací a zadáním názvu požadovaného souboru.

3. Pokud je na aplikačním serveru WebSphere povoleno zabezpečení, přihlaste se s použitím platného jména uživatele a hesla.
4. Vyberte volby, které chcete přizpůsobit, například informace o vyhledávacím serveru, typy zobrazovaných zpráv, volby pro dotazy a výsledky vyhledávání a názvy obrázků, které chcete ve výsledcích vyhledávání použít k odlišení různých typů dokumentů.  
Nápovědu k zadávání voleb zobrazíte klepnutím na volbu **Nápověda k modulu přizpůsobení**. Chcete-li si prohlédnout důsledky některých změn, například ve způsobu zobrazení výsledků vyhledávání, zadejte dotaz a klepněte na tlačítko **Vyhledat**.
5. Po výběru všech požadovaných voleb aktualizujte konfigurační soubor klepnutím na tlačítko **Uložit**. Klepnete-li na tlačítko **Obnovit**, budou v nástroji Přizpůsobení vyhledávacích aplikací obnoveny hodnoty voleb obsažené v poslední uložené verzi konfiguračního souboru.
6. Přihlaste se k vyhledávacímu serveru jako administrátor podnikového vyhledávání a zastavte a znovu spusíte aplikaci `ESSearchServer`.

### System AIX, Linux nebo Solaris

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

### System Windows

```
stopServer ESSearchServer  
startServer ESSearchServer
```

Tyto skripty jsou umístěny v adresáři `KOŘEN_INSTALLACE_WAS/AppServer/bin`:

- Pro produkt WebSphere Application Server verze 5 je výchozí instalační cesta v systémech AIX /usr/WebSphere, v systémech Linux nebo Solaris /opt/WebSphere a v systémech Windows C:\Program Files\WebSphere.
- Pro produkt WebSphere Application Server verze 6 je výchozí instalační cesta v systémech AIX /usr/IBM/WebSphere, v systémech Linux nebo Solaris /opt/IBM/WebSphere a v systémech Windows C:\Program Files\IBM\WebSphere.

### Související úlohy

“Konfigurování vyhledávacích serverů tak, aby přijímaly pouze zabezpečené požadavky (SSL)” na stránce 235

## Klonování ukázkové vyhledávací aplikace

Chcete-li použít ukázkovou vyhledávací aplikaci jako model pro vytváření vlastních vyhledávacích aplikací, můžete zkopírovat soubor `config.properties` nebo použít nástroj Přizpůsobení vyhledávacích aplikací.

### O této úloze

Vytváření vlastních vyhledávacích aplikací si můžete usnadnit naklonováním konfiguračních voleb, které jste zadali pro ukázkovou vyhledávací aplikaci, a následným přizpůsobením voleb, které chcete změnit.

Po naklonování ukázkové vyhledávací aplikace zahájíte vytváření nové vyhledávací aplikace zadáním názvu konfiguračního souboru. Název konfiguračního souboru zadáte také tehdy, chcete-li přizpůsobit novou vyhledávací aplikaci pomocí nástroje Přizpůsobení vyhledávacích aplikací.

Naklonováním ukázkové vyhledávací aplikace můžete rychle vytvořit vyhledávací aplikace určené pro konkrétní účely nebo uživatele. Můžete například vytvořit jednu vyhledávací aplikaci pro zaměstnance oddělení lidských zdrojů a druhou vyhledávací aplikaci pro oddělení prodeje.

Chcete-li, aby se změny projevíly, musíte zastavit a znovu spustit aplikaci ESSearchServer na serveru WebSphere Application Server.

### Postup

Chcete-li naklonovat ukázkovou vyhledávací aplikaci, postupujte takto:

1. Pokud chcete upravit konfigurační soubor:
  - a. Zkopírujte soubor `config.properties` ukázkové vyhledávací aplikace a přejmenujte jej.  
Soubor `config.properties` je nainstalován v následujícím adresáři, kde `ES_INSTALL_ROOT` je instalační adresář produktu OmniFind Enterprise Edition na vyhledávacím serveru:  
`ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war/WEB-INF/config.properties`  
  
Vytvořený soubor se musí nacházet v podadresáři `WEB-INF`.
  - b. Upravte vlastnosti, které chcete používat ve vlastní vyhledávací aplikaci, a uložte soubor. Musíte změnit alespoň vlastnost `applicationName` udávající název vaší vyhledávací aplikace.
2. Pokud chcete naklonovat ukázkovou vyhledávací aplikaci pomocí nástroje Přizpůsobení vyhledávacích aplikací:



- a. Spusíte nástroj Přizpůsobení vyhledávacích aplikací s použitím názvu konfiguračního souboru, který chcete vytvořit. V následujícím příkladu je vytvořen soubor s názvem myNewFile.properties:

`http://ESServer.com/ESSearchApplication/palette.do?configFile=/WEB-INF/myNewFile.properties`

Protože soubor dosud neexistuje, budou použity hodnoty uvedené v konfiguračním souboru ukázkové vyhledávací aplikace `config.properties`.

**Tip:** Konfigurační soubor vyhledávací aplikace můžete vytvořit také klepnutím na volbu **Načíst** po spuštění nástroje Přizpůsobení vyhledávacích aplikací a zadáním názvu souboru. Soubor bude vytvořen po uložení přizpůsobených voleb klepnutím na tlačítko **Uložit**.

- b. Pokud je na aplikačním serveru WebSphere povoleno zabezpečení, přihlaste se s použitím platného jména uživatele a hesla.
- c. Zadejte název vyhledávací aplikace a volby, které chcete používat při prohledávání kolekcí. Potom klepněte na tlačítko **Uložit**. Změny, které zadáte, budou uloženy do nového konfiguračního souboru v podadresáři `WEB-INF`.

Nápovědu k zadávání voleb zobrazíte klepnutím na volbu **Nápověda k modulu přizpůsobení**. Chcete-li si prohlédnout důsledky některých změn, například ve způsobu zobrazení výsledků vyhledávání, zadejte dotaz a klepněte na tlačítko **Vyhledat**.

3. Přihlaste se k vyhledávacímu serveru jako administrátor podnikového vyhledávání a zastavte a znovu spusíte aplikaci ESSearchServer.

#### System AIX, Linux nebo Solaris

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

#### System Windows

```
stopServer ESSearchServer  
startServer ESSearchServer
```

Tyto skripty jsou umístěny v adresáři `KOŘEN_INSTALLACE_WAS/AppServer/bin`:

- Pro produkt WebSphere Application Server verze 5 je výchozí instalační cesta v systémech AIX `/usr/WebSphere`, v systémech Linux nebo Solaris `/opt/WebSphere` a v systémech Windows `C:\Program Files\WebSphere`.
- Pro produkt WebSphere Application Server verze 6 je výchozí instalační cesta v systémech AIX `/usr/IBM/WebSphere`, v systémech Linux nebo Solaris `/opt/IBM/WebSphere` a v systémech Windows `C:\Program Files\IBM\WebSphere`.

## Analýza nejlepších výsledků

Pomocí konfigurace voleb analýzy nejlepších výsledků můžete uživatelům umožnit upřesnění sady výsledků vyhledávání.

Analýza nejlepších výsledků především klasifikuje aktuální sadu výsledků vyhledávání podle frekvence výskytu analyzovaných hodnot. Uživatelé mohou filtrovat výsledky vyhledávání pomocí výběru jedné z analyzovaných hodnot, a automaticky tak přidat vybranou hodnotu jako nové klíčové slovo do kritérií vyhledávání. Výhodou analýzy nejlepších výsledků je umožnění ladění výsledků vyhledávání bez nutnosti použití složité nebo rozšířené syntaxe vyhledávání.

V aplikaci podnikového vyhledávání je možné použít sloupcové grafy ke grafickému znázornění hodnot polí metadat, které se ve výsledcích vyhledávání vyskytují nejčastěji. Dále

můžete vytvořit vlastní třídu Java k zobrazení nejlepších výsledků, a to včetně výsledků z jiných zdrojů než jsou zdroje podnikového vyhledávání, v libovolném formátu HTML.

### Omezení

Zajistěte, že jsou v počítači, kde je spuštěn produkt WebSphere Application Server a vyhledávací aplikace, instalována vhodná písma pro používaný jazyk. Tento krok je nutný k zajištění, aby při generování sloupcového grafu bylo používané písmo nastaveno na písmo podporující znaky v popisku grafu. Tato podmínka je důležitá především pro asijské jazyky. Pokud instalujete písma až po instalaci produktu WebSphere Application Server, je nutné restartovat produkt WebSphere Application Server, aby se změny projevíly.

Spustíte-li vyhledávací aplikaci jako samostatnou aplikaci, můžete konfigurovat grafy nejlepších výsledků pomocí přizpůsobení vyhledávacích aplikací nebo pomocí úpravy konfiguračního souboru vyhledávací aplikace. Spustíte-li vyhledávací aplikaci jako portlet v produktu WebSphere Portal, musíte upravit vlastnosti a konfigurovat instanci portletu pomocí rozhraní správy produktu WebSphere Portal. Přizpůsobení vyhledávacích aplikací nelze použít ke konfiguraci voleb pro analýzu nejlepších výsledků.

### O této úloze

Pomocí zobrazení nejčastěji se vyskytujících hodnot polí metadat ve sloupcovém grafu můžete graficky reprezentovat nejlepší výsledky. Dále můžete použít třídu Java k rozšíření vyhledávací aplikace a zobrazení nejlepších výsledků v jiném formátu, například použít kód HTML k zobrazení nejlepších výsledků v netříděném seznamu.

### Sloupcové grafy pro pole metadat

Můžete zadat volby k analýze polí metadat a zobrazení výsledků této analýzy ve sloupcových grafech. Tyto grafy se zobrazí spolu s výsledky dotazu uživatele. Každý graf odpovídá jednomu poli metadat (jako například velikosti dokumentu, autorovi, datu atd.) a každý sloupec v grafu odpovídá jedné konkrétní hodnotě pole.

Délka sloupce indikuje počet dokumentů, které obsahují konkrétní hodnotu pole, v porovnání s jinými dokumenty, které v tomto poli obsahují jiné hodnoty. Čím delší je sloupec, tím větší je počet výskytů hodnoty pole. Uživatelé mohou ladit výsledky klepnutím na sloupec ve sloupcovém grafu. Hodnota pole reprezentovaná vybraným sloupcem je přidána jako další výraz dotazu a nové výsledky vyhledávání jsou omezeny pomocí dalších vyhledávacích kritérií.

Při konfiguraci prolézáního prostoru pro prolézací modul můžete pomocí výběru volby zadat volby vyhledávání pro pole metadat. Například můžete zadat, zda lze pole metadat prohledávat jako volný text, podle názvu pole zobrazeného ve výsledcích vyhledávání, jako parametrická data atd. Chcete-li zobrazit grafy polí metadat ve výsledcích vyhledávání, musíte konfigurovat volby polí metadat pro prolézací modul. Minimálně je nutné zadat, že se jedná o pole s možností prohledávání a že je lze zobrazit ve výsledcích vyhledávání.

Pokud jste při konfiguraci voleb pro pole metadat vybrali volbu **Úplná shoda**, bude při výběru sloupce z daného sloupcového grafu spuštěn dotaz úplné shody. Hledání úplné shody znamená, že výsledky mají být vráceny pouze v případě, že se výraz dotaz shoduje s celou hodnotou pole. Pokud pole obsahuje méně obsahu či více obsahu, nedochází ke shodě.

Pokud jste při konfiguraci voleb pro pole metadat nevybrali volbu **Úplná shoda**, bude při výběru sloupce z daného sloupcového grafu spuštěn dotaz s použitím pole. V tomto případě se další výraz dotazu musí v poli vyskytovat, ale nemusí se shodovat s celou hodnotou tohoto pole.

## Výsledky formátované pomocí kódu HTML

Rozhraní API `com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` můžete rozšířit pomocí vlastní třídy Java. Při použití tohoto přístupu k zobrazení výsledků můžete pomocí kódu HTML formátovat zobrazení libovolných výsledků vyhledávání, a ne pouze dotazů prohledávajících pole metadat. Výsledky mohou být vráceny pro dokumenty v kolekcích podnikového vyhledávání nebo z vyhledávání externích úložišť a webových stránek.

Každý graf odpovídá jedné třídě Java a tato třída určuje způsob prezentace nejlepších výsledků v grafu. Například třída

`com.ibm.es.searchui.charts.servlet.DynamicMostRecentDocuments`, která je ukázkovou třídou poskytovanou s ukázkovým kódem podnikového vyhledávání, prezentuje nejlepší výsledky v netříděném seznamu. Tyto dokumenty jsou řazeny podle data a zobrazují se pouze názvy dokumentů a data.

Další ukázková třída

`com.ibm.es.searchui.charts.servlet.DogearSearchResults`, poskytovaná spolu s podnikovým vyhledáváním, zobrazuje způsob poskytování seznamu záložek z produktu Lotus Connections Dogear souvisejících s dotazem uživatele.

Při klepnutí na výsledek vyhledávání ve vlastním formátovaném výstupu se dokument zobrazí v novém okně prohlížeče.

## Postup

Tato procedura zobrazuje způsob použití Přizpůsobení vyhledávacích aplikací k určení polí metadat, která mají být analyzována. Nejčastěji se vyskytující hodnoty polí metadat jsou graficky prezentovány ve sloupcovém grafu.

1. Otevřete Přizpůsobení vyhledávacích aplikací. Pokud ještě není zobrazen, načtete konfigurační soubor vyhledávací aplikace, kterou chcete přizpůsobit, a klepněte na tlačítko **Použít**.
2. Přejděte do části **Grafy nejlepších výsledků** a klepněte na tlačítko **Přidat graf**.
3. Při zobrazení seznamu voleb grafu ponechte políčko **Vlastní graf** nezaškrtnuté (tuto volbu vyberte pouze v případě, že vytváříte vlastní třídu Java k analýze a návratu nejlepších výsledků) a zaškrtnutím políčka **Povolit graf** se ujistěte, že uživatelé mohou tento graf používat při zobrazení výsledků vyhledávání.
4. Do pole **Název grafu** zadejte popis grafu. Zadaná hodnota nahradí zástupný text **Nový graf**.
5. Rozhodněte, zda chcete zaškrtnout políčko **Povolit výběr dynamických polí**. Při povolení této volby mohou uživatelé při zobrazení výsledků vyhledávání vybírat různá pole a zobrazit pro ně analýzu nejlepších výsledků.

Tuto volbu můžete vybrat a povolit tak uživatelům ladění sady výsledků vyhledávání. Například uživatel může v databázi vyhledávat informace o zaměstnankyních. Po zadání dotazu k vyhledání pole určujícího pohlaví zaměstnance (například pohlaví:Ž) se zobrazí hodnoty nejlepších výsledků jako sloupce ve sloupcovém grafu. Dále uživatel vybere volbu **pracovní pozice** ze seznamu polí, která existují v původní výsledné sadě. Ve sloupcovém grafu jsou nyní zobrazeny výsledky nejlepších pracovních pozic, které zastávají ženy. Nakonec uživatel vybere volbu **návrhář** ze seznamu dostupných polí. Ve výsledcích jsou nyní zobrazeny informace o zaměstnankyních pracujících jako návrháři.

6. Do pole **Název pole metadat** zadejte název pole metadat, jehož hodnoty budou v tomto grafu analyzovány.
7. Pokud pole obsahuje více hodnot oddělených pomocí znaku oddělovače, zadejte tento znak do pole **Oddělovač hodnot pole**. Pokud pole obsahuje například dvě hodnoty, jako jméno a příjmení zákazníka oddělené středníkem, můžete zadáním středníku do tohoto

pole přidat každou hodnotu do sloupcového grafu samostatně. Pokud nezadáte znak oddělovače, bude celá hodnota pole analyzována jako jedna položka.

8. Vyberte volby zobrazení sloupcového grafu, jako například velikost a barvu sloupců nebo počet sloupců zobrazených při rozbalení nebo sbalení grafu.
9. Zadejte volby pro řazení výsledků analýzy. K řazení můžete použít například hodnoty polí nebo frekvenci výskytu hodnot.
10. Chcete-li zajistit, aby se pro konkrétní hodnoty polí zobrazovaly informace i v případě, že se hodnota pole nevyskytuje ve výsledcích vyhledávání, klepněte na tlačítko **Přidat řádek** v oblasti **Vlastní popisky**. Zadejte hodnotu pole, která se má zobrazit i v případě, že ve výsledcích vyhledávání nejsou vráceny žádné výskyty této hodnoty, a zadejte popisek pro tento sloupec do sloupcového grafu.  

Chcete-li například, aby se ve výsledcích vyhledávání vždy zobrazila konkurence, například společnost JK Enterprises, i když by podle frekvence výskytů neměla být v nejlepších výsledcích zobrazena, zadejte řetězec JK Enterprises jako zdrojovou hodnotu k analýze a poté zadejte popisek tohoto sloupce ve sloupcovém grafu.
11. Klepnutím na tlačítko **Použít** vedle názvu grafu použijete volby zadané pro tento graf.
12. Pokud jsou výsledky vyhledávání již v oblasti vyhledávací aplikace zobrazeny, bude analýza nejlepších výsledků použita pro aktuální vyhledávání a zobrazí se nový graf. V opačném případě můžete zadáním dotazu otestovat zobrazení sloupcového grafu.
13. Jakmile budete s grafem spokojeni, můžete klepnutím na tlačítko **Uložit** aktualizovat konfigurační soubor vyhledávací aplikace.

#### Související pojmy



Nastavení vlastností dotazu



Třídy Java pro zobrazení nejlepších výsledků

---

## Přístup k vyhledávacím aplikacím

Přístup k vyhledávací aplikaci získáte zadáním adresy URL ve webovém prohlížeči.

### Než začnete

Vyhledávací aplikaci je třeba konfigurovat pro konkrétní prostředí webového serveru.

### O této úloze

Ukázková vyhledávací aplikace je instalována na vyhledávací servery podnikového vyhledávání. Prostřednictvím této aplikace v dodané podobě můžete testovat kolekce a externí zdroje před jejich zpřístupněním pro uživatele. Kromě toho ji lze využívat jako model při vytváření vlastních vyhledávacích aplikací.

### Postup

Chcete-li spustit vyhledávací aplikaci, postupujte takto:

1. Zadejte adresu URL vyhledávací aplikace ve webovém prohlížeči. Příklad:

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

*SearchServer.com* je název hostitele vyhledávacího serveru.

Není-li váš webový server konfigurován pro použití portu 80, je nutné zadat také správné číslo portu. Příklad:

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

2. Chcete-li spustit vlastní vyhledávací aplikaci, zadejte adresu URL ukázkové vyhledávací aplikace a připojte název konfiguračního souboru vaší vyhledávací aplikace. Příklad:

<http://SearchServer.com/ESSearchApplication/search.do?configFile=/WEB-INF/myConfig.properties>

Pokud zadaný soubor neexistuje, zobrazí se ukázková vyhledávací aplikace pro podnikové vyhledávání.

3. Pokud je na aplikačním serveru WebSphere povoleno globální zabezpečení, přihlaste se k aplikaci s použitím platného jména uživatele a hesla.

Jestliže je povoleno zabezpečení u kterékoli z kolekcí, k nimž má vyhledávací aplikace přístup, a v zabezpečených kolekcích se vyskytují prolézací moduly konfigurované pro ověřování údajů pověřeni uživatelů během zpracování dotazu, můžete konfigurovat profil uživatele. Na stránce Profil zadejte údaje pověření potřebné pro přístup k zabezpečeným doménám. Tyto domény pak budete moci prohledávat bez přihlašování k nim.

Pokud prolézací modul podporuje zabezpečení SSO (Single Sign-On), můžete zabezpečené domény prohledávat bez vytvoření profilu uživatele.

4. Na stránce Vyhledat zadejte dotaz. Budou prohledány všechny kolekce a externí zdroje, které jsou vybrány pro prohledávání na stránce Předvolby.

---

## Konfigurování vyhledávacích serverů tak, aby přijímaly pouze zabezpečené požadavky (SSL)

Je možné zakázat rozhraní protokolu HTTP vyhledávacích serverů a konfigurovat servery tak, aby přijímaly požadavky na vyhledávání pouze pomocí protokolu SSL a zabezpečeného rozhraní protokolu HTTPS.

### O této úloze

Chcete-li konfigurovat vyhledávací servery tak, aby při zpracování požadavků na vyhledávání používaly pouze protokol SSL (Secure Sockets Layer), musíte zakázat rozhraní protokolu HTTP. Musíte také zajistit, aby byl na obou vyhledávacích serverech i na všech klientských počítačích (například na serveru portálu Websphere, na němž je instalován vyhledávací portlet pro podnikové vyhledávání) uložen tentýž soubor úložiště klíčů. Soubor úložiště klíčů, který se také nazývá *důvěryhodné úložiště*, obsahuje veřejné klíče umožňující důvěryhodnou komunikaci pomocí protokolu SSL.

### Postup

Postup při konfiguraci vyhledávacích serverů, aby přijímaly pouze zabezpečené požadavky:

1. Přihlaste se jako administrátor podnikového vyhledávání. V případě instalace s více servery proveďte následující kroky na indexovém serveru:
  - a. Zastavte systém podnikového vyhledávání:  
esadmin system stopall
  - b. Upravte soubor ES\_NODE\_ROOT/master\_config/nodes.ini.
  - c. Hodnotu parametru *ID\_uzlu.searchserverport* změňte z hodnoty portu HTTP (obvykle 80) na hodnotu portu HTTPS (obvykle 443) a soubor uložte. V případě instalace pro více serverů aktualizujte hodnoty parametru *ID\_uzlu.searchserverport* na obou vyhledávacích serverech.
2. Aktualizujte vyhledávací server následujícím postupem. V případě instalace s více servery proveďte následující kroky na obou vyhledávacích serverech:
  - a. V případě instalace s více servery se přihlaste k vyhledávacímu serveru jako administrátor podnikového vyhledávání.
  - b. Upravte soubor ES\_NODE\_ROOT/nodeinfo/es.cfg.
  - c. Aktualizujte vlastnost TrustStore tak, aby určovala úplnou cestu k souboru úložiště klíčů protokolu SSL.

- d. Aktualizujte vlastnost HTTPProtocol tak, aby obsahovala hodnotu HTTPS, a poté soubor uložte.
- e. Zadejte následující příkaz, přičemž parametr *heslo\_důvěryhodného\_úložiště* určuje heslo pro soubor úložiště klíčů. Tento příkaz zašifruje hodnotu hesla a aktualizuje hodnotu parametru TrustStorePassword v souboru *es.cfg*.

**System AIX, Linux nebo Solaris**

`eschangetrustpw.sh heslo_úložiště`

**System Windows**

`eschangetrustpw heslo_úložiště`

- f. Zkontrolujte, zda vlastnosti *trustStore* a *trustPassword* v souboru *config.properties* určují správnou úplnou cestu a heslo pro soubor úložiště klíčů. Tyto údaje můžete ověřit nebo změnit upravením souboru *config.properties* nebo pomocí nástroje Přizpůsobení vyhledávacích aplikací.
  - g. Pokud používáte vyhledávací portlet pro podnikové vyhledávání, je třeba zkontrolovat, zda parametry portletu *trustStore* a *trustPassword* určují správnou úplnou cestu a heslo pro soubor úložiště klíčů. Tyto údaje můžete ověřit nebo změnit pomocí voleb správy portletu v rozhraní pro správu portálu WebSphere.
3. Restartujte systém podnikového vyhledávání:
- `esadmin system startall`
- Související úlohy**
- “Úpravy vlastností ukázkové vyhledávací aplikace” na stránce 227
  - “Přizpůsobení podnikových aplikací” na stránce 228
  - “Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 5.1” na stránce 325
  - “Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 6” na stránce 330

---

## Konfigurování vyhledávacích serverů tak, aby přijímaly požadavky prostřednictvím serveru proxy

Je možné konfigurovat vyhledávací servery tak, aby přijímaly požadavky prostřednictvím serveru proxy.

### Postup

Umožnění odesílat požadavky na vyhledávací servery prostřednictvím serveru proxy:

1. Přihlaste se jako administrátor podnikového vyhledávání a zastavte systém podnikového vyhledávání. V instalaci s více servery se přihlaste k indexovému serveru.
 

`esadmin system stopall`
2. Upravte soubor `ES_NODE_ROOT/master_config/nodes.ini`.
  - a. Změňte vlastnost *ProxyServer* tak, aby určovala úplný název hostitele serveru proxy.
  - b. Změňte vlastnost *ProxyServerPort* tak, aby určovala port serveru proxy.
  - c. Volitelné: Pokud server proxy vyžaduje ověření všech požadavků, aktualizujte vlastnost *ProxyServerUserName*, aby určovala platné jméno uživatele pro server proxy.
  - d. Volitelné: Pokud server proxy vyžaduje ověření všech požadavků, zadejte následující příkaz, ve kterém parametr *heslo\_serveru\_proxy* představuje heslo pro zadané jméno uživatele serveru proxy. Tento příkaz zašifruje hodnotu hesla a aktualizuje hodnotu parametru *ProxyServerUserPassword* v souboru *es.cfg*.

**Systém AIX, Linux, nebo Solaris**  
eschangeproxypw.sh *heslo\_serveru\_proxy*

**Windows**  
eschangeproxypw *heslo\_serveru\_proxy*

3. Zkontrolujte, zda vlastnosti proxyHost a proxyPort v konfiguračním souboru (soubor `.properties`) určují správný úplný název hostitele a číslo portu pro server proxy.  
Pokud server proxy vyžaduje ověřování, zkontrolujte, zda vlastnosti proxyUser a proxyPassword určují platné jméno uživatele a heslo pro server proxy. Tyto údaje můžete ověřit nebo změnit upravením konfiguračního souboru nebo pomocí nástroje Přizpůsobení vyhledávacích aplikací.
4. Pokud používáte vyhledávací portlet pro podnikové vyhledávání, je třeba zkontrolovat, zda parametry portletu proxyHost a proxyPort určují správný úplný název hostitele a číslo portu pro server proxy.  
Pokud server proxy vyžaduje ověřování, zkontrolujte, zda vlastnosti proxyUser a proxyPassword určují platné jméno uživatele a heslo pro server proxy. Tyto údaje můžete ověřit nebo změnit pomocí voleb správy portletu v rozhraní pro správu portálu WebSphere.
5. Restartujte systém podnikového vyhledávání:  
esadmin system startall





---

## Podpora externích zdrojů

*Externí zdroj* je zdroj dat, jehož prohledávání pomocí aplikace podnikového vyhledávání povolíte, aniž by bylo třeba prolézat, analyzovat nebo indexovat dokumenty v příslušném zdroji dat.

Následující typy zdrojů dat lze prohledávat jako externí zdroje:

- Databáze podporující protokol JDBC Java Database Connectivity Podporovány jsou pouze databáze IBM DB2, Oracle, Microsoft SQL Server 2000 a Microsoft SQL Server 2005. Pro každou tabulku, pro kterou je povoleno vyhledávání, je vytvořen samostatný externí zdroj.

**Omezení:** Podpora databází SQL Server 2005 je omezena na tabulky, které neobsahují datový typ Variant. V systémech AIX není podporován ovladač JDBC pro server SQL Server 2005.

- Servery LDAP (Lightweight Directory Access Protocol). Pro každý server LDAP je vytvořen jeden externí zdroj.

Po konfigurování informací o externím zdroji jej musíte asociovat alespoň s jednou vyhledávací aplikací. Uživatelé poté mohou prohledávat externí zdroj souběžně se zadáváním dotazů na kolekce vytvořené na základě prolézání, analyzování a indexování dat pro podnikové vyhledávání.

### Související pojmy



Federátory vyhledávacího a indexového rozhraní API

---

## Přidávání externích zdrojů do systému

Při přidávání externího zdroje do systému podnikového vyhledávání je třeba určit typ přidávaného zdroje. Informace o zdroji dat a způsob jeho prohledávání můžete určit pomocí průvodce.

### Než začnete

Chcete-li do systému přidat externí zdroj, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### Omezení

Chcete-li prohledávat databázi Oracle jako externí zdroj, musí být na vyhledávacích serverech pro podnikové vyhledávání instalován klientský program Oracle.

V systémech AIX není podporován ovladač JDBC pro server Microsoft SQL Server 2005.

### O této úloze

Po přidání informací o externím zdroji do systému můžete uživatelům povolit zadávání dotazů na zdroj prostřednictvím aplikace podnikového vyhledávání. Můžete povolit prohledávání serverů LDAP (Lightweight Directory Access Protocol) a databázových tabulek JDBC (Java Database Connectivity).

Při konfigurování serveru LDAP vám průvodce usnadní zadávání informací, které systému umožňují připojovat se k serveru, a zadávání voleb způsobu prohledávání serveru.

Při konfigurování databáze JDBC vám průvodce usnadní zadávání informací, které systému umožňují připojovat se k databázi, výběr tabulek, které chcete zpřístupnit pro prohledávání, a zadávání voleb způsobu prohledávání dat v tabulkách. Pro každou tabulku přidanou do systému je vytvořen samostatně prohledávatelný externí zdroj.

Informace o ovladačích pro server SQL 2000 naleznete v dokumentu <http://www.microsoft.com/downloads/details.aspx?familyid=07287B11-0502-461A-B138-2AA54BFDC03A&displaylang=en>. Informace o ovladačích pro server SQL 2005 naleznete v dokumentu <http://www.microsoft.com/downloads/details.aspx?familyid=e22bc83b-32ff-4474-a44a-22b6ae2c4e17&displaylang=en>.

## Postup

Chcete-li do systému přidat externí zdroj, postupujte takto:

1. Chcete-li do systému podnikového vyhledávání zahrnout databáze JDBC, proveďte před přidáním externího zdroje následující kroky. Tento krok, kterým systému umožníte vyhledat vhodné ovladače JDBC, provedete pouze jednou.
  - a. Přihlašte se k prolézacímu serveru jako administrátor podnikového vyhledávání.
  - b. Otevřete soubor `ES_INSTALL_ROOT/configurations/interfaces/discovery__interface.ini` pro úpravy a zadejte proměnné prostředí `CLASSPATH` a `LD_LIBPATH` obsahující cestu ke třídám ovladačů JDBC a cestu k souborům knihovny.
  - c. Otevřete soubor `ES_INSTALL_ROOT/configurations/interfaces/customcommunication__interface.ini` pro úpravy a zadejte proměnné prostředí `CLASSPATH` a `LD_LIBPATH` obsahující cestu ke třídám ovladačů JDBC a cestu k souborům knihovny.
  - d. Volitelné: Chcete-li používat ovladač Oracle JDBC pro lokální nebo katalogizované databáze, přidejte cestu ke knihovně Oracle do proměnné prostředí `LD_LIBPATH` (např. `LD_LIBPATH=../home/oracle/OraHome1/lib32`), otevřete soubor `escrset.sh` pro úpravy, zadejte cestu ke knihovně a exportujte proměnnou instalačního adresáře produktu Oracle. Příklad:

```
ORACLE_HOME=/home/oracle/OraHome1
export ORACLE_HOME
```
  - e. Restartujte systém podnikového vyhledávání včetně vrstvy CCL (Common Communication Layer):

### Systém AIX, Linux, nebo Solaris

```
esadmin stop
stopccl.sh
startccl.sh
esadmin start
```

### Příkazový řádek systému Windows

```
esadmin stop
stopccl
startccl
esadmin start
```

### Nástroj pro správu Služby systému Windows

Chcete-li spustit vrstvu CCL na pozadí, postupujte takto:

- 1) Zadejte příkaz `esadmin stop`.
- 2) Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.

- 3) Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Zastavit**. Po zastavení služby klepněte na tlačítko **Spustit**.
- 4) Zadejte příkaz `esadmin start`.
2. Klepnutím na volbu **Externí zdroje** otevřete zobrazení Externí zdroje.
3. Klepněte na volbu **Přidat externí zdroj**.
4. Vyberte typ externího zdroje, který chcete přidat (Server LDAP nebo Databáze JDBC).
5. Po klepnutí na tlačítko **Další** bude zahájeno konfigurování externího zdroje.

Otevře se průvodce pro typ vytvářeného zdroje. Podle pokynů v průvodci konfigurujte příslušný externí zdroj. Další informace o volbách, které můžete určit, zobrazíte klepnutím na volbu **Nápověda** na jednotlivých stránkách průvodce.

Následující výchozí názvy a umístění ovladačů JDBC vám mohou usnadnit konfigurování informací o připojení pro databáze DB2 a Oracle:

#### **DB2: Legacy JDBC Driver**

Název ovladače: `COM.ibm.db2.jdbc.app.DB2Driver`

Ukázka umístění: `kořenový_adresář_instalace_produkту_db2/java/db2java.zip`

#### **DB2: Universal JDBC Driver**

Název ovladače: `com.ibm.db2.jcc.DB2Driver`

Ukázky umístění:

`kořenový_adresář_instalace_produkту_db2/java/db2jcc.jar`

`kořenový_adresář_instalace_produkту_db2/java/db2jcc_license_cu.jar`

**Oracle** Název ovladače: `oracle.jdbc.driver.OracleDriver`

Ukázka umístění:

`domovský_adresář_oracle/jdbc/lib/ojdbc14.jar`

#### **Microsoft SQL Server 2000**

Název ovladače: `com.microsoft.jdbc.sqlserver.SQLServerDriver`

Ukázky umístění:

`mssql_jdbc_home/lib/mssqlserver.jar`

`mssql_jdbc_home/lib/msbase.jar`

`mssql_jdbc_home/lib/msutil.jar`

#### **Microsoft SQL Server 2005**

Název ovladače: `com.microsoft.sqlserver.jdbc.SQLServerDriver`

Ukázka umístění:

`instalační_adresář/sqljdbc_1.0/lokality/sqljdbc.jar`

Příklad: `instalační_adresář/sqljdbc_1.0/enu/sqljdbc.jar`

6. Po určení voleb pro prohledávání externího zdroje klepněte na tlačítko **Dokončit**.

Nový externí zdroj je uveden v zobrazení Externí zdroje spolu s ostatními externími zdroji přidanými do systému.

#### **Související pojmy**



Federátory vyhledávacího a indexového rozhraní API

---

## **Asociování vyhledávacích aplikací s externími zdroji**

Předtím, než bude možné prohledávat určitý externí zdroj, je třeba asociovat s ním alespoň jednu vyhledávací aplikaci.

#### **Než začnete**

Chcete-li asociovat vyhledávací aplikace s externími zdroji, které mohou prohledávat, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### Postup

Asociování vyhledávací aplikace s jedním či více externími zdroji:

1. Klepněte na panelu nástrojů konzoly pro správu na volbu **Zabezpečení**.
2. Na stránce Vyhledávací aplikace klepněte na volbu **Konfigurovat vyhledávací aplikace**.
3. Na stránce Konfigurovat vyhledávací aplikace klepněte na volbu **Přidat vyhledávací aplikaci**.
4. Zadejte název vyhledávací aplikace.
5. Vyberte externí zdroje, které může aplikace prohledávat:
  - Chcete-li, aby vyhledávací aplikace přistupovala ke všem externím zdrojům přidaným do systému, klepněte na volbu **Všechny kolekce a externí zdroje**.
  - Pokud chcete, aby vyhledávací aplikace měla přístup pouze k zadaným externím zdrojům, vyberte volbu **Specifické kolekce a externí zdroje**.

Po výběru této volby se zobrazí seznam názvů kolekcí a názvů externích zdrojů. U každého externího zdroje, který aplikace může prohledávat, zaškrtněte políčko **Vybrat**.
6. Klepněte na tlačítko **OK**.

#### Související pojmy



Federátory vyhledávacího a indexového rozhraní API

---

## Zabezpečení podnikového vyhledávání

Mechanismy zabezpečení v prostředí podnikového vyhledávání vám umožňují chránit zdroje před neautorizovaným prohledáváním a omezit možnost použití administrativních funkcí pouze na vybrané uživatele.

Pomocí podnikového vyhledávání mohou uživatelé prohledávat širokou škálu zdrojů dat. Aby bylo možné poskytnout přístup k obsahu a povolit práci s konzolou pro správu pouze uživatelům s příslušnou autorizací, systém podnikového vyhledávání koordinuje a vynucuje zabezpečení na několika úrovních.

### Tip:

Podrobné příklady konfigurace zabezpečení podnikového vyhledávání naleznete v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios. Scénáře uvádějí, jak se povoluje globální zabezpečení v produktu WebSphere Application Server s úložištěm LDAP, nastavuje komponenta správy identit a konfiguruje různé prolézací moduly k zajištění zabezpečení na úrovni dokumentů.

### Webový server

První úroveň zabezpečení tvoří webový server. Pokud na aplikačním serveru WebSphere povolíte globální zabezpečení, můžete přiřadit uživatele k administrativním rolím a zajistit ověřování uživatelů, kteří systém spravují. Když se uživatel přihlásí ke konzole pro správu, má k dispozici pouze funkce a kolekce, pro jejichž správu mu bylo uděleno oprávnění.

Mechanismy zabezpečení severu WebSphere Application Server k ověřování uživatelů prohledávajících kolekce mohou také používat vyhledávací aplikace.

### Zabezpečení na úrovni kolekce

Při vytvoření kolekce můžete povolit zabezpečení na její úrovni. Po vytvoření kolekce již toto nastavení nelze změnit. Pokud zabezpečení na úrovni kolekce nepovolíte, nemůžete později používat ovládací prvky zabezpečení na úrovni dokumentů.

Při povolení zabezpečení na úrovni kolekce:

- Procesy globální analýzy podnikového vyhledávání používají jiná pravidla pro indexování duplicitních dokumentů.
- Můžete nakonfigurovat volby k vynucení zabezpečení na úrovni dokumentu, jako je například přiřazení prvků zabezpečení k dokumentům při jejich prolézání, vyžadování ověření aktuálních údajů ověření během zpracování dotazu a určování, zda bude indexován kotvicí text ve webových dokumentech.
- Můžete vynutit zabezpečení mapováním vyhledávacích aplikací (nikoli jednotlivých uživatelů) na kolekce a externí zdroje, které mohou prohledávat. Potom se k povolení nebo odepření přístupu uživatelů k prohledávání aplikací používají standardní mechanismy řízení přístupu.

Povolení zabezpečení kolekce má určité negativní dopady na kvalitu vyhledávání. Povolení zabezpečení kolekce omezuje objem informací indexovaných pro jednotlivé dokumenty. V důsledku je proto pro některé dotazy nalezeno méně výsledků.

### Zabezpečení na úrovni dokumentu

Při konfigurování prolézacích modulů pro kolekci můžete povolit zabezpečení na úrovni dokumentů. Zadáním určitých voleb můžete například přiřadit prvky

zabezpečení k datům shromažďovaným prolézacími moduly. Vyhledávací aplikace mohou pomocí těchto prvků, které jsou ukládány spolu s dokumenty v indexu, vynutit řízení přístupu a omezit možnost zadávání dotazů na data a prohlížení výsledků vyhledávání pouze na uživatele s patřičným pověřením.

Pro určité typy zdrojů dat lze konfigurovat volby ověřování přihlašovacích údajů pověřením uživatele s použitím aktuálního řízení přístupu během zpracování dotazu. Tato dodatečná vrstva zabezpečení zaručuje ověření uživatelských oprávnění s nativním zdrojem dat v reálném čase. Tato funkce může chránit před instancemi, v nichž se údaje pověřením uživatele změni po provedení indexace dokumentu a příslušných prvků zabezpečení.

Fáze zpracování kotvicího textu v rámci globální analýzy za normálních okolností přiřazuje text obsažený v jednom dokumentu (zdrojový dokument) k jinému dokumentu (cílový dokument), v němž se tento text nemusí nutně objevit. Při konfiguraci webového prolézacího modulu můžete určit, zda chcete z indexu vyloučit kotvicí text, pokud je propojen s dokumentem, který daný webový prolézací modul nemůže prolézat.

Zabezpečení kolekcí se neomezuje jen na mechanismy ověřování a řízení přístupu, které může podnikové vyhledávání využít k ochraně indexovaného obsahu. Existují rovněž pojistky bránící neautorizovaným uživatelům a osobám, které mohou mít v úmyslu škodit, v přístupu k přenášeným datům. Vyhledávací servery například ke komunikaci s indexovým serverem a vyhledávací aplikací používají protokoly jako SSL (Secure Sockets Layer), SSH (Secure Shell) a HTTPS (Secure Hypertext Transfer Protocol).

Další zabezpečení je zajištěno šifrováním. V zašifrovaném formátu se ukládá například heslo administrátora podnikového vyhledávání zadávané při instalaci produktu. Hesla zadaná uživateli v profilech uživatelů jsou rovněž uchovávána v zašifrovaném formátu.

Pro zvýšení úrovně zabezpečení je nezbytné zajistit řádnou izolaci hardwaru serveru a jeho ochranu před neautorizovaným zásahem. Instalací ochranné bariéry firewall můžete servery podnikového vyhledávání chránit před útokem vedeným přes jinou část sítě. Ujistěte se také, že na serverech podnikového vyhledávání nejsou žádné otevřené porty. Konfigurujte systém tak, aby přijímal požadavky pouze prostřednictvím portů přiřazených explicitně k aktivitám a aplikacím podnikového vyhledávání.

---

## Zabezpečení instalace

Instalační program produktu OmniFind Enterprise Edition vytváří prostředí pro vynucení zabezpečení, pokud uživatelé spravují nebo prohledávají kolekce podnikového vyhledávání.

### ID administrátora podnikového vyhledávání

V průběhu instalace je instalační program vyzván k zadání jména uživatele a hesla pro administrátora podnikového vyhledávání. Použití zadaných údajů ověření v systému:

- Ověřování administrátora podnikového vyhledávání při provádění administrativních úloh
- Vytváření interní databáze podnikového vyhledávání
- Spouštění všech relací nebo procesů podnikového vyhledávání

Jméno uživatele zadané během instalace musí být platné jméno uživatele operačního systému s oprávněními pro administraci systému. Instalační program ukládá vhodně zašifrované údaje pověřením do souboru vlastností na každém serveru podnikového vyhledávání.

## Šifrování

K ochraně citlivých dat se používá šifrování, které zakóduje ověřovací část dat všech zpráv, které jsou přenášeny systémem podnikového vyhledávání. Tento proces představuje jen malé zatížení, protože jsou zašifrována pouze ID ověřování a hesla. Rovněž jsou zašifrována všechna hesla uložená systémem (v konfiguračních souborech, databázi podnikového vyhledávání, atd.).

## Zabezpečení serveru WebSphere Application Server

Pokud nebyl na vyhledávacím serveru dříve nainstalován server WebSphere Application Server, nainstaluje jej instalační program v tichém režimu se zakázaným globálním zabezpečením. Pokud je na serveru WebSphere Application Server později globální zabezpečení povoleno, bude za ověřování administrátora podnikového vyhledávání zodpovědný server WebSphere Application Server.

Při povolení globálního zabezpečení je třeba do registru uživatele přidat ID a heslo administrátora podnikového vyhledávání, které bylo zadáno během instalace na server WebSphere Application Server, například pomocí adresáře LDAP (Lightweight Directory Access Protocol).

Pokud globální zabezpečení povolíte po instalaci produktu OmniFind Enterprise Edition, je třeba aktualizovat hodnoty konfigurace a spustit příkaz **eschangewaspw**, který zašifruje a uloží údaje pověření serveru WebSphere Application Server do souboru vlastností podnikového vyhledávání.

---

## Ověřování a řízení přístupu

Pro zajištění ochrany obsahu před neautorizovanými uživateli a pro řízení přístupu k administrativním funkcím je v prostředí podnikového vyhledávání k dispozici ověřování uživatelů (řízení přístupů).

### Ověřování

Ověřování je jakýkoli proces, při kterém systém ověřuje identitu uživatele, který chce přistupovat do systému. Vzhledem k tomu, že je řízení přístupu k prostředkům obvykle založeno na identitě uživatelů, kteří tyto prostředky požadují, je ověřování základem efektivního zabezpečení.

Ověřování uživatelů podnikového vyhledávání je implementováno pomocí údajů pověření, které se skládají minimálně ze jména uživatele a hesla.

Při ověřování uživatelů, kteří přistupují ke konzole pro správu, využívá podnikové vyhledávání podporu zabezpečení poskytovanou aplikačním serverem WebSphere.

### Autorizace (řízení přístupu)

Autorizace je jakýkoli mechanismus, pomocí kterého systém uděluje nebo ruší práva pro přístup k určitým datům nebo pro provedení určité akce. Uživatel se často musí přihlašovat do systému pomocí určité formy ověřování. Mechanismy řízení přístupu určují porovnáním identity uživatele se seznamem přístupových práv (ACL), které operace může uživatel provádět a které nikoli. Řízení přístupu zahrnuje:

- Oprávnění k souborům, jako například práva pro vytvoření, čtení, úpravy nebo odstranění souboru.
- Oprávnění k programům, například práva ke spuštění programu.

- Oprávnění k datům, například práva k načtení nebo aktualizaci informací.

---

## Administrativní role

Podnikové vyhledávání využívá koncepci rolí pro řízení přístupu k různým funkcím konzoly pro správu.

Při instalaci produktu OmniFind Enterprise Edition (OmniFind Enterprise Edition) nastaví instalační program jméno uživatele a heslo administrátora podnikového vyhledávání. Při prvním přístupu ke konzole pro správu se musíte přihlásit jako tento uživatel. Pokud na aplikačním serveru WebSphere nepovolíte globální zabezpečení, můžete pro přístup ke konzole pro správu podnikového vyhledávání používat pouze toto jméno uživatele.

Povolíte-li na aplikačním serveru WebSphere globální zabezpečení, můžete definovat další uživatele s přístupem ke správě podnikového vyhledávání. Přiřazením uživatelů k rolím můžete omezit přístup k jednotlivým kolekcím a řídit funkce dostupné jednotlivým administrativním uživatelům. Jména uživatelů, která přiřadíte k administrativním rolím v prostředí podnikového vyhledávání, musí existovat v registru uživatelů aplikačního serveru WebSphere.

Při přihlášení administrativního uživatele podnikové vyhledávání ověří zadané jméno uživatele. V prostředí konzoly jsou dostupné pouze kolekce a funkce, k jejichž správě má uživatel oprávnění.

Uživatelům lze přidělit následující administrativní role:

### **Administrátor podnikového vyhledávání**

Tito uživatelé vytvářejí kolekce a mají oprávnění ke správě všech aspektů systému podnikového vyhledávání. Při instalaci produktu OmniFind Enterprise Edition můžete určit jméno uživatele a heslo pro prvního administrativního uživatele podnikového vyhledávání. Po prvním přihlášení může tento uživatel přiřadit ostatním uživatelům roli administrátora podnikového vyhledávání.

### **Administrátor kolekce**

Tito uživatelé mohou upravovat, monitorovat a ovládat činnost kolekcí, u kterých mají autorizaci pro správu. Tito uživatelé nemohou vytvářet kolekce. Administrátoři kolekcí mohou monitorovat a provádět činnosti na úrovni systému pouze v případě, že jim administrátor podnikového vyhledávání udělil příslušné oprávnění.

### **Operátor**

Tito uživatelé mohou monitorovat a ovládat činnost kolekcí, u kterých mají autorizaci pro správu. Mohou například spouštět a zastavovat činnosti kolekcí, nemohou však vytvářet kolekce ani je upravovat. Operátor může monitorovat a provádět činnosti na úrovni systému pouze v případě, že mu administrátor podnikového vyhledávání udělil příslušné oprávnění.

### **Monitor**

Tito uživatelé mohou monitorovat kolekce, u kterých mají autorizaci pro správu. Tito uživatelé nemohou ovládat činnosti (například spouštět a zastavovat servery), vytvářet kolekce ani je upravovat. Monitor může sledovat činnosti na úrovni systému pouze v případě, že mu administrátor podnikového vyhledávání udělil příslušné oprávnění, nemůže je však provádět.

## Konfigurování administrativních uživatelů

Konfigurováním administrativních rolí můžete omezit přístup ke kolekcím a řídit funkce, které jednotliví administrativní uživatelé mohou provádět.



## Než začnete

Před přiřazením uživatele k administrativní roli ověřte, zda je na serveru WebSphere Application Server povoleno zabezpečení. Zkontrolujte také, zda v registru uživatelů serveru WebSphere Application Server existuje dané jméno uživatele.

Chcete-li konfigurovat administrativní uživatele, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

## Postup

Přiřazení uživatelů k administrativním rolím:

1. Klepnutím na volbu **Zabezpečení** otevřete pohled Zabezpečení.
2. Na stránce Administrativní role klepněte na volbu **Přidat uživatele**.
3. Zadejte jméno uživatele, kterého chcete přidat, a vyberte příslušnou administrativní roli.
4. Pokud tohoto uživatele nepřidáváte jako administrátora podnikového vyhledávání, určete, zda má mít tento uživatel přístup ke stránkám z panelu nástrojů **Systém**.  
Může být například žádoucí povolit některým operátorům nebo administrátorům kolekce monitorování souborů žurnálu na úrovni systému.
5. Pokud tohoto uživatele nepřidáváte jako administrátora podnikového vyhledávání, vyberte kolekce a externí zdroje, které tento uživatel může spravovat.  
Můžete zaškrtnout políčka pro jednotlivé kolekce a externí zdroje nebo povolit uživateli správu všech kolekcí a externích zdrojů.

---

## Zabezpečení na úrovni kolekce

Pokud chcete použít zabezpečení na úrovni kolekce, musíte konfigurovat volby pro indexování obsahu a volby povolující vyhledávacím aplikacím prohledávat konkrétní kolekce.

Při vytvoření kolekce můžete vybrat volbu povolující zabezpečení této kolekce. Pokud tuto volbu vyberete, můžete později konfigurovat prvky zabezpečení na úrovni dokumentu. Je-li povoleno zabezpečení kolekcí, procesy globální analýzy podnikového vyhledávání rovněž používají jiná pravidla pro indexování duplicitních dokumentů.

Po vytvoření vyhledávací aplikace lze pomocí ID vyhledávací aplikace určit, které kolekce a externí zdroje může daná vyhledávací aplikace prohledávat a kteří uživatelé budou mít k této vyhledávací aplikaci přístup.

## Analýza duplicitních dokumentů a zabezpečení kolekce

Pokud povolíte zabezpečení kolekce, procesy globální analýzy v této kolekci neidentifikují duplicitní dokumenty.

V průběhu globální analýzy indexovací procesy identifikují skupiny totožných (nebo téměř totožných) dokumentů. Poté všechny tyto dokumenty přiřadí k jediné kanonické reprezentaci obsahu. Povolněním identifikace duplicitních dokumentů dosáhnete toho, že výsledky vyhledávání nebudou obsahovat více dokumentů se stejným (nebo velmi podobným) obsahem.

Pokud při vytvoření kolekce povolíte její zabezpečení, duplicitní dokumenty nebudou vyhledávány a tedy ani přiřazovány ke společné kanonické reprezentaci. Namísto toho bude každý z dokumentů indexován nezávisle. Tak bude zaručeno ohodnocení ovládacích prvků zabezpečení pro každý dokument, aby mohli uživatelé prohledávat pouze dokumenty s prvky zabezpečení odpovídajícími jejich pověřovacím údajům. Mohou existovat dva dokumenty s velmi podobným obsahem, avšak zabezpečené pomocí odlišných seznamů přístupových práv.

Například pro dva duplicitní dokumenty, dokument\_A a dokument\_B, předpokládejme, že uživatel má přístupové právo pouze k dokumentu\_B. Dojde-li k odebrání dokumentu\_B kvůli detekci duplikátu, nemůže uživatel zobrazit tento dokument ve výsledcích vyhledávání kvůli omezení přístupu k dokumentu\_A.

Zákaz analýzy duplicitních dokumentů může zlepšit zabezpečení dokumentů v kolekci, kvalita vyhledávání však může poklesnout, jestliže uživatelé obdrží ve výsledcích vyhledávání více kopií téhož dokumentu.

## Identifikátory vyhledávací aplikace

Možnost prohledávat různé kolekce je řízena vyhledávacími aplikacemi mapování v kolekci a externích zdrojích, které mohou prohledávat. Aplikace s názvem Výchozí umožňuje použití nezměněné ukázkové vyhledávací aplikace k prohledávání všech kolekcí a externích zdrojů.

U všech vyhledávacích aplikací je požadováno předání platného názvu aplikace (APPID) do rozhraní API podnikového vyhledávání. Vyhledávací aplikace může prohledávat pouze kolekce a externí zdroje asociované s tímto názvem APPID.

Aby mohla vyhledávací aplikace získat přístup ke kolekci nebo externímu zdroji, musí administrátor podnikového vyhledávání přiřadit k vyhledávací aplikaci konkrétní kolekce a zdroje, které tato aplikace smí prohledávat. Vyhledávací aplikace může prohledávat všechny kolekce a externí zdroje v systému podnikového vyhledávání nebo může prohledávat pouze určené kolekce a externí zdroje.



Ukázková vyhledávací aplikace (ESSearchApplication) má soubor vlastností, který určuje používaný název aplikace. Výchozí umístění tohoto souboru vlastností je `ES_INSTALL_ROOT\installedApps\ESSearchApplication.ear\ESSearchApplication.war\WEB-INF\config.properties`.

Počáteční hodnota názvu aplikace je Výchozí. Pokud tuto hodnotu změníte, bude změněn seznam kolekcí a externích zdrojů, které může aplikace ESSearchApplication prohledávat.

Chcete-li řídit, kteří uživatelé mohou prohledávat které kolekce, je třeba asociovat uživatele (nebo skupiny uživatelů) s aplikací klienta pomocí standardních funkcí serveru WebSphere Application Server pro řízení přístupu, podobně jako byste pomocí těchto funkcí omezovali přístup k adrese URL. Můžete například omezit přístup k adrese URL používané ke spuštění vyhledávací aplikace.

Další informace o ID vyhledávacích aplikací a o postupu při zahrnutí prvků zabezpečení do vlastních vyhledávacích aplikací naleznete v dokumentaci k rozhraní SI-API pro podnikové vyhledávání.

### Související pojmy

-  Vyhledávací a indexové rozhraní API - přehled
-  Zabezpečení vyhledávacího a indexového rozhraní API

---

## Zabezpečení na úrovni dokumentu

Pokud je při vytvoření kolekce povoleno její zabezpečení, lze konfigurovat ovládací prvky zabezpečení na úrovni dokumentů. Řízení přístupu na úrovni dokumentu zajišťuje, že výsledky vyhledávání obsahují pouze dokumenty, k jejichž zobrazení je uživatel zadávající požadavek na vyhledávání autorizován.

System podnikového vyhledávání podporuje mnoho postupů konfigurace ovládacích prvků zabezpečení na úrovni dokumentu.

- Dokumenty je možné před přidáním do indexu předfiltrovat a přiřadit k nim prvky zabezpečení.
- U některých typů dat lze výsledky vyhledávání dodatečně filtrovat k ověření údajů pověření přihlášení uživatele s aktuálními daty řízení přístupu. Pomocí komponenty správy identity podnikového vyhledávání je možné šifrovat různé údaje pověření, které uživatelé potřebují pro přístup k různým úložištím, a ukládat zašifrované údaje pověření do profilů. Jsou-li prohledávané zdroje chráněny produktem poskytujícím zabezpečení SSO, může komponenta správy identit řídit přístup k dokumentům bez požadavku na vytváření profilů uživateli.
- U většiny typů prolézacích modulů je možné pro přiřazení prvků zabezpečení k dokumentům v indexu použít třídu (modul plug-in) Java.
- U dokumentů prolézáných pomocí webového prolézacího modulu je možné z indexu vyloučit kotvicí text v dokumentech obsahujících odkazy na zakázané dokumenty.

### Související pojmy

 Zabezpečení aplikace

## Předběžné a dodatečné filtrování výsledků vyhledávání

Aby bylo zajištěno, že výsledky vyhledávání obsahují pouze dokumenty, k jejichž zobrazení je uživatel zadávající požadavek vyhledávání autorizován, existují dva odlišné přístupy k filtrování dokumentů.

- Prvním přístupem je replikace seznamů přístupových práv (ACL) v čase prolézání do indexu a použití vyhledávacího stroje k porovnání údajů pověření s replikovanými seznamy ACL dokumentů. Při předběžném filtrování dokumentů a řízení, které dokumenty budou přidány do indexu, je dosahováno nejlepšího výkonu. Je však složité v indexu modelovat všechny zásady zabezpečení z různých zdrojů back-end a implementovat logiku porovnávání jednotně. Tento přístup také není tak citlivý na změny ve zdrojových seznamech ACL.
- Druhým přístupem je dodatečné filtrování dokumentů v sadě výsledků pomocí konzultace aktuálních dat zabezpečení ve zdrojích back-end. Tento přístup umožňuje přispívajícím zdrojům back-end závěrečné posouzení dokumentů vrácených uživateli a zajišťuje, že sada výsledků odráží aktuální řízení přístupu. Tento postup má však za následek snížení výkonu vyhledávání, protože vyžaduje existenci připojení ke všem zdrojům back-end. Pokud zdroj není přístupný, je nutné odfiltrovat odkazy na dokumenty ze sady výsledků společně s dokumenty, k jejichž zobrazení nemá uživatel autorizaci.

**Důležité:** V konfiguraci s více servery se pro některé typy zdrojů dodatečné filtrování provádí na prolézacím serveru. Je-li prolézací server vypnut za účelem údržby, nezobrazí se uživatelům při dotazu na kolekce podnikového vyhledávání žádné výsledky. Dále nejsou žádné výsledky vráceny, pokud nejsou přístupné servery back-end požadované k řízení přístupu.

Při podnikovém vyhledávání závisí podpora vynucení řízení přístupu na kombinaci těchto dvou přístupů. Daný návrh poskytuje optimální výkon při zachování přesných zásad zabezpečení původních úložišť dokumentů. Uložením dat řízení přístupu na vysoké úrovni do indexu může systém poskytovat provizorní (potenciálně menší) sadu výsledků, kterou lze později dodatečně filtrovat k ověření aktuálního řízení přístupu. Předpokládá se, že pokud má uživatel přístup k úložišti, které vlastní daný dokument, má pravděpodobně přístup i k samotnému dokumentu.

Data řízení přístupu uložená v indexu se mění s typem prolézacího modulu. Například prolézací modul Notes může ukládat prvky řízení přístupu na úrovni serveru a databáze a prolézací modul QuickPlace může ukládat prvky řízení přístupu pro servery, místa a místnosti.

Všechny typy zdrojů dat v podnikovém vyhledávání podporují možnost indexování nativních seznamů řízení přístupu v čase prolézání. Některé typy zdrojů dat také podporují možnost dodatečného filtrování sady výsledků a ověření aktuálních pověření uživatele (tento typ podpory je poskytován pomocí mechanismů nativního zabezpečení nebo pomocí komponenty pro správu identit podnikového vyhledávání).

Tento dvoustranný návrh zabezpečení zahrnuje následující úlohy:

- Extrahování nativních informací seznamu ACL v čase prolézání.
- Ukládání informací seznamu ACL serverů a databázi do indexu.
- Vytváření kontextu zabezpečení uživatele po přihlášení uživatele nebo inicializaci relace. Tato úloha musí zahrnovat různé identifikátory, které musí uživatel použít pro přístup k různým zdrojům back-end.
- Zpracování vyhledávání s kontextem zabezpečení uživatele a poskytnutí provizorní sady výsledků, která obsahuje pouze dokumenty, k nimž má uživatel přístup na úrovni úložiště.
- Dodatečné filtrování provizorní sady výsledků pomocí konzultace zdrojů back-end, které poskytly dokumenty, se sadou výsledků aktuálních informací nativního seznamu ACL.

## Ověřování pomocí uložených prvků zabezpečení

Pokud je při vytvoření kolekce povoleno její zabezpečení, lze konfigurovat ovládací prvky zabezpečení na úrovni dokumentů uložením dat zabezpečení do indexu.

Standardně je ke každému dokumentu přiřazen veřejný prvek, který dokument zpřístupní všem. Je-li pro kolekci povoleno zabezpečení, je možné veřejný prvek nahradit hodnotou zadanou administrátorem nebo hodnotou, která byla extrahována z pole v prolézaném dokumentu. Při konfigurování prolézacího modulu určujete, zda chcete omezit přístup některých uživatelů k dokumentům zpracovávaným tímto prolézacím modulem.

Pokud administrátor kolekce nakonfiguruje prolézací modul, může zadat volby zabezpečení pro jednotlivé tabulky, souborové systémy, atd. (v prolézaném prostoru tedy lze pro různé zdroje dat nakonfigurovat různá pravidla zabezpečení). Administrátor může:

- Určit, že jsou dokumenty veřejné (mohou je prohledávat všichni uživatelé)
- Přiřadit ke každému dokumentu prvky zabezpečení definované uživatelem
- Extrahovat prvky zabezpečení z pole v prolézáných datech a přiřadit daný extrahovaný prvek ke každému dokumentu

Prvky zabezpečení (s výjimkou výchozího veřejného prvku) jsou zcela definovány uživateli. Prvek zabezpečení může představovat jméno uživatele, název skupiny, roli uživatele nebo libovolnou jinou hodnotu, kterou určíte jako platnou pro daný zdroj dat.

Administrátor může například určit, že bude k řízení přístupu k dokumentům prolézáných modulem Notes použito pole `hrDeptName`. Administrátor může také určit, že pokud dané pole v dokumentu neexistuje nebo neobsahuje data zabezpečení, budou pro řízení přístupu k dokumentům použity dva prvky definované uživatelem, `hrgroup1` a `hrgroup2`.

Prvky zabezpečení jsou prolézacímu modulu zpřístupněny pomocí příslušného konfiguračního souboru. Prolézací modul poskytuje každému dokumentu prvek zabezpečení jako metadata. Komponenta indexování přečte prvek zabezpečení a použije jej v odeslaných informacích o dokumentu v indexu. Pokud administrátor nativního zdroje dat aktualizuje

seznam přístupových práv, jsou aktualizované prvky zabezpečení zpřístupněny v nejbližším dalším hlavním nebo rozdílovém sestavení indexu.

Na základě použití vlastních obchodních pravidel můžete určit hodnotu prvků zabezpečení zakódováním pravidel ve třídě Java. Při konfigurování vlastností prolézacího modulu můžete určit název modulu plug-in, který má prolézací modulu při prolézání dokumentů používat. Prvky zabezpečení přidávané příslušným modulem plug-in jsou uchovávány v indexu a jejich prostřednictvím lze řídit přístup k dokumentům.

## Způsoby použití prvků zabezpečení ve vyhledávacích aplikacích

Za poskytnutí prvků zabezpečení v čase vyhledávání, aby došlo ke správnému filtrování dokumentů, je zodpovědná vyhledávací aplikace klienta. Pokud není dodán prvek zabezpečení, bude během zpracování vyhledávání automaticky použit výchozí veřejný prvek.

Ukázková vyhledávací aplikace pro podnikové vyhledávání demonstruje způsoby implementace zabezpečení na úrovni dokumentu. V tomto příkladu se předpokládá, že administrátor přiřadil hodnotu prvku zabezpečení skupině dokumentů (na rozdíl od extrahování prvku zabezpečení z pole v prolézaném dokumentu). Vyhledávací aplikace pomocí přihlašovacího jména uživatele určí, ke kterým dokumentům může uživatel přistupovat. Místo použití skutečného jména uživatele spoléhá vyhledávací aplikace na název skupiny, kam uživatel patří. Použitím názvu skupiny jako prvku zabezpečení mohou být uživatelé přidáváni a odebíráni bez požadavku na opakované sestavení indexu.

Prvek zabezpečení přiřazený administrátorem k sadě dokumentů představuje platný název skupiny operačního systému. V prolézaném prostoru jsou různým dokumentům přiřazovány různé názvy skupin. Příklad:

Dokument1-5: Prvek zabezpečení = Skupina1

Dokument6-10: Prvek zabezpečení = Skupina2

## Ověřování aktuálního pověření během zpracování dotazu

Pokud je pro kolekci při jejím vytvoření povoleno zabezpečení, určité typy domén umožňují ověřit aktuální údaje pověření uživatele během odesílání dotazu příslušného uživatele.

Než vyhledávací server odpoví na dotaz, ověří podle údajů nativních úložišť aktuální oprávnění uživatele a poté odebere z výsledků vyhledávání všechny dokumenty, k jejichž prohlížení nemá uživatel oprávnění.

Při konfigurování následujících typů prolézacích modulů můžete vybrat volbu určující, zda mají být údaje pověření uživatele ověřovány porovnáním s aktuálními prvky řízení přístupu spravovanými v nativním úložišti. Po prolezení a indexování dokumentů je komponenta pro správu identit podnikového vyhledávání použita k ověření uživatelů, kteří se pokoušejí prohledávat zabezpečené kolekce.

- Prolézací modul Content Edition (pouze typy úložišť Documentum, FileNet Panagon Content Services, Hummingbird DM, Portal Document Manager a SharePoint)
- Prolézací modul DB2 Content Manager
- Prolézací modul Domino Document Manager
- Prolézací modul Notes
- Prolézací modul QuickPlace
- Prolézací modul Souborový systém Windows

Pro následující typy prolézacích modulů mohou být aktuální údaje pověření uživatele ověřeny při použití vyhledávacího portletu v produktu WebSphere Portal k prohledávání kolekcí podnikového vyhledávání.

- Prolézací modul Web Content Management
- Prolézací modul WebSphere Portal

#### **Související pojmy**

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino”  
na stránce 267

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows” na stránce 271

#### **Související úlohy**

“Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů” na stránce 268

## **Správa identit podnikového vyhledávání**

Obecným problémem v podnicích je správa více údajů pověření uživatele. Systém podnikového vyhledávání tento problém řeší poskytnutím volitelné komponenty pro správu identit.

### **Tip:**

Podrobné příklady nastavení komponenty správy identit pro podnikové vyhledávání najdete v příručce IBM Redbook s názvem IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios.

Informace mohou v podnicích existovat v mnoha různých tvarech a formách. Mohou být v rámci podniku distribuovány a spravovány softwarem vhodným pro danou úlohu. Uživatelé v podniku mohou například pomocí aplikace SQL přistupovat k relačním databázím nebo pomocí systému správy dokumentů přistupovat k požadovaným dokumentům.

Řízení přístupu k citlivým informacím je v těchto úložištích obvykle vynucováno softwarem pro správu. Uživatelé se identifikují v systému hostitele pomocí kombinace jména uživatele a hesla. Po dokončení ověřování systémem řídí software pro správu, které dokumenty má uživatel povoleno zobrazit a pracovat s nimi podle definovaných přístupových práv uživatele.

Pro uživatele je běžné mít různá jména uživatelů a hesla přiřazená ke každému úložišti. Podobně jako se od uživatelů požaduje jejich identifikace v původních úložištích, musí uživatelé poskytnout údaje pověření před zobrazením dokumentů v kolekci podnikového vyhledávání, která vyžaduje ověřování aktuálních údajů pověření. Uživatelé, kteří mají více identit, musí předložit příslušné údaje pověření pro každou identitu.

Pokud určíte, že chcete pro správu identit používat podnikové vyhledávání v konzole pro správu, budou vyhledávací servery moci používat při ověřování aktuálních údajů pověření uživatele během zpracování dotazu následující přístupy:

- Vyhledávací aplikace může uživateli zobrazit výzvu k registraci údajů pověření potřebných pro přístup k různým doménám v profilu uživatele. Profil, který je šifrován a uložen v zabezpečeném úložišti dat, umožňuje uživatelům prohledávat zabezpečené domény. Nejsou-li určena pověření pro doménu, která vyžaduje ověření aktuálních pověření, budou dokumenty z příslušné domény vyloučeny z výsledků vyhledávání.
- Pokud byly dokumenty v kolekci prolezeny prolézacím modulem poskytujícím podporu zabezpečení SSO (single sign-on) a pokud jste určili, že chcete používat zabezpečení SSO pro řízení přístupu k dokumentům, bude systém používat metody SSO pro ověřování uživatelů v době trvání relace vyhledávání. Uživatel nemusí vytvořit profil, který by určoval údaje pověření, ani zadávat při prohledávání zabezpečených domén jméno uživatele a heslo.

Pokud uživatelé prohledávají kolekce vyžadující při odeslání dotazu ověření aktuálních údajů pověření, může systém pomocí profilu nebo metod zabezpečení SSO odeprít nebo povolit přístup k dokumentům.

## Získání informací o skupině uživatele

Chcete-li ověřit údaje pověření uživatele, je třeba pomocí komponenty pro správu identit získat informace o skupině uživatele pro každou identitu a přidat tyto informace do řetězce USC (user security context). Tyto informace o skupině slouží k filtrování výsledků podle dat řízení přístupu, která jsou uložena v indexu podnikového vyhledávání, nebo podle dat ověřování SSO. Komponenta správy identit tyto akce provádí použitím prvků SSO nebo údajů pověření uživatele k připojení k systému back-end a zadání požadavku na skupiny, kde je uživatel členem.

Při konfiguraci voleb správy identit v konzole pro zprávu můžete určit, jak často budou tyto informace o skupině aktualizovány. Nová data o skupině můžete extrahovat při každém přihlášení uživatele do vyhledávací aplikace nebo pravidelně, například každé tři dny.

## Zabezpečení bez komponenty pro správu identit

V některých podnicích není třeba spravovat více identit v komunitách uživatelů pomocí komponenty pro správu identit podnikového vyhledávání. Pokud komponentu pro správu identit v konzole pro správu podnikového vyhledávání zakážete, bude za generování kontextového řetězce zabezpečení uživatelů zodpovědná vyhledávací aplikace. Po vygenerování slouží řetězec USC k nastavení hodnoty omezení seznamu ACL v každém dotazu. Příklad:

```
Dotaz q = factory.createQuery("IBM");
q.setACLConstraints("Kontext zabezpečení uživatele v jazyce XML");
```

**Tip:** K zajištění možnosti napsat vlastní funkce správy identit poskytuje rozšíření rozhraní SIAPI (Search and Index API) programové řízení databáze správy identit. Toto rozhraní API umožňuje generování kontextu USC pomocí objektů Java, čímž je automaticky sestaven řetězec XML.

Řetězec dotazu XML musí být v následujícím formátu, kdy ... obsahuje úplný řetězec XML:  
@SecurityContext::'...'

Formát řetězce XML je následující:

```
<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="název_vlastnosti">hodnota_vlastnosti</property>
      ...
    </properties>
  </identity>
  ...
</identities>
```

### identities

Hodnota atributu id je jméno uživatele poskytnuté uživatelem při přihlašování do systému.

### **ssoToken**

Volitelné: Určuje prvek LTPA (Lightweight Third-Party Authentication), který je vytvořen pro uživatele na dobu trvání relace prohlížeče. Tento parametr se používá, pouze pokud je pro SSO povolena cílová doména a prolézací modul je nakonfigurován na zabezpečení SSO.

### **identity**

Obsahuje údaje pověření uživatele pro konkrétní zdroj dat. Hodnotou atributu id je doména, ve které jsou uloženy informace o pověření uživatele (v případě serveru Domino jde o název domény serveru Domino).

### **type**

Označuje typ dat a odpovídá typu prolézacího modulu (Notes, DB2, Exchange Server, atd.).

### **username**

Určuje jméno uživatele, které bude použito k prohledávání domény.

### **password**

Určuje heslo odpovídající určenému jménu uživatele. Atribut `encrypt` musí být nastaven na `no` (podnikové vyhledávání neposkytuje metodu šifrování mimo komponentu pro správu identit).

### **groups**

Určuje názvy skupin, ke kterým uživatel patří. Pro každý název skupiny se používá samostatný prvek skupiny.

### **properties**

Určuje seznam vlastností specifických pro připojení, například ID administrátora a šifrované heslo, které bylo použito k vytvoření prolézacího modulu, nebo zda bylo u zdroje povoleno zabezpečení SSO.

#### **název\_vlastnosti**

Název vlastnosti.

#### **hodnota\_vlastnosti**

Hodnota vlastnosti.

## **Ověření uživatelů pomocí profilů uživatelů**

Vyhledávací aplikace mohou uživatelům zobrazit výzvu k registraci údajů pověření potřebných pro přístup k různým doménám v profilu uživatele.

Pokud uživatelé chtějí prohledávat doménu vyžadující ověření údajů pověření uživatele při odeslání dotazu, musí vyhledávací aplikaci poskytnout údaje pověření, pomocí kterých se k doméně přihlašují. Prostřednictvím správy identit podnikového vyhledávání mohou uživatelé ukládat údaje pověření pro libovolný počet domén v profilu uživatele. Údaje pověření jsou šifrovány a uloženy v zabezpečené formě v systému podnikového vyhledávání.

Nejsou-li určena pověření pro doménu, která vyžaduje ověření aktuálních pověření, budou dokumenty z příslušné domény vyloučeny z výsledků vyhledávání.

Uživatelé pracující s vyhledávací aplikací mohou vytvořit uživatelský profil a zaregistrovat své údaje pověření. V ukázkové vyhledávací aplikaci pro podnikové vyhledávání je tato funkce poskytována prostřednictvím volby **Profil**. Vlastní vyhledávací aplikace mohou tuto funkci implementovat jinak.

Kolekce mohou obsahovat dokumenty z mnoha různých typů zdrojů. Kolekce může například obsahovat dokumenty, které byly prolézány ze souborového systému Windows a několika



databázi Lotus Notes. Komponenta pro správu identit rozlišuje mezi různými typy zdrojů a zobrazuje výzvy pouze k údajům pověření potřebným pro přístup k doménám vyžadujícím ověření.

Ve výchozím nastavení je u každého pověření povoleno vyhledávání, a proto musí uživatel zadat jméno uživatele a heslo odpovídající zabezpečeným doménám. Pokud uživatel zapomněl jméno uživatele nebo heslo pro konkrétní doménu, je zaškrtnutím políčka možné zakázat prohledávání této domény. Zakázáním domény bude zabráněno vracení zabezpečených dokumentů z těchto domén v sadě výsledků.

Po vytvoření profilu může uživatel zadat požadavek na vyhledávání. Komponenta pro správu identit má informace nezbytné k sestavení řetězce kontextu zabezpečení uživatele (USC), který bude použit v dalších požadavcích vyhledávání. Nepoužíváte-li komponentu pro správu identit, musí vyhledávací aplikace dodat řetězec kontextu zabezpečení uživatele při dotazech uživatelů na domény vyžadující ověření aktuálních údajů pověření.

Při dalším pokusu uživatele o prohledávání kolekci podnikového vyhledávání opakuje komponenta pro správu identit proces ověřování údajů pověření, ale tentokrát najde profil uživatele. Pokud se nic nezměnilo, bude uživatel automaticky umístěn do místa, kam mohou být odesílány požadavky na vyhledávání a nebude zobrazena výzva k vytvoření profilu.

Pokud najde komponenta pro správu identit změnu v údajích pověření uživatele, bude při přístupu k vyhledávací aplikaci uživateli automaticky zobrazena stránka profilu. K tomu například dochází, je-li změněno heslo pro kteroukoli doménu s povoleným vyhledáváním nebo je-li do kolekce přidána doména požadující ověřování.

Uživatelé mohou ignorovat doporučení k aktualizaci profilu, ale dané dokumenty tak budou vyloučeny z výsledků vyhledávání.

V ukázkové vyhledávací aplikaci dodané s podnikovým vyhledáváním mohou uživatelé aktualizovat profily kdykoli vybráním volby **Profil** na panelu nástrojů.

## **Ověření uživatelů pomocí zabezpečení SSO**

Pokud byly dokumenty v kolekci prolezeny prolézacím modulem poskytujícím podporu zabezpečení SSO (single sign-on), můžete při konfiguraci voleb správy identit určit, že chcete SSO používat pro řízení přístupu k dokumentům.

## **Povolení SSO**

Ověřování SSO (single sign-on) umožňuje uživateli jedním ověřováním získat přístup k mnoha prostředkům bez dalších výzev k předložení údajů pověření. V systému podnikového vyhledávání ulehčuje ověřování SSO správu mnoha jmen uživatelů a hesel, které musejí uživatelé zadávat pro přístup k dokumentům v zabezpečených kolekcích.

Servery IBM WebSphere Application Server a Lotus Domino podporují formu SSO známou jako LTPA (Lightweight Third-Party Authentication). Při pokusu o přístup k jednomu z těchto produktů je uživatel požádán o zadání jména uživatele a hesla. Toto jméno uživatele a heslo je ověřeno v úložišti LDAP sdíleném oběma produkty. Po dokončení ověřování uživatele je vytvořen soubor cookie relace, který obsahuje prvek LTPA. Uživatel může poté přistupovat k dalším prostředkům na libovolném serveru se stejnou konfigurací ověřování bez opakovaných výzev k zadání údajů pověření. Tento prvek trvá po dobu platnosti relace prohlížeče.

Povolení podpory SSO pro použití v kolekcích podnikového vyhledávání:

- Přesvědčte se, zda je pro podnikové vyhledávání na vyhledávacích serverech povoleno globální zabezpečení serveru WebSphere Application Server a platný registr LDAP. Registr LDAP může být libovolný platný produkt LDAP podporovaný serverem WebSphere Application Server.
- Zkontrolujte, zda je mechanismus ověřování produktu WebSphere nakonfigurován k použití aktivního mechanismu ověřování LTPA. Při konfiguraci LTPA zadejte platný flexibilní název domény, například tento.server.com.
- Zkontrolujte, zda byl ze serveru WebSphere Application Server exportován klíč LTPA a zda byl naimportován do jiných produktů ve stejné doméně, pro kterou chcete povolit podporu ověřování LTPA.

Poté, co pomocí prohlížeče ověříte, zda daná konfigurace zabezpečení pracuje správně, můžete pomocí konzoly pro správu podnikového vyhledávání nakonfigurovat prolézací moduly, které podporují ověřování SSO.

## Správa SSO a identit

Pokud uživatelé prohledávají kolekce vyžadující ověření aktuálních údajů pověření, může systém pomocí metod zabezpečení SSO odepřít nebo povolit přístup k dokumentům. Uživatelé nebudou vyzýváni k zadání údajů pověření při prohledávání zdrojů, které podporují ověřování SSO. Komponenta pro správu identit je použita, pokud jsou splněny všechny následující podmínky:

- Na serveru WebSphere Application Server a v cílových doménách je řádně povoleno ověřování SSO.
- Nejméně v jedné z kolekcí, které smí prohledávat vyhledávací aplikace, je povoleno zabezpečení.
- V konzole pro správu podnikového vyhledávání jsou povoleny volby použití komponenty pro správu identit a zabezpečení SSO.
- Při konfiguraci následujících typů prolézacích modulů byla vybrána volba použití SSO a volby vynucení zabezpečení na úrovni dokumentu (například indexování prvků řízení přístupu nebo ověření aktuálních údajů pověření při zpracování dotazu):
  - Content Edition (k dispozici pouze pro úložiště produktu Portal Document Manager)
  - Domino Document Manager (k dispozici pouze pro prolézací moduly, které používají protokol DIIOP)
  - Notes (k dispozici pouze pro prolézací moduly, které používají protokol DIIOP)
  - QuickPlace (k dispozici pouze pro prolézací moduly, které používají protokol DIIOP)

## Zabezpečení vyhledávacího portletu

Pokud uživatelé používají vyhledávací portlety pro podnikové vyhledávání k prohledávání kolekcí z produktu WebSphere Portal, je toto zabezpečení poskytováno i pro dokumenty prolézané pomocí prolézacích modulů Počáteční seznam, Web Content Management a WebSphere Portal.

Zabezpečené vyhledávání je pro tyto typy zdrojů podporováno pouze při použití portletu, ne při použití vyhledávací aplikace, která je spuštěna mimo produkt WebSphere Portal. V produktu WebSphere Portal jsou údaje pověření uživatele získávány prostřednictvím vyhledávacího portletu. Po přihlášení k produktu WebSphere Portal obsahují všechny požadavky na vyhledávání data zabezpečení uživatele (jméno uživatele, členství ve skupině atd.). Vzhledem k tomu, že jsou tyto informace vždy k dispozici, není vyžadováno přihlášení SSO prvku LTPA založeného na tokenech.

Pokud použijete k ochraně webových stránek a dokumentů na serveru produktu WebSphere Portal jiný produkt, musíte při konfiguraci prolézacího modulu zadat volby přihlášení SSO.

Pokud použijete například produkt IBM Tivoli Access Manager WebSEAL nebo CA SiteMinder SSO Agent for PeopleSoft, musíte zadat údaje pověření, které prolézacímu modulu umožní přístup k dokumentům na serveru prostřednictvím jednotného přihlášení. V tomto případě povolíte přihlášení SSO pro přístup prolézacího modulu k zabezpečenému obsahu, ale nepovolíte přihlášení SSO pro zabezpečené vyhledávání.

## Konfigurování správy identit

Komponentu pro správu identit pro podnikové vyhledávání můžete použít k zadání způsobů ověřování údajů pověření uživatele během zpracování dotazů.

### Než začnete

Chcete-li konfigurovat volby správy identit, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### O této úloze

Pokud uživatelé prohledávají kolekce vyžadující při zpracování dotazu ověření aktuálních údajů pověření, může komponenta pro správu identit pro podnikové vyhledávání pomocí profilu nebo metod zabezpečení SSO odeprít nebo povolit přístup k dokumentům.

### Postup

Konfigurování správy identit:

1. Klepnutím na volbu **Zabezpečení** otevřete pohled Zabezpečení.
2. Na stránce Vyhledávací aplikace klepněte na volbu **Konfigurovat správu identit**.
3. Na stránce Konfigurovat správu identit zaškrtněte políčko pro používání komponenty pro správu identit podnikového vyhledávání za účelem řízení způsobu ověřování aktuálních údajů pověření uživatelů během zpracování dotazu. Není-li toto políčko zaškrtnuto, musí vyhledávací aplikace předat řetězec kontextu zabezpečení uživatele při dotazech uživatelů na domény vyžadující ověření aktuálních údajů pověření během zpracování dotazu.
4. Určete, jak často má komponenta pro správu identit extrahovat údaje pověření uživatelů ze záznamů skupin v registru uživatelů produktu WebSphere Application Server. Údaje pověření můžete aktualizovat při každém přístupu uživatele k vyhledávací aplikaci nebo po určeném uplynulém počtu dnů.
5. Pro typy prolézacích modulů podporující ověřování SSO určete, zda má komponenta pro správu identit místo profilů uživatelů používat k ověřování uživatelů prvky zabezpečení SSO. Můžete zaškrtnout políčko pro použití ověřování SSO u všech typů prolézacích modulů nebo zaškrtnout políčka u jednotlivých typů prolézacích modulů.

**Důležité:** Komponenta pro správu identit používá mechanismy zabezpečení SSO, pouze pokud je toto zabezpečení řádně nakonfigurováno na serveru WebSphere Application Server a v cílových doménách.

## Analýza kotvicího textu

Pokud povolíte zabezpečení kolekce, procesy globální analýzy použijí pro indexování kotvicího textu v dokumentech zpracovávaných webovými prolézacími moduly speciální pravidla. Pokud zabezpečení kolekce nepovolíte, můžete při konfigurování jednotlivých webových prolézacích modulů určit, zda má být indexován kotvicí text v odkazech na zakázané dokumenty.

Kotvicí text je text obsažený v hypertextovém odkazu a popisující stránku, na niž daný odkaz směřuje. V následujícím odkazu je například text **Syntaxe dotazu** kotvicím textem odkazu směřujícího na stránku `syntax.htm`:

`<a href="../doc/syntax.htm">Syntaxe dotazu</a>`

Webový prolézací modul obvykle používá odkazy v dokumentech k prolézání dalších dokumentů a k zahrnutí odkazovaných stránek do indexu. Během globální analýzy indexovací procesy nepřirážují kotvicí text pouze k dokumentu, v němž se nachází (zdrojový dokument), ale také k cílovému dokumentu. Ve výše uvedeném příkladu je kotvicí text **Syntaxe dotazu** přiřazen k cílové stránce `syntax.htm` i ke zdrojové stránce obsahující kotvicí text. Cílový dokument je díky tomuto přiřazení možné načíst pomocí dotazů, v nichž je uveden text obsažený ve zdrojovém dokumentu. Toto přiřazení však představuje bezpečnostní riziko, pokud je uživatel oprávněn zobrazit cílový dokument, nikoli však zdrojový dokument.

Pokud při vytvoření kolekce povolíte její zabezpečení, je zpracování kotvicího textu zakázáno. Kotvicí text je tedy indexován spolu s dokumentem pouze tehdy, jestliže se v tomto dokumentu nebo v jeho metadatech skutečně objeví. Tento postup zabezpečení zaručuje, že uživatelé se neseťkají s informacemi z dokumentů, k nimž nemají povolen přístup; dokument bude ve výsledcích vyhledávání vrácen, pouze pokud danému dotazu odpovídá jeho vlastní obsah nebo metadata.

Povolením zabezpečení kolekce můžete zvýšit stupeň zabezpečení webových dokumentů díky tomu, že uživatelé budou moci prohledávat pouze dokumenty s prvky zabezpečení odpovídajícími jejich pověřovacím údajům. V důsledku nezpracování kotvicího textu však nemusí být ve výsledcích vyhledávání uvedeny všechny dokumenty vyhovující podmínkám zadaného dotazu.

Pokud zabezpečení kolekce nepovolíte, můžete při konfigurování rozšířených vlastností webového prolézacího modulu určit, zda má být indexován kotvicí text v odkazech na zakázané dokumenty.

## **Indexování kotvicího textu v odkazech na zakázané dokumenty**

Pokud dokument obsahuje odkazy na dokumenty, které webový prolézací modul nesmí zpracovávat, můžete při konfigurování webového prolézacího modulu určit, zda má být v indexu zachován kotvicí text těchto odkazů.

### **Než začnete**

Chcete-li konfigurovat volby indexování kotvicího textu, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo být administrátorem kolekce pro webový prolézací modul, který chcete konfigurovat.


### **O této úloze**

V přístupu k dokumentům na webovém serveru lze webovému prolézacímu modulu zabránit pomocí direktiv v souboru `robots.txt` nebo v metadatech webových dokumentů. Pokud dokument, který je webovému prolézacímu modulu povoleno prolézat, obsahuje odkazy na zakázané dokumenty, můžete určit, jakým způsobem mají být zpracovávány kotvicí texty těchto odkazů.

Při konfigurování webového prolézacího modulu můžete určit, zda má být indexován kotvicí text odkazů na zakázané dokumenty. Maximální úroveň zabezpečení dosáhnete, jestliže indexování kotvicího textu v odkazech na zakázané dokumenty nepovolíte. V důsledku nezahrnutí kotvicího textu do indexu však nemusí být ve výsledcích vyhledávání uvedeny všechny dokumenty vyhovující podmínkám zadaného dotazu.

### **Postup**

Chcete-li povolit nebo zakázat indexování kotvicího textu v odkazech na zakázané dokumenty, postupujte takto:

1. Upravte kolekci, vyberte stránku **Prolézání**, vyhledejte webový prolézací modul, který chcete konfigurovat, a klepněte na volbu  **Vlastnosti prolézacího modulu**.
2. Klepněte na volbu **Upravit rozšířené vlastnosti webového prolézacího modulu**.
3. Chcete-li indexovat kotvicí text ve všech dokumentech zpracovávaných daným prolézacím modulem, zaškrtněte políčko **Indexovat kotvicí text v odkazech na zakázané dokumenty**. Uživatelé budou moci získat informace o stránkách, které webový prolézací modul nesmí zpracovávat, vyhledají-li text, jenž tvoří kotvicí text odkazů na tyto stránky.

Chcete-li z indexu vyloučit kotvicí text odkazů na zakázané dokumenty, zrušte zaškrtnutí tohoto políčka. Uživatelé nebudou moci získat informace o stránkách, které webový prolézací modul nesmí zpracovávat. Kotvicí text bude vyloučen z indexu spolu se zakázanými dokumenty.

4. Klepněte na tlačítko **OK** a poté na stránce **Vlastnosti webového prolézacího modulu** klepněte znovu na tlačítko **OK**.
5. Změny se projeví po zastavení a restartu prolézacího modulu.

Chcete-li použít změny u dokumentů, které již byly indexovány, musíte provést nové prolezení těchto dokumentů, aby je bylo možné znovu indexovat. Pokud byly do indexu při předchozím prolézání přidány informace o zakázaných dokumentech, budou tyto informace z indexu poté odebrány.

---

## Povolení zabezpečení pro podnikové vyhledávání

Pokud plánujete vynucení zabezpečení v průběhu administrace nebo prohledávání systému podnikového vyhledávání, je třeba nakonfigurovat globální zabezpečení na serveru WebSphere Application Server. Informace o zabezpečení je třeba také nakonfigurovat v konfiguračních souborech podnikového vyhledávání, konzole pro správu a vyhledávacích aplikacích.

### Tip:

Podrobné příklady povolení zabezpečení produktu WebSphere Application Server s úložištěm LDAP, obsahující příklady konfigurace různých prolézacích modulů, aby bylo zajištěno zabezpečení na úrovni dokumentů, najdete v příručce IBM Redbook s názvem **IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios**.

### Postup

Povolení zabezpečení pro systém podnikového vyhledávání:

1. Rozhodněte se, který typ registru uživatelů chcete použít k ověřování uživatelů. Mnoho administrátorů serveru WebSphere Application Server například volí použití registru uživatelů LDAP (Lightweight Directory Access Protocol).
2. V konzole pro správu podnikového vyhledávání vyberte volbu **Zabezpečení** a přiřaďte alespoň jednomu uživateli v registru uživatelů serveru WebSphere Application Server roli **Administrátor podnikového vyhledávání**.

Jiná možnost je přidání ID administrátora podnikového vyhledávání zadaného při instalaci produktu OmniFind Enterprise Edition do registru uživatelů serveru WebSphere Application Server.

**Důležité:** Po povolení globálního zabezpečení mohou ke konzole pro správu za účelem administrace podnikového vyhledávání přistupovat pouze jména uživatelů, která jsou v registru uživatelů a byla jim přiřazena role administrátora podnikového vyhledávání.

3. Použitím procedur v “Konfigurace globálního zabezpečení a registru uživatelů LDAP na serveru WebSphere Application Server” povolíte globální zabezpečení a nakonfigurujete registr uživatelů.
4. Povolíte-li po instalaci produktu OmniFind Enterprise Edition globální zabezpečení, musíte systému podnikového vyhledávání poskytnout jméno uživatele a heslo serveru WebSphere Application Server. Tyto informace je možné zadat pomocí příkazu **eschangewaspw**:
  - Pokud jste produkt OmniFind Enterprise Edition nainstalovali na jeden server, použijte proceduru v části “Povolení zabezpečení pro systém podnikového vyhledávání s jediným serverem” na stránce 262.
  - Pokud jste produkt OmniFind Enterprise Edition nainstalovali na více než jeden server, použijte proceduru v části “Povolení zabezpečení pro systém podnikového vyhledávání s více servery” na stránce 262.
5. Dokončete úlohy vhodné pro typy dokumentů, které chcete prolézat a prohledávat. Podrobnosti naleznete v části “Požadavky nastavení prolézacího modulu pro podporu zabezpečení” na stránce 263.

## Konfigurace globálního zabezpečení a registru uživatelů LDAP na serveru WebSphere Application Server

Chcete-li povolit zabezpečení v produktu OmniFind Enterprise Edition, je třeba nejprve povolit na serveru WebSphere Application Server globální zabezpečení.

### O této úloze

V rámci povolování globálního zabezpečení je třeba nakonfigurovat registr uživatelů k ověřování jmen uživatelů. V této úloze je probrán způsob konfigurace registru uživatelů LDAP (Lightweight Directory Access Protocol) na serveru WebSphere Application Server současně s povolením globálního zabezpečení. Přestože server WebSphere Application Server podporuje i jiné typy registrů uživatelů, nelze k ověřování uživatelů podnikového vyhledávání použít místní registr operačního systému. Místní registr operačního systému lze použít, pouze pokud pro podnikové vyhledávání na vyhledávacím serveru existují uživatelské účty pro každého uživatele v podniku.

Tato úloha je založena na produktu WebSphere Application Server verze 6. Pokud používáte starší verzi serveru WebSphere Application Server, mohou se lišit některé výchozí cesty a popisy uživatelského rozhraní. Tato úloha také pro registr LDAP používá server IBM Tivoli Directory Server. Pokud použijete jiný typ registru nebo vlastní registr, musíte poskytnout příslušné informace o registru.

Tato úloha shrnuje kroky požadované ke konfiguraci globálního zabezpečení pro použití se systémem podnikového vyhledávání. Podrobné pokyny naleznete v informačním centru produktu *WebSphere Application Server verze 6.0.x* na následující adrese URL:  
<http://publib.boulder.ibm.com/infocenter/wasinfo/v6r0/index.jsp>

### Postup

Povolení globálního zabezpečení v produktu WebSphere Application Server:

1. Na vyhledávacím serveru podnikového vyhledávání pomocí následující adresy URL otevřete konzolu pro správu serveru WebSphere Application Server, kde *localhost* je lokální hostitel nebo název serveru, například *omnifind.search.xyz.com*.  
`http://localhost:9060/ibm/console`
2. Klepněte na volbu **Zabezpečení** a poté na **Globální zabezpečení**.
3. Nastavte produkt WebSphere k použití registru LDAP:

- a. V části **Registry uživatelů** klepněte na volbu **LDAP**.
  - b. Zadejte jméno uživatele a heslo používané ke spuštění aplikačního serveru.
  - c. Jako typ registru vyberte server IBM Tivoli Directory Server.
  - d. Zadejte název hostitele serveru LDAP jako adresu IP nebo název hostitele DNS. Výchozí číslo portu je 389.
  - e. Zadejte základní jednoznačný název (DN), který je počátečním bodem prohledávání registru, například `ou=sales,o=ibm,c=us`.
  - f. Protože některé servery LDAP nepodporují při prohledávání registru anonymní vazby, zadejte název DN aplikačního serveru, například `cn=searchuser,o=ibm,c=us`, a poté zadejte heslo aplikačního serveru. Aplikační server použije tento název DN a heslo k vytvoření vazby na registr.
  - g. Chcete-li mezi serverem WebSphere a LDAP použít komunikaci s protokolem SSL (Secure Sockets Layer), zaškrtněte políčko **SSL povoleno**.
  - h. Klepněte na tlačítko **Použít** a potom **OK**.
4. V části **Obecné vlastnosti** zaškrtněte políčka **Povolit globální zabezpečení** a **Vynutit zabezpečení Java 2**.
  5. Jako aktivní mechanismus ověřování vyberte SWAM (Simple WebSphere Authentication Mechanism).
  6. Jako aktivní registr uživatelů vyberte LDAP (Lightweight Directory Access Protocol).
  7. Klepněte na tlačítko **OK**.
  8. V horní části stránky klepněte na odkaz **Uložit**. Pokud budete vyzváni k uložení změn, klepněte na tlačítko **Uložit**.
  9. Na panelu nástrojů klepněte na volbu **Odhlásit**.
  10. Zastavte aplikaci ESSearchServer a znovu ji spusťte.

#### System AIX, Linux nebo Solaris

```
./stopServer.sh ESSearchServer
./startServer.sh ESSearchServer
```

#### System Windows

```
stopServer ESSearchServer
startServer ESSearchServer
```

Tyto skripty jsou umístěny v adresáři `KOŘEN_INSTALLACE_WAS/AppServer/bin`:

- Pro produkt WebSphere Application Server verze 5 je výchozí instalační cesta v systémech AIX `/usr/WebSphere`, v systémech Linux nebo Solaris `/opt/WebSphere` a v systémech Windows `C:\Program Files\WebSphere`.
  - Pro produkt WebSphere Application Server verze 6 je výchozí instalační cesta v systémech AIX `/usr/IBM/WebSphere`, v systémech Linux nebo Solaris `/opt/IBM/WebSphere` a v systémech Windows `C:\Program Files\IBM\WebSphere`.
11. Restartujte konzolu pro správu serveru WebSphere Application Server.
  12. Protože se nyní server spouští v zabezpečeném režimu, zadejte jméno uživatele a heslo serveru, které jste zadali při konfiguraci registru uživatelů LDAP, a přihlaste se do konzoly (viz krok 3b).

## Povolení zabezpečení pro systém podnikového vyhledávání s jediným serverem

Pokud po instalaci produktu OmniFind Enterprise Edition povolíte globální zabezpečení serveru WebSphere Application Server, musíte pomocí příkazu **eschangewaspw** aktualizovat konfigurační soubor podnikového vyhledávání **es.cfg** přidáním hesla pro uživatele serveru WebSphere Application Server.

### Než začnete

Presvědčte se, zda soubor `config.properties` pro aplikaci `ESearchApplication` specifikuje platné jméno uživatele a heslo serveru WebSphere Application Server. Výchozí umístění tohoto souboru je `ES_INSTALL_ROOT/installedApps/ESearchApplication.ear/ESearchApplication/ESearchApplication.war/WEB-INF`.

### O této úloze

Příkaz **eschangewaspw** dané heslo před uložením do souboru `es.cfg` zašifruje.

### Postup

Povolení použití globálního zabezpečení ve stávajícím systému podnikového vyhledávání s jediným serverem:

1. Přihlaste se k serveru podnikového vyhledávání jako administrátor podnikového vyhledávání a zastavte systém podnikového vyhledávání:  
`esadmin system stopall`
2. Presvědčte se, zda položka `WASUser` v souboru `ES_NODE_ROOT/nodeinfo/es.cfg` specifikuje platné jméno uživatele serveru WebSphere Application Server.
3. Spusťte následující skript, kde `heslo_WAS` představuje heslo ke jménu uživatele serveru WebSphere Application Server, které je uvedeno v souboru `ES_NODE_ROOT/nodeinfo/es.cfg` (viz krok 2).

#### Systém AIX, Linux nebo Solaris

```
eschangewaspw.sh heslo_WAS
```

#### Windows

```
eschangewaspw heslo_WAS
```

4. V systému Windows vyberte možnosti **Ovládací panel** → **Nástroje pro správu** → **Služby** a přidejte stejné jméno uživatele a heslo serveru WebSphere Application Server pro server WebSphere Application Server a pro služby `ESearchServer`.
5. Restartujte systém podnikového vyhledávání:  
`esadmin system startall`

## Povolení zabezpečení pro systém podnikového vyhledávání s více servery

Pokud po instalaci produktu OmniFind Enterprise Edition povolíte globální zabezpečení serveru WebSphere Application Server, musíte pomocí příkazu **eschangewaspw** aktualizovat konfigurační soubor podnikového vyhledávání **es.cfg** přidáním hesla pro uživatele serveru WebSphere Application Server.

### Než začnete

Presvědčte se, zda soubor `config.properties` pro aplikaci `ESearchApplication` specifikuje platné jméno uživatele a heslo serveru WebSphere Application Server. Výchozí umístění



tohoto souboru na vyhledávacích serverech je ES\_INSTALL\_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication/ESSearchApplication.war/WEB-INF.

### O této úloze

Příkaz **eschangewaspw** dané heslo před uložením do souboru **es.cfg** zašifruje.

### Postup

Povolení použití globálního zabezpečení ve stávajícím systému podnikového vyhledávání s více servery:

1. Na indexovém serveru podnikového vyhledávání proveďte následující kroky:
  - a. Přihlaste se jako administrátor podnikového vyhledávání a zastavte systém podnikového vyhledávání:  
`esadmin system stopall`
  - b. Přesvědčte se, zda položka **WASUser** v souboru **ES\_NODE\_ROOT/nodeinfo/es.cfg** specifikuje platné jméno uživatele serveru WebSphere Application Server.
  - c. Spusťte následující skript, kde **heslo\_WAS** představuje heslo uživatele serveru WebSphere Application Server, který je uveden v souboru **ES\_NODE\_ROOT/nodeinfo/es.cfg** (viz krok 1b).

#### Systém AIX, Linux, nebo Solaris

```
eschangewaspw.sh heslo_WAS
```

#### Windows

```
eschangewaspw heslo_WAS
```

2. Na druhém vyhledávacím serveru (pro konfiguraci se dvěma servery) nebo na prolézacím serveru a na obou vyhledávacích serverech (pro konfiguraci se čtyřmi servery) proveďte následující kroky:
  - a. Přihlaste se jako administrátor podnikového vyhledávání.
  - b. Spusťte následující skript, kde **heslo\_WAS** představuje heslo uživatele serveru WebSphere Application Server, který je uveden v souboru **ES\_NODE\_ROOT/nodeinfo/es.cfg** (viz krok 1b).

#### Systém AIX, Linux nebo Solaris

```
eschangewaspw.sh heslo_WAS
```

#### Příkazový řádek systému Windows

```
eschangewaspw heslo_WAS
```

3. V systému Windows vyberte možnosti **Ovládací panel** → **Nástroje pro správu** → **Služby** a přidejte stejné jméno uživatele a heslo serveru WebSphere Application Server pro server WebSphere Application Server a pro služby ESSearchServer.
4. Na indexovém serveru podnikového vyhledávání restartujte systém podnikového vyhledávání:  
`esadmin system startall`

---

## Požadavky nastavení prolézacího modulu pro podporu zabezpečení

Ke shromáždění informací umožňujících vynucení zabezpečení na úrovni dokumentu musí mít prolézací moduly oprávnění pro přístup k nativním datům zabezpečení. U některých typů dat je třeba ke konfiguraci zabezpečeného prostředí provést dodatečné kroky.

*Tabulka 7. Požadavky nastavení prolézacího modulu pro podporu zabezpečení*

---

Content Edition Prolézací moduly

---

Tabulka 7. Požadavky nastavení prolézacího modulu pro podporu zabezpečení (pokračování)

<p>Před vytvořením prolézacího modulu pro přístup k úložištím v přímém režimu nakonfigurujte systém WebSphere Information Integrator Content Edition na spuštění v přímém režimu a nakonfigurujte konektor pro prolézací server.</p> <p>Před vytvořením prolézacího modulu pro přístup k úložištím v režimu serveru nakonfigurujte prolézací server spuštěním skriptu (<code>escrvbr.sh</code> v systému AIX, Linux či Solaris nebo <code>escrvbr.vbs</code> v systému Windows).</p> <p>Při konfiguraci prolézacího modulu zadejte jméno uživatele a heslo, které umožňuje prolézacímu modulu přístup ke každému prolézanému úložišti. Pro každé úložiště v prolézaném prostoru můžete dle potřeb zadat jiné jméno uživatele a heslo.</p>	<p>Související témata:</p> <ul style="list-style-type: none"><li>• “Přístup k úložištím Content Edition v přímém režimu” na stránce 41</li><li>• “Přístup k úložištím WebSphere II Content Edition v režimu serveru” na stránce 42</li><li>• “Konfigurování prolézacího serveru v systému UNIX pro produkt WebSphere II Content Edition” na stránce 42</li><li>• “Konfigurování prolézacího serveru v systému Windows pro produkt WebSphere II Content Edition” na stránce 43</li></ul>
<b>Prolézací moduly DB2</b>	
<p>Před vytvořením prolézacího modulu nakonfigurujte prolézací server spuštěním skriptu <code>escrdb2.sh</code> v systému AIX, Linux či Solaris nebo <code>escrdb2.vbs</code> v systému Windows.</p> <p>Při konfiguraci prolézacího modulu na prolézání vzdálených nekatalogizovaných databází umožníte zadáním jména uživatele a hesla prolézání každé databáze na cílovém databázovém serveru. Pro každou databázi v prolézaném prostoru můžete dle potřeb zadat jiné jméno uživatele a heslo.</p>	<p>Související témata:</p> <ul style="list-style-type: none"><li>• “Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2” na stránce 46</li><li>• “Konfigurování prolézacího serveru v systému Windows pro prolézací moduly DB2” na stránce 47</li></ul>
<b>DB2 Content ManagerProlézací moduly</b>	
<p>Před vytvořením prolézacího modulu nakonfigurujte prolézací server spuštěním skriptu <code>escrcm.sh</code> v systému AIX, Linux či Solaris nebo <code>escrcm.vbs</code> v systému Windows.</p> <p>Při konfiguraci prolézacího modulu zadejte jméno uživatele a heslo, které umožňuje prolézacímu modulu přístup ke každému prolézanému serveru. Pro každý server v prolézaném prostoru můžete dle potřeb zadat jiné jméno uživatele a heslo.</p>	<p>Související témata:</p> <ul style="list-style-type: none"><li>• “Konfigurování prolézacího serveru v systému UNIX pro prolézací moduly DB2 Content Manager” na stránce 54</li><li>• “Konfigurování prolézacího serveru v systému Windows pro produkt DB2 Content Manager” na stránce 55</li></ul>
<b>Prolézací moduly Domino Document Manager, Notes a QuickPlace</b>	

Tabulka 7. Požadavky nastavení prolézacího modulu pro podporu zabezpečení (pokračování)

---

<p>Prolézání serverů Lotus Domino, které používají protokol NRPC (Notes Remote Procedure Call):</p> <ul style="list-style-type: none"><li>• V systému AIX se přesvědčte, zda je na prolézacím serveru nainstalován a k dispozici modul portu dokončení I/O.</li><li>• Před vytvořením prolézacího modulu nakonfigurujte prolézací server spuštěním skriptu <code>escrnote.sh</code> v systému AIX, Linux či Solaris nebo <code>escrnote.vbs</code> v systému Windows.</li><li>• Na prolézacím serveru podnikového vyhledávání musí být nainstalován server Domino a musí být členem domény Domino k prolézání.</li><li>• K ověření údajů pověření aktuálního uživatele při odeslání požadavku na vyhledávání musí být prolézán server Domino nakonfigurován jako Lotus Domino Trusted Server.</li><li>• Při konfiguraci prolézacího modulu zadejte cestu k souboru jména uživatele aplikace Lotus Notes, který je autorizován pro přístup k serveru, například <code>c:\Program Files\lotus\notes\data\jméno.id</code> nebo <code>/local/notesdata/jméno.id</code>, a heslo k tomuto souboru ID.</li></ul>	<p>Související témata:</p> <ul style="list-style-type: none"><li>• “Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 76</li><li>• “Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino” na stránce 72</li><li>• “Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 73</li><li>• “Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů” na stránce 268</li><li>• “Konfigurování serverů používajících protokol DIIOP” na stránce 75</li><li>• “Konfigurování serveru QuickPlace pro používání zabezpečení lokálních uživatelů” na stránce 269</li><li>• “Konfigurování služby Directory Assistance na serveru QuickPlace” na stránce 270</li></ul>
<p>Prolézání serverů Lotus Domino, které používají protokol DIIOP (Domino Internet Inter-ORB Protocol):</p> <ul style="list-style-type: none"><li>• V systému AIX se přesvědčte, zda je na prolézacím serveru nainstalován a k dispozici modul portu dokončení I/O.</li><li>• Nakonfigurujte prolézací server, aby mohl používat daný protokol.</li><li>• Při konfiguraci prolézacího modulu zadejte úplné jméno uživatele aplikace Lotus Notes, které je autorizováno pro přístup k serveru, například <code>Jméno uživatele/Město/Společnost</code>, a heslo k tomuto jménu uživatele.</li></ul>	
<p>Chcete-li prolézat servery QuickPlace, je třeba daný server nakonfigurovat na podporu zabezpečení místního uživatele nebo služby Directory Assistance v závislosti na požadovaném typu zabezpečení.</p>	

---

#### **Exchange ServerProlézací moduly**

---

<p>Při konfiguraci prolézacího modulu zadejte jméno uživatele, které je autorizováno pro přístup k veřejným složkám na prolézaném serveru Exchange Server, a heslo k tomuto jménu uživatele.</p>	<p>Související témata:</p> <ul style="list-style-type: none"><li>• “Ověřování přístupu k zabezpečeným dokumentům serveru Exchange Server” na stránce 267</li></ul>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------

Aby mohl prolézací modul používat při prolézání dat správu klíčů serveru Exchange Server a protokol SSL (Secure Sockets Layer), zadejte také úplnou cestu k souboru úložiště klíčů a heslo umožňující prolézacímu modulu přístup k tomuto souboru. Na prolézacím serveru podnikového vyhledávání musí existovat soubor úložiště klíčů.

#### **Databáze JDBC Prolézací moduly**

---

Tabulka 7. Požadavky nastavení prolézacího modulu pro podporu zabezpečení (pokračování)

---

Při konfiguraci prolézacího modulu můžete zadat jméno uživatele a heslo, které umožní prolézání tabulek v cílové databázi. Pro každou databázi v prolézaném prostoru můžete dle potřeb zadat jiné jméno uživatele a heslo.

---

#### **NNTP Prolézací moduly**

---

Servery NNTP určené k prolézání musí povolit prolézacímu serveru čtení dat.

---

#### **Souborový systém UNIX Prolézací moduly**

---

Podadresáře systému AIX, Linux nebo Solaris určené k prolézání musí povolit prolézacímu serveru čtení dat.

---

#### **Webové prolézací moduly**

---

Webový prolézací modul respektuje protokol vyloučení robotů. Pokud webový server obsahuje na nejvyšší úrovni adresářů serveru soubor `robots.txt`, prolézací modul soubor zanalyzuje a webové servery na tomto serveru proleze pouze tehdy, má-li to povoleno. Informace o tomto protokolu naleznete na stránce <http://www.robotstxt.org/wc/exclusion.html>.

Související témata:

- “Webové servery chráněné základním ověřováním HTTP” na stránce 95
- “Webové servery chráněné ověřováním založeným na formulářích” na stránce 96

Konfigurace webového prolézacího modulu:

- Pro prolézací modul je třeba zadat název uživatelského agenta. Pomocí pravidel v souborech `robots.txt` prolézaných serverů je možné určit, zda tento název povolí nebo odepře přístup.
  - Volitelné: Pokud webový server používá k omezení přístupu k webům základní ověřování HTTP, můžete zadat pověřovací údaje pro ověřování, které webovému prolézacímu modulu umožní přístup na stránky chráněné heslem.
  - Volitelné: Pokud webový server používá pro omezení přístupu k webům formuláře HTML, můžete zadat údaje pověření pro ověřování, které webovému prolézacímu modulu umožní přístup na stránky chráněné heslem.
- 

#### **Prolézací moduly Počáteční seznam, Web Content Management a WebSphere Portal**

---

Před vytvořením prolézacího modulu musíte spustit instalační skript pro integraci podnikového vyhledávání s produktem WebSphere Portal. Pro různé verze produktu WebSphere Portal jsou k dispozici různé skripty.

Související témata:

- “Instalační skripty integrace podnikového vyhledávání do portálu WebSphere Portal” na stránce 324

Při konfiguraci prolézacího modulu zadejte úplný jednoznačný název (DN) umožňující prolézacímu modulu načíst stránky z tohoto prolézaného serveru, například `uid=admin,cn=RegularEmployees,ou=Software Group,o=IBM,c=US`. Název DN se musí shodovat s názvem DN konfigurovaným v produktu WebSphere Portal.

Zkontrolujte, zda jsou pro zadané názvy DN uživatele v komponentě PAC (Portal Access Control) serveru WebSphere Portal definována oprávnění. Prolézací modul používá komponentu PAC k získání dat řízení přístupu pro prolézané dokumenty.

---

#### **Souborový systém Windows Prolézací moduly**

---

Tabulka 7. Požadavky nastavení prolézacího modulu pro podporu zabezpečení (pokračování)

Podadresáře určené k prolézání musí povolit prolézacímu serveru čtení dat. Při konfiguraci prolézacího modulu na prolézání vzdálených souborových systémů zadejte jméno uživatele, které umožní přístup prolézacího modulu ke vzdáleným datům, a zadejte heslo k tomuto jménu uživatele.

K ověření údajů pověření aktuálního uživatele při odeslání požadavku na vyhledávání zkontrolujte, zda jsou správně nakonfigurovány účty domény. Požadavky na nastavení účtů domény pro soubory, které byly prolezeny v lokálním počítači, se liší od požadavků pro soubory prolezené na vzdáleném serveru Windows.

Související témata:

- “Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows” na stránce 271
- “Bezpečné vyhledávání v důvěryhodných doménách systému Windows” na stránce 272

## Ověřování přístupu k zabezpečeným dokumentům serveru Exchange Server

Chcete-li pomocí prolézacího modulu Exchange Server prolézat dokumenty chráněné ochrannou bariérou firewall, musíte ověřit, že prolézací server může přistupovat k serveru veřejných složek Microsoft Exchange Server.

### O této úloze

Pokud prolézací server nemůže přistupovat k zabezpečenému serveru Exchange Server, obdržíte ze serveru kód HTTP 501 (Not Implemented - není implementováno). Může se rovněž zobrazit zpráva s informací o tom, že byla obdržena neočekávaná odezva HTTP.

### Postup

Chcete-li ověřit, že prolézací server může přistupovat k dokumentům za ochrannou bariérou firewall, postupujte takto:

1. Spusíte na prolézacím serveru webový prohlížeč.
2. Přejděte na adresu URL pro server veřejných složek Exchange Server, který chcete prolézat. Příklad: <http://exchange.spolecnost.com/public/>
3. Ověřte, že lze stránku serveru Exchange Server otevřít.

Pokud server Exchange Server není dostupný, obraťte se na administrátora serveru pro příslušnou organizaci.

## Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino

Pokud server Domino, který má být prolézán, používá protokol NRPC (Notes Remote Procedure Call), musíte prolézací server konfigurovat tak, aby bylo možné vynucovat řízení přístupu na úrovni dokumentu.

Chcete-li vynutit zabezpečení na úrovni dokumentu pro dokumenty na serveru Domino, který používá protokol NRPC, musíte na prolézacím serveru instalovat server Domino. Tento server Domino musí být členem domény produktu Domino. Instalujte server Domino a konfiguruje jej podle pokynů v dokumentaci k produktu Lotus Domino.

Musíte rovněž provést následující úlohy, aby mohly vyhledávací servery ověřit, zda je uživatel, který prohlédává zabezpečenou kolekci, oprávněn k zobrazování dokumentů vyhovujících kritériím vyhledávání:

- “Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů”.
- “Konfigurace globálního zabezpečení a registru uživatelů LDAP na serveru WebSphere Application Server” na stránce 260.

#### Související pojmy

“Ověřování aktuálního pověření během zpracování dotazu” na stránce 251

“Prolézací moduly Notes” na stránce 68

## Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů

Chcete-li vynutit zabezpečení pro dokumenty prolezené prolézacím modulem Notes, který používá protokol NRPC (Notes Remote Procedure Call), musí být servery Domino, které mají být prolézány, konfigurovány jako servery Lotus Domino Trusted Server.

### Než začnete

Tento postup je vyžadován, pokud chcete při prohledávání vzdálených databází vynutit zabezpečení na úrovni dokumentu. Chcete-li prohledávat databáze umístěné lokálně na prolézacím serveru, není tento postup nutný.

Chcete-li konfigurovat servery Trusted Server, musí být v prolézacím modulu instalován server Domino. Tento server Domino musí být členem domény produktu Domino.

### O této úloze

Při konfigurování voleb zabezpečení na úrovni dokumentu pro prolézací modul Notes můžete určit, zda chcete vynucovat řízení přístupu na základě ověřování aktuálních údajů pověření uživatele, když odešle dotaz. Chcete-li vynutit tento typ zabezpečení, musí být servery Domino, které mají být prolézány, konfigurovány jako servery Lotus Domino Trusted Server.

Když uživatelé prohledávají doménu vyžadující ověření jejich aktuálních údajů pověření, server Trusted Server povolí přepnutí identifikátoru serveru Domino na kontext aktuálního jména uživatele. Databáze Domino bude otevřena, jako by ji otevřel aktuální uživatel, a jsou vynuceny všechny informace týkající se seznamu řízení přístupu k databázi pro příslušného uživatele.

Možnost přepínání kontextů tímto způsobem je k dispozici typicky pouze pro databáze uložené v adresáři **data** na lokálním serveru Domino. Počínaje produktem Lotus Domino verze 6.5.1 je tato možnost poskytována prostřednictvím serveru Trusted Server. Při konfigurování serveru Trusted Server administrátor produktu Domino určuje, které servery Domino mají být důvěryhodné pro provádění citlivých operací, například vystupování jako jiný uživatel při přístupu k databázi ze vzdáleného počítače.

### Postup

Chcete-li konfigurovat server Trusted Server, proveďte následující kroky na všech serverech Domino prolézáných prolézacím modulem Notes:

1. Pomocí souboru s ID administrátora domény Domino na serveru Domino otevřete klienta pro správu produktu Lotus Domino.
2. V nabídce **File** (Soubor) vyberte volbu **Open server** (Otevřít server).
3. Zadejte název serveru Domino, pro který chcete povolit možnosti serveru Trusted Server.
4. Vyberte kartu **Configuration** (Konfigurace).
5. Rozbalte objekt **Server**, vyberte dokument **Current Server** (Aktuální server) a klepněte na volbu **Edit Server** (Upravit server).

6. Vyberte kartu **Security** (Zabezpečení), přejděte na konec dokumentu, vyhledejte položku **Trusted Servers** (Servery Trusted Server) a klepněte na šipku dolů.
7. Určete jednu z následujících voleb:

#### **LocalDomainServers**

Tuto volbu vyberte, mají-li být všechny servery v doméně Domino považovány za servery Trusted Server.

#### *název\_serveru*

Zadejte název serveru Domino, který chcete mít možnost prolézat a prohledávat jako server Trusted Server.

Pokud se server Domino, který má být prolézán, nalézá v jiné doméně Domino, musíte určit název serveru nebo vybrat skupinu **OtherDomainServers**. Musíte rovněž postupovat podle pokynů pro křížovou certifikaci souboru ID serveru Domino produktu s jinou doménou Domino. Informace o příslušných postupech naleznete v dokumentaci k serveru Domino.

8. Uložte změny klepnutím na odkaz **Save and Close** (Uložit a zavřít).
9. Restartujte vzdálené servery Domino, kterým jste povolili vystupovat jako servery Trusted Server.

#### **Související pojmy**

“Ověřování aktuálního pověření během zpracování dotazu” na stránce 251

“Prolézací moduly Notes” na stránce 68

## **Konfigurování serveru QuickPlace pro používání zabezpečení lokálních uživatelů**

Máte-li v úmyslu konfigurovat prolézací modul QuickPlace tak, aby používal pro implementaci zabezpečení volbu Lokální uživatel, musíte před vytvořením prolézacího modulu konfigurovat produkt Domino Directory na serveru Lotus QuickPlace.

### **O této úloze**

Při konfigurování prolézacího modulu QuickPlace můžete vybrat režim zabezpečení pro prolézací modul. Tento režim bude použit pro vynucení zabezpečení na úrovni dokumentu. Vyberete-li režim Lokální uživatel, musíte zaručit, že všechna jména lokálních uživatelů a lokální skupiny jsou registrovány pro službu Domino Directory (hierarchie služby Domino Directory musí odpovídat hierarchii produktu QuickPlace).

Musíte rovněž zaručit, že jméno uživatele a heslo zadané pro používání prolézacím modulem je registrováno pro službu Domino Directory a má oprávnění pro čtení databáze, která má být prolézána.

Pro použití produktu QuickPlace je vyžadováno pouze jméno uživatele. Pro prolézání zdrojů QuickPlace je však vyžadováno plně expandované jméno uživatele. Formát expandovaného jména uživatele je následující:

*jméno\_uživatele/název\_místa/QP/název\_domény*

Tímto postupem určete plně expandovanou verzi jména uživatele, zaručte, že je toto jméno uživatele oprávněno pro čtení databáze QuickPlace, a přidejte příslušné jméno uživatele do služby Domino Directory. Služba Domino Directory musí obsahovat jméno uživatele, které bude použito k prolézání databází QuickPlace, a všechny lokální uživatele a lokální skupiny produktu QuickPlace (hierarchie služby Domino Directory musí odpovídat hierarchii produktu QuickPlace).

### **Postup**

Chcete-li konfigurovat server QuickPlace pro používání zabezpečení lokálních uživatelů, postupujte takto:

1. Potvrďte oprávnění pro jméno uživatele:
  - a. Otevřete dokument Server na serveru QuickPlace.
  - b. Otevřete stránku Files (Soubory) a poté otevřete seznam řízení přístupu (ACL) pro databázi, kterou chcete prolézat.
  - c. Potvrďte, že jméno lokálního uživatele, pro jehož použití bude prolézací modul konfigurován, v seznamu řízení přístupu existuje a že má oprávnění pro čtení databáze.  
V kroku 2 je třeba zadat plně expandovaný tvar tohoto jména uživatele.
2. Přidejte uživatele do služby Domino Directory:
  - a. Otevřete dokument Server na serveru QuickPlace.
  - b. Na stránce People and Groups (Lidé a skupiny) v položce stromu osob přidejte plně expandované jméno uživatele, které jste potvrdili v kroku 1.
  - c. Do pole **Internet password** (Internetové heslo) zadejte heslo pro příslušné jméno uživatele.

## Konfigurování služby Directory Assistance na serveru QuickPlace

Máte-li v úmyslu konfigurovat prolézací modul QuickPlace tak, aby používal adresář LDAP pro implementaci zabezpečení, musíte před konfigurováním prolézacího modulu vytvořit databázi služby Directory Assistance na serveru Lotus QuickPlace.

### Omezení

Na serveru QuickPlace, který chcete prolézat, musí být spuštěny úlohy DIOP a HTTP.

### Postup

Chcete-li konfigurovat službu LDAP Directory Assistance na serveru QuickPlace, postupujte takto:

1. Vytvořte databázi služby Directory Assistance:
  - a. Otevřete dokument Server na serveru QuickPlace.
  - b. Vytvořte databázi s použitím šablony **Directory Assistance(6)**. Šablona je umístěna na serveru.
  - c. Klepnutím na volbu **Add Directory Assistance** (Přidat službu Directory Assistance) vytvořte v databázi dokument.
  - d. Otevřete kartu Basic (Základní) a do pole **DomainType** (Typ domény) zadejte hodnotu **LDAP**.
  - e. Otevřete kartu Naming Contexts (Kontexty pojmenovávání) a ověřte, že je políčko **Trusted for credentials** (Důvěryhodné pro pověření) zaškrtnuto.
  - f. Otevřete kartu LDAP a zadejte informace o serveru LDAP.
  - g. Uložte dokument Server a zavřete jej.
2. Konfigurujte server QuickPlace pro použití databáze služby Directory Assistance:
  - a. Otevřete dokument Server na serveru QuickPlace.
  - b. Otevřete kartu Basic (Základní) a do pole **Directory assistance database name** (Název databáze služby Directory Assistance) zadejte název databáze vytvořené v kroku 1.
  - c. Uložte dokument Server a zavřete jej.  
Server QuickPlace nyní může používat server LDAP jako sekundární adresář Domino.



## Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows

Chcete-li povolit ověřování aktuálních údajů pověření, když uživatel prohledává dokumenty prolezené prolézacím modulem Souborový systém Windows, musíte konfigurovat informace o účtu domény na prolézacím serveru i na serveru Microsoft Windows.

Při konfigurování prolézacího modulu Souborový systém Windows můžete určit, zda chcete prolézat podadresáře v lokálním počítači nebo podadresáře ve vzdáleném počítači. Je-li pro kolekci povoleno zabezpečení, můžete rovněž určit volby pro řízení přístupu k dokumentům v prolézáných podadresářích.

Rozhodnete-li se vynucovat řízení přístupu na základě ověřování aktuálních údajů pověření uživatele, když odešle dotaz, musíte zaručit, že jsou účty domény správně konfigurovány. Požadavky na nastavení účtů domény pro soubory, které byly prolezeny v lokálním počítači, se liší od požadavků pro soubory prolezené na vzdáleném serveru Windows.

**Důležité:** Údaje pověření uživatele nelze ověřit během zpracování dotazu, pokud jsou splněny obě z následujících podmínek:

- Server systému Windows určený k prolézání není členem domény.
- Adresář určený k prolézání je vzdáleným adresářem, například \\název\_serveru\název\_hostitele.

### Ověřování podle dat řízení lokálního přístupu

Při ověřování aktuálních údajů pověření uživatele systém používá informace o lokálním účtu uživatele i informace o účtu domény (pokud počítač náleží do domény systému Windows). Chcete-li ověřovat údaje pověření během zpracování dotazu, musí být obě jména uživatele uvedena v informacích o zabezpečení pro dokumenty, které mají být prohledávány.

#### Lokální účty

V případě lokálního účtu je formát jména uživatele následující:  
*NÁZEV\_POČÍTAČE\JMÉNO\_UŽIVATELE*

Při přihlašování uživatelé zadávají pouze jméno uživatele, pro přiřazení správně určených oprávnění uživatele systému Windows se však používá úplné jméno. Jméno uživatele pro lokální účet abcuser může například odpovídat úplný název účtu WINSERVER1\abcuser.

Když uživatelé používají vyhledávací aplikaci a konfigurují profil pro prohledávání zabezpečených dokumentů v lokálním systému, musí zadat jméno uživatele, které používají pro přihlašování k systému Windows (například abcuser).

#### Doménové účty

V případě doménového účtu je formát jména uživatele následující:  
*NÁZEV\_DOMÉNY\JMÉNO\_UŽIVATELE*

Při přihlašování uživatelé zadávají tyto informace v následujícím formátu:  
*JMÉNO\_UŽIVATELE@NÁZEV\_DOMÉNY*

Pokud například konfiguruje přiřazení uživatelských oprávnění pro soubor a vyberete doménu WIN1\abcuser, bude účet poté zobrazen ve tvaru abcuser@win1.spolecnost.com.

Když uživatelé používají vyhledávací aplikaci a konfigurují profil, který jim umožňuje prohledávat dokumenty v zabezpečené doméně, musí zadat jméno uživatele, které používají pro přihlašování k systému Windows (například abcuser@win1.spolecnost.com).

Chcete-li vynutit ověřování aktuálních údajů pověření v lokálních počítačích, musí mít uživatelské účty používané prolézacím serverem následující oprávnění uživatelů systému Windows. Uživatelská oprávnění lze přiřadit pomocí nástrojů pro správu systému Windows:

**Nástroje pro správu → Místní zásady zabezpečení → Místní zásady → Přiřazení oprávnění lokálního uživatele.**

- Jméno uživatele, pod kterým je spuštěn prolézací server, musí mít oprávnění **Jednat jako součást operačního systému**. Toto oprávnění je konfigurováno pro uživatele, která je administrátorem podnikového vyhledávání, při instalaci produktu OmniFind Enterprise Edition.
- Uživatelé musí mít přiřazeno uživatelské oprávnění **Přihlásit se místně**.

## Ověřování podle dat řízení přístupu ke vzdálené doméně

Pro operační systém Windows jsou všechny adresáře začínající textem `\\název_serveru` považovány za vzdálené adresáře. Příklad:

```
\\software\utilities\IBM
```

Pokud uživatelé chtějí přistupovat ke vzdálenému adresáři, musí zadat příslušná jména uživatelů v následujícím formátu:

```
JMÉNO_UŽIVATELE@NÁZEV_DOMÉNY
```

Když uživatelé používají vyhledávací aplikaci a konfigurují profil, který jim umožňuje prohledávat zabezpečené dokumenty ve vzdáleném systému, musí zadat jméno uživatele, které používají pro přístup ke vzdálenému systému Windows (například abcuser@win1.spolecnost.com).

Chcete-li vynutit ověřování aktuálních údajů pověření ve vzdálených počítačích, musí mít uživatelské účty následující oprávnění uživatelů systému Windows. Uživatelská oprávnění lze přiřadit pomocí nástrojů pro správu systému Windows: **Nástroje pro správu → Zásady zabezpečení domény**.

- Prolézací server a server se systémem Windows, který má být prolézán, musí být členy téže domény.
- Jméno uživatele, pod kterým je spuštěn prolézací server, musí mít oprávnění **Jednat jako součást operačního systému**. Toto oprávnění je konfigurováno pro uživatele, která je administrátorem podnikového vyhledávání, při instalaci produktu OmniFind Enterprise Edition.
- Uživatelé musí mít přiřazeno uživatelské oprávnění **Přihlásit jako dávkovou úlohu**.

### Související pojmy

“Ověřování aktuálního pověření během zpracování dotazu” na stránce 251

“Prolézací moduly souborového systému Windows” na stránce 107

“Bezpečné vyhledávání v důvěryhodných doménách systému Windows”

## Bezpečné vyhledávání v důvěryhodných doménách systému Windows

Chcete-li vynutit zabezpečení na úrovni dokumentů pro vzdálené systémy souborů systému Windows, systém podnikového vyhledávání podporuje verifikaci seznamu řízení přístupu (ACL) napříč důvěryhodnými doménami.

## Konfigurování prolézacího modulu

Chcete-li konfigurovat prolézací modul Souborový systém Windows k podpoře důvěryhodných domén, musíte zadat volby do nového konfiguračního souboru. Pro konfiguraci této možnosti v konzole pro správu podnikového vyhledávání neexistuje žádná podpora.

1. Chcete-li podporovat zabezpečení na úrovni dokumentů napříč důvěryhodnými doménami systému Windows, upravte následující soubor:

```
ES_NODE_ROOT/master_config/ID_relace/winfscrawler_ext.xml
```

**Tip:** Chcete-li určit ID relace pro prolézací modul Souborový systém Windows, který chcete konfigurovat, můžete monitorovat prolézací modul v konzole pro správu podnikového vyhledávání nebo použít příkaz **esadmin report collections**.

2. Zadejte název domény systému Windows a název NETBIOS služby Active Directory. Příklad:

```
<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. Změny se projeví po zastavení a restartu prolézacího modulu.

### Omezení

- Dokumenty nemohou obsahovat seznamy ACL z více domén systému Windows. Skupiny a uživatelé domén musí náležet do jedné domény Windows pro každou kolekci.
- Aby byla podporována verifikace přístupu ke vzdálenému systému souborů, musí být servery systému Windows spuštěny ve stejné doméně systému Windows nebo v důvěryhodných doménách systému Windows.
- Prolézací modul systému souborů systému Windows načítá název NETBIOS služby Active Directory přiřazené k serveru systému Windows, který má být prolézán, a používá název NETBIOS k filtrování seznamu ACL daného souboru. Služba Active Directory, ke které se server prolézacího modulu připojuje, důvěřuje jiné službě Active Directory, která definuje účty uživatelů a skupin.
- Uživatelský účet zadaný pro prolézací modul, který má být používán k přístupu ke vzdálenému serveru Windows, musí náležet do domény systému Windows, kde má být vynuceno a ověřeno řízení přístupu.
- Operační systém Windows umožňuje připojení k síťovým složkám na jednom souborovém serveru pouze jednomu účtu. Ostatní účty se v tuto dobu nemohou připojit ke stejnému souborovému serveru. Nelze tedy konfigurovat různé účty pro různé prolézací moduly k prolézání stejného serveru systému Windows ani v případě, že se prolézací moduly nachází v různých kolekcích.

#### Související pojmy

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows” na stránce 271

---

## Zákaz zabezpečení pro podnikové vyhledávání

Zabezpečení pro podnikové aplikace na aplikačním serveru WebSphere můžete zakázat. Pokud jste dříve nakonfigurovali ovládací prvky zabezpečení na úrovni dokumentu, můžete určit, že budou tyto ovládací prvky ignorovány. Nastavení zabezpečení také ovlivňuje způsob zobrazení sbalených výsledků ve výsledcích vyhledávání.

## Zákaz zabezpečení pro podnikové aplikace na aplikačním serveru WebSphere

Chcete-li určit, které aktivity podnikového vyhledávání vyžadují ověření uživatele, můžete na aplikačním serveru WebSphere zakázat globální zabezpečení pro jednotlivé podnikové aplikace.

### O této úloze

Instalační program produktu OmniFind Enterprise Edition implementuje v systému WebSphere Application Server tři podnikové aplikace:

- Aplikace ESAdmin obsahuje rozhraní konzoly pro správu podnikového vyhledávání.
- Aplikace ESSearchApplication obsahuje rozhraní ukázkové vyhledávací aplikace.
- Aplikace ESSearchServer zajišťuje veškerou vzdálenou komunikaci implementace rozhraní SIAPI podnikového vyhledávání a umožňuje rozhraním SIAPI komunikovat s vyhledávacími servery.

Standardně všechny tyto tři aplikace podporují globální zabezpečení systému WebSphere Application Server. Zjistí-li tyto aplikace, že je povoleno globální zabezpečení, začnou ověřovat veškeré přijaté požadavky.

V některých organizacích může být vhodné povolit nebo zakázat zabezpečení pro konkrétní podnikovou aplikaci. Můžete například ověřovat všechny uživatele, kteří požadují přístup ke konzole pro správu podnikového vyhledávání, ale neověřovat uživatele pracující s rozhraním SIAPI nebo ukázkovou vyhledávací aplikací.

### Postup

Chcete-li zakázat zabezpečení pro podnikovou aplikaci, postupujte takto:

1. Na vyhledávacím serveru spusíte konzolu pro správu produktu WebSphere Application Server.  
Konzolu pro správu lze otevřít následujícími způsoby:
  - Vyberte program z nabídky **Start** systému Windows.
  - V případě produktu WebSphere Application Server verze 5 otevřete okno webového prohlížeče a přejděte na stránku `http://název_hostitele:port/admin`, kde *název\_hostitele* je název hostitele vyhledávacího serveru a *port* je číslo portu konzoly pro správu produktu WebSphere Application Server. Konzola pro správu obvykle používá port 9090.
  - V případě produktu WebSphere Application Server verze 6 otevřete okno webového prohlížeče a přejděte na stránku `http://název_hostitele:port/ibm/console`, kde *název\_hostitele* je název hostitele vyhledávacího serveru a *port* je číslo portu konzoly pro správu produktu WebSphere Application Server. Konzola pro správu obvykle používá port 9060.
2. Když se zobrazí výzva k zadání jména uživatele a hesla, zadejte jméno administrátora a heslo použité při povolení globálního zabezpečení v produktu WebSphere Application Server.
3. Po přihlášení ke konzole pro správu klepněte na tlačítko **Applications** a poté na položku **Enterprise Applications**.
4. Zaškrtněte políčko u názvu podnikové aplikace, pro kterou chcete zakázat zabezpečení.
5. Posuňte obsah okna dolů a klepněte na odkaz **Map security roles to users/groups**.
6. Najděte roli **AllAuthenticated** a zaškrtněte políčko ve sloupci **Everyone?** .
7. Klepněte na tlačítko **OK**.
8. Uložte změny klepnutím na odkaz **Save**.

9. Pokud používáte nástroj WebSphere Network Deployment, zaškrtněte políčko **Synchronize changes with Nodes**.
10. Klepněte na tlačítko **Save**.
11. Zastavte aplikaci ESSearchServer a znovu ji spusťte.

#### **System AIX, Linux nebo Solaris**

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

#### **System Windows**

```
stopServer ESSearchServer  
startServer ESSearchServer
```

Tyto skripty jsou umístěny v adresáři KOŘEN\_INSTALACE\_WAS/AppServer/bin:

- Pro produkt WebSphere Application Server verze 5 je výchozí instalační cesta v systémech AIX /usr/WebSphere, v systémech Linux nebo Solaris /opt/WebSphere a v systémech Windows C:\Program Files\WebSphere.
- Pro produkt WebSphere Application Server verze 6 je výchozí instalační cesta v systémech AIX /usr/IBM/WebSphere, v systémech Linux nebo Solaris /opt/IBM/WebSphere a v systémech Windows C:\Program Files\IBM\WebSphere.

## **Zakázání zabezpečení na úrovni dokumentu**

Uživatelům lze povolit prohledávání kolekce bez ohledu na to, zda je s dokumenty v indexu asociováno řízení přístupu. Pro prolézací moduly podporující ověřování aktuálních údajů pověření lze také uživatelům povolit prohledávání kolekce bez ověřování řízení přístupu během zpracování dotazu.

### **Než začnete**

Chcete-li pro všechny dokumenty v kolekci povolit nebo zakázat zabezpečení na úrovni dokumentu, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### **Omezení**

Volby zabezpečení na úrovni dokumentu lze určit pouze v případě, že bylo pro kolekci při jejím vytváření povoleno zabezpečení.

### **O této úloze**

Prolézací moduly můžete nakonfigurovat tak, aby přiřazovaly prvky zabezpečení k dokumentům při prolézání. Vyhledávací aplikace mohou používat tyto prvky, které jsou uloženy v indexu, k vynucení řízení přístupu při prohledávání kolekce uživateli. Pro některé prolézací moduly lze rovněž určit, že chcete při odesílání dotazů uživateli ověřovat aktuální nastavení řízení přístupu asociované s dokumenty v příslušných nativních úložištích.

Chcete-li tato omezení zabezpečení odebrat, můžete určit, že vyhledávací servery mají ignorovat veškeré prvky zabezpečení předané s dotazem. Uživatelům lze také povolit dotazy na dokumenty bez nutnosti porovnávat jejich údaje pověření s aktuálním nastavením řízení přístupu.

Může se stát, že budete chtít zabezpečení na úrovni dokumentu dočasně zakázat, pokud testujete novou kolekci nebo pokud potřebujete vyřešit problém s vyhledávací aplikací.

## Postup

Zakázání řízení přístupu na úrovni dokumentu:

1. Zvolte úpravy kolekce, vyberte stránku Obecné a klepněte na volbu **Povolit nebo zakázat zabezpečení na úrovni dokumentu**.
2. Pokud nechcete, aby byly při zadávání uživatelských dotazů na kolekci použity prvky zabezpečení, které prolézací moduly asociovaly s dokumenty, zaškrtněte na stránce Zabezpečení na úrovni dokumentu pro všechny dokumenty políčko **Ignorovat řízení přístupu na úrovni dokumentu v indexu**.

Prolézací moduly nadále přidávají do dokumentů prvky zabezpečení, ale vyhledávací servery je ignorují a umožňují uživatelům vyhledávat dokumenty, které byly dříve chráněné.

3. Pokud nechcete, aby bylo při odesílání dotazů uživateli ověřováno řízení přístupu, které je aktuálně asociováno s dokumenty v jejich nativních úložištích, zaškrtněte políčko **Neověřovat aktuální pověření během zpracování dotazu**. Toto zaškrťovací políčko je k dispozici pouze u dokumentů zpracovaných prolézacími moduly, které tuto možnost podporují.

I když zaškrtnete toto políčko, ostatní volby zabezpečení na úrovni dokumentů zůstávají v platnosti. Jestliže jste například při konfigurování prolézacího modulu zadali volby ukládání dat řízení přístupu do indexu, tato data řízení přístupu budou používána i nadále, pokud nezaškrtnete rovněž políčko **Ignorovat řízení přístupu na úrovni dokumentu v indexu**.

## Zákaz zabezpečení pro sbalené výsledky vyhledávání

Je-li povoleno zabezpečení kolekcí, nelze sbalit výsledky hledání ze stejného webu, pokud neurčíte, že nechcete ověřovat údaje pověření uživatele během zpracování dotazu.

### Než začnete

Chcete-li pro všechny dokumenty v kolekci povolit nebo zakázat ověřování aktuálních údajů pověření, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Omezení

Volby zabezpečení na úrovni dokumentu lze určit pouze v případě, že bylo pro kolekci při jejím vytváření povoleno zabezpečení.

### O této úloze

Při ověřování aktuálních údajů pověření je zdroj každého dokumentu prozkoumán a v případě potřeby směřován k ověření. Chcete-li, aby byly dokumenty se stejnou předponou URI nebo dokumenty patřící ke dříve nakonfigurované sbalené skupině URI sbaleny ve výsledcích hledání, je třeba zakázat ověřování aktuálních údajů pověření.

## Postup

Zakázání ověřování aktuálních údajů pověření za účelem sbalení dokumentů ve výsledcích vyhledávání:

1. Zvolte úpravy kolekce, vyberte stránku Obecné a klepněte na volbu **Povolit nebo zakázat zabezpečení na úrovni dokumentu**.
2. Na stránce Zabezpečení na úrovni dokumentu pro všechny dokumenty zaškrtněte políčko **Neověřovat aktuální pověření během zpracování dotazu**.

3. V části monitorování kolekce vyberte stránku Vyhledávání a zastavte a restartujte procesy vyhledávacího serveru.

Pokud zadávají uživatelé do kolekce dotazy, budou dokumenty se stejnou předponou URI nebo dokumenty patřící k webům nakonfigurovaným jako sbalené ve výsledcích hledání sbaleny. V ukázkové vyhledávací aplikaci mohou uživatelé zobrazit sbalené výsledky klepnutím na odkaz **Další výsledky ze stejného zdroje**.





---

## Spouštění a zastavování systému podnikového vyhledávání

Po vytvoření kolekce je nutné spustit servery pro prolézání, analýzu, a indexování dat (vyhledávací servery se spouštějí automaticky). Provedete-li v kolekci nějaké změny, zastavte tyto servery a znovu je spusťte.

Většina serverů podnikového vyhledávání může pracovat nepřetržitě nebo podle zadaných plánů. Můžete například nastavit plány sestavení hlavního indexu a rozdílových indexů. Po spuštění systému podnikového vyhledávání je procesy serveru obvykle třeba zastavit a znovu spustit jen tehdy, změníte-li nastavení konfigurace (například po aktualizaci kategorií nebo po zvětšení vyhledávací mezipaměti).

Jestliže změníte obsah kolekce nebo pravidla určující, jakým způsobem mají prolézací moduly shromažďovat data ze zdrojů v podniku, změny vstoupí v platnost až po zastavení a novém spuštění prolézacích modulů. Pokud nezměníte pravidla prolézání, webový prolézací modul pracuje nepřetržitě a ostatní prolézací moduly se spouštějí podle zadaných plánů.

Chcete-li zlepšit dostupnost vyhledávacích serverů v situaci, kdy není k dispozici indexový server ani konzola pro správu, můžete zadat příkazy, které spustí vyhledávací servery pro kolekci v samostatném režimu.

---

## Spuštění systému podnikového vyhledávání

Chcete-li uživatelům povolit vyhledávání v kolekci, je nutné spustit systémové procesy a poté spustit servery, které kolekci prolézají, analyzují a indexují (vyhledávací servery se spouštějí automaticky).

### Než začnete

Konfigurujte zdroje dat, které chcete prolézat, a určete volby pro analýzu, indexování a prohledávání těchto dat. Chcete-li například uživatelům umožnit zobrazení podrobných informací o kategoriích ve výsledcích vyhledávání, konfigurujte kategorie před spuštěním analytického modulu.

Chcete-li spouštět servery podnikového vyhledávání, musíte mít přiřazenu roli administrátora podnikového vyhledávání, administrátora kolekce pro danou kolekci nebo operátora s oprávněním spravovat danou kolekci.

Servery podnikového vyhledávání je nutné pro kolekci spouštět ve správném pořadí. Nejprve je například nutné spustit prolézací modul a prolézt data a teprve poté lze prolezená data indexovat.

### Omezení

Ke spuštění systému podnikového vyhledávání je nutné použít uživatelský účet, který lze ověřit pomocí lokálních ověřovacích funkcí. Pokud se spustíte pod účtem AFS (Andrew File System), dojde k chybám.

### Postup

Při spuštění systému podnikového vyhledávání postupujte takto:

1. Pokud používáte podnikové vyhledávání v konfiguraci se dvěma nebo čtyřmi servery, přihlaste se jako administrátor podnikového vyhledávání a na všech serverech spusťte službu CCL (Common Communication Layer):

**Systém AIX, Linux, nebo Solaris**




```
startccl.sh -bg
```

**Příkazový řádek systému Windows**

```
startccl
```

**Nástroj pro správu Služby systému Windows**



Chcete-li spustit vrstvu CCL na pozadí, postupujte takto:

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.
2. Spusťte komponenty systému podnikového vyhledávání:
    - a. Přihlaste se k libovolnému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání.
    - b. Spusťte všechny komponenty systému: `esadmin system startall`  
Tento příkaz spustí následující procesy a aplikace:
      - webový server (v konfiguraci s více servery je webový server spuštěn na obou vyhledávacích serverech),
      - aplikace ESSearchServer a ESAdmin na aplikačním serveru WebSphere (v konfiguraci s více servery jsou tyto aplikace spuštěny na obou vyhledávacích serverech),
      - relaci ESAdmin na indexovém serveru,
      - službu CCL v počítači, z něž byl příkaz spuštěn, pokud tato služba již neběží,
      - databázový síťový server podnikového vyhledávání,
      - Informační centrum podnikového vyhledávání (v konfiguraci s více servery je Informační centrum spuštěno na obou vyhledávacích serverech).
  3. Spusťte konzolu pro správu podnikového vyhledávání a přihlaste se jako administrátor podnikového vyhledávání. Používáte-li administrativní role, můžete se přihlásit jako administrátor kolekce nebo operátor s oprávněním pro kolekci, kterou chcete spustit.
  4. V pohledu Kolekce vyhledejte kolekci, kterou chcete spravovat, a klepněte na ikonu  **Monitor**.
  5. Na stránce Prolézání klepněte pro každý prolézací modul, který chcete spustit, na volbu  **Spustit**.
    - Pokud spustíte webový prolézací modul, bude prolézání dat zahájeno ihned. Tyto typy prolézacích modulů souvisle a opakovaně prolézají webové dokumenty.
    - Pokud spustíte jeden z ostatních typů prolézacích modulů, bude zahájena relace prolézacího modulu. Prolézací modul začne prolézat data v naplánovanou dobu. Pokud jste spuštění prolézacího modulu nenaplánovali nebo chcete spustit prolézací modul dříve, monitorujte prolézací modul a klepněte na ikonu pro spuštění všech zdrojů dat, které chcete prolézat.  
Jakmile je prolézací modul spuštěn, můžete jej nechat pracovat bez přerušení. Pokud jste prolézací modul naplánovali, bude znovu spuštěn v naplánované dny a časy.
  6. Po prolezení dat otevřete stránku Analýza a klepnutím na volbu  **Spustit** spusťte modul analýzy.  
Modul analýzy můžete nechat spuštěný bez přerušení. Obvykle jej není třeba zastavovat, pokud nechcete změnit způsob analýzy dat (například aktualizováním kategorií nebo mapování polí XML).

7. Volitelné: Chcete-li vynutit spuštění procesů indexování a nečekat na zahájení indexování v naplánovanou dobu, otevřete stránku Indexovat a v oblasti **Hlavní** klepněte na volbu

 **Spustit.**

Procesy indexování můžete nechat spuštěné bez přerušení. Index bude sestaven v naplánovaných dnech a časech.

**Tip:** Vyhledávací servery se spouštějí automaticky a můžete je nechat běžet trvale. Obvykle je není třeba zastavovat, pokud nechcete provést změnu nastavení mezipaměti vyhledávání nebo shrnutí dokumentů. Chcete-li restartovat vyhledávací servery, otevřete stránku Vyhledat, klepněte na tlačítko  **Zastavit** a poté na tlačítko  **Spustit**.

Chcete-li zlepšit dostupnost vyhledávacích serverů v situaci, kdy konzola pro správu není k dispozici, můžete zadat příkazy, které spustí vyhledávací servery pro kolekci v samostatném režimu. Pokud není spuštěn indexový server, není konzola pro správu k dispozici.

**Související pojmy**

“Administrativní role” na stránce 246

**Související úlohy**

“Přihlášení ke konzole pro správu” na stránce 16

“Řízení spouštěných a zastavovaných komponent” na stránce 282

“Správa vyhledávacích serverů v samostatném režimu” na stránce 283

---

## Zastavení systému podnikového vyhledávání

Server podnikového vyhledávání může být nutné zastavit a znovu spustit, pokud provádíte změny v jeho konfiguraci nebo pokud potřebujete řešit problémy.

### Než začnete

Chcete-li zastavovat servery podnikového vyhledávání, musíte mít přiřazenu roli administrátora podnikového vyhledávání, administrátora kolekce pro danou kolekci nebo operátora s oprávněním spravovat danou kolekci.

### O této úloze

Servery podnikového vyhledávání můžete zastavovat nezávisle na ostatních serverech. Pokud například zastavíte a znovu spustíte prolézací modul, abyste se projevily změny provedené v jeho konfiguraci, není nutné zastavovat a znovu spouštět vyhledávací servery.

### Postup

Zastavování serverů podnikového vyhledávání:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete spravovat, a klepněte na ikonu

 **Monitor.**

2. Na stránce Prolézání vyhledejte prolézací modul, který chcete spravovat, a zastavte či pozastavte jej.

Pokud změníte prolézací prostor nebo vlastnosti prolézacího modulu, zastavte a znovu spusťte prolézací modul, aby se změny projevily. Pokud změníte prolézací prostor a chcete provést změny u dokumentů, které již byly indexovány, je nutné dokumenty také znovu prolézt.

**Tip:** Zpráva o vypršení časového limitu požadované operace se může zobrazit i v případě, že proces stále ještě běží na pozadí. Chcete-li zjistit, zda byla úloha dokončena,

klepněte na tlačítko **Aktualizovat** v konzole správy (nepoužívejte tlačítko **Aktualizovat** webového prohlížeče). Proces je dokončen, jestliže stavová ikona prolézacího modulu indikuje, že je modul zastaven.

3. Klepnutím na volbu  **Zastavit** na stránce Analýza zastavte modul analýzy.  
Změňte-li pravidla pro analýzu dat, projeví se změny po zastavení a opětném spuštění modulu analýzy. Změny se projeví pouze u nově prolézaných dokumentů. Chcete-li změny provést u dokumentů, které jsou již v indexu obsaženy, je nutné zahájit úplné prolézání, aby byly všechny tyto dokumenty znovu prolezeny a bylo je možné znovu analyzovat a indexovat.
4. Chcete-li zastavit index, jehož aktualizace nebo reorganizace probíhá, klepněte na stránce Indexovat na volbu  **Zastavit**.  
Sestavování indexu lze také zastavit při monitorování fronty indexů. Tuto akci můžete provést klepnutím na volbu **Systém** na panelu nástrojů, otevřením stránky Indexovat a klepnutím na volbu  **Zastavit** pro index, jehož sestavení chcete zastavit.
5. Klepnutím na volbu  **Zastavit** na stránce Vyhledávání zastavte vyhledávací servery. Vyhledávací servery je obvykle nutné zastavit a restartovat pouze v případě, že jste změnil nastavení mezí paměti vyhledávání nebo souhrnů dokumentů.
6. Chcete-li zastavit systém podnikového vyhledávání namísto jednotlivých serverů, postupujte takto:
  - a. Přihlašte se k libovolnému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání.
  - b. Zastavte všechny komponenty systému: `esadmin system stopall`  
Tento příkaz zastaví následující procesy a aplikace:
    - webový server (v konfiguraci s více servery je webový server zastaven na obou vyhledávacích serverech),
    - aplikace `ESSearchServer` a `ESAdmin` na aplikačním serveru `WebSphere` (v konfiguraci s více servery jsou tyto aplikace zastaveny na obou vyhledávacích serverech),
    - relaci `ESAdmin` na indexovém serveru,
    - službu `CCL` (`Common Communication Layer`) podnikového vyhledávání v počítači, z nějž byl příkaz spuštěn,
    - databázový síťový server podnikového vyhledávání,
    - Informační centrum podnikového vyhledávání (v konfiguraci s více servery je Informační centrum zastaveno na obou vyhledávacích serverech).

#### Související úlohy

“Přihlášení ke konzole pro správu” na stránce 16

“Řízení spouštěných a zastavovaných komponent”

“Správa vyhledávacích serverů v samostatném režimu” na stránce 283

---

## Řízení spouštěných a zastavovaných komponent

K řízení, které komponenty budou spuštěny či zastaveny, slouží příkazy `esadmin system startall` a `esadmin system stopall`.

#### O této úloze

Soubor `ES_INSTALL_ROOT/default_config/AutoRunComponents.properties` obsahuje seznam komponent podnikového vyhledávání, které lze spustit nebo zastavit pomocí příkazů `esadmin system startall` a `esadmin system stopall`. Ve výchozím nastavení lze pomocí uvedených příkazů spustit a zastavit všechny komponenty obsažené v seznamu.

Chcete-li znemožnit spouštění nebo zastavování některých komponent, upravte tento soubor vlastností.

### Postup

Postup při určení, které komponenty mají být spuštěny nebo zastaveny při spuštění nebo zastavení systému podnikového vyhledávání:

1. Přihlaste se jako administrátor podnikového vyhledávání k serveru, na kterém budete spouštět příkaz **esadmin system startall** nebo **esadmin system stopall**.
2. Upravte soubor `ES_INSTALL_ROOT/default_config/AutoRunComponents.properties`.
3. Chcete-li znemožnit spouštění některé komponenty, přidejte pro ni pole `Component.startable.ID_komponenty=false`, kde parametr `ID_komponenty` označuje komponentu, která nemá být spouštěna.
4. Chcete-li znemožnit zastavování některé komponenty, přidejte pro ni pole `Component.stopable.ID_komponenty=false`, kde parametr `ID_komponenty` označuje komponentu, která nemá být zastavována.
5. Uložte soubor a zavřete jej.

Při příštím použití příkazů **esadmin system startall** a **esadmin system stopall** již nebude uvedená komponenta spuštěna či zastavena, v souladu se změnami provedenými v souboru vlastností.

**Příklad:** V následujícím příkladu bude server HTTP spuštěn příkazem **esadmin system startall** (výchozí nastavení), avšak nebude zastaven příkazem **esadmin system stopall** (to je určeno zvýrazněným řádkem v příkladu).

```
#####  
# Podrobnosti týkající se komponenty 3  
#####  
Component.name.3=Server HTTP IBM  
Component.impl.class.3=com.ibm.es.control.util.component.impl.HTTPControlImpl  
Component.nodes.3=search  
# Ve výchozím nastavení se spouštějí všechny komponenty  
Component.stopable.3=false
```

#### Související úlohy

“Spuštění systému podnikového vyhledávání” na stránce 279

“Zastavení systému podnikového vyhledávání” na stránce 281

---

## Správa vyhledávacích serverů v samostatném režimu

Chcete-li zajistit vysokou dostupnost vyhledávacích serverů, můžete spustit vyhledávací servery pro jednotlivé kolekce i v případě, že není spuštěn indexový server.

### Omezení

Chcete-li zastavit nebo spustit vyhledávací servery, musíte být přihlášení jako administrátor podnikového vyhledávání.

Funkce spuštění a zastavení vyhledávacích serverů v samostatném režimu není k dispozici z konzoly pro správu podnikového vyhledávání. Pokud není spuštěn indexový server, není konzola pro správu přístupná.

Před spuštěním či zastavením vyhledávacích serverů v samostatném režimu zkontrolujte, že byly prolézací modul, analyzátor, index a vyhledávací servery pro kolekci již alespoň jednou spuštěny. To je nezbytné pro zajištění synchronizace vyžadovaných souborů na vyhledávacích serverech.

Je-li pro kolekci povoleno zabezpečení na úrovni dokumentů, zkontrolujte také, zda je spuštěn prolézací server. To je nezbytné k zajištění řízení zabezpečení na úrovni dokumentů. Není-li prolézací server spuštěn, budou ve výsledcích vyhledávání vráceny pouze dokumenty, které nevyžadují ověřování.

## O této úloze

Není-li indexový server dostupný, můžete přesto uživatelům umožnit další prohledávání systému tím, že spustíte vyhledávací servery v samostatném režimu. V instalaci s více servery můžete spustit příkaz pro spuštění nebo zastavení vyhledávacích serverů z kteréhokoli serveru podnikového vyhledávání. Tyto příkazy se pokusí spustit nebo zastavit vyhledávací servery pro danou kolekci na všech dostupných vyhledávacích serverech.

## Postup

Postup při spuštění nebo zastavení vyhledávacích serverů v samostatném režimu:

1. Chcete-li spustit vyhledávací servery pro kolekci v situaci, kdy není spuštěn indexový server, postupujte takto:

- a. Přihlaste se k prolézacímu serveru a k vyhledávacím serverům jako administrátor podnikového vyhledávání a spusíte službu CCL (Common Communication Layer):

### Systém AIX, Linux, nebo Solaris

```
startccl.sh -bg
```

### Příkazový řádek systému Windows

```
startccl
```

### Nástroj pro správu Služby systému Windows

Chcete-li spustit vrstvu CCL na pozadí, postupujte takto:

- 1) Spusíte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
- 2) Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.

- b. Na vyhledávacích serverech spusíte skript **startServer**, který je umístěn v adresáři **WAS\_INSTALL\_ROOT/AppServer/bin**, a spusíte aplikaci **ESSearchServer** na serveru WebSphere Application Server:

### Systém AIX, Linux nebo Solaris

```
./startServer.sh ESSearchServer
```

### Windows

```
startServer ESSearchServer
```

- c. Zadejte následující příkaz, ve kterém parametr *ID\_kolekce* určuje kolekci, která vlastní spouštěné vyhledávací servery:

```
esadmin startSearch -cid ID_kolekce
```

2. Chcete-li zastavit vyhledávací servery pro kolekci v situaci, kdy není spuštěn indexový server, postupujte takto:

- a. Přihlaste se k libovolnému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání.

- b. Zadejte následující příkaz, ve kterém parametr *ID\_kolekce* určuje kolekci, která vlastní zastavované vyhledávací servery:

```
esadmin stopSearch -cid ID_kolekce
```

## Související úlohy

“Spuštění systému podnikového vyhledávání” na stránce 279

“Zastavení systému podnikového vyhledávání” na stránce 281

---

## Monitorování aktivity podnikového vyhledávání

Při monitorování aktivity systému a kolekcí můžete zobrazit stav různých procesů, sledovat příznaky potenciálních problémů nebo zvýšit výkon úpravou nastavení konfigurace.

Pomocí konzoly pro správu podnikového vyhledávání můžete monitorovat systém a upravit jeho provozní parametry podle potřeby. Pro každý základní okruh aktivit (prolézáni, analýza, indexování a hledání) lze zobrazit podrobnou statistiku. Statistické údaje zahrnují průměrnou dobu odezvy a informace o průběhu, například počet dokumentů, které byly prolezeny nebo indexovány v průběhu relace.

Většinu aktivit lze zastavit nebo spustit klepnutím na příslušnou ikonu. Díky těmto funkcím můžete aktivitu pozastavit, změnit její konfiguraci nebo vyřešit problém, a znovu zpracování spustit, když jste připraveni povolit pokračování aktivity.

### Související úlohy

“Spuštění systému podnikového vyhledávání” na stránce 279

“Zastavení systému podnikového vyhledávání” na stránce 281

---

## Odhad počtu dokumentů v kolekci

Při vytvoření nebo úpravě kolekce podnikového vyhledávání nastavujte odhad počtu dokumentů, které bude tato kolekce obsahovat. Systém vychází z této hodnoty při odhadování paměťových a diskových prostředků, které budou po kolekci zapotřebí, velikost kolekce však nebude omezena.

### Než začnete

Chcete-li změnit odhad velikosti kolekce, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

Dosáhne-li velikost kolekce odhadu, který jste zadali, systém nepřestane do indexu přidávat dokumenty. Pokud pro kolekci konfiguruje výstrahy a vyberete funkci upozorňování na překročení limitu počtu dokumentů v indexu, bude výchozí limit shodný se zadaným odhadem počtu dokumentů v kolekci. Systém tento odhad monitoruje v kombinaci se zadaným procentuálním prahem výstrahy a odešle e-mail, pokud se skutečný počet dokumentů v kolekci blíží k maximálnímu počtu.

### Postup

Chcete-li zadat odhad potenciální velikosti kolekce, postupujte takto:

1. Upravte kolekci, vyberte stránku **Obecné** a klepněte na volbu **Konfigurovat obecné volby**.
2. Do pole **Odhadovaný počet dokumentů** zadejte číslo odpovídající odhadované budoucí velikosti kolekce. Výchozí hodnota je 1 000 000 dokumentů.

---

## Monitorování kolekce


Můžete zobrazit obecné informace o stavu jednotlivých komponent v kolekci nebo vybrat volbu pro zobrazení podrobných informací o jednotlivých komponentách a identifikátorech URI.


### Než začnete

Kolekce mohou monitorovat všichni uživatelé s oprávněním pro administraci podnikového vyhledávání. Chcete-li spouštět a zastavovat komponenty nebo povolovat a zakazovat časové plány, musíte mít přiřazenu roli administrátora podnikového vyhledávání, administrátora kolekce pro danou kolekci nebo operátora kolekce.

### Postup

Monitorování kolekce:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**. Zobrazí se informace o aktuálním stavu každé z komponent kolekce.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce General, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.

2. Podrobné informace o identifikátoru URI získáte po klepnutí na volbu  **Podrobnosti identifikátoru URI**.

Můžete například chtít zjistit, zda je určitý identifikátor URI v indexu nebo zda byl index, ve kterém se nachází daný identifikátor URI, zkopírován na vyhledávací servery.

3. Chcete-li monitorovat jednotlivé komponenty a zobrazit podrobnou statistiku o aktivitě dané komponenty, klepněte na ikonu **Stav**.

---

## Zobrazení podrobností o identifikátoru URI

Můžete zobrazit podrobné informace o identifikátoru URI. Tyto informace zahrnují aktuální a historické informace o tom, jak je dokument představovaný daným identifikátorem URI prolézán, indexován a vyhledáván.

### Než začnete

Před odesláním požadavku na zobrazení sestavy identifikátoru URI nebo odesláním sestavy na e-mailovou adresu ověřte, zda je komponenta, ze které chcete obdržet informace, aktivní. Chcete-li například zobrazit podrobnosti o tom, jak je dokument prolézán, indexován a vyhledáván, zkontrolujte, zda je spuštěn webový prolézací modul, indexový server a vyhledávací servery. Chcete-li sledovat odstraněný dokument, ověřte, že jsou konfigurovány volby protokolování pro sledování dokumentů.

### O této úloze

Shromáždění informací o identifikátoru URI může nějakou dobu trvat. Můžete vybrat volbu zobrazení požadovaných informací a potom počkat, než se tyto informace zobrazí. Efektivnější volbou je odeslání sestavy na zadanou e-mailovou adresu.

Než budete moci sestavu přijmout, je nutné zajistit, aby byly pro podnikové vyhledávání konfigurovány informace o poštovním serveru. Tyto informace můžete zadat při konfigurování voleb e-mailu na stránce Žurnál v pohledu Systém.

Indexový server a vyhledávací servery mohou poskytovat informace o všech identifikátorech URI (například zda je identifikátor URI v indexu a zda byl zkopírován na vyhledávací





servery). Chcete-li zobrazit informace o prolézání dokumentu, je nutné zadat identifikátor URI dokumentu, který byl prolézán webovým prolézacím modulem.

## Postup

Zobrazení podrobností o identifikátoru URI:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce General, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.

2. Klepněte na volbu  **Podrobnosti identifikátoru URI**.
3. Na stránce Podrobnosti identifikátoru URI zadejte identifikátor URI, pro který chcete zobrazit informace.
4. Zaškrtněte políčka odpovídající typu informací, které chcete zobrazit:

### **Podrobnosti prolézacího modulu (dostupné pouze u webových prolézacích modulů)**

Toto políčko zaškrtněte, chcete-li zobrazit informace o prolézání dokumentu webovým prolézacím modulem a informace o jeho aktuálním stavu v prolézaném prostoru.

### **Podrobnosti indexu**

Toto políčko zaškrtněte, chcete-li zobrazit, zda byl dokument indexován a zkopírován na vyhledávací servery.

### **Podrobnosti vyhledávání**

Toto políčko zaškrtněte, chcete-li zobrazit informace o tom, jak lze dokument vyhledávat a zda je dostupný pro vyhledávání.

### **Dokumenty odstraněné analyzátozem**

Toto políčko zaškrtněte, chcete-li zjistit, zda byl dokument odstraněn ze systému podnikového vyhledávání během analýzy, a příčinu jeho odstranění, pokud se tak stalo.

### **Dokumenty odstraněné z indexu**

Toto políčko zaškrtněte, chcete-li zjistit, zda byl dokument odstraněn ze systému podnikového vyhledávání během indexování nebo analýzy, a příčinu jeho odstranění, pokud se tak stalo.

5. Chcete-li počkat na zobrazení sestavy, klepněte na volbu **Zobrazit sestavu**.
6. Chcete-li sestavu odeslat na e-mailovou adresu, abyste ji mohli zobrazit později, klepněte na volbu **Odeslat sestavu**.
  - a. Na stránce Zaslát podrobnou sestavu o identifikátoru URI zadejte do pole **E-mailová adresa pro zaslání oznámení** e-mailovou adresu pro přijetí sestavy.
  - b. Klepněte na volbu **Odeslat sestavu**.

### **Související pojmy**

“Sledování dokumentů” na stránce 302

### **Související úlohy**

“Zobrazení sestav odstraněných dokumentů” na stránce 303

### **Související odkazy**

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 113

---

## Monitorování prolézacích modulů


Můžete zobrazit obecné informace o stavu jednotlivých prolézacích modulů v kolekci nebo vybrat volby pro zobrazení podrobných informací o aktivitě prolézacího modulu.


### Než začnete


Pokud je monitorování kolekcí omezeno administrativní rolí, lze zobrazit statistiku prolézacího modulu, ale nelze změnit jeho chování (například spustit nebo zastavit prolézací modul).

### Postup


Monitorování prolézacího modulu:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.
2. Otevřete stránku Prolézání.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce Prolézání, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování prolézacích modulů.

3. Pokud je prolézací modul spuštěn nebo pozastaven a chcete zobrazit podrobné informace o jeho stavu, klepněte na volbu  **Podrobnosti**. Zobrazené typy statistiky se liší podle typu prolézacího modulu.

Pokud vaše administrativní role umožňuje správu procesů pro kolekci, můžete zobrazit podrobnosti o aktivitě prolézacího modulu a přitom modul spustit, zastavit nebo pozastavit. Pokud lze prolézací modul naplánovat, můžete také povolit nebo zakázat časový plán prolézání.

4. Pokud je prolézací modul zastaven nebo pozastaven a chcete spustit jeho relaci, klepněte na volbu  **Spustit** nebo **Pokračovat**.

#### Webové prolézací moduly:

Pokud je prolézací modul zastaven, zahájí znovu prolézání a bude prolézat celý prolézáný prostor. Pokud je prolézací modul pozastaven, bude prolézání obnoveno od začátku cíle, ve kterém bylo pozastaveno.

Chcete-li vynutit okamžité zahájení úplného prolézání prolézacím modulem, klepněte na ikonu **Podrobnosti** a poté na ikonu **Zahájit úplné opětné prolézání**. Prolézací modul zahájí prolézání celého prolézaného prostoru včetně stránek, které se od posledního prolézání nezměnily. Pro opětné prolezení všech dokumentů se můžete rozhodnout například v případě, že jste změnili pravidla pro analýzu dokumentů a chcete tato pravidla použít pro dokumenty, které byly indexovány dříve.

#### Ostatní typy prolézacích modulů:

Pokud byl prolézací modul zastaven, bude prolézání obnoveno v naplánovanou dobu. Při prvním zpracování zdroje dat provádí prolézací modul operaci úplného prolézání. Při opakování naplánovaného prolézání bude modul prolézat buď všechny aktualizace zdroje dat (přidání, odstranění a úpravy dokumentů), nebo pouze přidané a upravené dokumenty. Typ prolézání je konfigurován v časovém plánu prolézacího modulu.

Pokud jste spuštění prolézacího modulu nenaplánovali nebo chcete zahájit prolézání dříve, klepněte na ikonu **Podrobnosti**. V oblasti podrobností pro prolézáný prostor poté klepněte na ikonu typu prolézání, které chcete zahájit: úplné prolézání, všechny aktualizace nebo pouze nové a upravené dokumenty. Pro každý zdroj dat, který chcete prolézat (například server, databázi nebo podsložku) musíte klepnout na příslušnou ikonu zahájení.

5. Pokud je prolézací modul spuštěn a chcete jej zastavit, klepněte na volbu **Zastavit** nebo **Pozastavit**. Prolézací modul přestane prolézat data až do dalšího restartu nebo pokračování.

Pokud má prolézací modul pokračovat v prolézání, bude prolézání obnoveno od začátku cíle, ve kterém bylo pozastaveno. Například prolézací modul DB2 pokračuje v prolézání u prvního řádku v tabulce, která byla prolézána, když jste prolézací modul pozastavili.

---

## Zobrazení podrobných informací o aktivitě webového prolézacího modulu

Zobrazíte-li podrobné informace o aktivitě webového prolézacího modulu, můžete odhadnout celkový výkon a podle potřeby upravit vlastnosti webového prolézacího modulu a definice prolézaného prostoru.

### Než začnete

Aktivita prolézacího modulu mohou monitorovat všichni uživatelé s oprávněním pro administraci podnikového vyhledávání. Chcete-li spouštět nebo zastavovat prolézací moduly, musíte mít přiřazenu roli administrátora podnikového vyhledávání, administrátora kolekce pro danou kolekci nebo operátora kolekce.

### Postup

Chcete-li zobrazit podrobné informace o aktivitě webového prolézacího modulu, postupujte takto:

1. V pohledu Kolekce vyhledejte kolekci vlastníci webový prolézací modul, který chcete prolézat, a klepněte na volbu **Monitor**.
2. Otevřete stránku Prolézání.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce Prolézání, můžete klepnutím na volbu **Monitor** přejít do pohledu pro monitorování prolézacích modulů.

3. Pokud webový prolézací modul, který chcete monitorovat, běží nebo je pozastaven, klepněte na ikonu **Podrobnosti**.
4. Na stránce s podrobnými údaji o webovém prolézacím modulu si prohlédněte nebo vyberte následující volby, které vám poskytnou podrobnou statistiku současné i dřívější aktivity prolézacího modulu.
  - Chcete-li zjistit, kolik podprocesů v současné době aktivně prolézá webové servery a kolik jich je ve stavu nečinnosti, klepněte na tlačítko **Podrobnosti o podprocesech**.
  - Klepnutím na tlačítko **Aktivní weby** zobrazíte informace o webových serverech, které prolézací modul aktivně zpracovává.
  - Klepněte na tlačítko **Naposledy prolézané adresy URL**. Tyto informace ukazují, které adresy prolézací modul naposledy prolézal. Pokud se položky v seznamu při aktualizaci zobrazení nemění, znamená to, že prolézání neprobíhá.
  - Klepnutím na tlačítko **Historie prolézacího modulu** zobrazíte zprávy o dřívější činnosti prolézacího modulu.
  - V oblasti **Stav adresy URL** zadejte adresu URL, o které chcete zobrazit informace.
    - a. Informace o adrese URL zobrazíte klepnutím na volbu **Podrobnosti adresy URL**. Podrobnosti adresy URL je možné požadovat pouze u adres URL, které byly dříve prolézány.
    - b. Klepnutím na volbu **Podrobnosti webu** zadáte informace, které chcete zahrnout v sestavě pro webové stránky dané adresy URL. Podrobnosti webu můžete požadovat u dříve prolézaných i neprolézaných webových stránek.

Pomocí této volby můžete například zjistit, zda se adresa URL nachází v prolézaném prostoru, zda již byla prolezena nebo zatím pouze nalezena, na kterou dobu je naplánováno její další prolézáání. Dále zde najdete informace o posledním pokusu o prolezení webového serveru. Můžete také požádat o zobrazení obsahu souboru robots.txt pro dané webové stránky a snadněji tak určit, proč nejsou prolézány.

## Podrobnosti o podprocesech webového prolézacího modulu

Monitorováním webového prolézacího modulu můžete získat informace o počtu podprocesů, které aktivně prolézají webové servery, a o tom, kolik z nich je ve stavu nečinnosti.

Zobrazíte-li při monitorování kolekce podrobné informace o webovém prolézacím modulu, můžete si prohlédnout údaje o stavu podprocesů prolézacího modulu. Nejčastěji se setkáte s následujícími stavy:

### Čekání

Informuje o tom, že podproces nemá informace o žádné adrese URL, kterou má prolézat. Tento stav vzniká, když podproces dokončí operaci prolézáání a prolézací modul dostatečně rychle nenalezne další adresy URL určené k prolézáání. Rychlost předávání informací o adresách URL může snižovat například příliš vysoká hodnota vlastnosti prolézacího modulu, která určuje, jak dlouho musí prolézací modul čekat, než smí načíst další stránku ze stejného webu.

### Načítání

Informuje o tom, že podproces stahuje stránku z webu.

### Dokončeno

Informuje o tom, že podproces odesílá prolezené stránky dalším částem prolézacího modulu, ale zatím není připraven k prolézáání dalších adres URL.

### Pozastaveno

Informuje o tom, že prolézací modul je pozastaven.

V ideálním případě všechny podprocesy neustále načítají stránky. Pokud se podprocesy často nacházejí ve stavu Dokončeno, může to signalizovat problém s propustností databáze.

Jestliže se podprocesy často nacházejí ve stavu čekání, zkontrolujte hodnotu v poli **Maximální počet aktivních hostitelů** ve vlastnostech prolézacího modulu. Pokud je tato hodnota nízká, znamená to, že prolézaný prostor neobsahuje dostatečný počet webů k tomu, aby bylo možné udržovat podprocesy neustále v činnosti, nebo neexistuje dostatečný počet identifikátorů URL vhodných k prolézáání. K podmínkám, které mohou způsobit nízkou aktivitu, patří selhání při vyhledávání v tabulkách DNS a výpadky vyhledávacích robotů.

## Aktivní weby webového prolézacího modulu

Monitorováním webového prolézacího modulu můžete získat informace o webových serverech, které prolézací modul aktivně zpracovává.

Zobrazíte-li při monitorování kolekce podrobné informace o webovém prolézacím modulu, můžete si prohlédnout statistické údaje o aktivních webech. K dispozici jsou následující statistiky:

- počet adres URL, které prolézací modul v daném okamžiku přesunul z interní databáze do paměti pro prolézáání,
- počet adres URL, o jejichž zpracování se prolézací modul dosud pokusil,
- zbývající doba do deaktivace webu a jeho odebrání z paměti v tomto cyklu prolézacího modulu,
- doba, po kterou byl web dosud načten v paměti.

Tyto informace se průběžně mění s tím, jak prolézací modul prochází pravidla prolézání, která pro něj byla nakonfigurována. V ideálním případě je počet aktivovaných adres URL blízký hodnotě nastavené v poli **Maximální počet aktivních hostitelů** ve vlastnostech paměti prolézacího modulu.

Pokud se počet aktivovaných adres URL blíží nule, znamená to, že prolézací modul nenachází použitelné adresy URL. K příčinám této nízké aktivity patří selhání při vyhledávání v tabulkách DNS, problémy se síťovým připojením, chyby databáze a problémy s definicí prolézaného prostoru. Příklad:

- Pokud je do paměti načten velký počet webů na dlouhou dobu a přitom bylo zpracováno jen několik adres URL, zkontrolujte, zda nedochází k problémům s připojením k síti.
- Pokud je v seznamu uveden malý počet webů, hledejte problémy s definicí prolézaného prostoru nebo s vyhledáváním ve službě DNS.
- Pokud jsou weby prolézány s přijatelnou frekvencí, ale v paměti zůstává velký počet neprolezených adres URL, otevřete vlastnosti paměti prolézacího modulu pro úpravy a upravte hodnotu časového limitu v poli **Doba, po kterou může každý hostitel zůstat aktivní** tak, aby weby zůstávaly v paměti déle.

## Frekvence prolézání webového prolézacího modulu

Monitorováním webového prolézacího modulu můžete získat informace o rychlosti, s jakou tento prolézací modul stahuje stránky z webových serverů.

Zobrazíte-li při monitorování kolekce podrobné informace o webovém prolézacím modulu, můžete si prohlédnout statistické údaje o rychlosti stahování dat prolézacím modulem (frekvenci prolézání). Dále můžete zobrazit statistické údaje o počtu adres URL zpracovaných prolézacím modulem od začátku aktuální relace.

Frekvence prolézání je definována jako počet prolezených stránek za sekundu. Tato hodnota závisí na několika vlastnostech, které můžete pro webový prolézací modul konfigurovat:

- počet podprocesů prolézacího modulu,
- počet aktivních webů,
- doba, po kterou musí prolézací modul čekat, než může načíst další stránku z téhož webového serveru.

Připadá-li na každý podproces prolézacího modulu jeden aktivní web a musí-li prolézací modul čekat dvě sekundy, než může načíst další stránku ze stejného webového serveru, nemůže rychlost prolézání překročit jednu stránku na podproces za dvě sekundy. Používá-li prolézací modul například výchozí počet podprocesů (200), může při plné aktivitě všech 200 podprocesů prolézt 100 stránek za sekundu.

Pokud je počet aktivních webů dvojnásobný oproti počtu podprocesů prolézacího modulu a musí-li prolézací modul čekat dvě sekundy, než může načíst další stránku ze stejného webového serveru, může dosáhnout až jednoho přístupu na stránku za sekundu na každý podproces. V této situaci se však stává limitujícím faktorem rychlost stahování v síti a propustnost databáze. Známkou dobrého výkonu prolézacího modulu je vyrovnání frekvence prolézání s počtem podprocesů prolézacího modulu, počtem aktivních webů a čekací dobou prolézacího modulu.

Dalším parametrem, který je třeba sledovat při monitorování výkonu webového prolézacího modulu, je počet adres URL zpracovaných prolézacím modulem od začátku aktuální prolézací relace. Vydělíte-li tuto hodnotu celkovou dobou běhu prolézacího modulu, získáte průměrnou dlouhodobou propustnost. Pokud se tato hodnota nezvyšuje, znamená to, že prolézací modul

již dokončil prolézání nebo že nemůže pokračovat. Práce prolézacího modulu může být blokována například chybami síťového připojení, databázovými chybami a výpadky při vyhledávání ve službě DNS.

## Vytváření sestav webového prolézacího modulu

Zobrazíte-li sestavy o dřívější aktivitě webového prolézacího modulu, můžete odhadnout celkový výkon a podle potřeby upravit vlastnosti webového prolézacího modulu a definice prolézaného prostoru.

### Než začnete

Pokud je monitorování kolekcí omezeno administrativní rolí, lze zobrazit statistiku prolézacího modulu a vytvořit sestavy s informacemi o činnosti prolézacího modulu, ale nelze změnit jeho chování (například spustit nebo zastavit prolézací modul).

### O této úloze


Informace o aktivitě webového prolézacího modulu vám mohou poskytnout různé typy sestav. U některých typů sestav jsou informace vráceny, jakmile se je podaří získat z interní databáze prolézacího modulu. Vytvoření sestavy webů a sestavy stavových kódů HTTP naproti tomu určitou dobu trvá. Vytvoříte-li tyto typy sestav, můžete zadat e-mailovou adresu, na kterou má být sestava zaslána, a nemusíte čekat na vrácení výsledků konzolou pro správu podnikového vyhledávání.


Chcete-li zobrazit informace o tom, jak interpretovat statistické údaje v sestavách, klepněte při monitorování webového prolézacího modulu a vytváření sestav na tlačítko **Nápověda**.

### Postup

Chcete-li vytvořit sestavy webového prolézacího modulu, postupujte takto:

1. V pohledu Kolekce vyhledejte kolekci vlastníci webový prolézací modul, který chcete prolézat, a klepněte na volbu  **Monitor**.
2. Otevřete stránku Prolézání.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce Prolézání, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování prolézacích modulů.

3. Pokud webový prolézací modul, pro který chcete vytvořit sestavy, běží nebo je pozastaven, klepněte na ikonu  **Podrobnosti**.
4. Na stránce podrobností webového prolézacího modulu vyberte volbu odpovídající typu sestavy, kterou chcete vytvořit:
  - Chcete-li vytvořit sestavy informující o prolézacím modulu a o všech serverech, které našel nebo navštívil, klepněte na položku **Historie prolézacího modulu** v oblasti **Souhrnný stav prolézacího modulu**.
  - V části **Stav adresy URL** zadejte adresu URL webu, pro který chcete vytvořit sestavu a poté klepněte na tlačítko **Podrobnosti webu**.
5. Chcete-li vytvořit sestavu historie prolézacího modulu nebo webu, zaškrtněte políčka u statistických údajů, které se mají v sestavě zobrazit, a poté klepněte na tlačítko **Zobrazit sestavu**.

U těchto typů statistických údajů prolézací modul vrací sestavu konzole pro správu, jakmile načte informace ze své interní databáze.
6. Pokud vytváříte sestavu historie prolézacího modulu, můžete zadat volby pro vytvoření sestavy webů a poté klepnout na tlačítko **Spustit sestavu**.

Tato sestava se vytváří na základě statistických údajů, které vyberete k zahrnutí, a ukládá se do určeného souboru (název souboru musí být zadán v absolutním tvaru). Můžete vybrat volbu zaslání vytvořené sestavy e-mailem.

7. Pokud vytváříte sestavu historie prolézacího modulu, můžete zadat volby pro vytvoření sestavy stavových kódů HTTP a poté klepnout na tlačítko **Spustit sestavu**.

Tato sestava obsahuje informace o počtu stavových kódů HTTP rozdělené podle webů. Sestava se ukládá do určeného souboru (název souboru musí být zadán v absolutním tvaru). Můžete vybrat volbu zaslání vytvořené sestavy e-mailem.

Z této sestavy můžete zjistit, které weby vracejí velký počet stavových kódů 4xx (informujících o tom, že stránka nebyla nalezena), 5xx (informujících o problémech se serverem), 6xx (informujících o problémech s připojením) atd.

Tato sestava je užitečná zejména po určité době činnosti prolézacího modulu (například u prolézacích modulů, které již pracují několik týdnů). Pomůže vám odhalit již neexistující weby, nově vzniklé weby, weby s vysokým počtem adres URL (možný příznak redundantního prolézání databáze Lotus Notes) a weby s rekurzivním souborovým systémem obsluhovaným serverem HTTP. Pokud weby s vysokým počtem stavových kódů HTTP nepřispívají do indexu, můžete zvýšit výkon prolézacího modulu odebráním těchto webů z prolézaného prostoru.

## Stavové kódy HTTP vrácené webovému prolézacímu modulu

Při monitorování webového prolézacího modulu můžete zobrazit informace o stavových kódech HTTP, které prolézacímu modulu vracejí prolézané stránky.

### Souhrn tabulky

Při monitorování historie webového prolézacího modulu nebo stavu konkrétní adresy URL si můžete prohlédnout informace o stavových kódech HTTP vrácených prolézacímu modulu. Tyto informace můžete využít při správě prolézaného prostoru a optimalizaci výkonu prolézacího modulu. Dostává-li prolézací modul například pro určitou adresu URL velký počet stavových kódů HTTP a informují-li tyto stavové kódy o tom, že stránky v daném umístění nelze prolézat, můžete zvýšit výkon odebráním příslušné adresy URL z prolézaného prostoru.

V následující tabulce jsou uvedeny stavové kódy HTTP spolu s údaji o tom, jakým způsobem tyto kódy interpretuje webový prolézací modul. Hodnoty v intervalu od 100 do 505 jsou standardní stavové kódy HTTP (další informace viz <http://www.w3.org/Protocols/rfc2616/rfc2616.html>). Ostatní stavové kódy HTTP jsou vlastní kódy podnikového vyhledávání a webového prolézacího modulu.

Tabulka 8. Stavové kódy HTTP pro webový prolézací modul

Kód	Popis	Kód	Popis	Kód	Popis	Kód	Popis
NULL	Neprolezeno	400	Chybný požadavek	500	Interní chyba serveru	693	Selhání výběru (URLFetcher)
100	Pokračovat	401	Neautorizováno	501	Není implementováno	694	Chyba zápisu (URLFetcher)
101	Přepínání protokolů	402	Je vyžadována platba	502	Chybná brána	695	Neúplné záhlaví bloku (URLFetcher)
200	Úspěch	403	Zakázáno	503	Služba není dostupná	699	Neočekávaná chyba (URLFetcher)

Tabulka 8. Stavové kódy HTTP pro webový prolézací modul (pokračování)

Kód	Popis	Kód	Popis	Kód	Popis	Kód	Popis
201	Vytvořeno	404	Nenalezeno	504	Časový limit brány	700	Chyba analýzy (chybí konec hlavičky)
202	Přijato	405	Nedovolená metoda	505	Verze HTTP není podporována	710	Chyba analýzy (hlavička)
203	Neautoritativní informace	406	Nepřijatelné	611	Chyba čtení	720	Chyba analýzy (bez kódu HTTP).
204	Žádný obsah	407	Je vyžadováno ověření na serveru proxy	612	Chyba připojení	730	Chyba analýzy (tělo)
205	Obnovení obsahu	408	Vypršení časového limitu požadavku	613	Vypršení časového limitu čtení	740 a 4044	Vyloučeno na základě souboru robots.txt.
206	Neúplný obsah	409	Konflikt	614	Selhání při navazování komunikace SSL	741	Roboti jsou dočasně nedostupní
300	Více voleb	410	Odstraněno	615	Jiná chyba čtení	760	Vyloučeno dle definice prolézáného prostoru
301	Trvale přesunuto	411	Požadována délka	616	Anomálie FBA	761	Zakázáno místním prolézáním prostorem; povoleno globálním
302	Nalezeno	412	Selhání nutné podmínky	617	Chyba kódování	770	Chybný protokol nebo nestandardní systémový port
303	Viz další	413	Příliš velká entita v požadavku	618	Přesměrování bez adresy URL přesměrování	780	Vyloučeno dle definice vyloučených typů souboru
304	Nezměněno	414	Příliš dlouhý identifikátor URI v požadavku	680	Chyba vyhledávání DNS	786	Neplatná adresa URL
305	Použijte server proxy	415	Nepodporovaný typ média	690	Chybně vytvořená adresa URL	2004	Chybí indexová značka META
306	(nepoužito)	416	Požadovaný rozsah není uspokojivý	691	Připojení ztraceno (URLFetcher)	3020	Softwarové přesměrování
307	Dočasné přesměrování	417	Požadavek nebyl splněn	692	Vypršení časového limitu zápisu (URLFetcher)		



## Poznámky k tabulce

### Stavové kódy 4xx

Kód 400 (chybný požadavek) se vyskytuje velmi zřídka. Podle standardu stavových kódů HTTP mají kódy 4xx indikovat selhání klienta (prolétacího modulu). Problém je však obvykle na straně serveru nebo v adrese URL, kterou prolétací modul obdržel jako odkaz. Některé webové servery například netolerují adresy URL, které se pokoušejí o přístup do vyšší než kořenové úrovně webu (např.

<http://xyz.ibm.com/../../sales>). Jiné webové servery tento druh navigace umožňují a operátor nadřazeného adresáře (..) ignorují, pokud se prolétací modul již pohybuje na kořenové úrovni.

Některé servery považují požadavek na kořen webu za chybu a některé zastaralé odkazy mohou požadovat operace, které již nejsou rozpoznávány jako implementované. Při přijetí požadavku na stránku, která již není podporována, aplikační server vygeneruje výjimku, která na webovém serveru vyvolá vrácení stavového kódu HTTP 400, protože požadavek již není považován za platný.

- 615** Informuje o tom, že na prolétacím serveru, který stahuje data z webových serverů, došlo k neočekávané výjimce. Mnohé ze stavových kódů tohoto typu mohou indikovat problém na straně prolétacího modulu.

### Stavové kódy 61x

S výjimkou stavového kódu 615 informují stavové kódy 61x o problémech, které lze při prolétání očekávat, například o vypršení časových limitů. Následující stavové kódy mohou vyžadovat provedení nápravných kroků:

#### 611, 612 a 613

Příčinou těchto problémů mohou být pomalé weby nebo slabý výkon sítě.

- 611** Indikuje, že při načítání dokumentu prolétacím modulem došlo k chybě.
- 612** Indikuje, že při připojování prolétacího modulu k webovému serveru došlo k chybě.
- 613** Indikuje, že došlo k vypršení časového limitu při načítání dokumentu prolétacím modulem.
- 614** Indikuje, že prolétací modul není schopen pracovat se zabezpečenými weby (HTTPS). Jestliže jste přesvědčeni, že tyto weby by měly být dostupné, zkontrolujte správnost konfigurace certifikátů na prolétacím serveru a na cílovém webovém serveru. Je-li například web certifikován uznávanými certifikačními úřady (CA), můžete přidat nové úřady do úložiště důvěryhodných certifikátů, které využívá prolétací modul.
- Zkontrolujte také konfiguraci certifikátů s vlastním podpisem na serverech, které se pokoušíte prolézat. Prolétací modul je konfigurován pro přijímání certifikátů s vlastním podpisem. Některé servery vytvářejí certifikáty s vlastním podpisem pro kořenovou adresu URL (např. <http://sales.ibm.com/>) a poté se pokoušejí používat stejný certifikát v podřízených doménách (např. <http://internal.sales.ibm.com/>). S certifikáty používanými tímto způsobem nemůže prolétací modul pracovat. Certifikáty s vlastním podpisem přijímá jen tehdy, odpovídá-li název domény subjektu ([sales.ibm.com](http://sales.ibm.com/)) a podpis na certifikátu názvu domény požadované stránky.
- 616** Indikuje, že i po opakovaném ověření se ve stažených datech stále objevuje přihlašovací formulář pro ověřování na základě formuláře (FBA).
- Pokud se nezdaří ověřování prolétacího modulu pomocí informací poskytnutých v konfiguračním souboru FBA (přihlašovací formulář a ověřovací údaje jako jméno uživatele a heslo), je všem stránkám závislým

na ověřování FBA přiřazen stavový kód 616. Příčiny chybné funkce konfigurace FBA by měl zjistit administrátor.

**617** Informuje o tom, že z bajtového obsahu dokumentu nelze vytvořit řetězec, protože kódovací řetězec (znaková sada) je neplatný nebo dokument obsahuje neplatné bajty.

**618** Indikuje, že je adresa URL přesměrování neplatná, pokud prolézací modul přijme následující stavové kódy HTTP. Je možné, že je neplatné umístění záhlaví odpovědi HTTP.

301 Trvale přesunuto  
302 nalezeno

**680** Informuje o tom, že prolézacímu modulu se nepodařilo získat adresy IP hostitelů v prolézaném prostoru, pravděpodobně v důsledku problémů s přístupem k síti. Chyba tohoto typu znamená, že prolézací modul není schopen prolézat celé weby a nikoli jen některé adresy URL. Vysoký počet stavových kódů tohoto typu výrazně zhoršuje propustnost.

#### **Stavové kódy 69x**

Stavové kódy 690 až 699 nejsou nikdy zaznamenávány v trvalé databázi prolézacího modulu. Tyto kódy představují výsledky, které neodrážejí skutečný výsledek stahování ze vzdáleného hostitele, ale spíše dočasný stav uvnitř prolézacího modulu, jako je například ukončení jedné komponenty, zatímco jiná čeká na výsledek nebo jej odesílá. Tyto stavové kódy se objevují v některých žurnálech, ale nikoli v trvalém záznamu, a proto by se neměly používat jako hodnoty nastavení výběrů.

#### **Stavové kódy 7xx**

Kódy 7xx nejčastěji souvisejí s pravidly prolézaného prostoru:

##### **710 - 730**

Informují o tom, že prolézací modul nemohl vzhledem k problémům dokončit celé stahování nebo že na webu našel neplatná data HTML. Pokud se setkáváte s vysokým počtem stavových kódů tohoto typu, požádejte o pomoc pracovníka podpory podnikového vyhledávání.

##### **740 a 4044**

Informuje o tom, že obsah souboru nelze indexovat, protože dokument byl z indexování vyloučen na základě omezení uvedených v souboru robots.txt daného webu.

**740** Informuje o tom, že index může obsahovat kotvící odkazy na vyloučené dokumenty.

**4044** Informuje o tom, že kotvící odkazy v dokumentech odkazující na vyloučené dokumenty jsou rovněž vyloučeny z indexu.

**741** Informuje o tom, že soubor robots.txt umístěný na webu prolézání umožňuje, došlo však k selhání při stahování. Pokud prolézací modul opakovaně není schopen prolézt adresu URL, je tato adresa URL odebrána z prolézaného prostoru. V případě, že se setkáváte s velkým počtem stavových kódů tohoto typu, zkontrolujte, zda není vzdálený web přechodně nebo trvale nedostupný. Není-li již cílový web k dispozici, odeberte jej z prolézaného prostoru.

Zbýlé stavové kódy 7xx se zpravidla vyskytují tehdy, provedete-li změny v prolézaném prostoru po určité době práce prolézacího modulu. Tyto stavové kódy obvykle nenaznačují problémy, které by bylo třeba řešit.

**3020** Informuje o tom, že dokument se stavovým kódem 200 obsahuje záhlaví umístění odkazující na uživatelského agenta s jinou adresou URL.

---

## Monitorování modulu analýzy

Modul analýzy monitorujte, jestliže potřebujete zobrazit informace o dokumentech analyzovaných modulem analýzy před jejich přidáním do indexu podnikového vyhledávání. Volby umožňují zkontrolovat statistiku a řídit aktivitu modulu analýzy.

### Než začnete

Pokud je monitorování kolekcí omezeno administrativní rolí, lze zobrazit stav modulu analýzy, ale nelze jej spustit nebo zastavit.

### O této úloze


Při monitorování podrobností modulu analýzy je zobrazen snímek aktivity modulu analýzy se statistikou aktivit modulu analýzy v určitém čase. V této statistice je zobrazen počet dokumentů, které byly prolézány a nyní jsou analyzovány nebo čekají na analýzu, a počet dokumentů, které byly analyzovány a čekají na uložení v indexu.


Je-li modul analýzy aktivní, poskytují zprávy další informace o stavu modulu analýzy. Příklad:




- Modu analýzy aktivně analyzuje dokumenty.
- Modul analýzy je pravděpodobně nečinný. Modul analýzy je ve stavu spánku, dokud nebude pro analýzu k dispozici více dokumentů. Dochází-li k chybám, modul analýzy čeká na restartování. Modul analýzy se sám restartuje, nejsou-li k dispozici žádné služby modulu analýzy (k automatickému restartování dojde například v případě, že nelze vytvořit připojení ke službě modulu analýzy nebo že jsou všechna prostředí JVM modulu analýzy zaneprázdněna zpracováváním jiných kolekcí).
- Činnost modulu analýzy může být pozastavena (činnost modulu analýzy může být například pozastavena, dokud nebude dokončeno sestavení indexu).

### Postup

Monitorování modulu analýzy pro kolekci:

1. V pohledu Kolekcevyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.
2. Otevřete stránku Analýza.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce Parse, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.

3. Pokud je modul analýzy spuštěn a chcete zobrazit podrobné informace o stavu analýzy, klepněte na volbu  **Podrobnosti**.  
Pokud vaše administrativní role umožňuje správu procesů pro kolekci, můžete zobrazit podrobnosti o aktivitě modulu analýzy a přitom modul spustit nebo zastavit.
4. Pokud je modul analýzy zastaven a chcete jej spustit, klepněte na volbu  **Spustit**.  
Při prvním vytvoření kolekce spusťte analytický modul až poté, co prolézací modul zahájí prolézání dat. Tak budou pro analytický modul zajištěna data pro analýzu a kategorizaci. Pokud neprovedete změnu pravidel analýzy, můžete modul analýzy nechat spuštěný bez přerušení.
5. Pokud je modul analýzy spuštěn a chcete jej zastavit, klepněte na volbu  **Zastavit**.  
Po provedení změn pravidel analýzy je nutné modul analýzy zastavit a znovu spustit. Pokud například změníte konfiguraci modulu analýzy, projeví se tyto změny až po zastavení a opětném spuštění modulu analýzy.

---

## Monitorování aktivity indexu pro kolekci


Index pro kolekci monitorujte, jestliže potřebujete zobrazit průběh zpracování sestavovaného indexu, povolit nebo zakázat časový plán indexu nebo spustit a zastavit indexování.


### Než začnete






Aktivity indexu mohou monitorovat všichni uživatelé s oprávněním pro administraci podnikového vyhledávání. Chcete-li spustit či zastavit sestavování indexu nebo povolit či zakázat časový plán indexu, musíte mít přiřazenu roli administrátora podnikového vyhledávání, administrátora kolekce pro danou kolekci nebo operátora kolekce.

### Postup

Monitorování indexu pro kolekci:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.
2. Otevřete stránku Index.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce Index, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.

3. Je-li index naplánován, ale nechcete, aby byl k plánovanému datu a času sestaven, klepněte na ikonu  **Zakázat časový plán**. Index nebude sestaven, dokud nepovolíte časový plán nebo nespustíte proces sestavení indexu.
4. Je-li index naplánován, ale časový plán pro sestavení je zakázán, klepněte na volbu  **Povolit časový plán**. Index bude zařazen do fronty pro sestavení k datu a času zadanému v jeho časovém plánu.
5. Pokud je index zastaven a chcete jej spustit, klepněte na volbu  **Spustit**.  
Obvykle indexování probíhá pravidelně na základě plánování. Pokud bylo sestavování indexu zastaveno nebo pokud byl zakázán časový plán pro index, můžete klepnutím na volbu **Spustit** vynutit zahájení procesu sestavování indexu.
6. Pokud je sestavování indexu aktivní a chcete je zastavit, klepněte na volbu  **Zastavit**.  
Potřeba zastavení sestavování rozdílového indexu může vzniknout například v případě, že je nutné vynutit sestavení hlavního indexu po změně typu kategorizace použité v kolekci.
7. Došlo-li během sestavování indexu k chybám, klepněte na volbu  **Chyba**.  
Zobrazí se stránka Obsah souboru žurnálu, kde si můžete prohlédnout další informace týkající se chyb indexování. Na této stránce můžete po výběru jednotlivých chybových zpráv zobrazit podrobnosti týkající se problému.

---

## Monitorování fronty indexů podnikového vyhledávání

Můžete zobrazit stav všech sestavení indexu ve frontě indexů, zastavit sestavování indexu nebo odstranit index z fronty.

### Než začnete

Chcete-li spravovat frontu indexů, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

### O této úloze


Současně lze sestavovat několik indexů, ale ve frontě může být vždy pouze jeden index pro každou kolekci. Při konfiguraci voleb indexu pro systém můžete určit, kolik indexů může současně sdílet frontu a prostředky indexování.

### Postup



Monitorování fronty indexů:

1. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
2. Vyberte stránku **Index**.

Zobrazí se seznam kolekcí, které mají indexy ve frontě indexů. U každého indexu je zobrazen typ sestavovaného indexu (rozdílový nebo úplný), čas, kdy byl index zařazen do fronty indexů, a čas, kdy začalo sestavování indexu (pokud probíhá sestavování).
3. Chcete-li spravovat konkrétní index, klepněte na ikonu **Stav**.

Můžete například chtít zjistit, za jak dlouho bude index dokončen nebo kolik dokumentů je v indexu, či zakázat časový plán indexu.
4. Chcete-li zastavit sestavovaný index, klepněte na volbu  **Zastavit**.

Pokud jste například změnili pravidla pro kategorie, může být nutné sestavování rozdílového indexu zastavit, aby bylo možné vynutit sestavení hlavního indexu.

Chcete-li spustit sestavování indexu po jeho zastavení, počkejte, než bude index zařazen do fronty indexů při příštím naplánovaném spuštění, nebo klepněte na ikonu **Stav** pro monitorování indexu a klepnutím na volbu  **Spustit** spustíte sestavování indexu.
5. Chcete-li odebrat index z fronty indexů, klepněte na volbu  **Odebrat**.

---

## Monitorování vyhledávacích serverů


Můžete zobrazit podrobné informace o stavu aktivity vyhledávacího serveru pro specifickou kolekci nebo zobrazit podrobné informace o stavu pro vyhledávací servery v celém systému podnikového vyhledávání.




### Než začnete

Všichni administrativní uživatelé podnikového vyhledávání mohou monitorovat vyhledávací servery pro kolekce, k jejichž správě mají oprávnění. Chcete-li monitorovat všechny vyhledávací servery, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

Chcete-li spouštět nebo zastavovat vyhledávací servery, musíte mít přiřazenu roli administrátora podnikového vyhledávání, administrátora kolekce pro danou kolekci nebo operátora kolekce.

### Postup

1. Monitorování vyhledávacích serverů pro jednu kolekci:
  - a. V pohledu **Kolekce** vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.
  - b. Otevřete stránku **Vyhledat**.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce **Search**, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.
2. Monitorování všech vyhledávacích serverů v systému podnikového vyhledávání:
  - a. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
  - b. Vyberte stránku **Vyhledávání**.
3. Pokud je vyhledávací server zastaven a chcete jej spustit, klepněte na volbu  **Spustit**.
4. Pokud je vyhledávací server spuštěn a chcete jej zastavit, klepněte na volbu  **Zastavit**.

Pokud povolíte nebo zakážete mezipaměť vyhledávání, změníte velikost mezipaměti vyhledávání nebo změníte rychlé odkazy, projeví se tyto změny až po zastavení a opětném spuštění vyhledávacích serverů.

5. Chcete-li zobrazit souhrn času, který vyhledávací server strávil zpracováním požadavků na vyhledávání, klepněte na volbu **Doba odezvy- historie**.

Tato sestava zobrazuje průměrnou dobu v milisekundách, kterou vyhledávacímu serveru trvala odezva na požadavky na vyhledávání ke konkrétnímu datu.

Průměrná doba odezvy je indikátorem výkonu systému a odpovídá kvalitě služby. Delší doba odezvy může znamenat, že je systém nadměrně zatížen. Systém může být například zahlcen vzhledem k počtu prohledávaných kolekcí nebo velikosti kolekce.

6. Chcete-li zobrazit seznam nejčastěji odesílaných dotazů, klepněte na volbu **Oblíbené dotazy**.

Tato sestava zobrazuje klíčová slova v 50 nejčastěji odesílaných dotazech a počet zadání příslušného dotazu uživateli.

Na základě seznamu nejčastějších dotazů můžete určit nejhodnější dotazy pro rychlé odkazy. Vytvoření rychlých odkazů může zvýšit kvalitu vyhledávání pro mnoho uživatelů. Zajistíte tím, že ve výsledcích vyhledávání budou vždy vráceny vysoce relevantní dokumenty.

Můžete také vytvořit v podnikovém portálu odkazy na prostředky, které poskytnou odpovědi na dotazy. Pokud například uživatelé často vyhledávají informace o úhradě nákladů, vytvořte na domovské stránce intranetu odkaz na stránku s informacemi o postupech při úhradě nákladů.

7. Chcete-li zobrazit seznam naposledy odeslaných dotazů, klepněte na volbu **Poslední dotazy**.

Tato sestava obsahuje klíčová slova v 50 naposledy odeslaných dotazech.

Na základě seznamu posledních dotazů můžete zjistit aktuální trendy a naléhavé situace v organizaci. Můžete například zaznamenat vlnu zájmu o určité téma. Na základě této vlny zájmu se můžete rozhodnout vytvořit pro dané téma rychlý odkaz nebo je zpřístupnit uživatelům jiným způsobem (například vytvořením odkazu na podnikovém portálu).

## Změna způsobu výpočtu statistik dotazů

Můžete změnit způsob, jakým systém vypočítává počet oblíbených dotazů a posledních dotazů.

### O této úloze



Při monitorování vyhledávacích serverů můžete vybrat volbu zobrazení seznamu 50 nejoblíbenějších dotazů a seznamu 50 naposledy zpracovaných dotazů. Ve výchozí konfiguraci vyhledávacího serveru jsou dotazy s ekvivalentním řetězcem dotazu a jiným nastavením rozsahu výsledků počítány jako nezávislé dotazy. Počet dotazů je tedy zvýšen například v případě, že uživatel klepne na volbu k zobrazení další stránky výsledků.

Systém výpočtu statistik dotazů pro vyhledávací server můžete změnit pomocí úpravy souboru `runtime-generic.properties`. Nastavíte-li parametr `distinctRecentQueryCheck` na hodnotu `true`, bude systém počítat jako samostatné dotazy pouze ty dotazy, které vrací úvodní stránku výsledků.

Chcete-li upravit vlastnosti vyhledávacího serveru, musíte se přihlásit jako administrátor podnikového vyhledávání. Chcete-li spouštět nebo zastavovat vyhledávací servery, musíte mít přiřazenu roli administrátora podnikového vyhledávání, administrátora kolekce pro danou kolekci nebo operátora kolekce.

### Postup

Chcete-li změnit způsob, jakým systém počítá statistiky dotazů, postupujte takto:

1. Přihlaste se jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition. V konfiguraci s více servery se přihlaste k vyhledávacímu serveru.
2. Upravte následující soubor, kde *ID\_kolekce* je ID kolekce, kterou chcete změnit, a *ID\_uzlu* určuje vyhledávací server, kde chcete změnu provést:  
`ES_NODE_ROOT/master_config/ID_kolekce.runtime.ID_uzlu/runtime-generic.properties`
3. Nastavte parametr **distinctRecentQueryCheck** na hodnotu `true` a poté soubor uložte a zavřete.
4. V konfiguraci s více servery opakujte předchozí kroky na druhém vyhledávacím serveru.
5. Chcete-li změnit chování statistik dotazů pro další kolekce, zopakujte předchozí kroky pro každou kolekci, kterou chcete změnit.
6. Chcete-li, aby se změny projevily, přihlaste se ke konzole pro správu podnikového vyhledávání a restartujte vyhledávací servery:
  - a. Klepnutím na volbu **Systém** otevřete pohled Systém.
  - b. Vyberte stránku Vyhledávání.
  - c. Pro každý změněný vyhledávací server klepněte na tlačítko  **Zastavit**.
  - d. Pro každý změněný vyhledávací server klepněte na tlačítko  **Spustit**.

---

## Monitorování modulu pro příjem dat

Při monitorování modulu pro příjem dat zobrazíte jeho stav a podrobnosti o aktivitě klientské aplikace modulu pro příjem dat.


### Než začnete

Chcete-li monitorovat modul pro příjem dat, musíte mít přiřazenu roli administrátora podnikového vyhledávání.


**Důležité:** Modul pro příjem dat nebude v budoucích verzích podporován. Při vývoji klientských aplikací pro podnikové vyhledávání používejte místo něj rozhraní API pro hledání a indexování (SI-API). Následující informace jsou určeny uživatelům, kteří dříve aplikace využívající modul pro příjem dat vytvořili.

### Postup

Monitorování modulu pro příjem dat:

1. Klepnutím na volbu **Systém** otevřete pohled Systém.
2. Ikona stavu na stránce Modul pro příjem dat zobrazuje, zda je modul pro příjem dat aktivní nebo zastaven.
3. Pokud je modul pro příjem dat spuštěn a chcete zobrazit podrobné informace o aktivitě klientské aplikace, klepněte na volbu  **Podrobnosti**.

Ikona stavu na stránce Podrobnosti modulu pro příjem dat označuje, zda je modul pro příjem dat aktivní nebo zastaven. Statistika zobrazuje, kolik požadavků čeká na zpracování, aktuální stav jednotlivých podprocesů, které jsou v činnosti v souvislosti s požadavky klientské aplikace, a počet aktivních podprocesů pro daný stav podprocesu.

4. Změníte-li číslo portu modulu pro příjem dat nebo maximální počet dokumentů, které mohou být uloženy v dočasném úložišti, klepněte na volbu  **Restartovat**.

Modul pro příjem dat je spuštěn při spuštění systému podnikového vyhledávání. Pokud nezměníte jednu z těchto voleb konfigurace nebo obě, není třeba modul pro příjem dat restartovat.

---

## Sledování dokumentů

Dokumenty lze ze systému odstranit v různých fázích zpracování. Můžete určit volby pro zjištění, že byl dokument odstraněn, a zjištění problémů, které jeho odstranění způsobily.

Pokud modul analýzy zjistí chybu, která brání analýze dokumentu, je pro odstranění dokument zaznamenána zpráva s kódem příčiny. (Tento typ chyby nezpůsobí odebrání starších verzí dokumentu z indexu.)

Dokumenty mohou být odstraněny během fází indexace; tyto informace jsou rovněž zaznamenány. Identifikátory URIs a vzory identifikátorů URI lze například explicitně odstraňovat. Je možné, že byl dokument prolezen prolézacím modulem, který byl později odstraněn. Může se stát, že zdrojový dokument již neexistuje (s dokumentem je asociován záporný kód HTTP) nebo že je s dokumentem asociován neznámý kód HTTP. Dokumenty mohou být odstraněny i v případě, že pro dokument vyžadující globální analýzu chybí informace o ohodnocení důležitosti.

Víte-li, že byl dokument prolezen, ale nezobrazuje se v indexu, můžete prostřednictvím konzoly pro správu podnikového vyhledávání sledovat proces zpracování dokumentu v systému. Podrobné sestavy mohou obsahovat informace o čase, místě a příčině odstranění dokumentu. Ze sestavy například může plynout, že byl dokument neočekávaně odstraněn během globální analýzy nebo že administrátor odebral identifikátor URI z indexu.

### Související úlohy

“Zobrazení podrobností o identifikátoru URI” na stránce 286

## Konfigurování souborů žurnálu pro sledování dokumentů

Chcete-li určit, kdy, kde a proč byl určitý dokument odstraněn ze systému, můžete konfigurovat soubory žurnálu pro sledování informací o odstraněných dokumentech.

### Než začnete

Chcete-li konfigurovat volby pro sledování odstraněných dokumentů, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### O této úloze

Aby soubory žurnálu nezabíraly příliš mnoho prostoru na disku, provádí systém rotování souborů žurnálu, a při změně aktuálního data vytvoří vždy nový soubor žurnálu. Pokud některý soubor žurnálu dosáhne maximální přípustné velikosti a nedojde ke změně data, vytvoří systém nový soubor žurnálu. Po dosažení maximálního počtu souborů žurnálu se odstraní nejstarší soubor žurnálu, aby mohl být vytvořen nový.

### Postup

Konfigurování souborů žurnálu pro sledování dokumentů:

1. Upravte kolekci, vyberte stránku Žurnál a klepněte na volbu **Konfigurovat sledování dokumentů**.
2. Na stránce Sledování dokumentů ověřte, že je políčko pro sledování dokumentů zaškrtnuto.
3. Určete počet souborů žurnálu, které mají být používány pro zaznamenávání informací o dokumentech, jež byly ze systému odstraněny. Tyto soubory žurnálu jsou sdíleny všemi relacemi, v nichž lze dokumenty odstranit.



## Zobrazení sestav odstraněných dokumentů

Můžete také zobrazit podrobné informace o dokumentech odstraněných ze systému podnikového vyhledávání. Tyto informace jsou k dispozici pouze v případě, že bylo kolekcí povoleno sledování dokumentů.

### Než začnete

Před odesláním požadavku na zobrazení sestavy odstraněných dokumentů nebo odeslání sestavy na e-mailovou adresu ověřte, zda jsou relace, ze kterých chcete obdržet informace, aktivní. Chcete-li například získat informace o dokumentech odstraněných během operace analýzy nebo indexování, ověřte, že jsou relace analýzy a indexování pro příslušnou kolekci spuštěny.

Před přijetím sestavy zajistěte nastavení konfigurace informací o vašem poštovním serveru pro podnikové vyhledávání. Tyto informace můžete zadat při konfigurování voleb e-mailu na stránce Žurnál v pohledu Systém.


### O této úloze


Shromažďování informací o odstraněných dokumentech může trvat delší dobu. Můžete vybrat volbu zobrazení informací a počkat, než se tyto informace zobrazí. Efektivnější volbou je odeslání sestavy na zadanou e-mailovou adresu.


Pokud byl dokument odstraněn, bude v sestavě uvedeno datum a čas odstranění, úroveň závažnosti chyby, komponenta a relace, v níž k problému došlo, a chybová zpráva.

### Postup

Zobrazení podrobností o odstraněných dokumentech:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce General, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.

2. Klepněte na volbu  **Podrobnosti identifikátoru URI**.
3. Na stránce Podrobnosti identifikátoru URI zadejte identifikátor URI, pro který chcete zobrazit informace.
4. Zaškrtněte políčka odpovídající typu informací, které chcete zobrazit:

#### Dokumenty odstraněné analyzátořem

Toto políčko zaškrtněte, chcete-li zjistit, zda byl dokument odstraněn během analýzy a příčinu jeho odstranění, pokud se tak stalo.

#### Dokumenty odstraněné z indexu

Toto políčko zaškrtněte, chcete-li zjistit, zda byl dokument odstraněn během indexování nebo analýzy, a příčinu jeho odstranění, pokud se tak stalo.

5. Určete, jakým způsobem chcete sestavu zobrazit:
  - Chcete-li počkat na zobrazení sestavy, klepněte na volbu **Zobrazit sestavu**.
  - Chcete-li sestavu odeslat na e-mailovou adresu, abyste ji mohli zobrazit později, klepněte na volbu **Odeslat sestavu**.

Na stránce Zaslát podrobnou sestavu o identifikátoru URI zadejte do pole **E-mailová adresa pro zaslání oznámení** e-mailovou adresu pro přijetí sestavy. Poté klepněte na volbu **Odeslat sestavu**.

### Související úlohy

“Zobrazení podrobností o identifikátoru URI” na stránce 286

### Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 113

## Zobrazení souborů žurnálu odstraněných dokumentů

Můžete zobrazit zprávy žurnálu o dokumentech odstraněných ze systému podnikového vyhledávání. Tyto informace jsou k dispozici pouze v případě, že bylo v kolekci povoleno sledování dokumentů.


### O této úloze


Chcete-li zobrazit sestavu o odstraněném dokumentu, je třeba znát identifikátor URI dokumentu. Zobrazením souboru žurnálu odstraněného dokumentu získáte datum a čas odstranění dokumentu, úroveň závažnosti chyby, komponentu a relaci, ve které došlo k problému, a podrobnou chybovou zprávu.

### Postup

Zobrazení souborů žurnálů pro odstraněné dokumenty:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce General, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.


2. Klepněte na volbu  **Soubory žurnálů odstraněných dokumentů**.
3. Na stránce Soubory žurnálů odstraněných dokumentů vyberte soubor žurnálu, který chcete zobrazit. Název každého souboru žurnálu udává, zda byl dokument odstraněn analyzátořem (pd) nebo během sestavování indexu (in) a obsahuje datum vytvoření souboru. Je-li ve stejný den vytvořeno více než jeden soubor žurnálu stejného typu, bude pořadí vytvoření souborů v daný den určeno číselnou příponou. Příklad:

```
dropped_doc_in_20060525.log  
dropped_doc_pd_20060524.log (obsahuje nejnovější položky pro toto datum)  
dropped_doc_pd_20060524.log.1  
dropped_doc_pd_20060524.log.2 (obsahuje nejstarší položky pro toto datum)
```

4. Klepněte na volbu **Zobrazit žurnál**.

U každé zprávy na stránce Obsah souboru žurnálu je zobrazeno datum a čas jejího zadání, úroveň závažnosti, název relace, která zprávu zadala, a číslo zprávy a chybový text.

Klepnutím na příslušné tlačítko můžete přejít na první stránku, poslední stránku, předchozí stránku nebo následující stránku souboru žurnálu. Můžete rovněž přejít na konkrétní stránku zadáním příslušného čísla stránky.

5. Chcete-li zobrazit podrobnější informace o zprávě, klepněte na volbu  **Podrobnosti**.

Na stránce Podrobnosti zprávy v žurnálu je zobrazen název hostitele serveru podnikového vyhledávání, na kterém došlo k chybě, název souboru, který chybu generoval, název funkce a číslo řádku, kde chyba vznikla, ID procesu a ID podprocesu.

Pomocí tlačítek se můžete přesunout na další a předchozí zprávy v souboru žurnálu.

---

## Soubory žurnálu a výstrahy

Pro kolekci a pro celý systém můžete vybrat typy zpráv, které chcete ukládat do žurnálu. Dále můžete zadat volby pro vytváření a prohlížení souborů žurnálu, pro zaslání výstrah a pro zaslání zpráv e-mailem.

Během běžných operací zapisují komponenty podnikového vyhledávání zprávy žurnálu do společného souboru žurnálu. Tento soubor žurnálu se nachází v adresáři `ES_NODE_ROOT/logs` na indexovém serveru. K zobrazení dat tohoto společného žurnálu slouží konzola pro správu.

Pokud dojde k problému, jako je selhání síťové komunikace, komponenty zapíší zprávy žurnálu do adresáře `logs` na serveru, na kterém je daná komponenta nainstalována. Chcete-li tyto lokální soubory žurnálu zobrazit, použijte prohlížeč souborů v daném počítači, například nástroj `tail` v systému UNIX. K zobrazení těchto typů souborů žurnálu nelze použít konzolu pro správu.

Při konfigurování souborů žurnálu můžete vybrat typy zpráv, které chcete ukládat do žurnálu (například chybové nebo varovné zprávy), určit způsob odstraňování starých souborů žurnálu v případě, že je třeba uvolnit místo pro nové soubory žurnálu, nastavit maximální velikost souborů žurnálu a vybrat jazyk zpráv. Můžete nastavit také volby zaslání e-mailů vždy, když dojde k určitém událostem, nebo když jsou do žurnálu zapsány určité zprávy či zprávy určitého typu.

Při monitorování souborů žurnálu můžete určit, který soubor žurnálu chcete otevřít. Filtrováním obsahu souboru žurnálu můžete zobrazit pouze zprávy určité úrovně závažnosti (například pouze chybové zprávy) nebo pouze zprávy vygenerované určitou relací podnikového vyhledávání. Při prohlížení souboru žurnálu můžete zobrazit podrobné informace o jednotlivých zprávách. V případě potřeby tak můžete zjistit například název funkce, která zprávu vygenerovala, a další informace, které vám pomohou odstranit případný problém.

### Související pojmy



Zprávy podnikového vyhledávání

---

## Výstrahy

Podnikové vyhledávání lze konfigurovat tak, aby byly při zjištění výskytu určitých událostí zapisovány zprávy do souboru žurnálu.

Zprávy spouštěné událostmi se nazývají výstrahy a informují vás o situacích, které je třeba řešit, například o tom, že dochází volný prostor určitého prostředku. Při konfigurování výstrah pro podnikové vyhledávání určujete podmínky, které má systém monitorovat. Vždy, když je sledovaná podmínka splněna, systém automaticky zapíše zprávu do souboru žurnálu.

Pokud chcete dostávat informace o určitých situacích přímo, můžete nastavit volby zaslání e-mailu vždy, když je do žurnálu zapsána některá z monitorovaných zpráv.

Výstrahy lze nastavit pro události na úrovni kolekce a pro události na úrovni systému. Na úrovni kolekce nabízí systém následující možnosti:

- Monitorování počtu dokumentů zpracovávaných jednotlivými prolézacími moduly a vygenerování výstražné zprávy těsně před dosažením maximálního povoleného počtu dokumentů.
- Monitorování počtu dokumentů přidávaných do indexu kolekcí a vygenerování výstražné zprávy těsně před dosažením maximálního povoleného počtu dokumentů.
- Informování o překročení nastaveného limitu doby, kterou systém potřebuje k reakci na vyhledávací požadavky.

Na úrovni systému může systém monitorovat diskový prostor v jednotlivých serverech podnikového vyhledávání a vygenerovat výstražnou zprávu při nedostatku volného prostoru.

## Konfigurování výstrah na úrovni kolekce

Konfigurováním výstrah zajistíte, že při každém výskytu určitých událostí na úrovni kolekce bude zapsána zpráva do souboru žurnálu. Při každém zaznamenání těchto událostí můžete také obdržet e-mail.

### Než začnete

Chcete-li konfigurovat výstrahy pro kolekci, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

### Postup

Konfigurování výstrah na úrovni kolekce:

1. Upravte kolekci, vyberte stránku Žurnál a klepněte na volbu **Konfigurovat výstrahy**.
2. Pokud chcete, aby systém monitoroval počet dokumentů prolézáných jednotlivými prolézacími moduly, postupujte následujícím způsobem:
  - a. Zaškrtněte políčko **Dosáhne-li počet dokumentů prolezených některým prolézacím modulem procentní částí povoleného maxima**.
  - b. Do pole **Procentní část** zadejte, kdy má být zpráva zaznamenána. Tuto hodnotu zadejte jako procentní část maximálního počtu dokumentů, které může prolézací modul prolézat (zadaného prostřednictvím volby **Maximální počet prolézáných dokumentů** při konfigurování vlastností prolézacího modulu). Výchozí hodnota je 90 procent.  
 Protože lze konfigurovat různá omezení pro různé prolézací moduly, jsou pro každý prolézací modul zaznamenávány samostatné zprávy. Pokud například použijete výchozí práh pro výstrahy a povolíte prolézacímu modulu DB2 prolézat 2 000 000 dokumentů a prolézacímu modulu Notes prolézat 1 000 000 dokumentů, bude zaprotokolována jedna zpráva, když prolézací modul DB2 proleze 1 800 000 dokumentů, a další zpráva bude zaprotokolována, když prolézací modul Notes proleze 900 000 dokumentů.
3. Pokud chcete, aby systém monitoroval počet dokumentů přidávaných do indexu, postupujte následujícím způsobem:
  - a. Zaškrtněte políčko **Dosáhne-li počet dokumentů v kolekci procentní částí odhadované velikosti**.
  - b. Do pole **Procentní část** zadejte, kdy má být zpráva zaznamenána. Tuto hodnotu zadejte jako procentní část odhadovaného počtu dokumentů, které může kolekce obsahovat. Výchozí hodnota je 85 procent.  
 V poli **Omezení** je uveden aktuální odhad velikosti kolekce. Chcete-li tuto hodnotu změnit, otevřete stránku Obecné pro danou kolekci, vyberte volbu pro nastavení obecných voleb a zadejte novou hodnotu do pole **Odhadovaný počet dokumentů**.

**Upozornění:** Toto omezení a odhad počtu dokumentů, které pro kolekci konfiguruje, slouží pouze k monitorování růstu kolekce. Nejde o absolutně platný limit velikosti, který by index nesměl překročit.

4. Chcete-li, aby vás systém informoval, že doba potřebná pro odezvu na požadavky na vyhledávání překračuje omezení, postupujte následujícím způsobem:
  - a. Zaškrtněte políčko **Překročí-li doba odezvy při vyhledávání povolené omezení**.
  - b. Do pole **Omezení** zadejte počet sekund, které považujete za maximální přijatelnou dobu odezvy při vyhledávání.

Při překročení této hodnoty запиše systém o této události zprávu do žurnálu. Pokud například zachováte výchozí hodnotu, systém vytvoří zprávu v žurnálu, kdykoli vyhledávací server dosáhne při odezvách na požadavky na vyhledávání průměru pět nebo více sekund.

Obvyklé doby odezvy jsou kratší než půl sekundy. Průměr vyšší než jedna sekunda může indikovat, že operační systém potřebuje vyladění na lepší výkon nebo že existuje problém v konfiguračním nastavení vyhledávacího serveru. Například můžete zvětšit velikost prostoru, který alokujete pro vyhledávací mezipaměť.

5. Klepněte na tlačítko **OK**.

Pokud chcete při zaznamenání zpráv o těchto událostech obdržet e-mail, otevřete stránku **Žurnál**, klepněte na možnost **Konfigurovat volby e-mailu pro zprávy** a zadejte e-mailovou adresu. Čísla zpráv pro povolené výstrahy jsou automaticky přidány do seznamu čísel zpráv, pro které mají být odesílány e-maily.

Než budete moci přijímat e-maily, je také nutné zajistit, aby byl nakonfigurován poštovní server. Chcete-li nakonfigurovat poštovní server, musí administrátor podnikového vyhledávání vybrat na panelu nástrojů volbu **Systém**, otevřít stránku **Žurnál** a klepnout na možnost **Konfigurovat volby e-mailu pro zprávy**.

#### **Související úlohy**

“Příjem e-mailů o protokolovaných zprávách” na stránce 310

## **Konfigurování výstrah na úrovni systému**


Konfigurováním výstrah zajistíte, že při každém výskytu určitých událostí na úrovni systému bude zapsána zpráva do souboru žurnálu. Při každém zaznamenání těchto událostí můžete také obdržet e-mail.

#### **Než začnete**

Chcete-li konfigurovat výstrahy na úrovni systému, musíte být administrátor podnikového vyhledávání.

#### **Postup**

Chcete-li konfigurovat výstrahy na úrovni systému, postupujte takto:

1. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
3. Na stránce **Žurnál** klepněte na volbu **Konfigurovat výstrahy**.
4. Pokud chcete, aby systém monitoroval množství volného prostoru, který je k dispozici na serverech podnikového vyhledávání, zaškrtněte políčko **Dosáhne-li velikost dostupného prostoru v souborovém systému určitého procentuálního podílu celkového prostoru**.
5. V poli **Procentní část** určete, kdy chcete být systémem upozorněni na nedostatek volného prostoru na serveru. Tuto hodnotu zadejte jako procentní část celkového prostoru souborového systému. Výchozí hodnota je 80 procent.

Pokud systém podnikového vyhledávání používáte na více serverech, systém vytvoří samostatnou zprávu žurnálu pro každý server. Jedna zpráva vás například informuje o tom, že na prolézacím serveru není dostatek místa; jiná samostatná zpráva informuje o prostorových omezeních na indexových a vyhledávacích serverech.

6. Klepněte na tlačítko **OK**.

Pokud chcete při každém zaznamenání zpráv o těchto událostech obdržet e-mail, otevřete stránku **Žurnál**, klepněte na možnost **Konfigurovat volby e-mailu pro zprávy** a zadejte e-mailovou adresu a informace o poštovním serveru.

#### **Související úlohy**

“Příjem e-mailů o protokolovaných zprávách” na stránce 310

---

## **Konfigurování souborů žurnálu**

Můžete určit typy zpráv, které chcete protokolovat, a volby pro vytváření souborů žurnálu.

### **Než začnete**

Chcete-li konfigurovat soubory žurnálu na úrovni kolekce, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci. Chcete-li konfigurovat soubory žurnálu na úrovni systému, musíte být administrátor podnikového vyhledávání.



### **O této úloze**

Aby soubory žurnálu nezabíraly příliš mnoho prostoru na disku, provádí systém rotování souborů žurnálu, a při změně aktuálního data vytvoří vždy nový soubor žurnálu. Pokud některý soubor žurnálu dosáhne maximální přípustné velikosti a nedojde ke změně data, vytvoří systém nový soubor žurnálu. Po dosažení maximálního počtu souborů žurnálu se odstraní nejstarší soubor žurnálu, aby mohl být vytvořen nový.

Chcete-li přijímat e-maily o protokolovaných zprávách, je nutné nejdříve zadat informace o způsobu doručování e-mailů. Potom určete zprávy, pro které chcete e-maily přijímat.

### **Postup**

Konfigurování souborů žurnálu podnikového vyhledávání:

1. Konfigurování voleb pro vytváření a cyklické použití souborů žurnálu na úrovni systému:
  - a. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
  - b. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
  - c. Na stránce **Žurnál** klepněte na volbu **Konfigurovat volby souboru žurnálu**. Zobrazí se stránka **Volby souboru žurnálu** na úrovni systému.
2. Konfigurování voleb pro vytváření a cyklické použití souborů žurnálu na úrovni kolekce:
  - a. V pohledu **Kolekce** vyhledejte kolekci, pro kterou chcete určit volby, a klepněte na volbu  **Upravit**.
  - b. Na stránce **Žurnál** klepněte na volbu **Konfigurovat volby souboru žurnálu**. Zobrazí se stránka **Volby souboru žurnálu** na úrovni kolekce.
3. V poli **Typ protokolovaných informací** vyberte typy zpráv, které chcete protokolovat:

#### **Pouze chybové zprávy**

Chybové zprávy indikují výskyt nežádoucí situace nebo neočekávané funkce a oznamují, že proces nemůže pokračovat. Problém je třeba vyřešit provedením příslušné akce.

### **Chybové zprávy a výstrahy**

Výstrahy indikují možný konflikt nebo nekonzistenci, nejsou však příčinou zastavení procesu. Tato volba je výchozí.

### **Všechny zprávy**

Informační zprávy poskytují obecné informace o systému nebo aktuální úloze a nevyžadují žádnou nápravnou akci.

**Důležité:** Výběr této volby může negativně ovlivnit výkon systému. Všechny zprávy protokolujte, pouze pokud potřebujete odstranit problémy nebo pokud vás k tomu vyzve softwarová podpora společnosti IBM.

4. Do pole **Maximální velikost jednotlivých souborů žurnálu** zadejte maximální počet megabajtů pro jednotlivé soubory žurnálu. Výchozí hodnota je 10 MB.

Jakmile soubor žurnálu dosáhne této velikosti, je vytvořen nový soubor žurnálu, a to až do maximálního povoleného počtu souborů žurnálu. Udržováním menších souborů žurnálu můžete dosáhnout větší přehlednosti při jejich zobrazení.

5. Do pole **Maximální počet souborů žurnálu** zadejte maximální počet souborů žurnálu, které chcete vytvořit. Výchozí hodnota je 16.

Chcete-li zajistit, aby byly pro účely kontroly dostupné i starší zprávy žurnálu, zvyšte tuto hodnotu. Pokud vás zajímají spíše novější zprávy a nepotřebujete udržovat dlouhou historii aktivity, snižte tuto hodnotu.

6. V poli **Výchozí lokalita** vyberte jazyk, který chcete použít pro protokolování zpráv. Výchozí hodnota je Angličtina.

7. Klepněte na tlačítko **OK**.

8. Aby byly změny použity, zastavte a restartujte systém podnikového vyhledávání pomocí následujících příkazů.

```
esdmin system stopall  
esdmin system startall
```

## **Konfigurování informací o serveru SMTP**

Než budete moci přijímat e-maily o aktivitách podnikového vyhledávání, je nutné konfigurovat informace o serveru SMTP (Simple Mail Transfer Protocol).

### **Než začnete**

Chcete-li konfigurovat informace o serveru SMTP, musíte mít přiřazenu roli administrátora podnikového vyhledávání.


### **O této úloze**

Příjem e-mailů je umožněn několika administrativními funkcemi podnikového vyhledávání. Než budete moci přijímat e-maily pomocí těchto funkcí, je nutné zadat informace o serveru SMTP:

- Pokud konfiguruje výstrahy na úrovni kolekce nebo na úrovni systému, můžete přijímat e-maily při každém zaznamenání těchto zpráv. E-maily lze přijímat také při zaznamenání jiných zpráv, nejen zpráv, které jsou spouštěny monitorovanými událostmi.
- Pokud chcete zobrazit podrobné informace o identifikátoru URI v indexu nebo dokumentu, který byl ze systému podnikového vyhledávání odstraněn, můžete požadovanou sestavu obdržet e-mailem.
- Pokud monitorujete webový prolézací modul a určíte, že chcete vytvářet sestavy historie webového prolézacího modulu, můžete oznámení o vytvoření sestavy obdržet e-mailem.

## Postup

Konfigurování informací o serveru SMTP:

1. Klepnutím na volbu **Systém** otevřete pohled Systém.
2. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
3. Na stránce Žurnál klepněte na volbu **Konfigurovat volby e-mailu pro zprávy**.
4. Na stránce Volby e-mailu pro systémové zprávy zadejte do pole **Poštovní server SMTP použitý pro doručování e-mailů** úplný název hostitele nebo adresu IP serveru SMTP, který chcete používat.  
Systém tento server použije k odesílání e-mailů na zadané adresy.
5. Do pole **Četnost kontroly došlé pošty** zadejte, jak často má systém vyhledávat příslušné zprávy a odesílat e-maily s informacemi o těchto zprávách.  
Systém sloučí všechny zprávy pro specifickou e-mailovou adresu do jedné zprávy a odešle tuto zprávu na základě určené četnosti.
6. Klepněte na tlačítko **OK**.

## Příjem e-mailů o protokolovaných zprávách

Můžete určit volby pro příjem e-mailů při každém zaznamenání určitých zpráv nebo zpráv určitých typů.

### Než začnete

Chcete-li konfigurovat volby e-mailu pro zprávy na úrovni systému, musíte mít přiřazenu roli administrátora podnikového vyhledávání. Chcete-li konfigurovat volby e-mailu pro zprávy na úrovni kolekce, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.


Než budete moci přijímat e-maily, je třeba konfigurovat informace o serveru SMTP (Simple Mail Transfer Protocol), aby bylo možné e-maily doručovat.

### O této úloze

Při konfigurování výstrah můžete zvolit volbu protokolování zpráv, dojde-li k určitým událostem. Pokud tyto volby povolíte, můžete potom konfigurovat volby pro automatický příjem e-mailů, kdykoli jsou tyto zprávy zaznamenány. Můžete také určit volby pro příjem e-mailů při zaznamenání jiných zpráv, nejen zpráv, které jsou spouštěny událostmi.


## Postup

Konfigurování voleb e-mailu pro zprávy:

1. Chcete-li přijímat e-maily o systémových zprávách, postupujte takto:
  - a. Klepnutím na volbu **Systém** otevřete pohled Systém.
  - b. Chcete-li změnit zobrazení pro úpravy systému, klepněte na volbu  **Upravit**.
  - c. Na stránce Žurnál klepněte na volbu **Konfigurovat volby e-mailu pro zprávy**.
  - d. Na stránce Volby e-mailu pro systémové zprávy zaškrtněte políčko **Odesílat e-maily se zprávami na úrovni systému**.
  - e. Do pole **E-mailová adresa pro příjem e-mailů** zadejte jednu či více e-mailových adres. Administrátor podnikového vyhledávání by obvykle měl přijímat informace o systémových zprávách.  
Jednotlivé adresy oddělte čárkami. Příklad:



novak@praha.spolecnost.cz, prochazka@brno.spolecnost.cz,  
svobodova@ostrava.spolecnost.cz.

- f. Pokud chcete přijímat e-maily o všech zaznamenaných chybových zprávách, zaškrtněte políčko **Odesílat e-maily se všemi chybovými zprávami**.
  - g. Pokud chcete přijímat e-maily pouze v případě zaznamenání určitých zpráv na úrovni systému, zadejte v oblasti **Odesílat e-maily s určitými zprávami** čísla těchto zpráv. Zadávejte jedno číslo zprávy na řádek. Příklad:  
FFQC4819E  
FFQ00005E  
Podle výchozího nastavení se vypisuje několik ID zpráv (popis k těmto zprávám získáte klepnutím na tlačítko **Nápověda**).
  - h. Klepněte na tlačítko **OK**.
2. Chcete-li přijímat e-maily o zprávách pro kolekci, postupujte takto:
- a. Klepnutím na volbu **Kolekce** otevřete pohled Kolekce.
  - b. V seznamu kolekcí vyhledejte kolekci, kterou chcete konfigurovat, a klepněte na volbu  **Upravit**.
  - c. Na stránce Žurnál klepněte na volbu **Konfigurovat volby e-mailu pro zprávy**.
  - d. Na stránce Volby e-mailu pro zprávy kolekci zaškrtněte políčko **Odesílat e-maily se zprávami na úrovni kolekce**.
  - e. Do pole **E-mailová adresa pro příjem e-mailů** zadejte jednu či více e-mailových adres. Administrátor kolekce by obvykle měl přijímat informace o zprávách na úrovni kolekce.  
Jednotlivé adresy oddělte čárkami. Příklad:  
novak@praha.spolecnost.cz, prochazka@brno.spolecnost.cz,  
svobodova@ostrava.spolecnost.cz.
  - f. Pokud chcete přijímat e-maily o všech zaznamenaných chybových zprávách, zaškrtněte políčko **Odesílat e-maily se všemi chybovými zprávami**.
  - g. Pokud chcete přijímat e-maily pouze v případě zaznamenání určitých zpráv na úrovni kolekce, zadejte v oblasti **Odesílat e-maily s určitými zprávami** čísla těchto zpráv. Zadávejte jedno číslo zprávy na řádek. Příklad:  
FFQC4819E  
FFQ00005E  
Podle výchozího nastavení se vypisuje několik ID zpráv (popis k těmto zprávám získáte klepnutím na tlačítko **Nápověda**).
  - h. Klepněte na tlačítko **OK**.

#### Související pojmy

 Zprávy podnikového vyhledávání

#### Související úlohy

“Konfigurování výstrah na úrovni kolekce” na stránce 306

“Konfigurování výstrah na úrovni systému” na stránce 307

## Změna velikosti protokolu dotazů

Velikost souborů protokolu vytvářených pro zpracování dotazů můžete zvýšit nebo snížit pomocí úpravy konfiguračního souboru. Konzola pro správu podnikového vyhledávání neposkytuje žádnou podporu této úlohy.

#### O této úloze

V průběhu zpracování dotazů jsou data protokolu zapisována do souboru *ID\_kolekce\_OmniFindQueryLog\_datum.log*, kde *ID\_kolekce* určuje kolekci, kterou

chcete konfigurovat, a *datum* je datem vytvoření souboru protokolu. Velikost tohoto souboru protokolu můžete zvýšit nebo snížit v závislosti na množství dat, která chcete protokolovat před vytvořením nového souboru protokolu.

### Postup

Chcete-li změnit velikost souboru protokolu zpracování dotazů, postupujte takto:

1. Přihlaste se jako administrátor podnikového vyhledávání. V systému podnikového vyhledávání s více servery se přihlaste k indexovému serveru.
2. Otevřete soubor `ES_ROOT_NODE/master_config/ID_kolekce.runtime.node1/runtime-generic.properties`.
3. Vyhledejte vlastnost **MaxFileSize**. Zvýšením nebo snížením její hodnoty zvýšte nebo snižte velikost souborů protokolu a uložte změny.
4. Pro systém podnikového vyhledávání s jedním serverem postupujte takto:
  - a. Použijte konzolu pro správu podnikového vyhledávání k monitorování změněné kolekce a zastavte vyhledávací servery.
  - b. Restartujte aplikaci `ESSearchServer` v produktu `WebSphere Application Server`.
  - c. V konzole pro správu restartujte zastavené vyhledávací servery.
  - d. Otevřete vyhledávací aplikaci v novém prohlížeči.
5. Pro systém podnikového vyhledávání s více servery postupujte takto:
  - a. Přihlaste se k indexovému serveru jako administrátor podnikového vyhledávání.
  - b. Zadáním následujících příkazů restartujte systém podnikového vyhledávání:

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

---


## Zobrazení souborů žurnálu


Můžete zobrazit zprávy žurnálu, které systémové komponenty a komponenty kolekce zapisují do společného souboru žurnálu. Lze také zobrazit zprávy specifické úrovně závažnosti a zprávy ze specifických relací podnikového vyhledávání, a to zadáním filtrů.

### Než začnete

Všichni administrativní uživatelé podnikového vyhledávání mohou zobrazit soubory žurnálu pro kolekce, k jejichž správě mají oprávnění. Chcete-li zobrazovat soubory žurnálu na úrovni systému, musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo mít oprávnění pro přístup k panelu nástrojů **Systém**.

### Postup


1. Zobrazení souborů žurnálu pro jednu kolekci:
  - a. Klepnutím na volbu **Kolekce** otevřete pohled Kolekce.
  - b. V seznamu kolekcí vyhledejte kolekci, kterou chcete zobrazit, klepněte na volbu  **Monitor**, a otevřete stránku Žurnál.

**Tip:** Pokud upravujete kolekci a nacházíte se již na stránce Log, můžete klepnutím na volbu  **Monitor** přejít do pohledu pro monitorování kolekce.

2. Zobrazení souborů žurnálu na úrovni systému:
  - a. Klepnutím na volbu **Systém** otevřete pohled Systém.
  - b. Vyberte stránku Žurnál.

3. V poli **Soubor žurnálu** vyberte soubor žurnálu, který chcete zobrazit. Název každého souboru žurnálu zahrnuje typ souboru žurnálu (například název systému nebo kolekce) a datum vytvoření souboru. Pokud byl vytvořen více než jeden soubor žurnálu stejného typu a se stejným datem, bude pořadí vytvoření souborů určeno číselnou příponou. Příklad:

*typ\_souboru\_žurnálu\_20060526.log (obsahuje nejnovější položky pro toto datum)*  
*typ\_souboru\_žurnálu\_20060526.log.1*  
*typ\_souboru\_žurnálu\_20060526.log.2 (obsahuje nejstarší položky pro toto datum)*  
*typ\_souboru\_žurnálu\_20060525.log (obsahuje nejnovější položky pro toto datum)*  
*typ\_souboru\_žurnálu\_20060525.log.1*  
*typ\_souboru\_žurnálu\_20060525.log.2*  
*typ\_souboru\_žurnálu\_20060525.log.3 (obsahuje nejstarší položky pro toto datum)*

4. Chcete-li zobrazit pouze zprávy s určitými úrovněmi závažnosti, zaškrtněte v poli **Závažnost** příslušná políčka.
5. Chcete-li zobrazit pouze zprávy z určitých relací, zaškrtněte v poli **Relace** příslušná políčka.
6. Klepněte na volbu **Zobrazit žurnál**.  
U každé zprávy na stránce Obsah souboru žurnálu je zobrazeno datum a čas jejího zadání, úroveň závažnosti, název relace, která zprávu zadala, a číslo zprávy a chybový text.  
Klepnutím na příslušné tlačítko můžete přejít na první stránku, poslední stránku, předchozí stránku nebo následující stránku souboru žurnálu. Můžete rovněž přejít na konkrétní stránku zadáním příslušného čísla stránky.
7. Chcete-li zobrazit podrobnější informace o zprávě, klepněte na volbu  **Podrobnosti**.  
Na stránce Podrobnosti zprávy v žurnálu je zobrazen název hostitele serveru podnikového vyhledávání, na kterém došlo k chybě, název souboru, který chybu generoval, název funkce a číslo řádku, kde chyba vznikla, ID procesu a ID podprocesu.  
Pomocí tlačítek se můžete přesunout na další a předchozí zprávy v souboru žurnálu.



---

## Zálohování a obnovování systému podnikového vyhledávání

Skripty zálohování a obnovování vám umožňují zálohovat a obnovovat systém podnikového vyhledávání.

### Co skripty zálohují

Skript zálohuje a obnovuje následující soubory:

- konfigurační soubory obsažené v adresáři `ES_NODE_ROOT/master_config`,
- databázové soubory prolézacích modulů včetně všech jejich metadat, například informací o tom, kdy byly zdroje dat naposledy prolézány,
- všechny soubory v adresáři `ES_NODE_ROOT/data`,
- indexové soubory pro kolekce konfigurované s použitím jiných než výchozích datových adresářů.

### Struktura záložních adresářů

Skript zálohování vytvoří následující podadresáře v adresáři určeném při spuštění skriptu. Jméno uživatele, který je administrátorem podnikového vyhledávání, musí být uděleno oprávnění pro zápis do určeného adresáře.

#### **master\_config**

Obsahuje konfigurační soubory z adresáře `ES_NODE_ROOT/master_config`.

#### **database**

Obsahuje databázové soubory z prolézacího serveru.

#### **data**

Obsahuje soubory indexu z indexového serveru.

### Pokyny pro použití

- Data zálohovaná v jednom počítači můžete obnovit v jiném počítači. Platí však následující omezení:
  - Soubory zálohované určitou verzí programu OmniFind Enterprise Edition nelze obnovit v systému, který používá jinou verzi programu OmniFind Enterprise Edition.
  - Data je možné obnovit pouze v systému, který obsahuje stejný nebo vyšší počet serverů podnikového vyhledávání. Zálohujete-li například systém podnikového vyhledávání běžící na jediném serveru, můžete záložní data obnovit v systému využívajícím dva nebo čtyři servery podnikového vyhledávání. Data zálohovaná v systému se čtyřmi servery však nelze obnovit v systému využívajícím pouze jeden nebo dva servery.
  - Soubory zálohované určitým operačním systémem nelze obnovit v systému, který používá jinou verzi operačního systému. Pokud jste například nainstalovali podnikové vyhledávání v systému AIX a nyní ho chcete provozovat v systému Linux, musíte instalovat nový systém podnikového vyhledávání na servery Linux.
- Před zálohováním sestavte hlavní index, aby bylo zaručeno, že záloha bude obsahovat nejaktuálnější indexovaná data.
- Veškeré položky nastavení instalačního adresáře (`ES_INSTALL_ROOT`), datového adresáře (`ES_NODE_ROOT`) a ID a hesla administrátora podnikového vyhledávání v systému, v němž data obnovujete, musí být totožné s nastavením zálohovaného systému.
- V případě konfigurace s více servery provádějte zálohování a obnovení systému z indexového serveru podnikového vyhledávání. Vzhledem k tomu, že veškerá data prolézacích modulů jsou umístěna v databázích na prolézacím serveru, skripty při zálohování a obnovování dat prolézacích modulů spouštějí vzdálené příkazy.

- Pro zálohování souborů systému podnikového vyhledávání do jiného adresáře potřebujete dostatek místa na disku. Skripty zálohování a obnovení neprovádějí kontrolu souborů.
- Během zpracování skriptů zálohování a obnovování jsou všechny relace systému zastaveny. Chcete-li zabránit zobrazování nesprávných nebo nekonzistentních informací o systému, nepoužívejte během zpracování skriptů konzolu pro správu podnikového vyhledávání.
- Dojde-li k selhání systému v důsledku nezotavitelné chyby, je nutné znovu instalovat produkt OmniFind Enterprise Edition a poté spustit skript obnovování.

---

## Zálohování systému podnikového vyhledávání

Systém podnikového vyhledávání lze zálohovat pomocí skriptu `esbackup.sh` pro systémy AIX, Linux a Solaris nebo pomocí skriptu `esbackup.bat` pro systémy Microsoft Windows.

### Omezení

Jméno uživatele, který je administrátorem podnikového vyhledávání, musí být uděleno oprávnění pro zápis do adresáře určeného při spouštění skriptu zálohování.

Během zpracování skriptů zálohování a obnovování jsou všechny relace systému zastaveny. Chcete-li zabránit zobrazování nesprávných nebo nekonzistentních informací o systému, nepoužívejte během zpracování skriptů konzolu pro správu podnikového vyhledávání.

**Upozornění:** Pokud zálohovací skript přerušíte stisknutím kombinace kláves Ctrl+C, systém se ocitne v nekonzistentním stavu. Chcete-li spustit všechny relace služeb a běžící relace, které byly během procesu zálohování zastaveny, musíte zadat následující příkaz:

```
esadmin system startall
```

### Postup

Chcete-li zálohovat systém podnikového vyhledávání, postupujte takto:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
2. Pokud není spuštěna služba CCL (Common Communication Layer) pro podnikové vyhledávání, spusťte ji:

#### Systém AIX, Linux nebo Solaris

```
startccl.sh -bg
```

#### Příkazový řádek systému Windows

```
startccl
```

#### Nástroj pro správu Služby systému Windows

Chcete-li spustit vrstvu CCL na pozadí, postupujte takto:

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.
3. V případě konfigurace s více servery se ujistěte, že je služba CCL spuštěna na všech serverech. Předchozí kroky opakujte podle potřeby.
  4. Spusťte zálohování následujícím příkazem, do něž místo textu *záložní\_adresář* doplňte název adresáře zálohovaných dat:

#### Systém AIX, Linux nebo Solaris

```
esbackup.sh záložní_adresář
```

**Příkazový řádek systému Windows**  
esbackup.bat *záložní\_adresář*

#### Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 347

---

## Obnovení systému podnikového vyhledávání

Po nové instalaci produktu OmniFind Enterprise Edition můžete obnovit systém podnikového vyhledávání pomocí skriptu esrestore.sh (AIX, Linux, Solaris) nebo esrestore.bat (Microsoft Windows).

### Omezení

Během zpracování skriptů zálohování a obnovování jsou všechny relace systému zastaveny. Chcete-li zabránit zobrazování nesprávných nebo nekonzistentních informací o systému, nepoužívejte během zpracování skriptů konzolu pro správu podnikového vyhledávání.

Soubory zálohované určitou verzí programu OmniFind Enterprise Edition nelze obnovit v systému, který používá jinou verzí programu OmniFind Enterprise Edition. Systém, do něž obnovená data ukládáte, musí navíc obsahovat stejný nebo vyšší počet serverů podnikového vyhledávání jako systém, v němž byla data zálohována.

### Postup

Chcete-li obnovit systém podnikového vyhledávání, postupujte takto:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
2. Pokud není spuštěna služba CCL (Common Communication Layer) pro podnikové vyhledávání, spusťte ji:

#### **Systém AIX, Linux nebo Solaris**

```
startccl.sh -bg
```

#### **Příkazový řádek systému Windows**

```
startccl
```

#### **Nástroj pro správu Služby systému Windows**

Chcete-li spustit vrstvu CCL na pozadí, postupujte takto:

- a. Spusťte ovládací panel Služby systému Windows: **Start** → **Programy** → **Nástroje pro správu** → **Služby**.
  - b. Klepněte pravým tlačítkem myši na položku **IBM OmniFind Enterprise Edition** a vyberte volbu **Spustit**.
3. V případě konfigurace s více servery se ujistěte, že je služba CCL spuštěna na všech serverech. Předchozí kroky opakujte podle potřeby.
  4. Zastavte vrstvu controller:  
esadmin stop
  5. Obnovte data podnikového vyhledávání zadáním následujícího příkazu, kde parametr *záložní\_adresář* určuje adresář, do kterého jste zálohovali soubory:

#### **Systém AIX, Linux nebo Solaris**

```
esrestore.sh záložní_adresář
```

#### **Příkazový řádek systému Windows**

```
esrestore.bat záložní_adresář
```

#### Související odkazy

## Export a import konfigurací kolekce

Je možné exportovat a importovat jednotlivé konfigurace kolekcí. Exportovány a importovány jsou pouze konfigurační soubory kolekcí, nikoliv data kolekcí.

Kolekci můžete exportovat z jednoho systému podnikového vyhledávání a importovat do jiného systému podnikového vyhledávání. Exportovat a importovat kolekce je však možné pouze mezi systémy, které používají stejnou verzi podnikového vyhledávání. Například je možné exportovat kolekci ze systému verze 8.4 a importovat ji do systému verze 8.4 nebo verze 8.4.0.150, ale není možné exportovat kolekci ze systému verze 8.3 a importovat ji do systému verze 8.4.

Exportovanou kolekci můžete importovat pouze do systému podnikového vyhledávání, který používá stejný operační systém. Například není možné exportovat kolekci ze systému podnikového vyhledávání nainstalovaného v operačním systému Linux a importovat ji do systému podnikového vyhledávání nainstalovaného v operačním systému Windows.

### Export kolekcí

Chcete-li exportovat kolekci, můžete ji exportovat pomocí příkazu **esadmin export**. Konzola pro správu podnikového vyhledávání neposkytuje žádnou podporu exportu kolekcí.

1. Přihlaste se jako administrátor podnikového vyhledávání. V konfiguraci s více servery se musíte přihlásit k indexovému serveru. (Při pokusu o spuštění tohoto příkazu z jiného serveru podnikového vyhledávání dojde k selhání příkazu.)
2. Zadejte následující příkazy:

```
esadmin export -help  
esadmin export -cid ID_kolekce [-fname název_souboru_exportu] [-verbose]
```

kde:

#### **-help**

Poskytuje nápovědu pro tento příkaz.

#### **-cid *ID\_kolekce***

Určuje ID exportované kolekce.

**Tip:** Chcete-li určit ID kolekce, kterou chcete exportovat, můžete použít konzolu pro správu podnikového vyhledávání nebo příkaz **esadmin report collections**.

#### **-fname *název\_souboru\_exportu***

Určuje cestu k souboru exportu. Pokud není uveden absolutní název souboru, předpokládá se použití adresáře `ES_NODE_ROOT/dump`. Pokud tuto volbu vynecháte, dojde k vytvoření souboru s následující konvencí pojmenování v adresáři `ES_NODE_ROOT/dump`:

```
_export_rrrrMMdd_HHmmsz.zip
```

kde:

*rrrrMMdd*

Jedná se o aktuální rok, měsíc a den spuštění příkazu exportu.

*HHmms*

Jedná se o aktuální hodinu, minutu a sekundu spuštění příkazu exportu.

*z* Jedná se o posun časového pásma od pásma GMT při spuštění příkazu exportu. Například pacifický standardní čas má hodnotu posunu GMT -0800.



### **-verbose**

Zobrazuje informace, které lze použít při odstraňování problémů operace exportu.

## **Import kolekcí**

Chcete-li importovat kolekci, můžete ji importovat pomocí příkazu **esadmin import**. Konzola pro správu podnikového vyhledávání neposkytuje žádnou podporu importu kolekcí.

1. Přihlaste se jako administrátor podnikového vyhledávání. V konfiguraci s více servery se musíte přihlásit k indexovému serveru. (Při pokusu o spuštění tohoto příkazu z jiného serveru podnikového vyhledávání dojde k selhání příkazu.)
2. Zadejte následující příkazy:

```
esadmin import -help
esadmin import -fname název_souboru_importu
                    [-cid ID_nové_kolekce]
                    [-name název_nové_kolekce]
                    [-colDataDir datový_adresář_nové_kolekce]
                    [-force]
                    [-verbose]
```

kde:

### **-help**

Poskytuje nápovědu pro tento příkaz.

### **-fname *název\_souboru\_importu***

Určuje cestu k souboru importu. Pokud není uveden absolutní název souboru, předpokládá se použití adresáře `ES_NODE_ROOT/dump`.

### **-cid *ID\_nové\_kolekce***

Určuje ID nové kolekce, pokud je kolekci nutné importovat s jiným ID.

**Tip:** Chcete-li určit ID kolekce, kterou chcete importovat, můžete použít konzolu pro správu podnikového vyhledávání nebo příkaz **esadmin report collections**.

### **-name *název\_nové\_kolekce***

Určuje název nové kolekce, pokud je kolekci nutné importovat s jiným názvem.

### **-colDataDir *datový\_adresář\_nové\_kolekce***

Určuje datový adresář kolekce. Při vynechání této volby je použit výchozí adresář.

### **-force**

Vynutí import kolekce, která již v cílovém systému existuje. Tento systém používá ID importované kolekce k určení kolekce, která má být v cílovém systému nahrazena.

### **Důležité:**

- Po importu kolekce nelze importovat kolekci se stejným ID kolekce znovu, pokud nepoužijete volbu vynucení importu kolekce.
- Kolekce, která byla importována pomocí vynucení, nezachovává stroje pro analýzu textu, slovníky a moduly plug-in prolézacího modulu. Tyto volby jsou přepsány informacemi obsaženými v importované kolekci.
- Dojde k odebrání prolézacích modulů přiřazených ke kolekci a jejich nahrazení prolézacími moduly zadanými v souboru importu. Vzhledem k odebrání prolézacích modulů dojde i k odebrání všech metadat a dokumentů prolézacího modulu, které ještě nebyly analyzovány. Po importu kolekce je nutné znovu prolézt všechny dokumenty a znovu tato data vytvořit pro importované prolézací moduly.

### **-verbose**

Zobrazuje informace, které lze použít při odstraňování problémů operace importu.

## **Pokyny pro použití**

### **Stroje pro analýzu textu a slovníky**

Libovolné stroje pro analýzu textu a slovníky (jako například slovníky synonym, ignorovaných slov a slov pro upřednostnění) přiřazené ke kolekci v systému exportu nejsou exportovány s konfiguračními daty kolekce. Exportována jsou pouze přidružení názvů ke strojům pro analýzu textu a slovníkům.

Pokud při importu existuje v cílovém systému stroj pro analýzu textu nebo slovník se stejným názvem, dojde k jeho přidružení k importované kolekci. V opačném případě se zobrazí varovná zpráva indikující, že nebylo možné vytvořit přidružení. Z tohoto důvodu byste měli odesílat stroje pro analýzu textu a slovníky na cílový systém a používat stejné názvy, které byly použity v importované kolekci.

Pokud v cílovém systému neexistuje stroj pro analýzu textu nebo slovník se stejným názvem, dojde k přerušení přidružení. Při přerušení přidružení bude kolekce v cílovém systému fungovat správně, ale nebude používat odpovídající stroje pro analýzu textu nebo slovníky.

### **Modely paměti**

Model paměti konfigurovaný pro exportovaný systém není zachovávan. Model paměti konfigurovaný pro exportovaný systém zůstává ve stavu, jak byl konfigurován. Při pokusu o import kolekce do systému konfigurovaného pro menší model paměti se zobrazí varovná zpráva. Kolekce nemusí pracovat podle očekávání a může dojít k významnému ovlivnění používání prostředků v menším systému. Chcete-li se vyhnout problémům, importujte kolekci do systému, který používá stejný nebo větší model paměti než systém exportu.

### **Prolézací moduly**

Je nutné se ujistit, že je v cílovém systému instalován a konfigurován předpokládaný software pro prolézací modul.

Importované prolézací moduly fungují pouze v případě, že jsou k dispozici zdroje dat, které tyto prolézací moduly prolézají. Například pokud kolekce zahrnuje prolézací modul systému souborů, který prolézá systém souborů, nemůže tento prolézací modul zahájit prolézání, pokud v cílovém systému neexistuje stejná struktura adresářů.

Moduly plug-in prolézacího modulu nejsou exportovány. Při importu kolekce se zobrazí varovná zpráva a import kolekce pokračuje. Po importu kolekce aktualizujte vlastnosti prolézacího modulu a přiřaďte mu moduly plug-in. Aby prolézací modul správně fungoval, je nutné zajistit, že jsou moduly plug-in prolézacího modulu v cílovém systému implementovány.

### **Související odkazy**

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 347

---

## Integrace s produktem Lotus Notes verze 8

Vyhledávací funkce implementací klientů produktu IBM Lotus Notes verze 8 můžete rozšířit pomocí implementace modulu plug-in vyhledávání produktu OmniFind Enterprise Edition Lotus Notes. Tento modul plug-in umožňuje uživatelům vyhledávat kolekce podnikového vyhledávání pomocí panelu vyhledávání klienta produktu Lotus Notes.

Chcete-li integrovat možnosti podnikového vyhledávání s produktem Lotus Notes, je nutné vytvořit web aktualizace modulu plug-in. Uživatelé poté mohou tento modul plug-in instalovat do lokálních klientských systémů produktu Lotus Notes verze 8.

---

### Vytváření webu aktualizace modulu plug-in podnikového vyhledávání

Chcete-li integrovat podnikové vyhledávání s produktem Lotus Notes verze 8, je nutné implementovat web aktualizace produktu Eclipse podnikového vyhledávání na webový server ve vaší organizaci. Web aktualizace umožňuje uživatelům implementovat modul plug-in podnikového vyhledávání do lokálních instalací klienta produktu Lotus Notes.

#### O této úloze

Soubory vyžadované k vytvoření webu aktualizace podnikového vyhledávání jsou poskytovány v souboru `com.ibm.es.notes.search.plugin_8.4.0.150.zip`.

#### Postup

Chcete-li vytvořit web aktualizace produktu Eclipse podnikového vyhledávání, postupujte takto:

1. Zkopírujte soubor `com.ibm.es.notes.search.plugin_8.4.0.150.zip` na webový server ve vaší organizaci.
2. Rozbalte obsah souboru typu ZIP do podadresáře sdíleného adresáře v cestě k instalaci webového serveru.

Chcete-li například implementovat web aktualizace na webový server, který používáte k podnikovému vyhledávání, rozbalte obsah souboru typu ZIP do adresáře `C:\Program Files\IBM\HTTP Server\htdocs\en_US\omnifind`.

---

### Instalace modulu plug-in podnikového vyhledávání v klientovi produktu Lotus Notes verze 8

Chcete-li použít funkce podnikového vyhledávání pro zpracování dotazů, můžete přidat modul plug-in podnikového vyhledávání do vyhledávacího panelu produktu Lotus Notes.

#### O této úloze

V pravém horním rohu klienta produktu Lotus Notes je umístěn vyhledávací panel. Vyhledávací panel obsahuje několik modulů plug-in vyhledávání poskytovaných se základní instalací klienta produktu Lotus Notes. Po instalaci modulu plug-in podnikového vyhledávání můžete prohledávat kolekce podnikového vyhledávání pomocí vyhledávacího panelu.

#### Postup

Chcete-li instalovat modul plug-in podnikového vyhledávání do klienta produktu Lotus Notes, postupujte takto:

1. Otevřete soubor NOTES\_INSTALL\_ROOT\notes.ini, kde NOTES\_INSTALL\_ROOT je obvykle c:\lotus\notes.
2. Na konec souboru přidejte vlastnost s názvem OMNIFIND\_ENTERPRISE\_EDITION\_SEARCH\_SERVER\_URL a zadejte adresu `http://název_hostitele:port/ESSearchApplication/search.do?q=` pro danou hodnotu, kde *název\_hostitele* je názvem hostitele vyhledávacího serveru podnikového vyhledávání a *port* je portem webového serveru.
3. Uložte soubor a zavřete jej.
4. Otevřete klientskou aplikaci produktu Lotus Notes.
5. Vyberte volbu **Soubor** → **Aplikace** → **Instalovat**.
6. Vyberte přepínač **Vyhledat nové funkce k instalaci**.
7. Klepněte na tlačítko **Přidat vzdálené umístění**.
8. Zadejte název nového umístění, jako například podnikové vyhledávání. Do pole Adresa URL zadejte kořenovou adresu URL webu aktualizace podnikového vyhledávání. Pokud jste například rozbalili soubor com.ibm.es.notes.search.plugin\_8.4.0.150.zip do podadresáře omnifind v kořenovém adresáři anglické verze serveru IBM HTTP Server, má zadaná adresa URL formát `http://název_hostitele:port/omnifind/`.
9. Klepnutím na tlačítko **Dokončit** uložte nový vzdálený web.
10. Zaškrtněte políčko vedle názvu webu a klepněte na tlačítko **Další**.
11. Přezkoumejte licenční smlouvu, vyberte volbu **Souhlasím s požadavky licenční smlouvy** a klepněte na tlačítko **Další**.
12. Klepnutím na tlačítko **Dokončit** nainstalujte modul plug-in. Pokud obdržíte varování, že instalujete nepodepsaný modul plug-in, vyberte přepínač **Instalovat tento modul plug-in** a klepněte na tlačítko **OK**.
13. Při zobrazení výzvy restartujte klienta produktu Lotus Notes klepnutím na tlačítko **Ano**.

Nyní můžete klepnout na nabídku vedle vyhledávacího panelu produktu Lotus Notes a vybrat produkt **OmniFind Enterprise Edition** jako volby vyhledávání. Pokud zadáte termín dotazu a klepnete na tlačítko **Vyhledat**, bude požadavek směrován na server podnikového vyhledávání a výsledky se zobrazí v okně webového prohlížeče produktu Lotus Notes.

---

## Integrace s produktem WebSphere Portal

Vyhledávací funkce produktu IBM WebSphere Portal můžete rozšířit implementací vyhledávacího portletu pro portlety podnikového vyhledávání v produktu WebSphere Portal a konfigurováním produktu WebSphere Portal tak, aby používal vyhledávací portlet jako výchozí vyhledávací stroj.

### Body integrace

Instalační program OmniFind Enterprise Edition poskytuje instalační skripty pro integraci podnikového vyhledávání do portálu WebSphere Portal. Po spuštění těchto skriptů může být váš systém podnikového vyhledávání integrován s portálem WebSphere několika způsoby:

#### Vyhledávací portlet pro podnikové vyhledávání

Produkt WebSphere Portal poskytuje uživatelům jednotný přístupový bod pro komunikaci s aplikacemi, obsahem, procesy a jinými osobami. Rámec tvořený produktem WebSphere Portal umožňuje integraci a implementaci nových aplikací, nazývaných portlety, bez ovlivnění dalších aplikací v portálu.

Pokud implementujete vyhledávací portlet podnikového vyhledávání v produktu WebSphere Portal, můžete pomocí rozhraní tohoto produktu prohledávat kolekce podnikového vyhledávání a pracovat s výsledky vyhledávání. Prostřednictvím nastavení konfigurace produktu WebSphere můžete u portletu podnikového vyhledávání docílit stejného vzhledu a stylu, jaký je použit u dalších portletů v prostředí produktu WebSphere Portal.

#### Centrum vyhledávání produktu WebSphere Portal

Centrum vyhledávání produktu WebSphere Portal představuje výchozí bod prohledávání všech zdrojů dostupných pro prohledávání pomocí produktu WebSphere Portal. Centrum vyhledávání a univerzální vyhledávací portlet vám umožňují prohledávat obsah produktu WebSphere Portal a všech dalších kolekcí, které byly zaregistrovány v centru vyhledávání.

Po spuštění instalačních skriptů, které integrují podnikové vyhledávání s portálem WebSphere verze 5.1, je na stránku v rozhraní vyhledávacího centra přidána stránka Podnikové vyhledávání. Tuto stránku můžete vybrat, chcete-li prohledávat pouze kolekce podnikového vyhledávání. Také můžete zadat dotaz prohledávající kolekce podnikového vyhledávání spolu s dalšími kolekcemi dostupnými v centru vyhledávání.

Po spuštění instalačních skriptů, které integrují podnikové vyhledávání s portálem WebSphere verze 6, jsou funkce podnikového vyhledávání integrovány jako sdružená služba, kterou můžete použít k prohledávání kolekcí podnikového vyhledávání a dalších kolekcí dostupných v centru vyhledávání.

#### Panel vyhledávání produktu WebSphere Portal

V pravém horním rohu všech motivů portálového rozhraní WebSphere je umístěn vyhledávací panel. Ve výchozím stavu tento panel směřuje veškeré vyhledávací požadavky na výchozí vyhledávací stroj centra vyhledávání. Chcete-li pro zpracování dotazů používat výkonnější funkce podnikového vyhledávání, můžete toto výchozí chování změnit tak, že budou všechny vyhledávací požadavky přesměrovány na vyhledávací portlet podnikového vyhledávání.

#### Prolézací moduly WebSphere Portal a Web Content Management

Chcete-li do indexu podnikového vyhledávání zahrnout portálové weby WebSphere a weby spravované produktem IBM Workplace Web Content Management, můžete

pomocí konzoly pro správu podnikového vyhledávání konfigurovat prolézací moduly WebSphere Portal a Web Content Management. Poté můžete prohledávat indexovaný obsah pomocí portletu podnikového vyhledávání nebo vyhledávací aplikace.

Prolézací modul WebSphere Portal může prolézat weby serveru WebSphere Portal verze 5.1 a 6. Prolézací modul Web Content Management může prolézat weby serveru WebSphere Portal verze 6.

#### **Dokumenty IBM Lotus Quickr**

Chcete-li do indexu podnikového vyhledávání zahrnout dokumenty produktu IBM Lotus Quickr documents in an enterprise search index, můžete pomocí konzoly pro správu podnikového vyhledávání konfigurovat prolézací modul Počáteční seznam. Poté můžete prohledávat indexovaný obsah pomocí portletu podnikového vyhledávání v rámci produktu WebSphere Portal nebo samostatné vyhledávací aplikace.

Prolézací modul Počáteční seznam může prolézat knihovny obsahu (dokumenty) produktu Lotus Quickr na serveru WebSphere Portal verze 6.

### **Výhody integrace**

Podnikové vyhledávání obohacuje vyhledávací prostředí produktu WebSphere Portal tím, že je doplňuje o podporu vyhledávání širší škály typů zdrojů dat. Pomocí vyhledávacího portletu pro podnikové vyhledávání můžete prohledávat weby i všechny ostatní typy zdrojů dat podporované systémem podnikového vyhledávání.

Podnikové vyhledávání rovněž nabízí výhody v oblasti rozšiřitelnosti. Portálový vyhledávací stroj je užitečný pro podniky malé a střední velikosti, v nichž ke zvládnutí zátěže spojené s vyhledáváním a načítáním postačuje jediný server. Podporu kapacit na úrovni podniku lze zajistit rozdělením zátěže spojené s podnikovým vyhledáváním mezi více serverů, přičemž dva servery poskytují podporu vyhledávání a načítání.

---

## **Instalační skripty integrace podnikového vyhledávání do portálu WebSphere Portal**

Pro integraci podnikového vyhledávání s produktem IBM WebSphere Portal je třeba spustit konfigurační skripty dodávané s instalačním programem OmniFind Enterprise Edition.

Soubor JAR obsahující konfigurační skripty pro vaši verzi portálu WebSphere musíte zkopírovat ze serveru podnikového vyhledávání na server, na kterém je nainstalován produkt WebSphere Portal. Konfigurační skripty provádějí následující činnosti:

- implementace souborů EAR, které vám umožní používat podnikové vyhledávání v rámci portálu WebSphere a vytvářet prolézací moduly pro přidání obsahu produktů WebSphere Portal a IBM Workplace Web Content Management do kolekcí podnikového vyhledávání,
- implementace souborů WAR vyžadovaných portletem podnikového vyhledávání,
- vytvoření stránek v portálu WebSphere a přiřazení souborů portletu podnikového vyhledávání k těmto stránkám,
- zkopírování všech vyžadovaných souborů JAR do instalačních adresářů portálu WebSphere (soubory JAR, které jsou již v instalačních adresářích obsaženy, jsou před zkopírováním souborů JAR používaných pro podnikové vyhledávání zálohovány),
- vytvoření integračního bodu pro prohledávání dokumentů správce PDM produktem WebSphere Information Integrator Content Edition.

Po spuštění skriptů je nutné aktualizovat vlastnosti vyhledávacího portletu pomocí administračního rozhraní portálu WebSphere a zadat informace o vyhledávacím serveru pro podnikové vyhledávání.

## Pokyny pro použití

- Skripty nastavují všechny integrační body mezi podnikovým vyhledáváním a portálem WebSphere. Nemůžete tedy například nainstalovat pouze portlet a nenainstalovat přitom soubory EAR, které zajišťují podporu prolézacích modulů WebSphere Portal a Web Content Management.
- Pokud nenainstalujete produkt WebSphere Information Integrator Content Edition a později se rozhodnete pro prohledávání dokumentů PDM používat portlet, musíte spustit skript, který odebere podnikové vyhledávání z portálu WebSphere. Poté můžete znovu spustit konfigurační skript a zadat instalační cestu produktu WebSphere Information Integrator Content Edition.
- Skripty zastaví a znovu spustí portál WebSphere. Proto byste je měli spouštět mimo běžnou pracovní dobu, aby uživatelé neměli problémy s nedostupností portálových služeb.
- Pokud při práci konfiguračních skriptů dojde k chybám, spusťte je znovu. Úlohy, které byly při prvním pokusu úspěšně dokončeny, mohou vrátit chyby, proces konfigurování však bude pokračovat a budou provedeny zbývající úlohy.
- Při prvním přístupu ke stránce s portletem podnikového vyhledávání po spuštění konfiguračního skriptu může být zobrazování stránky pomalejší, protože systém musí zkompilovat soubory JSP (Java Server Pages) portletu.

---

## Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 5.1

K integraci systému podnikového vyhledávání s produktem WebSphere Portal verze 5.1.0 nebo vyšší se používá skript `wp5_install`.

### O této úloze

Soubory potřebné pro integraci podnikového vyhledávání s produktem WebSphere Portal jsou dodávány v souboru `es.wp5.install.jar`. Po rozbalení tohoto souboru získáte následující soubory:

- `ESSearchPortlet.war`
- `ESSearchAdapterPortlet.war`
- `ESSearchAdapter.ear`
- `ESPACServer.ear`
- `esapi.jar`
- `siapi.jar`
- `es.security.jar`
- Soubory skriptů, dávek, XML a JACL potřebné pro instalaci

### Postup

Chcete-li integrovat podnikové vyhledávání se systémem WebSphere Portal verze 5.1, postupujte takto:

1. Zkopírujte soubor `es.wp5.install.jar` ze serveru podnikového vyhledávání na portálový server WebSphere a poté jej rozbalte příkazem **JAR** jazyka Java (nebo příkazem **TAR**).
2. Volitelné: Pokud chcete zavést podporu integrace se správcem WebSphere PDM (Portal Document Manager), proveďte některý z následujících kroků:
  - Spusťte instalační program WebSphere Information Integrator Content Edition, vyberte volbu provádějící pouze instalaci konektoru a nainstalujte konektor PDM na portálový server WebSphere.

- Vytvořte na portálovém serveru WebSphere adresářovou strukturu WebSphere Information Integrator Content Edition a zkopírujte na portálový server WebSphere následující soubory z existující instalace produktu WebSphere Information Integrator Content Edition, kde CE\_ROOT určuje kořenový instalační adresář produktu WebSphere Information Integrator Content Edition:

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

3. Z příkazového řádku spusíte příkaz **wp5\_install.bat** (v systémech Windows) nebo **wp5\_install.sh** (v systémech AIX, Linux a Solaris). V následujícím příkladu jsou parametry pro přehlednost uvedeny na samostatných řádcích, při zadávání příkazu je však musíte uvést na jednom řádku:

```
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**

Úplná cesta k instalačnímu adresáři produktu WebSphere Application Server.

#### **WASUser**

Jméno správce produktu WebSphere Application. Je vyžadováno, pouze pokud je na serveru WebSphere Application Server zapnuto globální zabezpečení.

#### **WASPassword**

Heslo správce produktu WebSphere Application Server, je-li zadán.

#### **WPSDir**

Úplná cesta k instalačnímu adresáři produktu WebSphere Portal.

#### **WPSUser**

Jméno správce produktu WebSphere Portal.

#### **WSPassword**

Heslo zadaného správce produktu WebSphere Portal.

#### **WPSHost**

Název hostitele a číslo portu serveru WebSphere Portal.

#### **IICEDir**

Úplná cesta k instalačnímu adresáři produktu WebSphere Information Integrator Content Edition. Tento parametr je vyžadován jen tehdy, jestliže jste na portálový server WebSphere předem nainstalovali konektor správce PDM.

4. Po spuštění skriptu (proběhne zastavení a nové spuštění produktu WebSphere Portal) aktualizujte portlet podnikového vyhledávání pro identifikaci vyhledávacího serveru:
  - a. Přihlaste se k portálu WebSphere s použitím jména a hesla administrátora portálu.
  - b. V pravém horním rohu klepněte na položku **Správa**.
  - c. V navigační oblasti vlevo klepněte na volbu **Správa portletů** a poté klepněte na volbu **Portlety**.
  - d. Volbu **Hledat podle** změňte na **Název obsahuje**.
  - e. Do pole **Vyhledat** zadejte text podnikové vyhledávání a poté klepněte na tlačítko **Vyhledat**.
  - f. Po zobrazení nových ikon vpravo klepněte na ikonu klíče. Tím otevřete okno pro úpravu vyhledávacího portletu podnikového vyhledávání.



g. V seznamu parametrů portletu upravte následující parametry:

**hostname**

Zadejte úplný název hostitele vyhledávacího serveru pro podnikové vyhledávání.

**port**

Zadejte číslo portu používané produktem WebSphere Application Server na vyhledávacím serveru podnikového vyhledávání. Výchozí hodnota je 80 (výchozí hodnota pro komunikaci SSL je 443).

**username**

Pokud je v produktu WebSphere Application Server na vyhledávacím serveru zapnuto globální zabezpečení, zadejte jméno uživatele, které je platné v registru produktu WebSphere Application Server.

**password**

Pokud jste zadali jméno uživatele produktu WebSphere Application Server, zadejte odpovídající heslo.

**protocol**

Zadejte protokol používaný pro komunikaci mezi produktem WebSphere Portal a vyhledávacím serverem. Výchozí hodnota je HTTP. Používáte-li zabezpečení SSL, zadejte hodnotu HTTPS.

**trustStore**

Používáte-li zabezpečení SSL, zadejte úplnou cestu (včetně názvu souboru) pro úložiště certifikátů SSL.

**trustPassword**

Používáte-li zabezpečení SSL, zadejte heslo pro zadaný soubor trustStore.

**ssoCookieName**

Zadejte název souboru cookie obsahujícího řetězec tokenu jednotného přihlášení (SSO). Výchozí hodnota je LtpaToken.

**proxyHost**

Pokud je pro přístup k vyhledávacímu serveru v rámci podnikového vyhledávání vyžadován server proxy, zadejte úplný název hostitele serveru proxy.

**proxyPort**

Pokud jste zadali server proxy, zadejte číslo portu serveru proxy.

**proxyUser**

Pokud server proxy vyžaduje základní ověřování, zadejte jméno uživatele, které se používá k přihlášení k serveru proxy.

**proxyPassword**

Pokud jste zadali jméno uživatele, zadejte odpovídající heslo.

h. Uložte změny klepnutím na tlačítko **OK**.

## Konfigurování vyhledávacího panelu produktu WebSphere Portal verze 5.1 pro práci s podnikovým vyhledáváním

Produkt WebSphere Portal verze 5.1.0 a vyšší můžete konfigurovat tak, aby bylo ke zpracování dotazů zaslých uživateli z vyhledávacího panelu použito podnikové vyhledávání namísto výchozího vyhledávacího stroje portálu WebSphere.

### Než začnete

Před přesměrováním vyhledávacích požadavků na podnikové vyhledávání je nutné spustit konfigurační skript **wp5\_install**, který zajistí integraci podnikového vyhledávání s portálem



6. Zastavte a znovu spusíte instanci aplikačního serveru portálu WebSphere.

## Odebrání podnikového vyhledávání z produktu WebSphere Portal verze 5.1

K odebrání podnikového vyhledávání se systému WebSphere Portal verze 5.1.0 nebo vyšší se používá skript `wp5_uninstall`.

### O této úloze

Odeberete-li podnikové vyhledávání z portálu WebSphere, nebudou uloženy parametry portletu podnikového vyhledávání, které jste zadali při instalaci.

Po spuštění skript zastaví portálový server WebSphere. Po odebrání softwaru podnikového vyhledávání skript znovu spustí portálový server WebSphere.

### Postup

Chcete-li odebrat podnikové vyhledávání ze systému WebSphere Portal verze 5.1, postupujte takto:

Z příkazového řádku spusíte příkaz **wp5\_uninstall.bat** (v systémech Windows) nebo **wp5\_uninstall.sh** (v systémech AIX, Linux a Solaris). V následujícím příkladu jsou parametry pro přehlednost uvedeny na samostatných řádcích, při zadávání příkazu je však musíte uvést na jednom řádku:

```
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

Úplná cesta k instalačnímu adresáři produktu WebSphere Application Server.

### WASUser

Jméno správce produktu WebSphere Application. Je vyžadováno, pouze pokud je na serveru WebSphere Application Server zapnuto globální zabezpečení.

### WASPassword

Heslo správce produktu WebSphere Application Server, je-li zadán.

### WPSDir

Úplná cesta k instalačnímu adresáři produktu WebSphere Portal.

### WPSUser

Jméno správce produktu WebSphere Portal.

### WSPassword

Heslo zadaného správce produktu WebSphere Portal.

### WPSHost

Název hostitele a číslo portu serveru WebSphere Portal.

---

## Nastavení podnikového vyhledávání v produktu WebSphere Portal verze 6

K integraci systému podnikového vyhledávání s produktem WebSphere Portal verze 6 se používá skript `wp6_install`.

### O této úloze

Soubory potřebné pro integraci podnikového vyhledávání s produktem WebSphere Portal jsou dodávány v souboru `es.wp6.install.jar`. Po rozbalení tohoto souboru získáte následující soubory:

- `ESSearchPortlet.war`
- `ESPACServer.ear`
- `esapi.jar`
- `es.search.provider.jar`
- `es.security.jar`
- Ikony typů zdroje vyhledávacích aplikací, používané na stránce výsledků poskytovatele vyhledávání
- Soubory skriptů, dávek, XML a JACL potřebné pro instalaci

### Postup

Chcete-li integrovat podnikové vyhledávání se systémem WebSphere Portal verze 6, postupujte takto:

1. Zkopírujte soubor `es.wp6.install.jar` ze serveru podnikového vyhledávání na portálový server WebSphere a poté jej rozbalte příkazem **JAR** jazyka Java (nebo příkazem **TAR**).
2. Volitelné: Pokud chcete zavést podporu integrace se správcem WebSphere PDM (Portal Document Manager), proveďte některý z následujících kroků:
  - Spusťte instalační program WebSphere Information Integrator Content Edition, vyberte volbu provádějící pouze instalaci konektoru a nainstalujte konektor PDM na portálový server WebSphere.
  - Vytvořte na portálovém serveru WebSphere adresářovou strukturu WebSphere Information Integrator Content Edition a zkopírujte na portálový server WebSphere následující soubory z existující instalace produktu WebSphere Information Integrator Content Edition, kde `CE_ROOT` určuje kořenový instalační adresář produktu WebSphere Information Integrator Content Edition:

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

3. Z příkazového řádku spusťte příkaz **wp6\_install.bat** (v systémech Windows) nebo **wp6\_install.sh** (v systémech AIX, Linux a Solaris). V následujícím příkladu jsou volby pro přehlednost uvedeny na samostatných řádcích, při zadávání příkazu je však musíte uvést na jednom řádku:

```
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

Úplná cesta k adresáři profilu produktu WebSphere Portal. Výchozí cestou je v systémech AIX cesta /usr/IBM/WebSphere/AppServer/profiles/wp\_profile, v systémech Linux a Solaris cesta /opt/IBM/WebSphere/AppServer/profiles/wp\_profile a v systémech Windows cesta C:\Program Files\IBM\WebSphere\profiles\wp\_profile.

### WASDir

Úplná cesta ke kořenovému adresáři produktu WebSphere Application Server; je vyžadována pouze v systémech AIX, Linux a Solaris. Výchozí cestou ke kořenovému adresáři je v systémech AIX cesta /usr/IBM/WebSphere/AppServer, v systémech Linux a Solaris cesta /opt/IBM/WebSphere/AppServer a v systémech Windows cesta C:\Program Files\IBM\WebSphere\AppServer.

### WASUser

Jméno správce produktu WebSphere Application. Je vyžadováno, pouze pokud je na serveru WebSphere Application Server zapnuto globální zabezpečení.

### WASPassword

Heslo správce produktu WebSphere Application Server, je-li zadán.

### WPSDir

Úplná cesta k instalačnímu adresáři produktu WebSphere Portal.

### WPSUser

Jméno správce produktu WebSphere Portal.

### WSPassword

Heslo zadaného správce produktu WebSphere Portal.

### WPSHost

Název hostitele a číslo portu serveru WebSphere Portal.

### IICEDir

Úplná cesta k instalačnímu adresáři produktu WebSphere Information Integrator Content Edition. Tento parametr je vyžadován jen tehdy, jestliže jste na portálový server WebSphere předem nainstalovali konektor správce PDM.

4. Po spuštění skriptu a zastavení a novém spuštění produktu WebSphere Portal aktualizujte portlet podnikového vyhledávání pro identifikaci vyhledávacího serveru:
  - a. Přihlaste se k portálu WebSphere s použitím jména a hesla administrátora portálu.
  - b. Klepněte na položku **Správa** v levém dolním rohu.
  - c. V navigační oblasti vlevo klepněte na volbu **Správa portletů** a poté klepněte na volbu **Portlety**.
  - d. Volbu **Hledat podle** změňte na **Název obsahuje**.
  - e. Do pole **Vyhledat** zadejte text podnikové vyhledávání a poté klepněte na tlačítko **Vyhledat**.
  - f. Po zobrazení nových ikon vpravo klepněte na ikonu klíče. Tím otevřete okno pro úpravy vyhledávacího portletu podnikového vyhledávání.
  - g. V seznamu parametrů portletu upravte následující parametry:

#### hostname

Zadejte úplný název hostitele vyhledávacího serveru pro podnikové vyhledávání.

#### port

Zadejte číslo portu používané produktem WebSphere Application Server na vyhledávacím serveru podnikového vyhledávání. Výchozí hodnota je 80 (výchozí hodnota pro komunikaci SSL je 443).

**username**

Pokud je v produktu WebSphere Application Server na vyhledávacím serveru zapnuto globální zabezpečení, zadejte jméno uživatele, které je platné v registru produktu WebSphere Application Server.

**password**

Pokud jste zadali jméno uživatele produktu WebSphere Application Server, zadejte odpovídající heslo.

**protocol**

Zadejte protokol používaný pro komunikaci mezi produktem WebSphere Portal a vyhledávacím serverem. Výchozí hodnota je HTTP. Používáte-li zabezpečení SSL, zadejte hodnotu HTTPS.

**trustStore**

Používáte-li zabezpečení SSL, zadejte úplnou cestu (včetně názvu souboru) pro úložiště certifikátů SSL.

**trustPassword**

Používáte-li zabezpečení SSL, zadejte heslo pro zadaný soubor trustStore.

**ssoCookieName**

Zadejte název souboru cookie obsahujícího řetězec tokenu jednotného přihlášení (SSO). Výchozí hodnota je LtpaToken.

**proxyHost**

Pokud je pro přístup k vyhledávacímu serveru v rámci podnikového vyhledávání vyžadován server proxy, zadejte úplný název hostitele serveru proxy.

**proxyPort**

Pokud jste zadali server proxy, zadejte číslo portu serveru proxy.

**proxyUser**

Pokud server proxy vyžaduje základní ověřování, zadejte jméno uživatele, které se používá k přihlášení k serveru proxy.

**proxyPassword**

Pokud jste zadali jméno uživatele, zadejte odpovídající heslo.

- h. Uložte změny klepnutím na tlačítko **OK**.

## Konfigurování vyhledávacího centra produktu WebSphere Portal verze 6 pro podnikové vyhledávání

Produkt WebSphere Portal verze 6 můžete konfigurovat tak, aby byly při zpracování dotazů zaslaných uživateli z vyhledávacího centra portálu WebSphere prohledávány kolekce podnikového vyhledávání.

### Omezení

Pokud chce uživatel prohledávat zabezpečené kolekce podnikového vyhledávání, musí spustit vyhledávací portlety pro podnikové vyhledávání a nastavit konfiguraci uživatelského profilu. Profil je zašifrovaný a uchovává se v zabezpečeném úložišti podnikového vyhledávání. Teprve po vytvoření profilu mohou uživatelé začít zasílat dotazy na zabezpečené kolekce z vyhledávacího centra portálu WebSphere.

### O této úloze

Centrum vyhledávání v produktu WebSphere Portal verze 6 podporuje funkce sdruženého vyhledávání ve více kolekcích. Kolekce mohou obsahovat různé typy dat, např. knihovny

dokumentů portálu a portálový obsah (stránky a portlety). Po spuštění instalačních skriptů, které integrují podnikové vyhledávání s portálem WebSphere, lze konfigurovat centrum vyhledávání tak, aby prohledávalo také kolekce podnikového vyhledávání.

## Postup

Chcete-li konfigurovat centrum vyhledávání tak, aby prohledávalo kolekce podnikového vyhledávání, postupujte takto:

1. Přihlaste se k portálu WebSphere s použitím jména a hesla administrátora portálu.
2. Klepněte na položku **Správa** v levém dolním rohu.
3. V navigační oblasti vlevo klepněte na volbu **Správa vyhledávání** a poté klepněte na volbu **Spravovat vyhledávání**.
4. Klepněte na volbu **Vyhledávací služby** a poté na volbu **Nová vyhledávací služba**.
5. V poli **Implementace vyhledávací služby** vyberte vyhledávací službu Podnikové vyhledávání a poté zadejte název, který chcete službě přidělit, do textového pole **Název služby**.
6. V seznamu parametrů upravte následující parametry:

### hostname

Zadejte úplný název hostitele vyhledávacího serveru pro podnikové vyhledávání.

### port

Zadejte číslo portu používané produktem WebSphere Application Server na vyhledávacím serveru podnikového vyhledávání. Výchozí hodnota je 80 (výchozí hodnota pro komunikaci SSL je 443).

### username

Pokud je v produktu WebSphere Application Server na vyhledávacím serveru zapnuto globální zabezpečení, zadejte jméno uživatele, které je platné v registru produktu WebSphere Application Server.

### password

Pokud jste zadali jméno uživatele produktu WebSphere Application Server, zadejte odpovídající heslo.

### protocol

Zadejte protokol používaný pro komunikaci mezi produktem WebSphere Portal a vyhledávacím serverem. Výchozí hodnota je HTTP. Používáte-li zabezpečení SSL, zadejte hodnotu HTTPS.

### trustStore

Používáte-li zabezpečení SSL, zadejte úplnou cestu (včetně názvu souboru) pro úložiště certifikátů SSL.

### trustPassword

Používáte-li zabezpečení SSL, zadejte heslo pro zadaný soubor trustStore.

### ssoCookieName

Zadejte název souboru cookie obsahujícího řetězec tokenu jednotného přihlášení (SSO). Výchozí hodnota je LtpaToken.

### proxyHost

Pokud je pro přístup k vyhledávacímu serveru v rámci podnikového vyhledávání vyžadován server proxy, zadejte úplný název hostitele serveru proxy.

### proxyPort

Pokud jste zadali server proxy, zadejte číslo portu serveru proxy.

### proxyUser

Pokud server proxy vyžaduje základní ověřování, zadejte jméno uživatele, které se používá k přihlášení k serveru proxy.

### proxyPassword

Pokud jste zadali jméno uživatele, zadejte odpovídající heslo.

7. Uložte změny klepnutím na tlačítko **OK**.

## Konfigurování vyhledávacího panelu produktu WebSphere Portal verze 6 pro práci s podnikovým vyhledáváním

Produkt WebSphere Portal verze 6 můžete konfigurovat tak, aby bylo ke zpracování dotazů zaslaných uživateli z vyhledávacího panelu použito podnikové vyhledávání namísto výchozího vyhledávacího stroje portálu WebSphere.

### Než začnete

Před přesměrováním vyhledávacích požadavků na podnikové vyhledávání je nutné spustit konfigurační skript **wp6\_install**, který zajistí integraci podnikového vyhledávání s portálem WebSphere. Dále je nutné aktualizovat parametry portletu podnikového vyhledávání tak, aby identifikovaly název hostitele, port a další informace o vyhledávacím serveru podnikového vyhledávání.

### O této úloze

V pravém horním rohu všech motivů portálového rozhraní WebSphere je umístěn vyhledávací panel. Ve výchozím stavu tento panel směřuje veškeré vyhledávací požadavky na portlet Centrum vyhledávání. Chcete-li pro zpracování dotazů používat výkonnější funkce podnikového vyhledávání, můžete toto výchozí chování změnit tak, že budou všechny vyhledávací požadavky přesměrovány na vyhledávací portlet podnikového vyhledávání.

Přesměrování vyhledávacího panelu ovlivní stránky, které využívají stejný motiv portálu WebSphere jako vyhledávací portlet pro podnikové vyhledávání a volají soubor `banner_searchControl.jspf`. Stránky využívající jiný motiv a stránky, které nevolají soubor `banner_searchControl.jspf`, budou i nadále používat výchozí portlet Centra vyhledávání.

Po provedení této úlohy nebudete moci Centrum vyhledávání používat, pokud provedené změny nezrušíte (například obnovením původního souboru `banner_searchControl.jspf`).

### Postup

Chcete-li ke zpracování dotazů odeslaných uživateli z panelu vyhledávání portálu WebSphere používat portlet podnikového vyhledávání, postupujte takto:

1. Zastavte instanci aplikačního serveru portálu WebSphere.
2. Na portálovém serveru WebSphere přejděte do adresáře `WPS_PROFILE_ROOT/installedApps/název_uzlu/wps.ear/wps.war/themes/html/název_aktuálního_tématu`, kde `název_uzlu` je název uzlu portálového serveru WebSphere a `název_aktuálního_tématu` je aktuálně použité téma pro portálový server WebSphere. Výchozím názvem tématu pro portálový server WebSphere je IBM.
3. Vytvořte zálohu souboru `banner_searchControl.jspf` tím, že tento soubor zkopírujete pod jiným názvem (například `banner_searchControl.jspf.BACKUP`).
4. Upravte soubor `banner_searchControl.jspf` a nahraďte jeho obsah následujícím textem. V atributu `action=` prvku formuláře nahraďte volbu `localhost:10038` názvem hostitele a číslem portu serveru WebSphere Portal.

```
<%@ taglib uri="/WEB-INF/tld/SearchMenuControl.tld" prefix="searchmenu" %>
<%String ic = (bidiImageRTL == null) ? "icons/scope_search_submit.gif" :
"icons/scope_search_submit"+bidiImageRTL+".gif";%>
```

```
<searchmenu:adminlinkinfo name="SEARCH_CENTER">
<div class="searchControl">
```



```

<form name="SearchForm" style="margin: 0px;" method="GET"
  action="http://localhost:10038/wps/omnifind/portalSearchBar.jsp">
<table border="0" cellpadding="0" cellspacing="0">
<tr>
<td><span class="wpsInstructionText">
  <portal-fmt:text key="search.theme.control.label" bundle="nls.engine"/></span></td>
<td valign="middle" style="padding: 0px 4px 0px 4px;">
  <input type="text" name="q"></input></td>
<td valign="middle"><input tabIndex="4" valign="middle"
  title="<portal-fmt:text key='search.theme.searchresultsicon.alttext' bundle='nls.engine' />"
  alt="<portal-fmt:text key='search.theme.searchresultsicon.alttext' bundle='nls.engine' />"
  src="<portal-logic:urlFindInTheme file=">" />" type="image"></input></td>
</tr>
</table>
</form>
</div>
</searchmenu:adminlinkinfo>

```

5. Otevřete soubor `banner.jspf` a uložte jej. Tento krok, který aktualizuje datum změny souboru, aby byl soubor znovu zkompileován, není povinný, pokud místo výchozího motivu portálu WebSphere používáte vlastní motiv.
6. Otevřete soubor `Default.jsp` a uložte jej.
7. Znovu spusťte instanci aplikačního serveru portálu WebSphere.

## Nastavení portletu podnikového vyhledávání pro produkt Lotus Quickr

Portlet podnikového vyhledávání v produktu WebSphere Portal verze 6 můžete nastavit k vyhledávání zdrojů produktu Lotus Quickr.

### Než začnete

Spusťte příkaz `wp6_install.bat` (v systémech Windows) nebo příkaz `wp6_install.sh` (v systémech AIX, Linux nebo Solaris) a pomocí uvedených postupů nastavte podnikové vyhledávání v produktu WebSphere Portal verze 6.

### Postup

Chcete-li nastavit portlet podnikového vyhledávání v produktu WebSphere Portal verze 6 k vyhledávání zdrojů produktu Lotus Quickr, postupujte takto:

1. Aktualizujte parametry portletu v konfiguraci produktu WebSphere Portal:
  - a. Přihlaste se k portálu WebSphere s použitím jména a hesla administrátora produktu Lotus Quickr.
  - b. Klepněte na volbu **Administrace webu** a poté klepněte na volbu **Rozšířená administrace**.
  - c. V navigační oblasti vlevo klepněte na volbu **Správa portletů** a poté klepněte na volbu **Portlety**.
  - d. Volbu **Hledat podle** změňte na **Název obsahuje**.
  - e. Do pole **Vyhledat** zadejte text podnikové vyhledávání a poté klepněte na tlačítko **Vyhledat**.
  - f. Po zobrazení nových ikon vpravo klepněte na ikonu klíče. Tím otevřete okno pro úpravy vyhledávacího portletu podnikového vyhledávání.
  - g. V seznamu parametrů portletu upravte následující parametry:

#### hostname

Zadejte úplný název hostitele vyhledávacího serveru pro podnikové vyhledávání.

#### port

Zadejte číslo portu používané produktem WebSphere Application Server na

vyhledávacím serveru podnikového vyhledávání. Výchozí hodnota je 80 (výchozí hodnota pro komunikaci SSL je 443).

**username**

Pokud je v produktu WebSphere Application Server na vyhledávacím serveru zapnuto globální zabezpečení, zadejte jméno uživatele, které je platné v registru produktu WebSphere Application Server.

**password**

Pokud jste zadali jméno uživatele produktu WebSphere Application Server, zadejte odpovídající heslo.

**protocol**

Zadejte protokol používaný pro komunikaci mezi produktem WebSphere Portal a vyhledávacím serverem. Výchozí hodnota je HTTP. Používáte-li zabezpečení SSL, zadejte hodnotu HTTPS.

**trustStore**

Používáte-li zabezpečení SSL, zadejte úplnou cestu (včetně názvu souboru) pro úložiště certifikátů SSL.

**trustPassword**

Používáte-li zabezpečení SSL, zadejte heslo pro zadaný soubor trustStore.

**ssoCookieName**

Zadejte název souboru cookie obsahujícího řetězec tokenu jednotného přihlášení (SSO). Výchozí hodnota je `LtpaToken`.

**proxyHost**

Pokud je pro přístup k vyhledávacímu serveru v rámci podnikového vyhledávání vyžadován server proxy, zadejte úplný název hostitele serveru proxy.

**proxyPort**

Pokud jste zadali server proxy, zadejte číslo portu serveru proxy.

**proxyUser**

Pokud server proxy vyžaduje základní ověřování, zadejte jméno uživatele, které se používá k přihlášení k serveru proxy.

**proxyPassword**

Pokud jste zadali jméno uživatele, zadejte odpovídající heslo.

h. Uložte změny klepnutím na tlačítko **OK**.

2. Chcete-li přistupovat k portletu po jeho nastavení, postupujte takto:

a. Přihlaste se k serveru Lotus Quickr.

b. V okně prohlížeče změňte adresu URL na následující adresu: `http://název_hostitele:port/lotus/myquickr/ESSearchPortlet`

## Odebrání podnikového vyhledávání z produktu WebSphere Portal verze 6

K odebrání podnikového vyhledávání se systému WebSphere Portal verze 6 se používá skript `wp6_uninstall`.

### O této úloze

Odeberete-li podnikové vyhledávání z portálu WebSphere, nebudou uloženy parametry portletu podnikového vyhledávání, které jste zadali při instalaci.

Po spuštění skript zastaví portálový server WebSphere. Po odebrání softwaru podnikového vyhledávání skript znovu spustí portálový server WebSphere.

### Postup

Chcete-li odebrat podnikové vyhledávání ze systému WebSphere Portal verze 6, postupujte takto:

Z příkazového řádku spusťte příkaz **wp6\_uninstall.bat** (v systémech Windows) nebo **wp6\_uninstall.sh** (v systémech AIX, Linux a Solaris). V následujícím příkladu jsou parametry pro přehlednost uvedeny na samostatných řádcích, při zadávání příkazu je však musíte uvést na jednom řádku:

```
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

Úplná cesta k adresáři profilu produktu WebSphere Portal.

### WASDir

Úplná cesta ke kořenovému adresáři produktu WebSphere Application Server; je vyžadována pouze v systémech AIX, Linux a Solaris.

### WASUser

Jméno správce produktu WebSphere Application. Je vyžadováno, pouze pokud je na serveru WebSphere Application Server zapnuto globální zabezpečení.

### WASPassword

Heslo správce produktu WebSphere Application Server, je-li zadán.

### WPSDir

Úplná cesta k instalačnímu adresáři produktu WebSphere Portal.

### WPSUser

Jméno správce produktu WebSphere Portal.

### WSPassword

Heslo zadaného správce produktu WebSphere Portal.

### WPSHost

Název hostitele a číslo portu serveru WebSphere Portal.

---

## Integrace podnikového vyhledávání s klastrovanými systémy produktu WebSphere Portal

V klastrovaných systémech produktu WebSphere Portal verze 6 můžete nastavit spuštění podnikového vyhledávání.

### Pokyny pro použití

- Před spuštěním konfiguračních skriptů podnikového vyhledávání se ujistěte, že je spuštěn správce síťové implementace produktu WebSphere Application Server a každý uzel v klastru.
- Skripty nastavují všechny integrační body mezi podnikovým vyhledáváním a portálem WebSphere. Nemůžete tedy například nainstalovat pouze portlet a nenainstalovat přitom soubory EAR, které zajišťují podporu prolézacích modulů WebSphere Portal a Web Content Management.

- Skripty zastaví a restartují všechny instance serveru WebSphere Portal v klastru. Proto byste je měli spouštět mimo běžnou pracovní dobu, aby uživatelé neměli problémy s nedostupností portálových služeb.
- Pokud při práci konfiguračních skriptů dojde k chybám, spusťte je znovu. Úlohy, které byly při prvním pokusu úspěšně dokončeny, mohou vrátit chyby, proces konfigurování však bude pokračovat a budou provedeny zbývající úlohy.
- Při prvním přístupu ke stránce s portletem podnikového vyhledávání po spuštění konfiguračního skriptu může být zobrazování stránky pomalejší, protože systém musí zkompilovat soubory JSP (Java Server Pages) portletu.

## Nastavení podnikového vyhledávání v klastrovém systému WebSphere Portal

K integraci systému podnikového vyhledávání s produktem WebSphere Portal verze 6 se používá skript `wp6_cluster_install`.

### O této úloze

Soubory potřebné pro integraci podnikového vyhledávání s produktem WebSphere Portal jsou dodávány v souboru `es.wp6.install.jar`. Po rozbalení tohoto souboru získáte následující soubory:

- `ESSearchPortlet.war`
- `ESPACServer.ear`
- `esapi.jar`
- `es.search.provider.jar`
- `es.security.jar`
- Soubory skriptů, dávek, XML a JACL potřebné pro instalaci

### Postup

Chcete-li integrovat podnikové vyhledávání s klastrovým systémem WebSphere Portal verze 6, postupujte takto:

1. Zkopírujte soubor `es.wp6.install.jar` ze serveru podnikového vyhledávání na každý uzel v klastru, kde je instalován produkt WebSphere Portal, a poté jej rozbalte příkazem **JAR** jazyka Java (nebo příkazem **TAR**).
2. Volitelné: Pokud chcete zavést podporu integrace se správcem WebSphere PDM (Portal Document Manager), proveďte na každém uzlu v klastru některý z následujících kroků:
  - Spusťte instalační program WebSphere Information Integrator Content Edition, vyberte volbu provádějící pouze instalaci konektoru a nainstalujte konektor PDM na portálový server WebSphere.
  - Vytvořte na portálovém serveru WebSphere adresářovou strukturu WebSphere Information Integrator Content Edition a zkopírujte na portálový server WebSphere následující soubory z existující instalace produktu WebSphere Information Integrator Content Edition, kde `CE_ROOT` určuje kořenový instalační adresář produktu WebSphere Information Integrator Content Edition:
 

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```
3. Z příkazového řádku spusťte příkaz `wp6_cluster_copyFiles.bat` (v systémech Windows) nebo příkaz `wp6_cluster_copyFiles.sh` (v systémech AIX, Linux nebo Solaris) na každém uzlu v klastru, kde je nainstalován produkt WebSphere Portal. V následujících

příkladech jsou volby pro přehlednost uvedeny na samostatných řádcích, při zadávání příkazu je však musíte uvést na jednom řádku:

```
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. Z příkazového řádku spusíte příkaz **wp6\_cluster\_install.bat** (v systémech Windows) nebo příkaz **wp6\_cluster\_install.sh** (v systémech AIX, Linux nebo Solaris). V následujícím příkladu jsou volby pro přehlednost uvedeny na samostatných řádcích, při zadávání příkazu je však musíte uvést na jednom řádku:

```
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 -WASPassword wpsadmin
-WPSHost "portalserver.ibm.com"
-webServerName webserver1
-webServerNodeName node1
-IICEDir "C:\\Program Files\\IBM\\Content Edition"
```

#### **WPSClusterName**

Jedná se o název klastru, ve kterém je nainstalován produkt WebSphere Portal.

#### **WSPProfileDir**

Úplná cesta k adresáři profilu produktu WebSphere Portal. Výchozí cestou je v systémech AIX cesta /usr/IBM/WebSphere/AppServer/profiles/wp\_profile, v systémech Linux a Solaris cesta /opt/IBM/WebSphere/AppServer/profiles/wp\_profile a v systémech Windows cesta C:\\Program Files\\IBM\\WebSphere\\profiles\\wp\_profile.

#### **WASDir**

Úplná cesta ke kořenovému adresáři produktu WebSphere Application Server; je vyžadována pouze v systémech AIX, Linux a Solaris. Výchozí cestou ke kořenovému adresáři je v systémech AIX cesta /usr/IBM/WebSphere/AppServer a v systémech Linux a Solaris cesta /opt/IBM/WebSphere/AppServer.

#### **WASUser**

Jméno správce produktu WebSphere Application. Je vyžadováno, pouze pokud je na serveru WebSphere Application Server zapnuto globální zabezpečení.

#### **WASPassword**

Heslo správce produktu WebSphere Application Server, je-li zadán.

#### **WPSDir**

Úplná cesta k instalačnímu adresáři produktu WebSphere Portal.

#### **WPSUser**

Jméno správce produktu WebSphere Portal.

#### **WSPassword**

Heslo zadaného správce produktu WebSphere Portal.

#### **WPSHost**

Název hostitele a číslo portu serveru WebSphere Portal.

#### **webServerName**

Jedná se o název definice webového serveru, kam patří server WebSphere Portal.

### **webServerNodeName**

Jedná se o název uzlu produktu WebSphere Application Server, ke kterému patří definice webového serveru.

### **IICEDir**

Úplná cesta k instalačnímu adresáři produktu WebSphere Information Integrator Content Edition. Tento parametr je vyžadován jen tehdy, jestliže jste na portálový server WebSphere předem nainstalovali konektor správce PDM.

5. Po dokončení skriptu otevřete webový prohlížeč a přihlaste se ke konzole pro správu produktu WebSphere na serveru síťové implementace. Obvyklý formát adresy: `http://název_hostitele:9060/ibm/console`.
6. Rozbalte sekci **Servery** a vyberte volbu **Webové servery**.
7. Zaškrtněte políčko **Vybrat** vedle webového serveru a poté klepněte na tlačítko **Generovat modul plug-in**.
8. Zaškrtněte políčko **Vybrat** vedle webového serveru a poté klepněte na tlačítko **Rozšířit modul plug-in**.
9. Odhlaste se od konzoly pro správu.
10. Pomocí aktualizace portletu podnikového vyhledávání identifikujte vyhledávací server:
  - a. Přihlaste se k portálu WebSphere s použitím jména a hesla administrátora portálu.
  - b. Klepněte na položku **Správa** v levém dolním rohu.
  - c. V navigační oblasti vlevo klepněte na volbu **Správa portletů** a poté klepněte na volbu **Portlety**.
  - d. Volbu **Hledat podle** změňte na **Název obsahuje**.
  - e. Do pole **Vyhledat** zadejte text podnikové vyhledávání a poté klepněte na tlačítko **Vyhledat**.
  - f. Po zobrazení nových ikon vpravo klepněte na ikonu klíče. Tím otevřete okno pro úpravy vyhledávacího portletu podnikového vyhledávání.
  - g. V seznamu parametrů portletu upravte následující parametry:

#### **hostname**

Zadejte úplný název hostitele vyhledávacího serveru pro podnikové vyhledávání.

**port** Zadejte číslo portu používané produktem WebSphere Application Server na vyhledávacím serveru podnikového vyhledávání. Výchozí hodnota je 80 (výchozí hodnota pro komunikaci SSL je 443).

#### **username**

Pokud je v produktu WebSphere Application Server na vyhledávacím serveru zapnuto globální zabezpečení, zadejte jméno uživatele, které je platné v registru produktu WebSphere Application Server.

#### **password**

Pokud jste zadali jméno uživatele produktu WebSphere Application Server, zadejte odpovídající heslo.

#### **protocol**

Zadejte protokol používaný pro komunikaci mezi produktem WebSphere Portal a vyhledávacím serverem. Výchozí hodnota je HTTP. Používáte-li zabezpečení SSL, zadejte hodnotu HTTPS.

#### **trustStore**

Používáte-li zabezpečení SSL, zadejte úplnou cestu (včetně názvu souboru) pro úložiště certifikátů SSL.

#### **trustPassword**

Používáte-li zabezpečení SSL, zadejte heslo pro zadaný soubor trustStore.

**ssoCookieName**

Zadejte název souboru cookie obsahujícího řetězec tokenu jednotného přihlášení (SSO). Výchozí hodnota je LtpaToken.

**proxyHost**

Pokud je pro přístup k vyhledávacímu serveru v rámci podnikového vyhledávání vyžadován server proxy, zadejte úplný název hostitele serveru proxy.

**proxyPort**

Pokud jste zadali server proxy, zadejte číslo portu serveru proxy.

**proxyUser**

Pokud server proxy vyžaduje základní ověřování, zadejte jméno uživatele, které se používá k přihlášení k serveru proxy.

**proxyPassword**

Pokud jste zadali jméno uživatele, zadejte odpovídající heslo.

h. Uložte změny klepnutím na tlačítko **OK**.

## Odebrání podnikového vyhledávání z klastrovaného systému WebSphere Portal

K odebrání podnikového vyhledávání z klastrovaného systému WebSphere Portal verze 6 se používá skript `wp6_cluster_uninstall`.

**O této úloze**

Odeberete-li podnikové vyhledávání z portálu WebSphere, nebudou uloženy parametry portletu podnikového vyhledávání, které jste zadali při instalaci.

Po spuštění skript zastaví portálový server WebSphere. Po odebrání softwaru podnikového vyhledávání skript znovu spustí portálový server WebSphere.

**Postup**

Chcete-li odebrat podnikové vyhledávání z klastrovaného systému WebSphere Portal verze 6, postupujte takto:

1. Z příkazového řádku spusíte příkaz **wp6\_cluster\_uninstall.bat** (v systémech Windows) nebo příkaz **wp6\_cluster\_uninstall.sh** (v systémech AIX, Linux nebo Solaris) na jednom z uzlů v klastru. V následujícím příkladu jsou parametry pro přehlednost uvedeny na samostatných řádcích, při zadávání příkazu je však musíte uvést na jednom řádku:

```
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**

Jedná se o název klastru, ve kterém je nainstalován produkt WebSphere Portal.

**WSPProfileDir**

Úplná cesta k adresáři profilu produktu WebSphere Portal. Výchozí cestou je v systémech AIX cesta `/usr/IBM/WebSphere/AppServer/profiles/wp_profile`, v

systemech Linux a Solaris cesta /opt/IBM/WebSphere/AppServer/profiles/wp\_profile a v systémech Windows cesta C:\Program Files\IBM\WebSphere\profiles\wp\_profile.

#### **WASDir**

Úplná cesta ke kořenovému adresáři produktu WebSphere Application Server; je vyžadována pouze v systémech AIX, Linux a Solaris. Výchozí cestou ke kořenovému adresáři je v systémech AIX cesta /usr/IBM/WebSphere/AppServer a v systémech Linux a Solaris cesta /opt/IBM/WebSphere/AppServer.

#### **WASUser**

Jméno správce produktu WebSphere Application. Je vyžadováno, pouze pokud je na serveru WebSphere Application Server zapnuto globální zabezpečení.

#### **WASPassword**

Heslo správce produktu WebSphere Application Server, je-li zadán.

#### **WPSDir**

Úplná cesta k instalačnímu adresáři produktu WebSphere Portal.

#### **WPSUser**

Jméno správce produktu WebSphere Portal.

#### **WPSPassword**

Heslo zadaného správce produktu WebSphere Portal.

#### **WPSHost**

Název hostitele a číslo portu serveru WebSphere Portal.

#### **webServerName**

Jedná se o název definice webového serveru, kam patří server WebSphere Portal.

#### **webServerNodeName**

Jedná se o název uzlu produktu WebSphere Application Server, ke kterému patří definice webového serveru.

2. Po dokončení skriptu otevřete webový prohlížeč a přihlaste se ke konzole pro správu produktu WebSphere na serveru síťové implementace. Obvyklý formát adresy: `http://název_hostitele:9060/ibm/console`.
3. Rozbalte sekci **Servery** a vyberte volbu **Webové servery**.
4. Zaškrtněte políčko **Vybrat** vedle webového serveru a poté klepněte na tlačítko **Generovat modul plug-in**.
5. Zaškrtněte políčko **Vybrat** vedle webového serveru a poté klepněte na tlačítko **Rozšířit modul plug-in**.
6. Odhlaste se od konzoly pro správu.



---

## Migrace z produktu WebSphere Portal do podnikového vyhledávání

Podnikové vyhledávání nabízí průvodce migrací, kterého můžete použít k migraci kolekcí a taxonomií založených na pravidlech z produktu IBM WebSphere Portal do prostředí podnikového vyhledávání.

V prostředí podnikového vyhledávání se namísto pojmu taxonomie používá pojem *strom kategorií*. Po migraci taxonomie můžete použít konzolu správy podnikového vyhledávání k úpravám stromu kategorií a pravidel kategorií. Po migraci kolekce můžete provádět její správu pomocí konzoly pro správu.

Chcete-li provést migraci taxonomií a kolekcí, spusťte průvodce migrací na indexovém serveru podnikového vyhledávání.

---

## Migrace kolekce z produktu WebSphere Portal

Chcete-li migrovat kolekce a taxonomie založené na pravidlech z produktu WebSphere Portal do prostředí podnikového vyhledávání, připravte je nejprve v produktu WebSphere Portal a poté proveďte jejich migraci pomocí průvodce migrací.

### Než začnete

Chcete-li migrovat taxonomie a kolekce, je třeba před použitím tohoto postupu pro migraci kolekcí nejprve migrovat soubory taxonomie. Tímto postupem zajistíte, že budou migrovaná pravidla kategorizace funkční v migrovaných kolekcích.

### Postup

Při migraci kolekce (a případně taxonomie) z produktu WebSphere Portal do prostředí podnikového vyhledávání postupujte takto:

1. Ve vyhledávacím stroji produktu WebSphere Portal zastavte všechny prolézací procesy v kolekcích, které chcete migrovat, a schvalte nebo odmítněte všechny nevyřízené dokumenty. (Podnikové vyhledávání nepodporuje koncepci nevyřízených dokumentů.)
2. Pomocí portletů portálového vyhledávacího stroje exportujte nastavení všech kolekcí, které chcete migrovat, do souborů XML.
3. Pokud je indexový server podnikového vyhledávání nainstalován na samostatném serveru, zkopírujte exportované soubory XML na indexový server.
4. Přihlaste se k indexovému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu OmniFind Enterprise Edition.
5. Přejděte do instalačního adresáře podnikového vyhledávání:

```
UNIX: cd $ES_INSTALL_ROOT/bin  
Windows: cd %ES_INSTALL_ROOT%\bin
```

6. Chcete-li migrovat kolekce s povoleným zabezpečením, spusťte průvodce migrací zadáním následujícího příkazu a poté klepněte na tlačítko **Další**.

```
UNIX:  
./eswpsmigrate.sh  
Windows: eswpsmigrate.bat
```

7. Chcete-li zakázat zabezpečení na úrovni migrované kolekce, spusťte průvodce migrací zadáním následujícího příkazu a poté klepněte na tlačítko **Další**.

UNIX: `./eswpsmigrate.sh disable.security`

Windows: `eswpsmigrate.bat disable.security`

8. Vyberte volbu **Migrovat nastavení vyhledávání z portálového vyhledávacího stroje portálu WebSphere** a poté klepněte na tlačítko **Další**.
9. Přejděte do adresáře obsahujícího exportované konfigurační soubory portálového vyhledávacího stroje, vyberte soubory, které chcete migrovat, a poté klepněte na tlačítko **Další**. Proběhne analýza a vyhodnocení vybraných konfiguračních souborů.
10. Pro každou kolekci zadejte následující informace a poté spusíte migraci kolekcí do prostředí podnikového vyhledávání klepnutím na tlačítko **Další**:
  - Název, pod kterým chcete kolekci používat v prostředí podnikového vyhledávání
  - Kritérium, podle něhož se v kolekci určuje důležitost dokumentu. Faktor statického pořadí může být prázdný, založený na datech dokumentů nebo založený na počtu odkazů na webové dokumenty z jiných webových dokumentů.
  - Typ kategorizace, kterou chcete v této kolekci používat. Nezadáte-li žádnou, nebudou do prostředí podnikového vyhledávání migrovány žádné informace o taxonomii. Pokud vyberete kategorie založené na pravidlech, bude do prostředí podnikového vyhledávání migrována taxonomie společně s kolekcí.

V případě výskytu chyb během migrace vyhledejte bližší informace v souboru `MigrationWizard.log` umístěném v instalačním adresáři průvodce migrací.

Nyní můžete pomocí konzoly pro správu podnikového vyhledávání konfigurovat další nastavení migrovaných kolekcí.

**Požadavek:** Při konfigurování vlastností webového prolézacího modulu pro migrovanou kolekci je nutné zadat e-mailovou adresu, na kterou budou zasílány poznámky k prolézacímu modulu, a název uživatelského agenta (pokud potřebujete s konfigurováním vlastností webového prolézacího modulu pomoci, klepněte na tlačítko **Nápověda**).

11. Na konzole správy podnikového vyhledávání spusíte procesy prolézání, analýzy a indexování pro migrovanou kolekci.
12. Jakmile dosáhnete stavu, kdy lze migrovanou kolekci prohledávat v prostředí podnikového vyhledávání, odstraňte původní kolekci z portálového vyhledávacího jádra.
13. Volitelné: Chcete-li jako administrátor produktu WebSphere Portal umožnit uživatelům prohledávání migrované kolekce z portálu v prostředí WebSphere Portal, proveďte následující kroky.
  - a. Implementujte portlet podnikového vyhledávání v instalaci produktu WebSphere Portal.

V serverovém klastru produktu WebSphere Portal je tento krok nutné provést na serveru, na kterém je instalován správce implementace aplikačního serveru WebSphere. Správce implementace provede distribuci portletu podnikového vyhledávání na ostatní servery v serverovém klastru produktu WebSphere Portal.
  - b. Přidejte portlet podnikového vyhledávání na příslušné stránky portálu.

V produktu WebSphere Portal je řízení přístupu vyhledávacího portletu modelováno přístupností konkrétních stránek a portletů. I když je nastavení kolekce migrováno, umístění portletu musí provést ručně administrátor serveru WebSphere Portal.

---

## Migrované nastavení kolekce

Při migraci kolekcí z produktu IBM WebSphere Portal vytvoří průvodce migrací výchozí nastavení pro kolekce a prolézací moduly.

Pokud existuje stejné nastavení pro kolekce portálového vyhledávacího stroje i pro kolekce podnikového vyhledávání, průvodce při migraci kolekcí do prostředí podnikového vyhledávání použije nastavení portálového vyhledávacího stroje. U nastavení existujících pouze v prostředí podnikového vyhledávání použije průvodce nastavení, které zadáte při migraci kolekce, nebo výchozí nastavení definované pro kolekce v prostředí podnikového vyhledávání.

## **Nastavení existující v portálovém vyhledávacím stroji i v prostředí podnikového vyhledávání**

Průvodce migrací převádí pro každou migrovanou kolekci následující nastavení:

- servery portálového vyhledávacího stroje uvedené v kolekci portálového vyhledávacího stroje,
- jazyk kolekce,
- taxonomii (neboli strom kategorií) a pravidla pro kategorie založené na pravidlech, pokud kolekce podnikového vyhledávání používá kategorizaci založenou na pravidlech.

Do webového prolézacího modulu podnikového vyhledávání jsou konsolidovány všechny servery portálového vyhledávacího stroje v kolekci. Průvodce migrací provádí migraci následujících nastavení prolézacího modulu:

- počáteční adresy URL,
- počet paralelních prolézacích procesů,
- hloubka prolézání,
- časový limit (v sekundách) pro načtení dokumentu,
- výchozí znaková sada,
- pravidla pro prolézání webových stránek (zahrnutí nebo vyloučení).

## **Nastavení existující pouze v prostředí podnikového vyhledávání**

Při migraci kolekce je nutné zadat informace o této kolekci. Průvodce migrací toto nastavení migruje a ke konfigurování jednotlivých migrovaných kolekcí použije výchozí nastavení pro kolekce podnikového vyhledávání.

Konfiguraci kolekce a webového prolézacího modulu lze změnit pomocí konzoly pro správu podnikového vyhledávání. Hodnoty uvedené v závorkách ( ) představují výchozí nastavení pro migrovaná data.

- Název kolekce,
- strategie přiřazování statických tříd důležitosti pro dokumenty,
- použitý typ kategorizace (kategorizace založená na pravidlech nebo žádná kategorizace),
- zda má být použita mezipaměť vyhledávání a kolik odpovědí na dotazy může obsahovat (ano, 5000),
- zda mají být monitorovány doby odezvy vyhledávání a při překročení limitu vygenerována výstraha (ano, 5 sekund),
- zda má být použito řízení přístupu (ne),
- plán sestavení rozdílových indexů,
- plán sestavení hlavního indexu,
- úroveň podrobností žurnálu (všechny zprávy).

Průvodce migrací rovněž vytvoří pro každý webový prolézací modul následující nastavení:

- název prolézacího modulu,
- popis prolézacího modulu,

- maximální délka stránky,
- nastavení zabezpečení dokumentu,
- typy MIME (Multipurpose Internet Mail Extension), které mají být prolézány, pokud má takové nastavení u daného typu zdroje dat smysl.

Než spustíte nově migrovaný webový prolézací modul, zkontrolujte všechny vlastnosti tohoto prolézacího modulu a nastavení prolézaného prostoru a ujistěte se, že byly zadány všechny požadované hodnoty (povinná pole jsou označena červenou hvězdičkou). Zejména nezapomeňte uvést e-mailovou adresu pro zasilání poznámek k prolézacímu modulu a název uživatelského agenta prolézacího modulu. Pokud potřebujete asistenci, klepněte při konfigurování vlastností webového prolézacího modulu na tlačítko **Nápověda**.

---

## Soubor žurnálu průvodce migrací

Průvodce migrací zapisuje všechny zprávy do souboru `WpsMigratorLog.log` ve svém instalačním adresáři.

Pro každou migrovanou kolekci je v souboru žurnálu `WpsMigratorLog.log` uvedeno nastavení všech hodnot, které byly načteny z vyhledávacího stroje produktu WebSphere Portal, spolu s informacemi o tom, zda bylo toto nastavení importováno do kolekci podnikového vyhledávání.

---

## Příkazy podnikového vyhledávání, návratové kódy a ID relací

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

V případě instalace na více serverech můžete ve svém systému spouštět příkazy ze kteréhokoli serveru. Příkazy by však měly být spouštěny z indexového serveru. Indexový server (neboli řídicí server) může přistupovat k informacím ze všech ostatních serverů v systému.

Formát většiny příkazů je následující:

```
esadmin název_příkazu argumenty  
esadmin ID_relace akce -volba
```

Chcete-li zobrazit více informací o všech příkazech, zadejte příkaz `esadmin help`. Chcete-li zobrazit další informace o konkrétním příkazu, zadejte příkaz `esadmin akce help`.

### Příkazy `esadmin` podnikového vyhledávání

Následující příkazy zadávejte na jeden řádek.

Tabulka 9. Příkazy `esadmin` podnikového vyhledávání

Příkaz	Popis
<code>esadmin system startall</code>	<p>Spustí komponenty podnikového vyhledávání na všech serverech podnikového vyhledávání, včetně webového serveru, aplikace <code>ESSearchServer</code> a informačního centra na vyhledávacích serverech, relace prolézacích modulů na prolézacím serveru a relace indexu na indexovém serveru. Spustí službu CCL (Common Communication Layer) pouze na místním serveru. Chcete-li znovu použít službu CCL, je ji třeba ručně zastavit a restartovat na každém vzdáleném serveru podnikového vyhledávání.</p> <p><b>Ukázka použití příkazu:</b> <code>esadmin system startall</code></p>
<code>esadmin system stopall</code>	<p>Zastaví komponenty podnikového vyhledávání na všech serverech podnikového vyhledávání, včetně webového serveru, aplikace <code>ESSearchServer</code> a informačního centra na vyhledávacích serverech, relace prolézacích modulů na prolézacím serveru a relace indexu na indexovém serveru. Zastaví službu CCL pouze na místním serveru. Chcete-li znovu použít službu CCL, je ji třeba ručně zastavit a restartovat na každém vzdáleném serveru podnikového vyhledávání.</p> <p><b>Ukázka použití příkazu:</b> <code>esadmin system stopall</code></p>
<code>esadmin system checkall</code>	<p>Zkontroluje stav všech komponent podnikového vyhledávání na všech serverech podnikového vyhledávání.</p> <p><b>Ukázka použití příkazu:</b> <code>esadmin system checkall</code></p>

Tabulka 9. Příkazy **esadmin** podnikového vyhledávání (pokračování)

Příkaz	Popis
<code>esadmin ID_relace_prolézacího_modulu start</code>	<p>Spustí relaci prolézacího modulu. Tento příkaz nespouští žádné prolézání.</p> <p><b>Ukázka použití příkazu:</b>  <code>esadmin col1.WEB1.esadmin start</code></p> <p><b>Ukázka zpráv a návratového kódu:</b>            FFQC5310I Proces WEBCrawler1 (sid: col1.WEB1.esadmin) není spuštěn.            FFQC5314I Výsledek: 0</p>
<code>esadmin ID_relace_prolézacího_modulu startCrawl</code>	<p>Zahájí prolézání.</p> <p><b>Ukázka použití příkazu:</b>  <code>esadmin col3.DB21.esadmin startCrawl</code></p> <p><b>Ukázka zpráv a návratového kódu:</b>            FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650            FFQC5314I Výsledek: 0</p>
<code>esadmin ID_relace_prolézacího_modulu pause</code>	<p>Pozastaví prolézání.</p> <p><b>Ukázka použití příkazu:</b>  <code>esadmin col3.DB21.esadmin pause</code></p> <p><b>Ukázka zpráv a návratového kódu:</b>            FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650            FFQC5314I Výsledek: 0</p>
<code>esadmin ID_relace_prolézacího_modulu resume</code>	<p>Obnoví prolézání.</p> <p><b>Ukázka použití příkazu:</b>  <code>esadmin col3.DB21.esadmin resume</code></p> <p><b>Ukázka zpráv a návratového kódu:</b>            FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650            FFQC5314I Výsledek: 0</p>
<code>esadmin ID_relace_prolézacího_modulu stopCrawl</code>	<p>Zastaví prolézání.</p> <p><b>Ukázka použití příkazu:</b>  <code>esadmin col3.DB21.esadmin stopCrawl</code></p> <p><b>Ukázka zpráv a návratového kódu:</b>            FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650            FFQC5314I Výsledek: 0</p>
<code>esadmin ID_relace_prolézacího_modulu stop</code>	<p>Zastaví relaci prolézacího modulu.</p> <p><b>Ukázka použití příkazu:</b>  <code>esadmin col3.DB21.esadmin stop</code></p> <p><b>Ukázka zpráv a návratového kódu:</b>            FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650            FFQC5314I Výsledek: 0</p>

Tabulka 9. Příkazy **esadmin** podnikového vyhledávání (pokračování)

Příkaz	Popis
<pre>esadmin ID_relace_prolézacího_modulu getCrawlerStatus</pre>	<p>Zjistí stav prolézacího modulu. Vracené informace se liší podle toho, zda jde o webový prolézací modul nebo o prolézací modul pro všechny ostatní zdroje dat.</p> <p><b>Příklad pro webový prolézací modul:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlerStatus</pre> <p><b>Možné návratové kódy a zprávy pro webový prolézací modul:</b></p> <pre>FFQC5303I Proces WebCrawler1 (sid: col1.WEB1.esadmin) je již spuštěn. PID: 23650</pre> <p><b>Příklad pro jiný než webový prolézací modul:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlerStatus</pre> <p><b>Možné návratové kódy a zprávy pro jiný než webový prolézací modul:</b></p> <pre>FFQC5303I Proces db2crawler (sid: db2col.DB2_96945) je již spuštěn. PID: 5936</pre> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech” na stránce 352.</p>
<pre>esadmin ID_relace_jiného_prolézacího_modulu getCrawlSpaceStatus esadmin ID_relace_webového_prolézacího_modulu getCrawlStatus -selections hodnota</pre>	<p>Zjistí obecný stav prolézaného prostoru pro každý prolézací modul, který není webovým prolézacím modulem.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatus</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650</pre> <p>Zjistí obecný stav prolézaného prostoru pro webový prolézací modul.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlStatus</pre> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech” na stránce 352.</p>

Tabulka 9. Příkazy **esadmin** podnikového vyhledávání (pokračování)

Příkaz	Popis
<pre>esadmin ID_relace_jiného_prolézacího_modulu getCrawlSpaceStatusDetail -ts ID_cílového_serveru esadmin ID_relace_webového_prolézacího_modulu getCrawlDetailsPerSite -url řetězec -selections počet -threshold počet</pre>	<p>Zjistí podrobný stav prolézaného prostoru pro každý prolézací modul, který není webovým prolézacím modulem. Neurčíte-li volbu cílového serveru, budou vrácena data pro všechny cílové servery. Pokud například prolézací modul DB2 prolézá databáze FOUNTAIN a SAMPLE a není určena volba cílového serveru, bude vrácen stav všech tabulek v databázích FOUNTAIN a SAMPLE.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatusDetail -ts FOUNTAIN</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650</pre> <p>Zjistí podrobný stav prolézaného prostoru pro webový prolézací modul.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlDetailsPerSite</pre> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech” na stránce 352.</p>
<pre>esadmin monitor getCollectionParserMonitorStatus -cid ID_kolekce</pre>	<p>Zjistí stav modulu analýzy.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin monitor getCollectionParserMonitorStatus -cid col1</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Monitor (node1) (sid: monitor) je již spuštěn. PID: 12543</pre> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech” na stránce 352.</p>
<pre>esadmin startMain -cid ID_kolekce</pre>	<p>Spustí sestavení hlavního indexu.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin startMain -cid col1</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 25917 FFQC5314I Výsledek: 1117671147056</pre>
<pre>esadmin startDelta -cid ID_kolekce</pre>	<p>Spustí sestavení rozdílového indexu.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin startDelta -cid col1</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 4548 FFQC5314I Výsledek: 1117670603408</pre>



Tabulka 9. Příkazy **esadmin** podnikového vyhledávání (pokračování)

Příkaz	Popis
<pre>esadmin controller startIndexBuild -cid <i>ID_kolekce</i> -buildType <i>typ</i> -detectChanges</pre>	<p>Spustí sestavení hlavního nebo rozdílového indexu a určuje, že sestavení má být zpracováno pouze tehdy, pokud nejsou zjištěny žádné změny, které by bylo nutné použít na index.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin controller startIndexBuild -cid col_1 -buildType main -detectChanges</pre>
<pre>esadmin monitor getCollectionIndexMonitorStatus -cid <i>ID_kolekce</i> -buildType [main   delta] -numrecords <i>posledních_N_záznamů</i></pre>	<p>Získá stav sestavení hlavního nebo rozdílového indexu. Volba <code>numrecords</code> zobrazuje posledních <i>N</i> záznamů stavu sestavování indexu. Je-li parametr <code>numrecords</code> vynechán, bude vrácen stav pro posledních 20 sestavení indexu.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin monitor getCollectionIndexMonitorStatus -cid col1 -buildType main -numrecords 4</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Monitor (node1) (sid: monitor) je již spuštěn. PID: 12649</pre> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech” na stránce 352.</p>
<pre>esadmin startSearch -cid <i>ID_kolekce</i></pre>	<p>Spustí procesy vyhledávacího serveru.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin startSearch -cid col1</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 25917 FFQC5314I Výsledek: 0</pre>
<pre>esadmin stopSearch -cid <i>ID_kolekce</i></pre>	<p>Zastaví procesy vyhledávacího serveru.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin stopSearch -cid col1</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 15292 FFQC5314I Výsledek: 0</pre>

Tabulka 9. Příkazy **esadmin** podnikového vyhledávání (pokračování)

Příkaz	Popis
<pre>esadmin monitor getCollectionSearchMonitorStatus -cid ID_kolekce esadmin ID_relace_správce_vyhledávání getStatus -cid ID_kolekce</pre>	<p>Zjistí stav vyhledávacího serveru.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin monitor getCollectionSearchMonitorStatus -cid coll</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Monitor (node1) (sid: monitor) je již spuštěn. PID: 12649</pre> <p>Vrátí podrobné informace o stavu indexu vyhledávání pro kolekci na daném vyhledávacím serveru. Pro každý server existuje jedna relace správce vyhledávání. Každá relace správce vyhledávání odpovídá za monitorování a manipulaci s indexy vyhledávání na konkrétním vyhledávacím serveru.</p> <p><b>Ukázka použití příkazu:</b></p> <pre>esadmin searchmanager.node1 getStatus -cid coll</pre> <p><b>Ukázka zpráv a návratového kódu:</b></p> <pre>FFQC5303I Proces Search Manager (node1) (sid: searchmanager.node1) je již spuštěn. PID: 15711 FFQC5314I Výsledek: PID=18390 CacheHits=3 QueryRate=1 Port=44008 SessionId=coll.runtime.node1 CacheHitRate=0.333 ResponseTime=70 Status=1 SessionName=coll.runtime.node1.1</pre> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech”.</p>

## Podrobné informace o stavových příkazech

Některé příkazy mohou vracet rozsáhlé informace. V této části jsou popsány informace, které mohou být vráceny jako stav prolézacího modulu a stav prolézaného prostoru. V tabulce v části “Příkazy esadmin podnikového vyhledávání” na stránce 347 jsou uvedeny možné informace, které mohou být vráceny jednotlivými příkazy esadmin. V této části jsou popsány informace vrácené následujícími příkazy:

- Stav webového prolézacího modulu
- Stav jiného než webového prolézacího modulu
- Stav prolézaného prostoru pro webový prolézací modul
- Stav prolézaného prostoru pro jiný než webový prolézací modul
- Podrobný stav prolézaného prostoru pro webový prolézací modul
- Podrobný stav prolézaného prostoru pro jiný než webový prolézací modul
- Stav modulu analýzy
- Stav sestavování indexu
- Stav vyhledávacího serveru
- Podrobný stav vyhledávacího serveru

**Stav webového prolézacího modulu:** Po spuštění příkazu pro zjištění stavu webového prolézacího modulu jsou vráceny informace ve formátu dokumentu XML. Příkaz pro zjištění stavu webového prolézacího modulu může vrátit následující informace:

```
FFQC5314I Výsledek: <?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>
```

Následující tabulka popisuje každý z prvků XML a jejich možných atributů, které mohou být vráceny příkazem pro zjištění stavu webového prolézacího modulu:

Tabulka 10. Informace o stavu webového prolézacího modulu

Prvek	Atributy	Popis
CrawlerStatus	<ul style="list-style-type: none"> <li>CrawlerThreadStateDist</li> <li>ActiveBucketList</li> <li>CrawlRate</li> <li>RecentlyCrawledURLList</li> <li>NumURLsThisSession</li> </ul>	Stav prolézacího modulu.
CrawlerRunLevel hodnota	<ul style="list-style-type: none"> <li>Řetězec (v angličtině) “Not started” (není spuštěn): Relace prolézacího modulu existuje, dosud však neobdržela zprávu s pokynem k zahájení zpracování dokumentů.</li> <li>“Started” (spuštěn): Probíhá spuštění prolézacího modulu.</li> <li>“Running” (běží): Prolézací modul dokončil fázi inicializace a spuštění a provádí aktivní prolézání.</li> <li>“Paused” (pozastaven): Prolézací modul obdržel pokyn k pozastavení aktivního prolézání, nikoli však k ukončení činnosti.</li> <li>“Stopping” (zastavování): Prolézací modul obdržel signál k zastavení a připravuje se na zastavení.</li> <li>“Error” (chyba): Prolézací modul je v nezotavitelném stavu a je třeba zastavit jej a znovu spustit, aby mohlo být prolézání obnoveno.</li> </ul>	Informace o aktivitě, kterou prolézací modul provádí.
CrawlerThreadState State	Řetězec (v angličtině)	Aktivita podprocesu prolézacího modulu. Toto pole obsahuje informace o aktuální aktivitě podprocesů.

Tabulka 10. Informace o stavu webového prolézacího modulu (pokračování)

Prvek	Atributy	Popis
ActiveBucket	<ul style="list-style-type: none"> <li>URL: Řetězec (specifikace adresy URL) Protokol, hostitel a port prolézáných adres URL.</li> <li>NumActURLs: Celé číslo (kladné) Počet adres URL ve fondu při jeho zpřístupnění pro prolézání (aktivaci).</li> <li>NumProcURLs: Celé číslo (nezáporné) Počet adres URL ve fondu, které byly dosud zpracovány (prolezeny nebo zamítnuty).</li> <li>TimeRem: Celé číslo Počet sekund zbývajících do vypršení časového limitu pro fond.</li> <li>Duration: Celé číslo (nezáporné) Počet sekund od aktivace fondu.</li> </ul>	Aktuální aktivita určeného webu.
CrawlRate	<p>Hodnota: Celé číslo (nezáporné)</p> <p>Počet prolézáných stránek za sekundu (souhrnně pro všechny fondy).</p>	Míra propustnosti prolézacího modulu.
RecentlyCrawledURL	<p>URL: Řetězec (specifikace adresy URL)</p> <p>Řetězec určující protokol, hostitele, port a prolezený soubor.</p>	Stránka, která byla nedávno prolezena.
NumURLsThisSession	Hodnota: Celé číslo (nezáporné)	Počet adres URL prolezených od doby, kdy tato instance prolézacího modulu (procesu) zahájila prolézání.

**Stav jiného než webového prolézacího modulu:** Po spuštění příkazu pro zjištění stavu prolézacího modulu, který není webovým prolézacím modulem, jsou vráceny informace ve formátu dokumentu XML. Příkaz **getCrawlerStatus** pro jiné než webové prolézací moduly může vrátit následující informace:

```
FFQC5314I Výsledek: <?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>
```

V následujících tabulkách jsou popsány prvky a atributy XML pro jednotlivé prolézací moduly podnikového vyhledávání s výjimkou webového prolézacího modulu. Tyto informace jsou vráceny příkazem pro zjištění stavu prolézacího modulu.

Tabulka 11. Informace o stavu prolézacího modulu pro prolézací moduly NNTP, DB2, JDBC a Notes

Název prvku a atributu	Prolézací moduly DB2 a		
	Prolézací modul NNTP	JDBC	Prolézací modul Notes
Status	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)

Tabulka 11. Informace o stavu prolézacího modulu pro prolézací moduly NNTP, DB2, JDBC a Notes (pokračování)

Název prvku a atributu	Prolézací moduly DB2 a		
	Prolézací modul NNTP	JDBC	Prolézací modul Notes
StatusMessage	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)
NumberOfServers	Počet serverů NNTP v prolézaném prostoru.	Počet databází v prolézaném prostoru.	Počet databází v prolézaném prostoru.
NumberOfCompletedServers	Počet prolezených serverů NNTP.	Počet prolezených databází.	Počet prolezených databází.
NumberOfTargets	Počet diskusních skupin v prolézaném prostoru.	Počet databází v prolézaném prostoru.	Počet pohledů a složek v prolézaném prostoru.
NumberOfCompletedTargets	Počet prolezených diskusních skupin.	Počet prolezených tabulek.	Počet prolezených pohledů a složek.
NumberOfCompletedRecords	Počet prolezených článků.	Počet prolezených záznamů.	Počet prolezených dokumentů.
RunningThreads	Počet podprocesů prolézacího modulu.	Počet podprocesů prolézacího modulu.	Počet podprocesů prolézacího modulu.

Tabulka 12. Informace o stavu prolézacího modulu pro prolézací moduly Exchange Server, DB2 Content Manager a Content Edition

Název prvku a atributu	Prolézací moduly Exchange Server, DB2 Content Manager a Content Edition		
	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Status	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)
StatusMessage	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)
NumberOfServers	Počet serverů Exchange Server v prolézaném prostoru.	Počet serverů Content Manager v prolézaném prostoru.	Počet úložišť v prolézaném prostoru.
NumberOfCompletedServers	Počet prolezených serverů Exchange Server.	Počet prolezených serverů Content Manager.	Počet prolezených úložišť.
NumberOfTargets	Počet podsložek v prolézaném prostoru.	Počet typů položek v prolézaném prostoru.	Počet tříd v prolézaném prostoru.
NumberOfCompletedTargets	Počet prolezených podsložek.	Počet prolezených typů položek.	Počet prolezených tříd položek.
NumberOfCompletedRecords	Počet prolezených dokumentů.	Počet prolezených dokumentů.	Počet prolezených dokumentů.
RunningThreads	Počet podprocesů prolézacího modulu.	Počet podprocesů prolézacího modulu.	Počet podprocesů prolézacího modulu.

Tabulka 13. Informace o stavu prolézacího modulu pro prolézací moduly QuickPlace, Domino Document Manager a pro prolézací moduly souborového systému UNIX a Windows

Název prvku a atributu	Prolézací moduly QuickPlace, Domino Document Manager a pro souborového systému UNIX a Windows		
	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX a Windows
Status	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)

Tabulka 13. Informace o stavu prolézacího modulu pro prolézací moduly QuickPlace, Domino Document Manager a pro prolézací moduly souborového systému UNIX a Windows (pokračování)

Název prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
StatusMessage	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)
NumberOfServers	Počet míst v prolézaném prostoru.	Počet knihoven v prolézaném prostoru.	Pevná hodnota 1.
NumberOfCompletedServers	Počet prolezených míst.	Počet prolezených knihoven.	0 nebo 1, jsou-li prolézány všechny podadresáře.
NumberOfTargets	Počet databází míst a databází místností v prolézaném prostoru.	Počet kabinetů v prolézaném prostoru.	Počet podadresářů v prolézaném prostoru.
NumberOfCompletedTargets	Počet prolezených databází míst a databází místností.	Počet prolezených kabinetů.	Počet prolezených podadresářů.
NumberOfCompletedRecords	Počet prolezených dokumentů.	Počet prolezených dokumentů.	Počet prolezených souborů.
RunningThreads	Počet podprocesů prolézacího modulu.	Počet podprocesů prolézacího modulu.	Počet podprocesů prolézacího modulu.

Tabulka 14. Informace o stavu prolézacího modulu pro prolézací moduly WebSphere Portal a Web Content Management

Název prvku a atributu	Prolézací modul WebSphere Portal	Prolézací modul Web Content Management
Status	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)
StatusMessage	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)	Stav: 0 - Idle (nečinný), 1 - Running (běží), 2 - Paused (pozastaven), -1 - Error (chyba)
NumberOfServers	Počet serverů v prolézaném prostoru.	Počet webů v prolézaném prostoru.
NumberOfCompletedServers	Počet prolezených serverů.	Počet prolezených webů.
NumberOfTargets	Počet serverů v prolézaném prostoru.	Počet webů v prolézaném prostoru.
NumberOfCompletedTargets	Počet prolezených serverů.	Počet prolezených webů.
NumberOfCompletedRecords	Počet prolezených dokumentů.	Počet prolezených dokumentů.
RunningThreads	Počet podprocesů prolézacího modulu.	Počet podprocesů prolézacího modulu.

**Stav prolézaného prostoru pro webový prolézací modul:** Po spuštění příkazu pro zjištění stavu prolézaného prostoru pro webový prolézací modul jsou vráceny informace ve formátu dokumentu XML. Příkaz pro zjištění stavu webového prolézaného prostoru může vrátit následující informace:

Tabulka 15. Hodnoty masek výběru pro příkaz pro zjištění stavu prolézaného prostoru webového prolézacího modulu

Bit masky	Výběr
1	Počet stránek v úložišti nezpracovaných dat.
2	Počet nalezených webů.
4	Počet webů se záznamem DNS.

Tabulka 15. Hodnoty masek výběru pro příkaz pro zjištění stavu prolézaného prostoru webového prolézacího modulu (pokračování)

Bit masky	Výběr
8	Počet webů bez záznamu DNS.
16	Počet nalezených adres URL.
32	Počet jedinečných uložených stránek.
64	Počet prolezených adres URL.
128	Počet adres URL, které nebyly prolezeny.
256	Počet prošlých adres URL.
512	Distribuce stavového kódu HTTP.

Všechny hodnoty reprezentují kumulativní součty pro všechny relace, které používají aktuální interní databázi:

```
<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"/>
  </HTTPCodeDist>
</CrawlStatus>
```

Vrácená data obsahují některé, všechny nebo žádné z následujících prvků:

Tabulka 16. Informace o stavu prolézaného prostoru pro webový prolézací modul

Prvek	Atribut	Popis
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>	Informace o kumulativním stavu prolézání (všechny relace), které lze rychle získat.
NumPagesInRDS	<p>Hodnota: Nezáporné celé číslo</p> <p>Počet stránek, které jsou aktuálně umístěny v oblasti pro dočasné ukládání v úložišti nezpracovaných dat (pouze pro tento prolézací modul).</p>	Míra zaplňování úložiště nezpracovaných dat (pouze z příspěvků tohoto prolézacího modulu).

Tabulka 16. Informace o stavu prolézaného prostoru pro webový prolézací modul (pokračování)

Prvek	Atribut	Popis
NumSitesDiscovered	Hodnota: Nezáporné celé číslo  Počet hostitelů nalezených při prolézání (nebo z počátečních položek).	Míra pokrytí prolézané domény prolézacím modulem (počet hostitelů).
NumSitesWithDNS	Hodnota: Nezáporné celé číslo  Počet hostitelů s asociovanými adresami IP (jsou převáděny prolézacím modulem na pozadí).	Míra efektivity prolézacího modulu při získávání adres IP pro hostitele, kteří jsou nalezeni podle názvů DNS v adresách URL.
NumSitesWithoutDNS	Hodnota: Nezáporné celé číslo  Počet hostitelů bez asociovaných adres IP (jsou převáděny prolézacím modulem na pozadí).	Míra efektivity prolézacího modulu při získávání adres IP pro hostitele, kteří jsou nalezeni podle názvů DNS v adresách URL.
NumURLsDiscovered	Hodnota: Nezáporné celé číslo  Počet jedinečných adres URL navštívených příslušným prolézacím modulem.	Míra pokrytí prolézané domény prolézacím modulem (počet adres URL).
NumUniquePagesSaved	Hodnota: Nezáporné celé číslo  Počet jedinečných stránek zapsaných do úložiště nezpracovaných dat k dalšímu zpracování ostatními komponentami podnikového vyhledávání.	Podíl tohoto prolézacího modulu na velikosti indexu.
NumURLsCrawled	Hodnota: Nezáporné celé číslo  Počet jedinečných adres URL prolezených příslušným prolézacím modulem.	Míra schopnosti prolézacího modulu zpracovávat data mezi koncovými body. Tento počet se liší od počtu stránek zapsaných do úložiště nezpracovaných dat, protože do úložiště nezpracovaných dat nejsou zapisovány všechny prolezené stránky.
NumURLsOverdue	Hodnota: Nezáporné celé číslo  Počet jedinečných adres URL, které jsou k dispozici pro prolézání.	Míra schopnosti prolézacího modulu procházet webový prostor.

**Stav prolézaného prostoru pro jiný než webový prolézací modul:** Po spuštění příkazu pro zjištění stavu prolézaného prostoru pro jiný než webový prolézací modul jsou vráceny informace ve formátu dokumentu XML. Příkaz **getCrawlSpaceStatus** pro jiné než webové prolézací moduly může vrátit následující informace:

```
FFQC5314I Výsledek: <?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>
```

V následujících tabulkách jsou popsány prvky a atributy XML pro jednotlivé prolézací moduly podnikového vyhledávání s výjimkou webového prolézacího modulu. Tyto informace jsou vráceny příkazem pro zjištění stavu prolézaného prostoru. Pro prolézací moduly Notes v



případě úrovně agregace 0 má položka Server@Name hodnotu název serveru + název databáze. V případě úrovně agregace 1 má položka Server@Name hodnotu název serveru + název adresáře.

Tabulka 17. Informace o stavu prolézaného prostoru pro prolézací moduly NNTP, DB2, JDBC a Notes

Název prvku a atributu	Prolézací modul NNTP	Prolézací moduly DB2 a JDBC	Prolézací modul Notes
Server@Name	Název serveru diskusních skupin	Název databáze	Název databáze nebo název adresáře
Server/Status	Stav: (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>
Server/NumberOfTargets	Počet diskusních skupin v prolézaném prostoru.	Počet databází v prolézaném prostoru.	Počet pohledů a složek nebo adresářů v prolézaném prostoru.
Server/NumberOfCompletedTargets	Počet prolezených diskusních skupin.	Počet prolezených tabulek.	Počet prolezených pohledů a složek nebo adresářů.
Server/NumberOfErrors	Počet chyb.	Počet chyb.	Počet chyb.
Server/StartTime	Čas spuštění, je-li k dispozici.	Čas spuštění, je-li k dispozici.	Čas spuštění, je-li k dispozici.
Server/EndTime	Čas dokončení, je-li k dispozici.	Čas dokončení, je-li k dispozici.	Čas dokončení, je-li k dispozici.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>• 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>• 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>• 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>• 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>• 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>• 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>
Server/ScheduleTime	Časový plán, je-li k dispozici.	Časový plán, je-li k dispozici.	Časový plán, je-li k dispozici.
Server/TotalTime	Celkový čas, je-li k dispozici.	Celkový čas, je-li k dispozici.	Celkový čas, je-li k dispozici.

Tabulka 17. Informace o stavu prolézaného prostoru pro prolézací moduly NNTP, DB2, JDBC a Notes (pokračování)

Název prvku a atributu	Prolézací modul NNTP	Prolézací moduly DB2 a JDBC	Prolézací modul Notes
Server/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.	0, 1: <ul style="list-style-type: none"> <li>0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)</li> <li>1: Prolézací modul Notes prolézá dokumenty v režimu adresářů.</li> </ul>

Tabulka 18. Informace o stavu prolézaného prostoru pro prolézací moduly Exchange Server, DB2 Content Manager a Content Edition

Název prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Server@Name	Název serveru Exchange Server.	Servery DB2 Content Manager.	Název úložiště.
Server/Status	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>
Server/NumberOfTargets	Počet podsložek v prolézaném prostoru.	Počet typů položek v prolézaném prostoru.	Počet tříd položek v prolézaném prostoru.
Server/NumberOfCompletedTargets	Počet prolezených podsložek.	Počet prolezených typů položek.	Počet prolezených tříd položek.
Server/NumberOfErrors	Počet chyb.	Počet chyb.	Počet chyb.
Server/StartTime	Čas spuštění, je-li k dispozici.	Čas spuštění, je-li k dispozici.	Čas spuštění, je-li k dispozici.
Server/EndTime	Čas dokončení, je-li k dispozici.	Čas dokončení, je-li k dispozici.	Čas dokončení, je-li k dispozici.

Tabulka 18. Informace o stavu prolézaného prostoru pro prolézací moduly Exchange Server, DB2 Content Manager a Content Edition (pokračování)

Název prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>
Server/ScheduleTime	Časový plán, je-li k dispozici.	Časový plán, je-li k dispozici.	Časový plán, je-li k dispozici.
Server/TotalTime	Celkový čas, je-li k dispozici.	Celkový čas, je-li k dispozici.	Celkový čas, je-li k dispozici.
Server/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.

Tabulka 19. Informace o stavu prolézaného prostoru pro prolézací moduly QuickPlace, Domino Document Manager a pro prolézací moduly souborového systému UNIX a Windows

Název prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
Server@Name	Adresář míst	Databáze knihoven	Pevná hodnota localhost.
Server/Status	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>
Server/NumberOfTargets	Počet databází míst a databází místností v prolézaném prostoru.	Počet kabinetů v prolézaném prostoru.	Počet podadresářů v prolézaném prostoru.

Tabulka 19. Informace o stavu prolézaného prostoru pro prolézací moduly QuickPlace, Domino Document Manager a pro prolézací moduly souborového systému UNIX a Windows (pokračování)

Název prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
Server/NumberOfCompletedTargets	Počet prolezených databází míst a databází místností.	Počet prolezených kabinetů.	Počet podadresářů v prolézaném prostoru.
Server/NumberOfErrors	Počet chyb.	Počet chyb.	Počet chyb.
Server/StartTime	Čas spuštění, je-li k dispozici.	Čas spuštění, je-li k dispozici.	Čas spuštění, je-li k dispozici.
Server/EndTime	Čas dokončení, je-li k dispozici.	Čas dokončení, je-li k dispozici.	Čas dokončení, je-li k dispozici.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>
Server/ScheduleTime	Časový plán, je-li k dispozici.	Časový plán, je-li k dispozici.	Časový plán, je-li k dispozici.
Server/TotalTime	Celkový čas, je-li k dispozici.	Celkový čas, je-li k dispozici.	Celkový čas, je-li k dispozici.
Server/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.

Tabulka 20. Informace o stavu prolézaného prostoru pro prolézací moduly WebSphere Portal a Web Content Management

Název prvku a atributu	Prolézací modul WebSphere Portal	Prolézací modul Web Content Management
Server@Name	Server WebSphere Portal	Základní adresa URL prohledávání modulu Web Content Management
Server/Status	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>5: Naplánován</li> <li>-1: Chyba</li> </ul>

Tabulka 20. Informace o stavu prolézaného prostoru pro prolézací moduly WebSphere Portal a Web Content Management (pokračování)

Název prvku a atributu	Prolézací modul WebSphere Portal	Prolézací modul Web Content Management
Server/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• 5: Naplánován</li> <li>• -1: Chyba</li> </ul>
Server/NumberOfTargets	Počet serverů v prolézaném prostoru.	Počet webů v prolézaném prostoru.
Server/NumberOfCompletedTargets	Počet prolezených serverů.	Počet prolezených webů.
Server/NumberOfErrors	Počet chyb.	Počet chyb.
Server/StartTime	Čas spuštění, je-li k dispozici.	Čas spuštění, je-li k dispozici.
Server/EndTime	Čas dokončení, je-li k dispozici.	Čas dokončení, je-li k dispozici.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>• 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>• 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>• 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu.</li> <li>• 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno.</li> <li>• 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.</li> </ul>
Server/ScheduleTime	Časový plán, je-li k dispozici.	Časový plán, je-li k dispozici.
Server/TotalTime	Celkový čas, je-li k dispozici.	Celkový čas, je-li k dispozici.
Server/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.

**Podrobný stav prolézaného prostoru pro webový prolézací modul:** Po spuštění příkazu pro zjištění podrobného stavu prolézaného prostoru pro webový prolézací modul jsou vráceny informace ve formátu dokumentu XML. Příkaz pro zjištění podrobného stavu prolézaného prostoru může vrátit následující informace:

Tabulka 21. Hodnoty masek výběru pro příkaz pro zjištění podrobného stavu prolézaného prostoru webového prolézacího modulu

Bit masky	Výběr
1	Počet stránek v úložišti nezpracovaných dat.
2	Počet nalezených webů.
4	Počet webů se záznamem DNS.
8	Počet webů bez záznamu DNS.
16	Počet nalezených adres URL.
32	Počet jedinečných uložených stránek.

Tabulka 21. Hodnoty masek výběru pro příkaz pro zjištění podrobného stavu prolézáného prostoru webového prolézacího modulu (pokračování)

Bit masky	Výběr
64	Počet prolezených adres URL.
128	Počet adres URL, které nebyly prolezeny.
256	Počet prošlých adres URL.
512	Distribuce stavového kódu HTTP.

Ukázka vrácených informací:

```
<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>
```

V následující tabulce jsou popsána všechna pole vrácená pro podrobný stav prolézacího prostoru webového prolézacího modulu:

Tabulka 22. Podrobné informace o stavu prolézáného prostoru pro webový prolézací modul

Prvek	Atributy	Popis
CrawlDetailsPerSite	<ul style="list-style-type: none"> <li>LastActivationTime:</li> <li>LastActivationDuration:</li> <li>IPAddressList:</li> <li>RobotsContent:</li> <li>HTTPCodeDist:</li> </ul>	Informace o podrobném stavu konkrétního webu, které lze rychle získat.
Web	Adresa URL	Adresa URL kořenové stránky webu.
NumURLsDiscovered	Hodnota	Počet adres URL nalezených z tohoto webu.
NumURLsOverdue	Hodnota	Počet adres URL k opětovnému prolézání z tohoto webu.
NumURLsCrawled	Hodnota	Počet adres URL prolezených pro příslušný web.
NumURLsUncrawled	Hodnota	Počet dosud neprolezených adres URL pro příslušný web.

Tabulka 22. Podrobné informace o stavu prolézaného prostoru pro webový prolézací modul (pokračování)

Prvek	Atributy	Popis
NumURLsOverdueBy	Threshold, Value: Celé číslo (kladné nebo záporné)  Parametr Value reprezentuje počet adres URL, které jsou k dispozici k opětovnému prolézání. Parametr Threshold určuje množství času, po který adresy URL čekaly na opětovné prolezení. Hodnota parametru Threshold se udává v sekundách jako časové posunutí oproti aktuálnímu času. Záporná hodnota parametru Threshold znamená, že opětovné prolezení adres URL je opožděno. Kladná hodnota parametru Threshold znamená, že opětovné prolezení adres URL je naplánováno k provedení.	Počet adres URL, které byly k dispozici pro opětovné prolezení alespoň před určitým počtem sekund nebo které budou k dispozici během tohoto počtu sekund.
NumURLsActivated	Hodnota	Počet adres uložených do paměti během posledního procházení tohoto webu a zpřístupněných pro podprocesy prolézacího modulu.
LastActivationTime	Hodnota	Počet sekund od posledního uložení adres URL z tohoto webu do paměti.
LastActivationDuration	Hodnota	Počet sekund od okamžiku, kdy byly adresy URL z tohoto webu naposledy uloženy v paměti a k dispozici pro podprocesy prolézacího modulu.
IPAddressList	IPAddress	Všechny známé adresy IP pro hostitele serveru tohoto webu.
IPAddress	Hodnota	Adresa IPv4 v notaci s tečkami pro hostitele serveru příslušného webu.
RobotsContent	Text	Text ze souboru robots, existuje-li.
HTTPCodeDist	HTTPCode	Distribuce kódů HTTP z pokusů o stahování tohoto webu.
HTTPCode	Code: Integer  Stavový kód HTTP nebo jiný interní kód.	Počet výskytů konkrétního stavového kódu HTTP během prolézání tohoto webu.

**Podrobný stav prolézaného prostoru pro jiný než webový prolézací modul:** Po spuštění příkazu pro zjištění podrobného stavu prolézaného prostoru pro jiné než webové prolézací moduly jsou vráceny informace ve formátu dokumentu XML. Příkaz **getCrawlSpaceStatusDetail** pro jiné než webové prolézací moduly může vrátit následující informace:

```
FFQC5314I Výsledek: <?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>

```

Tabulka 23. Podrobné informace o stavu prolézaného prostoru pro prolézací moduly NNTP, DB2, JDBC a Notes

Název prvku a atributu	Prolézací modul NNTP	Prolézací moduly DB2 a JDBC	Prolézací modul Notes
Target@Name	Název diskusní skupiny	Název tabulky	Název pohledu nebo složky
Target@CrawlType	Nelze použít.	0,1 (DB2); 0 (databáze JDBC)	0
		<ul style="list-style-type: none"> <li>0: Aktivní prolézání (normální)</li> <li>1: Pasivní prolézání (publikování událostí produktu DB2)</li> </ul>	
Target/Status	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>
Target/NumberOfRecords	Číslo posledního článku na serveru.	Počet prolezených záznamů.	Počet prolezených dokumentů.
Target/NumberOfCompletedRecords	Počet prolezených článků.	Počet prolezených záznamů.	Počet prolezených dokumentů.
Target/NumberOfInsertedRecords	Počet nově odeslaných článků.	Počet vložených záznamů.	Počet vložených záznamů.
Target/NumberOfUpdatedRecords	Nelze použít.	Počet aktualizovaných záznamů.	Počet aktualizovaných záznamů.
Target/NumberOfDeletedRecords	Nelze použít.	Počet odstraněných záznamů.	Počet odstraněných záznamů.
Target/StartTime	Datum a čas posledního spuštění prolézacího modulu.	Datum a čas posledního spuštění prolézacího modulu.	Datum a čas posledního spuštění prolézacího modulu.
Target/EndTime	Datum a čas dokončení prolézání.	Datum a čas dokončení prolézání.	Datum a čas dokončení prolézání.
Target/TotalTime	Čas, který prolézací modul strávil prolézáním.	Čas, který prolézací modul strávil prolézáním.	Čas, který prolézací modul strávil prolézáním.



Tabulka 23. Podrobné informace o stavu prolézaného prostoru pro prolézací moduly NNTP, DB2, JDBC a Notes (pokračování)

Název prvku a atributu	Prolézací modul NNTP	Prolézací moduly DB2 a JDBC	Prolézací modul Notes
Target/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.	0, 1: <ul style="list-style-type: none"> <li>0: Prolézací modul prolézá dokumenty v normálním režimu.</li> <li>Prolézací modul prolézá dokumenty v režimu adresářů.</li> </ul>
Target/LastUpdatedTime	Nelze použít.	Čas poslední aktualizace: <ul style="list-style-type: none"> <li>0: Aktivní prolézání (normální)</li> <li>1: Pasivní prolézání (publikování událostí produktu DB2)</li> </ul>	Nelze použít.
Target/LastResetTime	Nelze použít.	Čas posledního vynulování statistiky: <ul style="list-style-type: none"> <li>0: Aktivní prolézání (normální)</li> <li>1: Pasivní prolézání (publikování událostí produktu DB2)</li> </ul>	Nelze použít.

Tabulka 24. Podrobné informace o stavu prolézaného prostoru pro prolézací moduly Exchange Server, DB2 Content Manager a Content Edition

Název prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Target@Name	Název podsložky	Název typu položky	Název třídy položek
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>0: Neprolezen</li> <li>1: Prolézání</li> <li>2: Dokončen (není plánován)</li> <li>3: Čekání</li> <li>4: Pozastaven</li> <li>-1: Chyba</li> </ul>
Target/NumberOfRecords	Nelze použít.	Nelze použít.	Nelze použít.
Target/NumberOfCompletedRecords	Počet prolezených dokumentů.	Počet prolezených dokumentů.	Počet prolezených dokumentů.

Tabulka 24. Podrobné informace o stavu prolézaného prostoru pro prolézací moduly Exchange Server, DB2 Content Manager a Content Edition (pokračování)

Název prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Target/NumberOfInsertedRecords	Počet vložených záznamů.	Počet vložených záznamů.	Počet vložených záznamů.
Target/NumberOfUpdatedRecords	Nelze použít.	Počet aktualizovaných záznamů.	Počet aktualizovaných záznamů.
Target/NumberOfDeletedRecords	Nelze použít.	Počet odstraněných záznamů.	Počet odstraněných záznamů.
Target/StartTime	Datum a čas posledního spuštění prolézacího modulu.	Datum a čas posledního spuštění prolézacího modulu.	Datum a čas posledního spuštění prolézacího modulu.
Target/EndTime	Datum a čas dokončení prolézání.	Datum a čas dokončení prolézání.	Datum a čas dokončení prolézání.
Target/TotalTime	Čas, který prolézací modul strávil prolézáním.	Čas, který prolézací modul strávil prolézáním.	Čas, který prolézací modul strávil prolézáním.
Target/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.
Target/LastUpdateTime	Nelze použít.	Nelze použít.	Nelze použít.
Target/LastResetTime	Nelze použít.	Nelze použít.	Nelze použít.

Tabulka 25. Podrobné informace o stavu prolézaného prostoru pro prolézací moduly QuickPlace, Domino Document Manager a pro prolézací moduly souborového systému UNIX a Windows

Název prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
Target@Name	Název databáze míst nebo název databáze místností	Název databáze kabinetů	Název podadresáře
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>
Target/NumberOfRecords	Nelze použít.	Nelze použít.	Nelze použít.
Target/NumberOfCompletedRecords	Počet prolezených dokumentů.	Počet prolezených dokumentů.	Počet prolezených souborů.
Target/NumberOfInsertedRecords	Počet vložených záznamů.	Počet vložených záznamů.	Počet vložených záznamů.

Tabulka 25. Podrobné informace o stavu prolézaného prostoru pro prolézací moduly QuickPlace, Domino Document Manager a pro prolézací moduly souborového systému UNIX a Windows (pokračování)

Název prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
Target/NumberOfUpdatedRecords	Počet aktualizovaných záznamů.	Počet aktualizovaných záznamů.	Počet aktualizovaných záznamů.
Target/NumberOfDeletedRecords	Počet odstraněných záznamů.	Počet odstraněných záznamů.	Počet odstraněných záznamů.
Target/StartTime	Datum a čas posledního spuštění prolézacího modulu.	Datum a čas posledního spuštění prolézacího modulu.	Datum a čas posledního spuštění prolézacího modulu.
Target/EndTime	Datum a čas dokončení prolézání.	Datum a čas dokončení prolézání.	Datum a čas dokončení prolézání.
Target/TotalTime	Čas, který prolézací modul strávil prolézáním.	Čas, který prolézací modul strávil prolézáním.	Čas, který prolézací modul strávil prolézáním.
Target/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.
Target/LastUpdatedTime	Nelze použít.	Nelze použít.	Nelze použít.
Target/LastResetTime	Nelze použít.	Nelze použít.	Nelze použít.

Tabulka 26. Podrobné informace o stavu prolézaného prostoru pro prolézací moduly WebSphere Portal a Web Content Management

Název prvku a atributu	Prolézací modul WebSphere Portal	Prolézací modul Web Content Management
Target@Name	Název serveru WebSphere Portal	Základní adresa URL prohledávání, která reprezentuje daný web
Target@CrawlType	0	0
Target/Status	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Neprolezen</li> <li>• 1: Prolézání</li> <li>• 2: Dokončen (není plánován)</li> <li>• 3: Čekání</li> <li>• 4: Pozastaven</li> <li>• -1: Chyba</li> </ul>
Target/NumberOfRecords	Nelze použít.	Nelze použít.
Target/NumberOfCompletedRecords	Celkový počet prolezených záznamů.	Celkový počet prolezených záznamů.
Target/NumberOfInsertedRecords	Počet vložených záznamů.	Počet vložených záznamů.
Target/NumberOfUpdatedRecords	Počet aktualizovaných záznamů.	Počet aktualizovaných záznamů.
Target/NumberOfDeletedRecords	Počet odstraněných záznamů.	Počet odstraněných záznamů.

Tabulka 26. Podrobné informace o stavu prolézáného prostoru pro prolézací moduly WebSphere Portal a Web Content Management (pokračování)

Název prvku a atributu	Prolézací modul WebSphere Portal	Prolézací modul Web Content Management
Target/StartTime	Datum a čas posledního spuštění prolézacího modulu.	Datum a čas posledního spuštění prolézacího modulu.
Target/EndTime	Datum a čas dokončení prolézání.	Datum a čas dokončení prolézání.
Target/TotalTime	Čas, který prolézací modul strávil prolézáním.	Čas, který prolézací modul strávil prolézáním.
Target/AggregationLevel	0: Prolézací modul prolézá dokumenty v normálním režimu.	0: Prolézací modul prolézá dokumenty v normálním režimu.
Target/LastUpdatedTime	Nelze použít.	Nelze použít.
Target/LastResetTime	Nelze použít.	Nelze použít.

**Stav modulu analýzy:** Po spuštění příkazu pro zjištění stavu modulu analýzy jsou vráceny informace ve formátu dokumentu XML. Příkaz pro zjištění stavu modulu analýzy může vrátit následující informace:

```
FFQC5314I Výsledek:
<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>
```

V následující tabulce jsou popsány prvky XML pro informace vrácené příkazem pro zjištění stavu modulu analýzy:

Tabulka 27. Prvky pro příkaz pro zjištění stavu modulu analýzy

Prvek	Popis
Status	<ul style="list-style-type: none"> <li>• 0: Relace modulu analýzy je pro tuto kolekci zastavena.</li> <li>• 1: Relace modulu analýzy je pro tuto kolekci spuštěna.</li> </ul>

Tabulka 27. Prvky pro příkaz pro zjištění stavu modulu analýzy (pokračování)

Prvek	Popis
State	<p>Možné stavy: Initializing, Idle, Restart, Parsing, Stopped, Paused, Resuming, NoParserServiceIsAvailable.</p> <p>Stav Initializing (inicializace) indikuje, že probíhá spouštění modulu analýzy a inicializace jeho stavu.</p> <p>Stav Idle (nečinný) indikuje, že je modul analýzy ve stavu spánku na <math>N</math> minut při čekání na více dokumentů, které mají být předány z prolézacích modulů v této kolekci. Výchozí doba spánku je 300 sekund.</p> <p>Stav Restart (restartování) indikuje, že modul analýzy čeká na restartování prostředí JVM pro analýzu nebo převod na prvky. Prostedí JVM pro analýzu nebo převod na prvky je spouštěno v rámci samostatné relace; v tomto prostředí jsou dokumenty skutečně analyzovány.</p> <p>Stav Parsing (provádění analýzy) indikuje, že modul analýzy zpracovává dokumenty.</p> <p>Stav Pause (pozastaveno) indikuje, že činnost modulu analýzy byla pozastavena relací sestavování indexu pro tuto kolekci.</p> <p>Stav Resuming (pokračování) indikuje, že činnost modulu analýzy byla změněna ze stavu Pause do stavu Parsing a pokračuje na základě pokynu relace sestavování indexu pro tuto kolekci.</p> <p>Stav NoParserServiceIsAvailable (k dispozici není žádná služba modulu analýzy) indikuje, že pro zpracování dokumentů pro tuto kolekci nejsou k dispozici žádná prostředí JVM pro analýzu nebo převod na prvky. Tento stav znamená, že jsou všechna prostředí JVM pro analýzu nebo převod na prvky používána jinými kolekcemi.</p>
NumberOfDocsToBeIndexed	Počet dokumentů v úložišti pro tuto kolekci. Tento počet zahrnuje i dokumenty označené pro odstranění z příštího sestavení indexu.
ParseRate	Rychlost analýzy v počtu dokumentů za sekundu.
ParseRateMBPerHour	Rychlost analýzy v počtu MB za hodinu.
NumberOfCpmThreads	Počet podprocesů CPM, které jsou využívány prostředím JVM pro analýzu nebo převod na prvky při zpracovávání dokumentů pro tuto kolekci.
ParserServiceSession	Název prostředí JVM pro analýzu nebo převod na prvky, které zpracovává dokumenty pro tuto kolekci. Toto pole je k dispozici pouze v případě, že je modul analýzy ve stavu Parsing.
Name	Název prolézacího modulu.
Crawlerid	ID vytvořené systémem pro tento prolézací modul.
Type	Typ prolézacího modulu (webový prolézací modul, prolézací modul NNTP, prolézací modul DB2 atd.).
ParserStatus	<ul style="list-style-type: none"> <li>• 0: Dokumenty z tohoto prolézacího modulu nejsou analyzovány (relace analýzy byla zastavena).</li> <li>• 1: Dokumenty z tohoto prolézacího modulu jsou analyzovány (relace analýzy je spuštěna).</li> </ul>
NubmerOfDocsAlreadyParsed	Počet dokumentů z tohoto prolézacího modulu, které byly analyzovány.

**Stav sestavování indexu:** Po spuštění příkazu pro zjištění stavu sestavování indexu jsou vráceny informace ve formátu dokumentu XML. Příkaz pro zjištění stavu sestavování indexu může vrátit následující informace:

```

<?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>

```

V následující tabulce jsou popsány všechny prvky XML pro informace vrácené příkazem pro zjištění stavu sestavování indexu:

Tabulka 28. Prvky pro příkaz pro zjištění stavu sestavování indexu

Prvek	Popis
IndexStatusId	ID stavu indexu.
StartTime	Doba tohoto sestavování indexu v sekundách od roku 1970. Chcete-li spočítat současný čas představovaný touto dobou, použijte vzorec <code>January 1, 1970 %2B StartTime</code> . Více o čase epochy se dozvíte na stránkách <a href="http://en.wikipedia.org/wiki/Unix_epoch">http://en.wikipedia.org/wiki/Unix_epoch</a> .
Progress	Procentní část dokončení tohoto sestavování indexu.
CurrentPhase	<ul style="list-style-type: none"> <li>1: fáze přepisování úložiště</li> <li>2: fáze globální analýzy</li> <li>3: fáze sestavování indexu</li> </ul>
TotalPhase	Počet fází pro tento proces sestavování indexu. Aktuálně jde o hodnotu 3.
IndexCopyProgress	Procentní část dokončení kopírování indexu. Proces kopírování indexu kopíruje sestavený index ze serveru pro sestavování indexu na vyhledávací servery.
CurrentServer	Vyhledávací server, na který server pro kopírování indexu kopíruje index.
TotalServer	Počet vyhledávacích serverů, na které má být index zkopírován.
IndexCopyTime	Celkový čas pro nakopírování indexu na všechny vyhledávací servery.
IndexBuildTime	Celkový čas všech fází sestavování indexu.
Status	<ul style="list-style-type: none"> <li>0: Index sestaven a zkopírován</li> <li>-1: Selhání požadavku na sestavení indexu</li> <li>1: Probíhá sestavování či kopírování indexu nebo obojí</li> </ul>
JobID	Jedinečné ID přidružené ke každému požadavku na sestavení indexu.
MessagesAvailable	Logická hodnota označující, zda jsou k dispozici chybové zprávy (v případě selhání).
StopTime	Čas dokončení pro sestavování indexu (všechny fáze) a kopírování indexu.

Tabulka 28. Prvky pro příkaz pro zjištění stavu sestavování indexu (pokračování)

Prvek	Popis
TotalTime	Doba mezi časem spuštění a časem dokončení.
NumberOfDocuments	Počet dokumentů v indexu.
CurrentIndexWildcardSupport	Nastavení podpory zástupných znaků, které má být použito pro příští sestavování indexu. Možno hodnoty: None (žádná), QueryExpansion (expanze dotazu) nebo IndexExpansion (expanze indexu).
ScheduleStatus	<ul style="list-style-type: none"> <li>0, není-li pro tuto kolekci a typ indexu povolen časový plán.</li> <li>1, je-li pro tuto kolekci a typ indexu povolen časový plán.</li> </ul>
ScheduledTimeEnabled	Čas v sekundách od roku 1970, kdy bude spuštěno sestavování dalšího indexu pro tuto kolekci a typ indexu. Chcete-li spočítat současný čas představovaný uvedeným časem, použijte vzorec January 1, 1970 %2B ScheduledTimeEnabled. Více o čase epochy se dozvíte na stránkách <a href="http://en.wikipedia.org/wiki/Unix_epoch">http://en.wikipedia.org/wiki/Unix_epoch</a> .

**Stav vyhledávacího serveru:** Po spuštění příkazu pro zjištění stavu vyhledávacího serveru jsou vráceny informace ve formátu dokumentu XML. Příkaz pro zjištění stavu vyhledávacího serveru může vrátit následující informace:

```
FFQC5314I Výsledek: <?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>
```

V následující tabulce jsou popsány prvky XML pro informace vrácené příkazem pro zjištění stavu vyhledávacího serveru:

Tabulka 29. Prvky pro příkaz pro zjištění stavu vyhledávacího serveru

Prvek	Popis
SearchStatusName	Název a ID relace správce vyhledávání, která monitoruje a udržuje index vyhledávání pro tuto kolekci.
HostName	Název hostitele serveru, na kterém je spuštěn index vyhledávání.
Status	<ul style="list-style-type: none"> <li>0, není-li index vyhledávání pro tuto kolekci spuštěn.</li> <li>1, je-li index vyhledávání pro tuto kolekci spuštěn.</li> </ul>

**Podrobný stav vyhledávacího serveru:** Příkaz pro vrácení stavu vyhledávacího serveru můžete vracet následující informace:

```
FFQC5303I Proces Search Manager (node1) (sid: searchmanager.node1)
je již spuštěn. PID: 15711
FFQC5314I Výsledek: PID=18390
CacheHits=3
QueryRate=1
Port=44008
SessionId=coll.runtime.node1
CacheHitRate=0.333
ResponseTime=70
Status=1
SessionName=coll.runtime.node1.1
```

V následující tabulce jsou popsány položky v informacích vrácených příkazem pro zjištění podrobného stavu vyhledávacího serveru:

Tabulka 30. Položky pro příkaz pro zjištění podrobného stavu vyhledávacího serveru

Položka	Popis
CacheHits	Počet výsledků načtených z mezipaměti vyhledávání.
QueryRate	Počet dotazů obdržených během posledního časového intervalu. Při výchozím nastavení je délka časového intervalu pět minut.
Port	Číslo portu, který index vyhledávání používá pro přijímání dotazů.
SessionId	ID relace pro index vyhledávání této kolekce.
CacheHitRate	Počet výsledků načtených z mezipaměti vyhledávání jako procentní hodnota všech výsledků vyhledávání.
ResponseTime	Průměrná doba odezvy v milisekundách pro určený časový interval. (Výchozí nastavení je pět minut.)
Status	<ul style="list-style-type: none"> <li>• 0, není-li index vyhledávání pro tuto kolekci spuštěn.</li> <li>• 1, je-li index vyhledávání pro tuto kolekci spuštěn.</li> </ul>
SessionName	Název relace pro index vyhledávání této kolekce.

## Návratové kódy pro příkazy esadmin

Pro příkazy **esadmin** mohou být vráceny následující kódy:

Tabulka 31. Návratové kódy pro příkazy esadmin

Kód	Název	Popis
0	CODE_ERROR_NONE	Zpracování příkazu bylo úspěšně dokončeno.
102	CODE_ERROR_INSTANTIATION_EXCEPTION	Při vytváření instance manipulátoru pro příkaz došlo k chybě.
103	CODE_ERROR_ACCESS_EXCEPTION	Při vytváření instance manipulátoru pro příkaz došlo k chybě nepovoleného přístupu.
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	

## Získávání ID relací

Pomocí příkazu **esadmin check** můžete zobrazit seznam komponent podnikového vyhledávání a příslušná ID relací. Následující tabulka obsahuje seznam běžných relací, příslušná ID, informace o serveru, na němž jsou spuštěny, a informace o stavu relace.

Tabulka 32. Příklady názvů relací, serverů původce, ID relací a stavů relací

Relace	Server, na němž je relace spuštěna	ID relace	Stav relace
configmanager	indexový server	10433	Spuštěno
controller	indexový server	10464	Spuštěno
customcommunication	indexový server	Nelze použít	Nelze použít
discovery	indexový server	10649	Spuštěno



Tabulka 32. Příklady názvů relací, serverů původce, ID relací a stavů relací (pokračování)

Relace	Server, na němž je relace spuštěna	ID relace	Stav relace
monitor	indexový server	10682	Spuštěno
parserservice	indexový server	10718	Spuštěno
resource.node1	indexový server	10759	Spuštěno
samplecpp	indexový server	10827	Spuštěno
sampletest	indexový server	10857	Spuštěno
scheduler	indexový server	10889	Spuštěno
searchmanager.node1	indexový server	10927	Spuštěno
utilities.node1	indexový server	10384	Spuštěno

### Související pojmy

“Monitorování aktivity podnikového vyhledávání” na stránce 285

“Zálohování a obnovování systému podnikového vyhledávání” na stránce 315

 Zprávy podnikového vyhledávání

 Zprávy podnikového vyhledávání

### Související úlohy

“Monitorování prolézacích modulů” na stránce 288

“Spuštění systému podnikového vyhledávání” na stránce 279

“Zastavení systému podnikového vyhledávání” na stránce 281

“Správa vyhledávacích serverů v samostatném režimu” na stránce 283



---

## Rozlišování malých a velkých písmen v podnikovém vyhledávání

Komponenty podnikového vyhledávání, například syntaxe dotazu, rychlé odkazy nebo názvy polí, pracují s malými a velkými písmeny různě.

### Syntaxe dotazů

Vyhledávání nerozlišuje malá a velká písmena kromě následujících případů:

#### Názvy prvků XML v názvech atributů

Rozlišují se malá a velká písmena. U výrazů a hodnot atributů se malá a velká písmena nerozlišují ani v dotazech XML. Například v následujícím 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>
```

Následující dotazy nevrátí daný dokument:

- `@xmlxp::'author[Name ftcontains ("Ferdinand")]`
- `@xmlxp::'//contact[@type="eMail"]'`
- `@xmlf2::'<author><name>Ferdinand</name></author>`
- `@xmlf2::'<CONTACT TYPE="email">ferdi</contact>`

Tyto dotazy jej však vrátí:

- `@xmlxp::'Author[Name ftcontains ("ferdinand")]`
- `@xmlxp::'//Contact[@Type="email"]'`
- `@xmlf2::'<Author><Name>ferdinand</Author><Name>`
- `@xmlf2::'<Contact Type="email">ferdi</Contact>`

#### Seznamy přístupových práv (ACL)

Rozlišují se malá a velká písmena.

#### Adresy URL ve výrazech docid: a samegroupas:

Rozlišují se malá a velká písmena. V částech adres URL ve výrazech site: nebo url: se ale malá a velká písmena nerozlišují. Příklad: v dokumentu s adresou URL `http://www.here.com/HR/`:

- Dotaz `docid:http://www.here.com/hr` nevrátí daný dokument.
- Dotazy `url:hr` a `url:HERE` vrátí daný dokument.
- Dotaz `site:HERE.com` vrátí daný dokument.

#### ID kategorií a ID taxonomií

Nerozlišují se malá a velká písmena. V dotazu `taxonomy_id::category_id` například nehrají malá a velká písmena v `taxonomy_id` ani `category_id` roli. Dotaz `RuleBased::c42` odpovídá kategorie c42 v taxonomii založené na pravidlech a také odpovídá dotazu `rulebased::C42`. Podle názvu kategorie nelze vyhledávat, ale je to možné pomocí ID kategorie.

**Obory** Nerozlišují se malá a velká písmena. Například dotazy `Scope:RESEARCH` a `scope::research` vrátí dokumenty z oboru s názvem Research.

**Zástupné výrazy**

Nerozlišují se malá a velká písmena. Výraz Fer\*n\*d je například rovnocenný výrazu fer\*n\*d.

**Názvy polí**

Nerozlišují se malá a velká písmena. Například dotazy Title:Expenses, TITLE:expenses a title:expenses jsou všechny rovnocenné. Žádné názvy polí nerozlišují malá a velká písmena, ani když pocházejí ze souboru mapování XML. Externí zdroje však mohou u názvů polí rozlišovat malá a velká písmena podle vlastní sémantiky dotazů.

**Rychlé odkazy**

Nerozlišují se malá a velká písmena.

**Kategorizační funkce na základě pravidel**

Pravidla adres URL rozlišují malá a velká písmena, ale pravidla obsahu dokumentů nikoli.

**Sbalené identifikátory URI a upřednostněné definice identifikátorů URI na základě vzorů**

Identifikátory URI rozlišují malá a velká písmena, ale názvy sbalených skupin URI nikoli. Není možné definovat dva názvy sbalených skupin URI, které se liší pouze velikostí písmen.

**Názvy polí**

Nerozlišují se malá a velká písmena. Toto pravidlo platí pro názvy polí v dotazech, mapováních XML, definicích tříd upřednostněni a pro všechny další administrativní rozhraní, kde jsou zadávány názvy polí. Pokud v rozhraní nakonfigurujete název pole s velkými nebo smíšenými písmeny, je možné, že bude tento název v systému převeden do malých písmen, aby tak byl znázorněn při následujícím zobrazení konfigurace. V případě externího zdroje může také název pole být interpretován s rozlišením malých a velkých písmen.

**Slovníky**

Nerozlišují se malá a velká písmena. Toto pravidlo platí pro slovníky synonym, zakázaných slov, kontroly pravopisu a slov pro upřednostnění.

**Kolekce názvů a ID**

Rozlišují se malá a velká písmena. Zadáte-li v rozhraní API pro správu nebo vyhledávání název kolekce nebo ID kolekce, musí v daném názvu přesně odpovídat malá a velká písmena. Přestože však ID kolekcí rozlišují malá a velká písmena, nelze zadat dvě ID kolekcí, která se liší pouze velikostí písmen. Stejně omezení platí pro názvy a ID prolézacích modulů a zdrojů dat.

**Vyhledávací aplikace**

Rozlišují se malá a velká písmena. Není možné zadat dva názvy nebo ID aplikací, které se liší pouze velikostí písmen.

---

## Dokumentace k podnikovému vyhledávání

Dokumentaci k produktu OmniFind Enterprise Edition je k dispozici ve formátu PDF a HTML.

Instalační program produktu OmniFind Enterprise Edition automaticky nainstaluje Informační centrum produktu, které zahrnuje dokumentaci k podnikovému vyhledávání ve formátu HTML. V případě instalace s více servery je Informační centrum instalováno na oba vyhledávací servery. Pokud Informační centrum nebylo instalováno, po klepnutí na odkaz na nápovědu se otevře Informační centrum na webu společnosti IBM.

Instalované verze dokumentů PDF naleznete v adresáři `ES_INSTALL_ROOT/docs/lokalita/pdf`. Anglická verze dokumentů je například uložena v adresáři `ES_INSTALL_ROOT/docs/en_US/pdf`.

Přístup k verzi dokumentace ve formátu PDF ve všech dostupných jazycích získáte na webu Dokumentace k produktu OmniFind Enterprise Edition verze 8.5.

Z webu Podpora produktu OmniFind Enterprise Edition můžete také získat přístup k položkám ke stažení, k balíčkům oprav, technickým poznámkám a k Informačnímu centru.

V následující tabulce je uvedena dostupná dokumentace, názvy souborů a příslušná umístění.

*Tabulka 33. Dokumentace pro podnikové vyhledávání*

Název	Název souboru	Umístění
Informační centrum		<a href="http://publib.boulder.ibm.com/infocenter/discover/v8r5/">http://publib.boulder.ibm.com/infocenter/discover/v8r5/</a>
<i>Instalační příručka podnikového vyhledávání</i>	iiysi.pdf	ES_INSTALL_ROOT/docs/lokalita/pdf/
<i>Quick Start Guide</i> - tento dokument je pro angličtinu, francouzštinu a japonštinu k dispozici také v tištěné verzi	OmniFindEE850_qsg_dvoupísmenný_identifikátor_lokality.pdf	ES_INSTALL_ROOT/docs/lokalita/pdf/
<i>Správa podnikového vyhledávání</i>	iiysa.pdf	ES_INSTALL_ROOT/docs/lokalita/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/lokalita/pdf/
<i>Text Analysis Integration</i>	iiyst.pdf	ES_INSTALL_ROOT/docs/lokalita/pdf/
<i>Plug-in for Google Desktop Search</i>	iiysg.pdf	ES_INSTALL_ROOT/docs/lokalita/pdf/



---

## Funkce usnadnění přístupu

Funkce usnadnění přístupu pomáhají uživatelům s postižením, například s omezenou pohyblivostí nebo poruchou zraku, úspěšně využívat produkty informačních technologií.

Společnost IBM usiluje o to, aby její produkty byly dostupné a použitelné pro každého, bez ohledu na věk a schopnosti.

### Funkce usnadnění přístupu

V následujícím seznamu jsou uvedeny hlavní funkce usnadnění přístupu produktu OmniFind Enterprise Edition:

- ovládání pouze pomocí klávesnice,
- rozhraní běžně používaná čtecími zařízeními.

Informační centrum produktu OmniFind Enterprise Edition a příručky, které se k němu vztahují, podporují usnadnění přístupu. Funkce usnadnění přístupu Informačního centra jsou popsány na adrese [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).

### Ovládání pomocí klávesnice

Tento produkt používá standardní navigační klávesy systému Microsoft Windows.

Pro navigaci a postup instalačním programem produktu OmniFind Enterprise Edition můžete dále použít následující klávesové zkratky.

*Tabulka 34. Klávesové zkratky pro instalační program*

Akce	Zkratka
Označení přepínače	Šipka
Výběr přepínače	Klávesa Tab
Označení tlačítka	Klávesa Tab
Výběr tlačítka	Klávesa Enter
Přechod do následujícího nebo předchozího okna nebo zrušení akce	Označení tlačítka pomocí klávesy Tab a stisknutí klávesy Enter
Deaktivace aktivního okna	Ctrl + Alt + Esc

### Informace o rozhraní

Uživatelská rozhraní konzoly pro správu, ukázkové vyhledávací aplikace a modulu pro přizpůsobení vyhledávacích aplikací jsou rozhraní založená na prohlížeči, která lze zobrazit v prohlížečích Microsoft Internet Explorer či Mozilla FireFox. Seznam klávesových zkratk a dalších funkcí podporujících usnadnění přístupu pro prohlížeč Internet Explorer nebo FireFox naleznete v nápovědě online k příslušnému prohlížeči.

### Související informace o usnadnění přístupu

Publikace týkající se produktu OmniFind Enterprise Edition lze zobrazit ve formátu Adobe PDF (Portable Document Format) pomocí produktu Adobe Acrobat Reader. Dokumenty PDF jsou dodávány na disku CD, který je přibaleno k produktu, a jsou rovněž k dispozici na adrese

<http://www.ibm.com/support/docview.wss?rs=63&uid=swg27010938>.

## **Společnost IBM a usnadnění přístupu**

Další informace o závazcích společnosti IBM týkajících se usnadnění přístupu naleznete na webu IBM Human Ability and Accessibility Center.



---

## Slovníček pojmů z oblasti podnikového vyhledávání

Tento slovníček definuje pojmy používané v rozhraních a dokumentaci podnikového vyhledávání.

### **administrativní role**

Klasifikace uživatele, která určuje jeho přístup.

### **administrátor podnikového vyhledávání**

Administrativní role, která uživateli umožňuje spravovat celý systém podnikového vyhledávání.

### **adresa IP**

Jedinečná adresa zařízení nebo logické jednotky v síti odpovídající standardu IP.

### **analýza odkazů**

Metoda založená na analýze hypertextových odkazů mezi dokumenty a používaná ke zjišťování důležitosti jednotlivých stránek v kolekci pro uživatele.

### **analýza textu**

Proces extrakce sémantických a dalších informací z textu za účelem zlepšení dostupnosti a možností načítání dat v kolekci. Viz také sémantické vyhledávání.

### **analýzátor**

Program, který interpretuje dokumenty přidané do datového skladu podnikového vyhledávání. Analýzátor extrahuje informace z dokumentů a připravuje je pro indexování, vyhledávání a načítání.

### **anotace**

Informace o úseku textu. Anotace může například udávat, že určitý úsek textu reprezentuje název společnosti. V rámci architektury UIMA (Unstructured Information Management Architecture) je anotace speciálním typem struktury funkcí.

### **anotační modul**

Softwarová komponenta, která provádí specifické úlohy lingvistické analýzy a vytváří a zaznamenává anotace. Anotační modul je logická komponenta pro analýzu obsažená ve stroji pro analýzu.

### **anotační modul regulárních výrazů**

Softwarová komponenta, která detekuje entity či jednotky informací v textovém dokumentu, například čísla produktů, na základě regulárních výrazů popisujících přesné vzory vyhledávané v textu dokumentu. Pokud jeden z regulárních výrazů odpovídá určitým částem textu dokumentu, anotační modul regulárních výrazů vytvoří příslušné anotace pokrývající oblast shody nebo její část. Tyto dokumenty s anotacemi jsou poté uloženy do indexu podnikového vyhledávání (s použitím souboru mapování indexů) nebo do databáze s podporou technologie JDBC (s použitím souboru mapování databází).

### **archiv stroje pro zpracování**

Archivní soubor PEAR typu ZIP obsahující stroj pro analýzu architektury UIMA (Unstructured Information Management Architecture) a veškeré prostředky, jejichž použití je vyžadováno při vlastní analýze v rámci podnikového vyhledávání.

### **booleovské vyhledávání**

Hledání tvořené kombinací jednoho nebo více vyhledávacích výrazů s operátory AND, NOT a OR.

**certifikační autorita**

Důvěryhodná externí organizace nebo společnost, která vydává digitální certifikáty. Ty slouží k vytváření digitálních podpisů a dvojic veřejných a soukromých klíčů. Certifikační autorita zaručuje pravost identity jednotlivců, kterým jsou jedinečné certifikáty uděleny.

**certifikát**

Digitální dokument, který svazuje veřejný klíč s identitou vlastníka certifikátu, a tím umožňuje ověření vlastníka certifikátu. Certifikát je vydán a digitálně podepsán certifikační autoritou.

**cesta funkce**

Cesta používaná pro přístup k hodnotě funkce v rámci struktury funkcí architektury UIMA (Unstructured Information Management Architecture).

**diakritika**

Značka označující změnu fonetické hodnoty znaku nebo kombinace znaků.

**DIIOP (Domino Internet Inter-ORB Protocol), protokol**

Serverová úloha, která je spouštěna na serveru, pracuje s modulem Domino Object Request Broker a umožňuje komunikace mezi aplety Java, které jsou vytvářeny pomocí tříd Java Notes, a serverem Domino. Uživatelé, kteří využívají prohlížeč, a servery Domino používají protokol DIIOP ke komunikaci a k výměně objektových dat.

**DOM (Document Object Model)**

Systém, ve kterém lze strukturovaný dokument, například soubor XML, zobrazit jako strom objektů, ke kterým lze přistupovat a které lze aktualizovat prostřednictvím programů.

**dynamické řazení podle důležitosti**

Typ řazení, u kterého jsou výrazy v dotazu analyzovány s ohledem na prohledávané dokumenty a na základě získaných informací se určuje pořadí výsledků. Viz také hodnocení na základě textu. Porovnejte s termínem statické třídy důležitosti.

**dynamické shrnutí**

Typ shrnutí, v němž jsou hledané výrazy zvýrazněny a výsledky vyhledávání obsahují fráze, které nejlépe reprezentují obsah dokumentu hledaného uživatelem. Porovnejte s termínem statické shrnutí.

**externí zdroj dat**

Datový zdroj pro federování, který není prolézán, analyzován ani indexován produktem OmniFind Enterprise Edition. Prohledávání externích zdrojů dat je delegováno rozhraní API pro zpracování dotazů příslušných zdrojů dat.

**extrakce informací**

Typ extrakce pojmů, při níž jsou v textových dokumentech automaticky rozpoznávány důležité slovníkové položky, například názvy, termíny a výrazy.

**extrakce pojmů**

Funkce analýzy textu, která v textových dokumentech hledá důležité slovníkové položky (například jména osob či názvy míst a produktů) a generuje seznam těchto položek. Viz také extrakce témat.

**extrakce témat**

Typ extrakce pojmů, při níž jsou v textových dokumentech automaticky rozpoznávány důležité slovníkové položky a z nich jsou extrahovány informace o tématech a zaměření dokumentů. Viz také extrakce pojmů.

**federované vyhledávání**

Možnost vyhledávání v rámci více vyhledávacích služeb, přičemž výsledky vyhledávání jsou vráceny v konsolidovaném seznamu.

**federování**

Proces sloučení systémů názvů, díky kterému může agregovaný systém zpracovávat složené názvy platné v rámci systémů názvů.

**fronta indexů**

Seznam požadavků na zpracování sestavení hlavního indexu a rozdílových indexů.

**hledání volné shody**

Hledání vracející slova, jejichž psaný tvar se podobá hledanému výrazu.

**hodnocení na základě textu**

Proces, v němž jsou k dokumentům přiřazeny celočíselné hodnoty vyjadřující důležitost dokumentu z hlediska výrazů v dotazu. Vyšší celočíselná hodnota odpovídá přesnější shodě s dotazem. Viz také dynamické řazení.

**hybridní hledání**

Kombinace booleovského a volného textového vyhledávání.

**identifikace jazyka**

Vyhledávací funkce podnikového vyhledávání zjišťující, v jakém jazyku je dokument napsán.

**index** Viz plnotextový index.

**JavaScript**

Webový skriptovací jazyk používaný v prohlížečích a na webových serverech.

**JDBC (Java Database Connectivity), rozhraní**

Průmyslový standard pro propojení nezávislé na databázi mezi platformou Java a nejrůznějšími databázemi. Rozhraní JDBC poskytuje rozhraní API na úrovni volání pro přístup k databázi založený na jazyku SQL.

**jedinečný název**

Název, který jednoznačně identifikuje položku v adresáři. Jedinečný název sestává z dvojic atribut:hodnota oddělených čárkami. Může také jít o sadu dvojic název-hodnota (například CN=jméno osoby a C=země nebo oblast), která jednoznačně identifikuje entitu v digitálním certifikátu.

**JSP (JavaServer Page), stránky**

Serverová skriptovací technologie, která umožňuje dynamické vkládání kódu Java do webových stránek (soubory HTML) a jeho spuštění při obsluhování stránky. Klientovi je poté vrácen dynamický obsah.

**JVM (Java Virtual Machine), prostředí**

Softwarová implementace procesoru, ve kterém je spuštěn kompilovaný kód Java (aplety a aplikace).

**kabinet Domino Document Manager**

Databáze produktu Domino Document Manager používaná pro uspořádávání dokumentů. Kabinety obsahují databáze Domino.

**katakana**

Znaková sada sestávající ze symbolů používaných v jedné ze dvou běžných fonetických abeced japonštiny. Tato sada je primárně používána k fonetickému zápisu cizích slov.

**kategorie založené na pravidlech**

Kategorie vytvářené podle pravidel určujících dokumenty, které jsou asociovány s konkrétními kategoriemi. Můžete například definovat pravidla přiřazující ke speciálním kategoriím dokumenty odpovídající určitému vzoru URI (Uniform Resource Identifier) či dokumenty obsahující nebo neobsahující určitá slova.

**kmen** Viz převod slova na základní tvar.

**knihovna**

Systémový objekt, který slouží jako adresář pro jiné objekty. Viz také knihovna Domino Document Manager.

**knihovna Domino Document Manager**

Databáze produktu Domino Document Manager, která je vstupním bodem pro aplikaci Domino Document Manager.

**kód shingle**

Retězec po sobě jdoucích prvků (slov) extrahovaných z věty. Věta "Tato ukázková věta je velmi krátká." obsahuje následující kódy shingle o třech slovech (neboli trigramy):

Tato ukázková věta  
ukázková věta je  
věta je velmi  
je velmi krátká

Kódy shingle lze využívat ve statistické lingvistice. Pokud je například pro dva různé texty mnoho kódu shingle společných, texty spolu pravděpodobně do určité míry souvisejí.

**kolekce**

Sada zdrojů dat a voleb pro prolézání, analýzu, indexování a prohledávání těchto zdrojů dat.

**koncový znak**

Znak na poslední pozici ve slově.

**LDAP (Lightweight Directory Access Protocol), protokol**

Otevřený protokol, který využívá protokol TCP/IP k poskytování přístupu k adresářům podporujícím model X.500 a nemá takové požadavky na prostředky jako komplexnější protokol X.500 DAP (Directory Access Protocol). Protokol LDAP lze například použít pro vyhledání lidí, organizací nebo dalších prostředků v adresáři v Internetu nebo v intranetu.

**lemma** Základní tvar slova. Lemmata jsou významná v jazycích s vysokou mírou skloňování, například v češtině.

**lemmatizace**

Proces určující kořen a různé gramatické tvary určitého slova. Příklad: Vyhledávání slova *myš* nalezne také dokumenty obsahující slovo *myši*. Podobně vyhledávání slova *jít* nalezne i dokumenty obsahující slova *šel*, *jde* nebo *půjde*.

**lexikální afinita**

Vzájemný vztah vyhledávacích slov, která jsou si významově blízká, v dokumentu. Lexikální afinita se používá k výpočtu významnosti výsledků.

**lingvistické hledání**

Typ vyhledávání, při kterém jsou dokumenty procházeny, načítány a indexovány s použitím výrazů redukovaných na základní tvar (slovo *mice* je například indexováno ve tvaru *mouse*) nebo expandovaných s použitím základního tvaru (jako v případě složených slov).

**lokální federátor**

Objekt klienta v aplikaci podnikového vyhledávání vytvořený rozhraními API pro vyhledávání a indexaci, který umožňuje uživatelům prohledávat sadu heterogenních kolekcí a získat unifikovanou sadu výsledků vyhledávání.

**maskovací znak**

Znak používaný k reprezentaci nepovinných znaků na začátku, ve středu a na konci

vyhledávaného výrazu. Maskovací znaky se obvykle používají pro vyhledání různých tvarů výrazu v indexu. Viz také zástupný znak.

#### **méně závažné chyby, stránka**

Typ webové stránky poskytující informace o důvodech, proč nelze vrátit požadovanou webovou stránku. Server HTTP může například namísto jednoduchého stavového kódu vrátit stránku, která stavový kód podrobně vysvětluje.

#### **mezipaměť vyhledávání**

Mezipaměť, do níž se ukládají data a výsledky dřívějších vyhledávacích požadavků.

#### **místnost**

Program, který uživatelům umožňuje vytvářet dokumenty, které ostatní mohou číst, reagovat na poznámky ostatních uživatelů a prohlížet informace o stavu projektu a termínech. Uživatelé mohou rovněž vést konverzace s ostatními uživateli ve stejné místnosti. Viz také místnost Lotus QuickPlace.

#### **místnost Lotus QuickPlace**

Rozdělená oblast místa Lotus QuickPlace, k níž je přístup omezen pouze na oprávněné členy se společným zájmem a potřebou kolektivní práce.

**místo** Virtuální umístění viditelné v portálu, ve kterém se při spolupráci setkávají jednotlivci a skupiny. V portálu má každý uživatel osobní místo pro soukromou práci; kromě toho mají jednotlivci a skupiny přístup k řadě sdílených míst, která mohou být veřejná nebo k nim může být přístup omezen. Viz také místo Lotus QuickPlace.

#### **místo Lotus QuickPlace**

Webové umístění, které je k dispozici prostřednictvím produktu Lotus QuickPlace. Tento produkt umožňuje geograficky rozptýleným účastníkům spolupracovat na projektech a komunikovat online prostřednictvím strukturovaného a zabezpečeného pracovního prostoru.

#### **monitor**

Uživatel podnikového vyhledávání s oprávněním ke sledování procesů na úrovni kolekce.

#### **n-gramová segmentace**

Metoda analýzy, která považuje překrývající se posloupnosti daného počtu znaků za jedno slovo (na rozdíl od používání mezerových znaků k oddělování slov, jako například v případě segmentace pomocí mezer založené na kódu Unicode).

#### **no-follow, direktiva**

Direktiva v rámci webové stránky, která je pokynem pro roboty (například webový prolézací modul), aby neprolézaly odkazy nalezené na příslušné stránce.

#### **no-index, direktiva**

Direktiva v rámci webové stránky, která je pokynem pro roboty (například webový prolézací modul), aby obsah příslušné stránky nezahrnovaly do indexu.

#### **normalizace znaku**

Proces, při kterém jsou variantní tvary znaku (například velké a malé písmeno nebo varianty s diakritickými znaménky) převáděny na společný tvar.

#### **NRPC, vzdálené volání procedur Notes**

Komunikační mechanismus produktu Lotus Notes používaný pro veškerou komunikaci mezi produkty Notes.

#### **obecná struktura analýzy (CAS)**

Struktura, v níž je uložen obsah a metadata dokumentu spolu s veškerými výsledky analýzy vygenerovanými strojem pro analýzu textu. Veškerá výměna dat během analýzy dokumentu je řízena prostřednictvím obecné struktury analýzy.

**odebírání zakázaných slov**

Proces odebírání zakázaných slov z dotazu. Díky tomu jsou ignorována běžná slova a vráceny mohou být relevantnější výsledky.

**operátor**

Uživatel podnikového vyhledávání s oprávněním ke sledování, spouštění a zastavování procesů na úrovni kolekce.

**ovladač analyzátoru**

Služba podnikového vyhledávání, která službě analyzátoru předává dokumenty. Každé kolekci odpovídá jeden ovladač analyzátoru. Služba ovladače analyzátoru kolekce odpovídá analyzátoru kolekce v konzole pro správu podnikového vyhledávání.

**parametrické vyhledávání**

Typ hledání, při kterém jsou vyhledávány objekty obsahující číselnou hodnotu nebo atribut, například data, celá čísla nebo jiné číselné typy dat, v určeném rozsahu.

**plnotextový index**

Datová struktura s odkazy na datové položky, která slouží při hledání k nalezení dokumentů obsahujících výrazy dotazu.

**počáteční adresa URL (Uniform Resource Locator)**

Počáteční bod prolézání.

**pole** Oblast, do níž je zadávána konkrétní kategorie dat nebo řídicích informací.

**Portal Document Manager (PDM)**

Umožňuje uživatelům využívat pro spolupráci v rámci týmu jedno centrální úložiště dokumentů. Administrátoři mají možnost efektivně spravovat dokumenty a mohou řídit způsob interakce uživatelů s informacemi.

**pořadí popularity**

Typ ohodnocení důležitosti, při kterém se stávající hodnocení dokumentu zvyšuje úměrně k jeho popularitě.

**prolézací modul**

Softwarový program, který načítá dokumenty ze zdrojů dat a shromažďuje informace, které lze použít k vytvoření vyhledávacích indexů.

**prolézáný prostor**

Sada zdrojů odpovídajících určitým vzorům (například adresy URL, názvy databází, cesty v souborovém systému, názvy domén nebo adresy IP), kterou načítá prolézací modul a získává z ní položky pro indexování.

**prvek** Základní textová jednotka indexovaná v rámci podnikového vyhledávání. Prvky mohou být slova v jazyku nebo jiné textové jednotky, které lze indexovat.

**prvek zabezpečení**

Informace o identitě a zabezpečení používané pro autorizaci přístupu k dokumentům v kolekci. Různé typy zdrojů dat podporují různé typy prvků zabezpečení. Patří k nim například role uživatelů, jména uživatelů, ID skupin a další informace, které lze použít pro řízení přístupu k obsahu.

**prvkový analyzátor**

Program pro segmentaci textu, který prochází text a zjišťuje, zda a kdy lze posloupnost znaků považovat za prvek.

**převod na prvky**

Proces analýzy převádějící vstup na prvky.

**převod slova na základní tvar**

Proces lingvistické normalizace, při níž jsou různé gramatické tvary slova převáděny

na kmen - základní složku nesoucí významový obsah slova. Slova *connections*, *connective* a *connected* jsou například redukována na tvar *connect*.

### **příklonka**

Slovo, které je z hlediska syntaxe samostatné, foneticky je však připojeno k jinému slovu. Příklonka může být v psané podobě připojena ke slovu, k němuž se váže, nebo může stát samostatně. Běžným příkladem příklonek je poslední část staženého tvaru v angličtině (*wouldn't* nebo *you're*).

### **REP (Robots Exclusion Protocol), protokol**

Protokol, který umožňuje administrátorům webu určit, které části příslušného webu nemají být navštěvovány prolézajícími roboty.

### **rychlý odkaz**

Asociace mezi identifikátorem URI (Uniform Resource Identifier) a klíčovými slovy nebo frázemi.

### **řazení podle důležitosti**

Přiřazení celočíselných hodnot k jednotlivým dokumentům ve výsledcích vyhledávání vrácených dotazem. Pořadí dokumentů ve výsledcích vyhledávání je založeno na jejich významnosti z hlediska zadaného dotazu. Vyšší úroveň dat znamená větší shodu se zadanými podmínkami. Viz také dynamické řazení a statické třídy důležitosti.

### **segmentace**

Rozdělení textu do oddělených lexikálních jednotek. Zpracování, které není založeno na použití slovníků, zahrnuje segmentaci pomocí mezer a n-gramovou segmentaci, zatímco podpora založená na použití slovníků zahrnuje segmentaci podle slov, vět a odstavců a lemmatizaci.

### **segmentace pomocí mezer založená na kódu Unicode**

Metoda převodu na prvky, která na základě vlastností znaků kódu Unicode rozlišuje mezi prvky a oddělujícími znaky.

### **segmentace textu**

Viz segmentace.

### **sémantické vyhledávání**

Typ vyhledávání klíčových slov, který zahrnuje lingvistickou a kontextovou analýzu. Viz také analýza textu.

### **server proxy**

Server, který vystupuje jako prostředník pro webové požadavky HTTP, jejichž hostitelem je aplikace nebo webový server. Server proxy vystupuje jako zástupce pro servery obsahu v rámci podniku.

**servlet** Program v jazyku Java spouštěný na webovém serveru; tento program rozšiřuje funkčnost serveru generováním dynamického obsahu jako odpověď na požadavky webových klientů. Servlety se běžně používají pro připojování databází k webu.

### **sestavení hlavního indexu**

Proces sestavení celého indexu v systému podnikového vyhledávání. Porovnejte s pojmem sestavení rozdílového indexu.

### **sestavení rozdílového indexu**

Proces přidávání nových informací do existujícího indexu v systému podnikového vyhledávání. Porovnejte s pojmem sestavení hlavního indexu.

### **seznam řízení přístupu (ACL)**

V kontextu zabezpečení počítače se jedná o seznam přiřazený objektu identifikující všechny subjekty, které mohou k objektu přistupovat, a jejich přístupová práva.

**shrnutí**

Proces zahrnutí neredundantních vět stručně popisujících obsah dokumentu do výsledků vyhledávání. Viz také dynamické shrnutí a statické shrnutí.

**slitek** Dva nebo více znaků, které jsou spojeny a jeví se jako jeden znak. Příkladem slitku je spojení znaků ff a fi.

**slovník synonym**

Slovník, který uživatelům umožňuje vyhledávat při prohledávání kolekce synonyma výrazů uvedených v dotazu.

**slovo pro upřednostnění**

Slovo, které může ovlivnit relativní pořadí dokumentu ve výsledcích vyhledávání. Během zpracování dotazu může být důležitost dokumentu obsahujícího slovo pro upřednostnění zvýšena nebo snížena v závislosti na skóre, které je pro příslušné slovo předdefinováno.

**služba analyzátoru**

Služba podnikového vyhledávání, která řídí veškerou analýzu dokumentů a zpracování analýzy textu ve více kolekcích dokumentů. Vždy je spuštěna alespoň jedna služba analyzátoru.

**soubor databáze klíčů**

Viz svazek klíčů.

**soubor úložiště klíčů**

Svazek klíčů, který obsahuje jak veřejné klíče uložené jako certifikáty podepisujícího subjektu, tak soukromé klíče uložené v osobních certifikátech.

**soubory vyhledávacího indexu**

Sada souborů, do kterých se v prostředí vyhledávacího stroje ukládá index.

**správa identit**

Sada rozhraní API podnikového vyhledávání, která řídí přístup k zabezpečeným datům a umožňuje uživatelům prohledávat kolekci bez toho, že by uživatel musel zadávat své jméno a heslo pro každé úložiště v kolekci.

**SSL (Secure Sockets Layer), zabezpečení**

Protokol zabezpečení, který poskytuje ochranu soukromí při komunikaci. Pomocí protokolu SSL mohou aplikace klienta a serveru komunikovat způsobem, který brání odposlouchávání či falšování zpráv a zasahování do nich.

**statické shrnutí**

Typ shrnutí, u kterého výsledky vyhledávání obsahují předem zadaný a uložený stručný popis obsahu dokumentu. Porovnejte s termínem dynamické shrnutí.

**statické třídy důležitosti**

Typ řazení, u něhož se skóre dokumentu zvyšuje na základě různých vlastností hodnocených dokumentů, například kalendářního data, počtu odkazů na daný dokument apod. Porovnejte s pojmem dynamické řazení.

**stránka s počátečním seznamem**

V produktu WebSphere Portal jde o stránku s počátečním seznamem je stránka ve formátu XML obsahující odkazy na stránky, které jsou v rámci portálu k dispozici. Prolézací moduly používají počáteční seznam k identifikaci dokumentů, které mají být prolézány. Stránka s počátečním seznamem obsahuje také metadata uložená spolu s prolezenými dokumenty v indexu podnikového vyhledávání.

**stroj pro analýzu**

Viz stroj pro analýzu textu.



**stroj pro analýzu textu**

Softwarová komponenta, která vyhledává a reprezentuje kontext a sémantický obsah v textu.

**strom kategorií**

Hierarchie kategorií.

**struktura funkcí**

Základní datová struktura reprezentující výsledek analýzy textu. Struktura funkcí je struktura dvojic atribut-hodnota. Každá struktura funkcí je určitého typu a pro každý typ je určena sada platných funkcí nebo atributů, podobně jako v třídy Java.

**svazek klíčů**

V kontextu zabezpečení počítače se jedná o soubor obsahující veřejné a soukromé klíče, důvěryhodné kořeny a certifikáty. Viz také soubor úložiště klíčů.

**systém typů**

Systém typů definuje typy objektů (struktury funkcí), které mohou být strojem pro analýzu textu vyhledány v dokumentu. Systém typů definuje všechny možné struktury funkcí s hlediska typů a funkcí. V rámci systému typů můžete definovat libovolný počet různých typů. Systém typů je specifický pro konkrétní doménu a aplikaci.

**taxonomie**

Rozdělení objektů do skupin na základě vzájemné podobnosti. V prostředí podnikového vyhledávání taxonomie dělí data do kategorií a podkategorií. Viz také strom kategorií.

**text ve volném formátu**

Nestrukturovaný text skládající se ze slov nebo vět.

**třída upřednostnění**

Objekt obsahující specifikace, které mohou ovlivnit relativní pořadí dokumentu ve výsledcích vyhledávání.

**typ MIME**

Internetový standard pro identifikaci typu objektu, který je přenášen prostřednictvím sítě Internet.

**typ zdroje dat**

Seskupení zdrojů dat podle protokolu používaného pro přístup k datům.

**údaj pověření**

Podrobné informace získané během ověřování; tyto informace popisují uživatele, přiřazení ke skupinám a další atributy identity související se zabezpečením. Údaje pověření lze použít pro řadu služeb, například ověřování, auditování nebo delegování. Příklad: Informace pro přihlašování (jméno uživatele a heslo) uživatele jsou pověřeními, která danému uživateli umožní přístup k účtu.

**UIMA (Unstructured Information Management Architecture), architektura**

Architektura společnosti IBM definující strukturu pro implementaci systémů pro analýzu nestrukturovaných dat.

**úložiště dat**

Datová struktura, v níž jsou dokumenty uchovávány v analyzovaném tvaru.

**úložiště nezpracovaných dat**

Datová struktura, v níž jsou uloženy prolézané dokumenty před jejich odesláním analyzátoru. Prolézací moduly do úložiště nezpracovaných dat zapisují a analyzátor z něj čte. Po provedení analýzy jsou dokumenty z úložiště nezpracovaných dat odebrány. Nezaměňujte úložiště nezpracovaných dat s úložištěm dat.

**URI (Uniform Resource Identifier)**

Souvislý znakový řetězec identifikující abstraktní nebo fyzicky existující prostředek.

**URL (Uniform Resource Locator)**

Jedinečná adresa zdroje informací přístupného v síti typu Internet. Adresa URL zahrnuje zkrácený název protokolu použitého k přístupu k informačnímu zdroji a informace použité protokolem k nalezení tohoto zdroje.

**uživatelský agent**

Aplikace, která prochází web a na serverech, které navštíví, o sobě zanechává informace. V prostředí podnikového vyhledávání se jako uživatelský agent chová webový prolézací modul.

**uživatel struktury CAS**

Uživatel, který provádí konečné zpracování výsledků analýzy uložených v obecné struktuře analýzy. Uživatel například může provádět indexování obsahu obecné struktury analýzy ve vyhledávacím stroji nebo naplnit relační databázi konkrétními výsledky analýzy.

**vlastní stroj pro analýzu textu**

Stroj pro analýzu textu, který byl vytvořen s použitím sady SDK (Software Development Kit) pro architekturu UIMA (Unstructured Information Management Architecture) a lze je přidat do sady standardních strojů pro analýzu textu podnikového vyhledávání (označovaných také výrazem základní anotační moduly podnikového vyhledávání). Viz také stroj pro analýzu textu.

**volné vyhledávání textu**

Hledání, u kterého je hledaný výraz vyjádřen volně utvořeným textem.

**vrstva CCL (Common Communication Layer)**

Komunikační infrastruktura, která sjednocuje různé komponenty produktu OmniFind Enterprise Edition (řadič, analyzátor, prolézací modul, indexový server).

**vyhledávací aplikace**

Program, který zpracovává dotazy, prohledává index, vrací výsledky vyhledávání a načítá zdrojové dokumenty v podnikovém vyhledávání.

**vyhledávací stroj**

Program, který přijímá požadavky na vyhledávání a vrací uživateli seznamy dokumentů.

**vyhledávání blízkých slov**

Textové vyhledávání, které vrací výsledek, pokud se dva nebo více odpovídajících termínů vyskytuje do určité vzájemné vzdálenosti, například v téže větě nebo v témže odstavci.

**vyhledávání s použitím polí**

Dotaz omezený na určité pole.

**vyhledávání vážených výrazů**

Dotaz, v němž je některým výrazům přidělena vyšší důležitost.

**vyřadit z fronty**

Odebrat položky z fronty.

**výsledky analýzy**

Informace zpracovávané anotačními moduly. Výsledky analýzy jsou zapisovány do datové struktury nazývané obecná struktura analýzy. Výsledky analýzy, které byly vygenerovány vlastními stroji pro analýzu textu (anotačními moduly), lze zpřístupnit pro vyhledávání jejich zahrnutím do indexu podnikového vyhledávání.

**výsledky vyhledávání**

Seznam dokumentů odpovídajících vyhledávacímu požadavku.

**vzdálený federátor**

Serverový federátor, který provádí federování sady prohledávatelných objektů.

**webový prolézací modul**

Typ prolézacího modulu, který prozkoumává web načtením webového dokumentu a prolezením odkazů v rámci tohoto dokumentu.

**XPath (XML Path Language), jazyk**

Jazyk, navržený pro jedinečnou identifikaci či adresování částí zdrojových dat v jazyce XML pro použití s technologiemi, které se týkají jazyka XML, jako například analyzátoři XSLT, XQuery či XML. Jazyk XPath je standardem konsorcia World Wide Web Consortium.

**zakázané slovo**

Běžně používané slovo, například *a*, *i*, *nebo*, které aplikace při zpracování dotazu ignoruje.

**základní anotační moduly podnikového vyhledávání**

Sada standardních strojů pro analýzu textu používaných v rámci podnikového vyhledávání pro výchozí zpracování analýzy dokumentů.

**zařadit do fronty**

Zařadit zprávu nebo položku do fronty.

**zástupný znak**

Znak používaný k reprezentaci nepovinných znaků na začátku, uprostřed nebo na konci vyhledávaného výrazu.

**zdroj dat**

Libovolné datové úložiště, ze kterého lze načítat dokumenty, jako například web, relační nebo nerelační databáze či systém správy obsahu.

**zjišťovací modul**

Funkce prolézacího modulu zjišťující, které zdroje dat má prolézací modul k dispozici a může z nich načítat informace.

**znak nového řádku (newline)**

Řídící znak, který způsobí posunutí pozice při tisku nebo zobrazení o jeden řádek dolů.

**znak změny významu**

Znak, který potlačuje nebo aktivuje speciální význam jednoho nebo více následujících znaků.



---

## Doložky a ochranné známky

---

### Poznámky

Tyto informace byly vytvořeny pro produkty a služby poskytované v USA.

Společnost IBM nemusí produkty, služby nebo funkce uvedené v tomto dokumentu nabízet v ostatních zemích. Informace o produktech a službách, které jsou ve vaší oblasti aktuálně dostupné, získáte od místního zástupce společnosti IBM. Odkazy na produkty, programy nebo služby společnosti IBM v této publikaci nejsou míněny jako vyjádření nutnosti použití pouze uvedených produktů, programů či služeb společnosti IBM. Místo produktu, programu nebo služby společnosti IBM lze použít libovolný funkčně ekvivalentní produkt, program nebo službu, která neporušuje intelektuální vlastnická práva společnosti IBM. Ověření funkčnosti produktu, programu nebo služby pocházející od jiného výrobce je však povinností uživatele.

K jednotlivým subjektům popisovaným v tomto dokumentu se mohou vztahovat patenty nebo nevyřízené patentové přihlášky společnosti IBM. Vlastnictví tohoto dokumentu uživateli neposkytuje žádná licenční práva k těmto patentům. Dotazy týkající se licencí můžete posílat písemně na adresu:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive Armonk, NY  
10504-1785  
U.S.A.

Odpovědi na dotazy týkající se licencí pro dvoubajtové znakové sady (DBCS) získáte od oddělení IBM Intellectual Property Department ve vaší zemi, nebo tyto dotazy můžete zasílat písemně na adresu:

IBM World Trade Asia Corporation Licensing  
2-31 Roppongi 3-chome,  
Minato-ku  
Tokyo 106-0032, Japan

Následující odstavec se netýká Spojeného království ani jiných zemí, ve kterých je takové vyjádření v rozporu s místními zákony: **SPOLEČNOST INTERNATIONAL BUSINESS MACHINES CORPORATION TUTO PUBLIKACI POSKYTUJE "TAK, JAK JE" BEZ JAKÉKOLI ZÁRUKY, AŽ UŽ PŘÍMÉ ČI ODVOZENÉ, VČETNĚ, ALE NE VÝHRADNĚ, ODVOZENÝCH ZÁRUK TÝKAJÍCÍCH SE PORUŠOVÁNÍ ZÁKONŮ, PRODEJNOSTI ČI VHODNOSTI K URČITÉMU ÚČELU.** V některých státech nejsou prohlášení týkající se přímých či odvozených záruk v určitých případech dovolena, a proto se vás toto prohlášení nemusí týkat.

Uvedené údaje mohou obsahovat technické nepřesnosti nebo typografické chyby. Údaje zde uvedené jsou pravidelně upravovány a tyto změny budou zahrnuty v nových vydáních této publikace. Společnost IBM může kdykoli bez upozornění provádět vylepšení nebo změny v produktech či programech popsaných v této publikaci.

Veškeré uvedené odkazy na stránky WWW, které nespravuje společnost IBM, jsou uváděny pouze pro referenci a v žádném případě neslouží jako záruka funkčnosti těchto stránek. Materiály uvedené na těchto stránkách WWW nejsou součástí materiálů pro tento produkt IBM a použití uvedených stránek je pouze na vlastní nebezpečí.

Společnost IBM může použít nebo distribuovat jakékoli informace, které jí sdělíte, libovolným způsobem, který společnost považuje za odpovídající, bez vyžádání vašeho svolení.

Vlastníci licence k tomuto programu, kteří chtějí získat informace o možnostech (i) výměny informací s nezávisle vytvořenými programy a jinými programy (včetně tohoto) a (ii) oboustranného využití vyměňovaných informací, mohou kontaktovat informační středisko na adrese:

IBM Corporation  
J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003  
U.S.A.

Poskytnutí takových informací může být podmíněno dodržением určitých podmínek a požadavků zahrnujících v některých případech uhrazení stanoveného poplatku.

Licencovaný program popsáný v tomto dokumentu a veškerý licencovaný materiál k němu dostupný jsou společností IBM poskytovány na základě podmínek uvedených ve smlouvách IBM Customer Agreement, IBM International Program License Agreement nebo v jiné ekvivalentní smlouvě.

Jakékoli údaje o výkonnosti obsažené v této publikaci byly zjištěny v řízeném prostředí. Výsledky získané v jakémkoli jiném operačním prostředí se proto mohou výrazně lišit. Některá měření mohla být prováděna na vývojových verzích systémů a není zaručeno, že tato měření budou stejná i na běžně dostupných systémech. Některé údaje mohly být navíc zjištěny pomocí extrapolace. Skutečné výsledky mohou být jiné. Čtenáři tohoto dokumentu by měli zjistit použitelné údaje pro své specifické prostředí.

Informace týkající se produktů jiných výrobců pocházejí od dodavatelů těchto produktů, z jejich veřejných oznámení nebo z jiných veřejně dostupných zdrojů. Společnost IBM tyto produkty netestovala a nemůže potvrdit jejich správnou výkonnost, kompatibilitu ani žádné jiné výroky týkající se produktů jiných výrobců než IBM. Otázky týkající se kompatibility produktů jiných výrobců by měly být směřovány dodavatelům těchto produktů.

Veškerá tvrzení týkající se budoucího směru vývoje nebo záměrů společnosti IBM se mohou bez upozornění změnit nebo mohou být zrušena a reprezentují pouze cíle a plány společnosti.

Všechny uvedené ceny produktů společnosti IBM jsou aktuální maloobchodní ceny doporučené společností IBM, které se mohou změnit bez předchozího upozornění. Ceny jednotlivých prodejců se mohou lišit.

Tyto informace slouží pouze pro účely plánování. Zde uvedené informace se mohou změnit před tím, než budou popsány produkty k dispozici.

Tyto údaje obsahují příklady dat a sestav používaných v běžných obchodních operacích. Aby byla představa úplná, používají se v příkladech jména osob a názvy společností, značek a produktů. Všechna tato jména a názvy jsou fiktivní a jejich podobnost se jmény, názvy a adresami používanými ve skutečnosti je zcela náhodná.

#### LICENČNÍ INFORMACE:

Tyto informace obsahují ukázkové aplikační programy ve zdrojovém jazyce ilustrující programovací techniky na různých operačních platformách. Tyto ukázkové programy můžete bez závazků vůči společnosti IBM jakýmkoli způsobem kopírovat, měnit a distribuovat za

účelem vývoje, používání, odbytu či distribuce aplikačních programů odpovídajících rozhraní API pro operační platformu, pro kterou byly ukázkové programy napsány. Tyto příklady nebyly plně testovány za všech podmínek. Společnost IBM proto nemůže zaručit spolehlivost, upotřebitelnost nebo funkčnost těchto programů.

Každá kopie nebo část těchto ukázkových programů nebo jakákoli práce z nich odvozená musí obsahovat následující copyrightovou doložku:

© (název vaší společnosti)( rok). Části tohoto kódu jsou odvozeny z ukázkových programů společnosti IBM Corp. © Copyright IBM Corp. \_zadejte rok nebo roky\_. Všechna práva vyhrazena.

Tento produkt se skládá z následujících částí:

- Oracle® Outside In Content Access, Copyright © 1992, 2008, Oracle. Všechna práva vyhrazena.
- IBM XSLT Processor Licencované materiály - vlastnictví společnosti IBM ©Copyright IBM Corp., 1999-2008. Všechna práva vyhrazena.

## Ochranné známky

V dokumentu <http://www.ibm.com/legal/copytrade.shtml> najdete seznam ochranných známek společnosti IBM.

Následující výrazy jsou ochrannými známkami nebo registrovanými ochrannými známkami jiných společností.

Adobe, Acrobat, Portable Document Format (PDF), PostScript a všechny ochranné známky založené na termínu Adobe jsou buď registrované ochranné známky nebo ochranné známky společnosti Adobe Systems Incorporated ve Spojených státech a případně v dalších jiných zemích.

Intel, logo Intel, Intel Inside, logo Intel Inside, Intel Centrino, logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium a Pentium jsou ochranné společnosti Intel Corporation a jejich dceřinných společností ve Spojených státech a případně v dalších jiných zemích.

Java a všechny ochranné známky založené na termínu Java jsou ochrannými známkami společnosti Sun Microsystems, Inc. ve Spojených státech a případně v dalších jiných zemích.

Linux je registrovaná ochranná známka Linuse Torvaldse ve Spojených státech a případně v dalších jiných zemích.

Microsoft, Windows, Windows NT a logo Windows jsou ochrannými známkami společnosti Microsoft Corporation ve Spojených státech a případně v dalších jiných zemích.

UNIX je registrovaná ochranná známka skupiny The Open Group ve Spojených státech a případně v dalších jiných zemích.

Jména dalších společností, produktů nebo služeb mohou být ochrannými známkami nebo značkami služeb ostatních společností.





# Rejstřík

## A

- administrativní role
  - administrátor kolekce 246, 247
  - administrátor podnikového vyhledávání 246, 247
  - konfigurování 247
  - monitor 246, 247
  - operátor 246, 247
  - popis 246
- administrátor kolekce
  - konfigurace role 247
  - popis 246
- administrátor podnikového vyhledávání
  - konfigurace role 247
  - popis 246
  - změna hesla na jednom serveru 17
  - změna hesla na více serverech 18
- administrátor, heslo
  - změna na jednom serveru 17
  - změna na více serverech 18
- AdminLinkBarInclude.jsp, soubor 327
- adresová pravidla pro webové prolézací moduly 88
- adresy IP
  - adaptér zpětné smyčky 24
  - podpora duálních 24
  - podpora IPv6 25
- adresy IP, změna 23
- aktivní weby, monitorování 289, 290
- analýza kotvícího textu
  - globální analýza 257
  - indexování dokumentů 258
  - popis 243
  - zabezpečení kolekce 257
- analýza textu
  - mapování prvků XML 137
  - obecná struktura analýzy 138, 139
  - stroj pro analýzu textu 135, 136
- Analýza, popis stránky 13
- analýzátor ASCII 147
- analýzátor Stellent
  - analýza typů dokumentů 144
  - popis 143
  - přiřazení typů dokumentů 150
  - výchozí typy dokumentů 153
- analýzátory
  - typy dokumentů podporované analyzátory Stellent 153
- anotační moduly 134
- aplikace ESSearchApplication
  - config.properties, soubor 211, 227
  - spuštění 234
  - zakázání zabezpečení 274
- automatická detekce
  - jazyky 158
  - kódové stránky 159
- autorizace, popis 245

## B

- banner.jspf, soubor 334
- banner\_searchControl.jspf, soubor 334
- bos.iocp.rte, modul 76

## C

- CCLServer\_datum.log, soubor 21
- config.properties, soubor 262
  - klonování 230
  - popisy vlastností 211
  - přizpůsobení 228
  - úpravy 227
- crawl.rules, soubor 99

## Č

- čínština
  - n-gramová segmentace 160
  - odebrání znaků nového řádku z rozsahů znaků 160
- číslo portu, podnikové vyhledávání 21

## D

- data prolezených dokumentů
  - konfigurování pro webové prolézací moduly 102
- databáze DB2
  - přístup jako k externímu zdroji 239
  - přístup prolézacích modulů databáze JDBC 60
  - přístup prolézacích modulů DB2 44
- databáze Oracle
  - přístup prolézacích modulů DB2 44
- databáze Oracle,
  - přístup jako k externímu zdroji 239
  - přístup prolézacích modulů databáze JDBC 60
- databáze SQL Server
  - přístup prolézacích modulů databáze JDBC 60
  - přístup prolézacích modulů DB2 44
- datový tok systému podnikového vyhledávání 11
- DB2 Content Manager, prolézací moduly
  - nastavení v operačním systému AIX 54
  - nastavení v operačním systému Linux 54
  - nastavení v provozním prostředí Solaris 54
- DB2, prolézací moduly
  - WebSphere II Classic Federation 52
- Default.jsp, soubor 327, 334
- detekce duplicitních dokumentů
  - globální analýza 173, 247
  - konfigurace tříd upřednostnění 202
  - popis 173, 243
  - povolení zabezpečení 247

- DIOP, protokol, konfigurace prolézacího modulu 75
- direktivy no-follow
  - konfigurování 101
  - popis 101
- direktivy no-index
  - konfigurování 101
  - popis 101
- doba odezvy vyhledávání
  - monitorování 299
  - výstrahy 306
- dokumentace
  - hledání 379
  - HTML 379
  - PDF 379
- dokumentace ve formátu HTML pro podnikové vyhledávání 379
- dokumentace ve formátu PDF pro podnikové vyhledávání 379
- dokumenty HTML
  - analýza 148, 149
  - pravidla náhrady 148, 149
  - vyhledávání 132, 133
- dokumenty XML
  - nativní vyhledávání XML 142
  - vyhledávání 129
- dokumenty, ohodnocení důležitosti
  - obnovení výchozích hodnot 197
- doménová pravidla pro webové prolézací moduly 88
- domény produktu Lotus Domino 267, 268
- domény systému Windows 271
- dotaz, ověřování 251
- dotazy s nízkou návratností
  - popis 202
  - výchozí faktory upřednostnění 205
- dotazy s vysokou návratností
  - popis 202
  - výchozí faktory upřednostnění 205
- Dublin Core, prvky 133
- důležitost dokumentu
  - obnovení výchozích hodnot 197
  - povolení pro kolekci 30
  - slovníky slov pro upřednostnění 200
  - statické 196
  - třídy upřednostnění 202, 205
  - v migrovaných kolekcích 343
  - vzory identifikátorů URI 201
- dynamické řazení podle důležitosti 195
- dynamické shrnutí 190

## E

- EAR, soubory
  - aplikace ESSearchApplication 274
  - ESAdmin, aplikace 274
  - ESSearchServer, aplikace 274
- es.cfg, soubor 17, 18, 23, 24, 235, 262
- es.search.provider.jar, soubor 330, 338
- es.security.jar, soubor 325, 330, 338
- es.wp5.install.jar, soubor 325

- es.wp6.install.jar, soubor 330, 338
- ES\_INSTALL\_ROOT, popis 17, 18
- ES\_NODE\_ROOT, popis 17, 18
- es\_special\_field.default\_field, vyhrazené pole 205
- es\_special\_field.default\_metadata\_field, vyhrazené pole 205
- es\_special\_field.regular\_text, vyhrazené pole 205
- esadmin stopIndex, příkaz 168
- esadmin system startall, příkaz 347
- esadmin system stopall, příkaz 347
- ESAdmin, aplikace
  - přihlášení 16
  - zakázání zabezpečení 274
- esadmin, příkaz 347
- esapi.jar, soubor 325, 330, 338
- esacdb2.sh, skript 46
- esacdb2.vbs, skript 47
- escnote.sh, skript 72
- escnote.vbs, skript 74
- escvbr.sh, skript 42
- escvbr.vbs, skript 43
- exchangeproxypw, příkaz 236
- exchangepw, skript 17, 18
- ESPACServer.ear, soubor 325, 330, 338
- ESSearchAdapter.ear, soubor 325
- ESSearchPortlet.war, soubor 325, 330, 338
- ESSearchRegistrationPortlet.war, soubor 325
- ESSearchServer, aplikace
  - restart 227, 228
  - zakázání zabezpečení 274
- expanze dotazu
  - popis 174
  - vliv na dobu sestavování indexu 175
  - vliv na velikost indexu 175
- expanze indexu
  - popis 174
  - vliv na dobu sestavování indexu 175
  - vliv na velikost indexu 175
- externí prolézací moduly
  - aplikace modulu pro příjem dat 109
  - konfigurování 109
- externí zdroje
  - asociace s vyhledávacími aplikacemi 241
  - konfigurace 239
  - popis 239
  - zabezpečení pomocí ID aplikací 248
- externí zdroje JDBC
  - konfigurace 239
  - odstranění 239
  - ovladače JDBC 239
  - úpravy 239
- externí zdroje LDAP
  - konfigurování 239
  - odstranění 239
  - úpravy 239

**F**

- faktory upřednostnění
  - konfigurace tříd upřednostnění 202, 205
  - pro slovníky slov pro upřednostnění 198
  - pro třídy upřednostnění 204, 205
  - pro vzory identifikátorů URI 201
- federované kolekce 29
- frekvence prolézání, monitorování 291

- fronta indexů 298
- funkce usnadnění přístupu pro tento produkt 381

## G

- global.rules, soubor 99
- globální analýza
  - analýza kotvicího textu 243, 257
  - detekce duplicitních dokumentů 173, 243, 247
  - popis 6
- globální prolézání webový prostor 99
- globální zabezpečení produktu WebSphere
  - vyhledávací aplikace, přizpůsobení 228
  - vyhledávací aplikace, vlastnosti 227
  - zakázání 274

## H

- heslo administrátora podnikového vyhledávání 17, 18
- historie doby odezvy, monitorování 299
- hlavní indexy
  - plánování 164
  - popis 6, 163
  - souběžná sestavení 166
  - zjišťování změn 167
- Hloubka cesty URL 88
- hodnocení na základě textu 195
- HTTP, servery proxy 97
- HTTPS, konfigurace vyhledávacího serveru 235

## CH

- chybové zprávy
  - konfigurace serveru SMTP 309
  - příjem e-mailů 308, 310
  - zobrazení souborů žurnálu 312
  - zobrazení souborů žurnálů odstraněných dokumentů 304

## I

- I/O Completion Port, modul, konfigurace prolézacího modulu 76
- ID aplikací 248
- ID kolekce, pravidla syntaxe 30
- ID kolekcí 33
- ID relací, podnikové vyhledávání 347
- identifikátory URI
  - formáty v prostředí podnikového vyhledávání 113
  - obory 177, 178
  - odebrání z indexu 180
  - ovlivňování statického skóre 201
  - pravidla kategorií 124, 127
  - rychlé odkazy 192
  - sbalené ve výsledcích vyhledávání 178, 179
  - zobrazení podrobností 286
- ideografické jazyky 157
- implementace vyhledávacího portletu
  - es.wp5.install.jar, soubor 325

- implementace vyhledávacího portletu (*pokračování*)

- es.wp6.install.jar, soubor 330, 338
- wp5\_install, skript 325
- wp5\_uninstall, skript 329
- wp6\_cluster\_install, skript 338
- wp6\_cluster\_uninstall 341
- wp6\_install, skript 330
- wp6\_uninstall, skript 336
- Index, popis stránky 13
- indexové servery
  - spuštění 279
  - zastavení 281
- indexy
  - formáty identifikátorů URI 113
  - kotvicí text 258
  - monitorování 298
  - obory 169, 177
  - odebrání identifikátorů URI 169, 180
  - odstranění z fronty 298
  - paralelní sestavení 166
  - plánování 164
  - popis 6, 163
  - povolení časového plánu 165, 298
  - sbalené identifikátory URI 169, 178, 179
  - souběžná sestavení 166
  - startIndexBuild, příkaz 167
  - vliv zástupných znaků 175
  - výstrahy 306
  - zakázání časového plánu 165, 298
  - zástupné znaky 169, 174, 176
  - zjišťování změn 167
  - změna plánu 165
- integrace s produktem WebSphere Portal
  - es.wp5.install.jar, soubor 325
  - es.wp6.install.jar, soubor 330, 338
  - klastrované systémy 337
  - konfigurační skripty 324
  - Lotus Quickr 323
  - popis 323
  - Web Content Management 323
  - wp5\_install, skript 325
  - wp6\_cluster\_install, skript 338
  - wp6\_install, skript 330
- IOCP, konfigurace prolézacího modulu 76
- IP adresy, pravidla pro webové prolézací moduly 88

**J**

- japonština
  - n-gramová segmentace 160
  - odebrání znaků nového řádku z rozsahů znaků 160
- JavaScript, podpora ve webových prolézacích modulech 87
- jazyky
  - automatická detekce 158
  - dvouznačkové kódy 157
  - podporované 157, 158
  - vyhledávání 157
- jednobařtové kódování 159
- jednoduchý text, jazyky 157

## K

### kategorie

- formáty identifikátorů URI 113
- migrace z produktu WebSphere Portal 343
- popis 124
- stromy kategorií 126
- typ kategorizace 127
- vnořování podkategorií 126
- vyhledávání 124
- vytvoření 127
- založené na pravidlech 124, 127

### kategorie založené na pravidlech

- popis 124
- výběr typu kategorizace 127
- vytvoření 127

### klastry

- WebSphere Portal 337
- klastry produktu WebSphere Portal
  - integrační skripty 338
  - integrované pokyny 337
- odebrání podnikového vyhledávání 341

### klíčová slova v rychlých odkazech 192

### klonování

- prolézací moduly 35
- vyhledávací aplikace 230

### kódové stránky

- automatická detekce 159
- podporované 159

### kolekce

- analýza 123
- asociace s vyhledávacími aplikacemi 210
- detekce duplicitních dokumentů 173
- federování 29
- koncept 29
- migrace z produktu WebSphere Portal 343
- MigrationWizard.log, soubor 343
- monitorování 286
- odhad prostředků 285
- odhad velikosti 30
- odstranění 33
- popis 3
- stav systému 286
- úpravy 32
- určení ID 33
- vyhledávací servery 183
- vyhledávání 169
- výchozí nastavení migrace 345
- vynechání řízení přístupu na úrovni dokumentu 275
- vytvoření pomocí Průvodce kolekci 29
- vytvoření s použitím pohledu Kolekce 30
- zabezpečení 247
- zabezpečení duplicitních dokumentů 247
- zabezpečení kotvicího textu 257
- zabezpečení pomocí ID aplikací 248
- způsoby vytvoření 29
- konektor Java pro produkt DB2 Content Manager 54, 55
- konfigurace adaptéru zpětné smyčky 24
- konfigurace protokolu dotazů 311
- konfigurace serveru SMTP 309
- konfigurace serveru Trusted Server 268
- konfigurace služby Directory Assistance 270
- konfigurace uživatelů produktu Domino, prolézací moduly QuickPlace 269

### konfigurační skripty

- WebSphere Portal 324
- konzola pro správu
  - popis 8
  - přihlášení 16
  - rozhraní 13
  - úlohy, souhrnné informace 13
- korejština
  - analýza složených výrazů 141
  - n-gramová segmentace 160

## L

### lingvistická podpora

- detekce jazyka 158
- detekce kódové stránky 159
- kódy jazyků 157
- lokality 157
- mezerové znaky, odebrání 160
- n-gramová segmentace 160
- nativní vyhledávání XML 142
- sémantické vyhledávání 134, 142
- slovníky slov pro upřednostnění 198
- slovníky synonym 184
- slovníky zakázaných slov 187
- vlastní analýza textu 134

### lokality

- analýza 157
- vyhledávání 157

### Lotus Domino Trusted Server, servery 268

### Lotus Notes

- instalace modulu plug-in 321
- integrace s podnikovým vyhledáváním 321
- panel vyhledávání 321
- web aktualizace modulu plug-in 321

### Lotus Quickr

- integrace s produktem WebSphere Portal 323, 335
- konfigurace prolézacího modulu počátečního seznamu 81
- konfigurace prolézacího modulu QuickPlace 78

## M

### mapování

- obecná struktura analýzy na index 138
- obecné struktury analýzy na relační databáze 139
- pole na třídě upřednostnění 204
- prvků XML na obecnou strukturu analýzy 137
- vyhledávací pole HTML 133
- vyhledávací pole XML 129
- výsledků analýzy na relační databáze 139
- maximální interval opakovaného prolézání 92
- méně závažné chyby, zpracování stránek webovými prolézacími moduly 93
- mezerové znaky, odebrání 160
- mezipaměť vyhledávání
  - konfigurování 184
  - popis 183
- migrace
  - kolekce 343

### migrace (pokračování)

- taxonomie založená na pravidlech 343
- MIME, zahrnutí typů do prolézáných webových prostorů 88
- minimální interval opakovaného prolézání 92
- modul plug-in produktu Lotus Notes
  - instalace 321
  - web aktualizace 321
- modul plug-in více strukturovaných tabulek 62, 64
- modul pro příjem dat
  - konfigurování 109
  - monitorování 301
  - restart 109, 301
- moduly analýzy
  - analýza složených výrazů 141
  - analýza typů dokumentů 144
  - analýzátor ASCII 147
  - datová analýza, úlohy 5
  - detekce formátu dokumentu 143
  - detekce jazyka 158
  - detekce kódové stránky 159
  - lingvistické zpracování 157
  - mezerové znaky, odebrání 160
  - monitorování 297
  - n-gramová segmentace 160
  - nativní vyhledávání XML 142
  - neznámé typy dokumentů 147
  - nový řádek, odebrání znaku 160
  - podporované jazyky 157
  - podprocesy 140
  - popis 5, 123
  - pravidla náhrady kódu HTML 148, 149
  - soubory bez přípon 147
  - spuštění 297
  - stav systému 297
  - typy dokumentů pro analyzátor Stellent 150
  - typy dokumentů pro službu analyzátoru 145, 147
  - výběr typu analyzátoru 143
  - zastavení 297
- moduly plug-in
  - Prolézací moduly databáze JDBC 62, 64
  - prolézání více strukturovaných tabulek 62, 64
- moduly plug-in prolézacího modulu 110
- moduly plug-in, pro prolézací moduly 110
- monitor
  - konfigurace role 247
  - popis 246
- monitorování
  - aktivní weby webového prolézacího modulu 290
  - frekvence prolézání webového prolézacího modulu 291
  - historie doby odezvy 299
  - kolekce 286
  - modul pro příjem dat 301
  - moduly analýzy 297
  - oblíbené dotazy 299, 300
  - odstraněné dokumenty 303, 304
  - podnikové vyhledávání 285
  - podrobnosti identifikátoru URI 286
  - podrobnosti o podprocesech webového prolézacího modulu 290
  - poslední dotazy 299, 300

monitorování (*pokračování*)  
  prolézací moduly 288  
  soubory žurnálu 312  
  vyhledávací servery 299, 300  
  webové prolézací moduly 289  
Monitorování, popis pohledu 13

## N

n-gramová segmentace 160  
naposledy prolézané adresy URL,  
  monitorování 289  
nativní vyhledávání XML 142  
návrátové kódy, podnikové vyhledávání 347  
návštěva adresy URL, co nejdřívější 93  
nejlepší výsledky  
  sloupcové grafy 231  
neznámé typy dokumentů 147  
NNTP, konfigurování prolézacích modulů 68  
nový řádek, odebrání znaku 160  
NRPC, protokol, konfigurace prolézacího  
  modulu 72, 74

## O

obecná struktura analýzy  
  mapování na index 138  
  mapování na relační databáze 139  
  mapování prvků XML 137  
  popis 134  
oblíbené dotazy  
  výpočet 300  
oblíbené dotazy, monitorování 299  
obnovování podnikového vyhledávání 315  
obnovování systému 315, 317  
obory  
  formáty identifikátorů URI 113  
  popis 177  
  vyhledávání 177  
  vytvoření 178  
  vzory identifikátorů URI 177, 178  
obsah dokumentu, popis 169  
odebrání identifikátorů URI z indexu 180  
odhad systémových prostředků 285  
odstraněné dokumenty  
  konfigurace souboru žurnálu 302  
  popis 302  
  sestavy 303  
  ukládat soubory žurnálu pro 304  
odstranění  
  indexy z fronty 298  
  kolekce 33  
  prolézací moduly 38  
ochranné bariéry firewall, prolézání  
  dokumentů Exchange Server 267  
omezení prolézaného webového prostoru 88  
OmniFind Enterprise Edition  
  datový tok, schéma 11  
  ID relací 347  
  indexové servery 6  
  integrace s produktem Lotus Notes 321  
  integrace s produktem WebSphere  
  Portal 323  
  komponenty 3  
  konfigurace čísla portu 21  
  konzola pro správu 8

OmniFind Enterprise Edition (*pokračování*)  
  moduly analýzy 5  
  návrátové kódy 347  
  podpora duálních adres IP 24  
  podpora protokolu IPv6 25  
  prolézací servery 4  
  přehled 1  
  příkazy 347  
  rozhraní API 10  
  vyhledávací aplikace 10  
  vyhledávací servery 7  
  změna adres IP 23  
  změna hesla na jednom serveru 17  
  změna hesla na více serverech 18  
opakovaná návštěva adresy URL, co  
  nejdřívější 93  
opakované prolézání, intervaly pro webové  
  prolézací moduly 92  
operační systém AIX  
  DB2 Content Manager, prolézací modul,  
  konfigurování 54  
  konfigurace prolézacího modulu DB2 46  
  konfigurace prolézacího modulu Domino  
  Document Manager 72  
  konfigurace prolézacího modulu  
  Notes 72  
  konfigurace prolézacího modulu  
  QuickPlace 72  
  konfigurace publikování událostí 46  
  prolézací modul Content Edition,  
  konfigurace 42  
operační systém Linux  
  DB2 Content Manager, prolézací modul,  
  konfigurování 54  
  konfigurace prolézacího modulu DB2 46  
  konfigurace prolézacího modulu Domino  
  Document Manager 72  
  konfigurace prolézacího modulu  
  Notes 72  
  konfigurace prolézacího modulu  
  QuickPlace 72  
  konfigurace publikování událostí 46  
  prolézací modul Content Edition,  
  konfigurace 42  
  provozní prostředí Solaris  
  konfigurace publikování událostí 46  
operační systém Solaris  
  konfigurace prolézacího modulu DB2 46  
operační systém Windows  
  DB2 Content Manager, prolézací modul,  
  konfigurování 55  
  konfigurace prolézacího modulu 74  
  konfigurace prolézacího modulu DB2 47  
  konfigurace publikování událostí 47  
  prolézací modul Content Edition,  
  konfigurace 43  
operátor  
  konfigurace role 247  
  popis 246  
ověření aktuálních pověřovacích údajů 251,  
  267, 268, 271  
ověřování  
  popis 245  
  zakázání pro podnikové aplikace 274  
ověřování založené na použití formulářů 95,  
  96

ovladače JDBC  
  pro externí zdroje JDBC 239  
  pro prolézací moduly databáze JDBC 60

## P

paralelní sestavení indexu 166  
parametr distinctRecentQueryCheck 300  
parametr newHtmlTagReplacement 148  
parametrická pole  
  numerické řazení 169  
  popis 169  
plánování  
  prolézací moduly 35, 39  
  sestavení indexu 164, 165  
počáteční adresy URL pro webové prolézací  
  moduly 88, 93  
podnikové aplikace  
  aplikace ESSearchApplication 274  
  ESAdmin, aplikace 274  
  ESSearchServer, aplikace 274  
podnikové vyhledávání  
  administrativní role 246  
  datový tok, schéma 11  
  formáty identifikátorů URI 113  
  ID relací 347  
  indexové servery 6, 163  
  integrace s produktem Lotus Notes 321  
  integrace s produktem WebSphere  
  Portal 323  
  komponenty 3  
  konfigurace čísla portu 21  
  konzola pro správu 8  
  moduly analýzy 5, 123  
  monitorování 285  
  návrátové kódy 347  
  obnovování ze zálohy 317  
  prolézací servery 4, 35  
  přehled 1  
  příkazy 347  
  rozhraní API 10  
  skripty obnovování 315  
  skripty zálohování 315  
  soubory žurnálu 305  
  spuštění serverů 279  
  spuštění vyhledávacích serverů 283  
  vyhledávací aplikace 10  
  vyhledávací servery 7, 183  
  zabezpečení 243  
  zabezpečení na úrovni dokumentu 249  
  zabezpečení na úrovni kolekce 247  
  zálohování 316  
  zastavení vyhledávacích serverů 283  
  zastavování serverů 279, 281  
podprocesy  
  moduly analýzy 140  
  webový prolézací modul 290  
podrobnosti identifikátoru URI  
  monitorování 286  
  odstraněné dokumenty 303  
podrobnosti o podprocesech,  
  monitorování 289  
Pohled Kolekce  
  popis 13  
  vytváření kolekcí 30  
pole metadat, sloupcové grafy nejlepších  
  výsledků 231

- pole, mapování na třídy upřednostnění 204
- portlety
  - integrace s klastrem produktu WebSphere Portal 338
  - integrace s produktem WebSphere Portal 5.1 325
  - integrace s produktem WebSphere Portal 6 330, 335
  - nastavení zdrojů produktu Lotus Quickr 335
  - odebrání z klastrů produktu WebSphere Portal 341
  - odebrání z produktu WebSphere Portal 5.1 329
  - odebrání z produktu WebSphere Portal verze 6 336
  - podnikové vyhledávání 323
  - popis 323
- poslední dotazy
  - výpočet 300
- poslední dotazy, monitorování 299
- povolení časových plánů indexování 165
- pravidla kategorií
  - konfigurování 127
  - obsah dokumentu 124, 127
  - vzory identifikátorů URI 124, 127
- pravidla náhrady kódu HTML 148, 149
- profily uživatelů
  - konfigurace 257
  - popis 254
- prolézací moduly
  - aplikace modulu pro příjem dat 109
  - Content Edition 39, 41, 42
  - Databáze JDBC 60, 62, 64
  - DB2 44
  - DB2 Content Manager 53
  - Domino Document Manager 57
  - Exchange Server 59, 267
  - formáty identifikátorů URI 113
  - kombinování typů prolézacích modulů 35
  - moduly plug-in 110
  - monitorování 288
  - NNTP 68
  - Notes 68, 71
  - odstranění 38
  - plánování 35, 39
  - počáteční hodnoty 36
  - Počáteční seznam 81
  - podpora externích 109
  - popis 4
  - přehled konfigurace 35
  - QuickPlace 78
  - souborový systém UNIX 82
  - soubory archivu 112
  - stav systému 288
  - úprava vlastností prolézacího modulu 37
  - úpravy prolézacího prostoru 38
  - výchozí nastavení migrace 345
  - vytvoření 36
  - Web Content Management 103, 106
  - webový prolézací modul 83
  - WebSphere Portal 105, 106
  - Windows, souborový systém 107
  - zabezpečení na úrovni dokumentu 35, 249
  - základní hodnoty 35
- prolézací moduly Content Edition
  - formáty identifikátorů URI 113
  - konfigurace 39
  - nastavení v operačním systému AIX 42
  - nastavení v operačním systému Linux 42
  - nastavení v provozním prostředí Solaris 42
  - nastavení v systému Windows 43
  - přímý režim 41
  - režim serveru 42
- Prolézací moduly databáze JDBC
  - formáty identifikátorů URI 113
  - konfigurace 60
  - modul plug-in k prolézání více tabulek 62, 64
  - podporované ovladače 60
  - prolézání více tabulek 62, 64
- prolézací moduly DB2
  - formáty identifikátorů URI 113
  - instalace produktu WebSphere MQ v systému AIX 46
  - instalace produktu WebSphere MQ v systému Linux 46
  - instalace produktu WebSphere MQ v systému Solaris 46
  - instalace produktu WebSphere MQ v systému Windows 47
  - konfigurace 44
  - konfigurace produktu WebSphere MQ 50
  - konfigurace publikování událostí 46, 47
  - nastavení v systému AIX 46
  - nastavení v systému Linux 46
  - nastavení v systému Solaris 46
  - nastavení v systému Windows 47
  - publikování událostí 44
  - WebSphere II Event Publisher Edition, konfigurace 48
- prolézací moduly DB2 Content Manager
  - formáty identifikátorů URI 113
  - konfigurace 53
  - nastavení v systému Windows 55
- prolézací moduly Domino Document Manager
  - DIIOP, konfigurace protokolu 75
  - formáty identifikátorů URI 113
  - IOCP, konfigurace 76
  - konfigurace 57
  - nastavení v operačním systému AIX 72
  - nastavení v operačním systému Linux 72
  - nastavení v provozním prostředí Solaris 72
  - nastavení v systému Windows 74
  - NRPC, protokol 72, 74
- prolézací moduly Exchange Server
  - formáty identifikátorů URI 113
  - konfigurace 59
  - zabezpečené dokumenty 267
- prolézací moduly Notes
  - DIIOP, konfigurace protokolu 75
  - formáty identifikátorů URI 113
  - IOCP, konfigurace 76
  - konfigurace 68
  - Lotus Domino Trusted Server 268
  - nastavení v operačním systému AIX 72
  - nastavení v operačním systému Linux 72
  - nastavení v provozním prostředí Solaris 72
  - nastavení v systému Windows 74
- prolézací moduly Notes (*pokračování*)
  - NRPC, protokol 72, 74
  - ověření aktuálních pověřovacích údajů 268
  - pravidla pro mapování polí 71
  - radly pro použití 71
  - zabezpečení na úrovni dokumentu 267
- prolézací moduly počátečního seznamu
  - formáty identifikátorů URI 113
  - integrace s produktem WebSphere Portal 323, 335
  - konfigurace 81
- prolézací moduly QuickPlace
  - DIIOP, konfigurace protokolu 75
  - formáty identifikátorů URI 113
  - IOCP, konfigurace 76
  - konfigurace 78
  - konfigurace služby Directory Assistance 270
  - konfigurace uživatelů produktu Domino 269
  - nastavení v operačním systému AIX 72
  - nastavení v operačním systému Linux 72
  - nastavení v provozním prostředí Solaris 72
  - nastavení v systému Windows 74
  - NRPC, protokol 72, 74
  - zabezpečení lokálních uživatelů 269
- prolézací moduly souborového systému UNIX
  - formáty identifikátorů URI 113
  - konfigurace 82
- prolézací moduly souborového systému Windows
  - formáty identifikátorů URI 113
  - konfigurace 107
  - zabezpečení na úrovni dokumentu 271
- Prolézací moduly Web Content Management
  - formáty identifikátorů URI 113
  - konfigurace 103
  - kopírování adres URL webů 106
- prolézací moduly WebSphere Portal
  - formáty identifikátorů URI 113
  - konfigurace 105
- prolézací servery
  - spuštění 279, 288
  - zastavení 281, 288
- Prolézání, popis stránky 13
- prolézací prostor
  - popis 4
  - úpravy 38
  - výstrahy 306
  - webový prolézací modul, konfigurace 88
- protokol IPv6 25
- provozní prostředí Solaris
  - DB2 Content Manager, prolézací modul, konfigurování 54
  - konfigurace prolézacího modulu Domino Document Manager 72
  - konfigurace prolézacího modulu Notes 72
  - konfigurace prolézacího modulu QuickPlace 72
  - prolézací modul Content Edition, konfigurace 42
- proxy, servery 97
- konfigurace vyhledávacího serveru 236
- Průvodce kolekcí 29

- průvodce migrací
  - kolekce 343
  - popis 343
  - spuštění 343
  - taxonomie založené na pravidlech 343
  - výchozí nastavení kolekce 345
  - výchozí nastavení prolézacího modulu 345
  - žurnál, soubor 346
- prvky XML
  - mapování na obecnou strukturu analýzy 137
  - mapování na vyhledávací pole 129
  - vyhledávání 129, 137
- prvky zabezpečení
  - konfigurace prolézacího modulu 250
  - zabezpečení na úrovni dokumentu 250
  - zakázání pro kolekci 275
- předpony, pravidla pro webové prolézací moduly 88
- přihlášení ke konzole pro správu 16
- příkaz esadmin startSearch 283
- příkaz esadmin stopSearch 283
- příkaz eschangelustpw 235
- příkaz eschangewaspw
  - konfigurace pro jediný server 262
  - konfigurace pro více serverů 262
- příkazy, podnikové vyhledávání 347
- přímý režim, úložiště Content Edition 41
- přípony souborů
  - podporované analyzátoři kolekcí 145, 147
  - podporované analyzátoři Stellent 150, 153
  - vyloučení z prolézáných webových prostorů 88
- přizpůsobení vyhledávacích aplikací 227, 228
- Přizpůsobení vyhledávacích aplikací
  - config.properties, soubor 228
  - spuštění 228
- publikování událostí
  - konfigurace prolézacího modulu DB2 48, 50
  - nastavení v operačním systému AIX 46
  - nastavení v operačním systému Linux 46
  - nastavení v provozním prostředí Solaris 46
  - nastavení v systému Windows 47
  - popis 44

## R

- registr uživatelů LDAP 260
- removeCjNewlineChars, volba 160
- removeCjNewlineCharsMode, volba 160
- režim serveru, úložiště Content Edition 42
- Robots Exclusion, protokol
  - dodržení protokolu webovým prolézacím modulem 85
  - uživatelský agent, identifikace 84
- robots.txt, soubory
  - dodržení protokolu webovým prolézacím modulem 85
  - uživatelský agent, identifikace 84
- rozdílové indexy
  - plánování 164

- rozdílové indexy *(pokračování)*
  - popis 6, 163
  - souběžná sestavení 166
  - zjišťování změn 167
- rozhraní API
  - popis 10
  - vyhledávání a indexování 10, 209
- rychlé odkazy
  - formáty identifikátorů URI 113
  - popis 192
  - vyhledávání 192
  - vytvoření 192

## Ř

- řaditelná pole
  - numerické řazení 169
  - řazení řetězců 169
- řazení výsledků vyhledávání
  - dynamické 195
  - hodnocení na základě textu 195
  - obnovení výchozích hodnot 197
  - popis 195
  - runtime.properties, soubor 197
  - slovníky slov pro upřednostnění 200
  - statické 196, 197
  - třídy upřednostnění 202, 204, 205
  - vzory identifikátorů URI 201
- řetězec kontextu zabezpečení uživatele 252
- řetězec USC 252
- řízení přístupu
  - ověření aktuálních uživatelů 251
  - popis 245
  - požadavky pro produkt Lotus Domino 267, 268
  - požadavky pro souborové systémy Windows 271
  - správa identit 252, 254
  - zabezpečení na úrovni dokumentu 250
  - zabezpečení SSO (single sign-on) 255
  - zakázání pro kolekci 275

## S

- sbalené identifikátory URI
  - konfigurování 179
  - omezení zabezpečení 276
  - popis 178
- sbalené výsledky vyhledávání
  - konfigurování 179
  - omezení zabezpečení 276
  - popis 178
- SearchBarInclude.jsp, soubor 327
- sémantické vyhledávání 134, 137, 142
- servery modulu analýzy
  - konfigurace podprocesů 140
- servery podnikového vyhledávání
  - podpora duálních adres IP 24
  - podpora protokolu IPv6 25
  - změna adres IP 23
- sestavění indexu
  - paralelní 166
  - plánování 165
  - popis 163
  - souběžná 166
  - spuštění 298

- sestavění indexu *(pokračování)*
  - startIndexBuild, příkaz 167
  - stav systému 298
  - zastavení 168, 298
  - zjišťování změn 167
- sestavění indexů 163
- sestavy historie prolézacího modulu
  - popis 289
  - sestavy stavových kódů HTTP 292
  - vytvoření 292
  - webová sestava 292
- sestavy podrobnosti webu
  - popis 289
  - vytvoření 292
- SI-API (rozhraní API pro vyhledávání a indexaci) 10, 209
- siapi.jar, soubor 325
- skript esbackup.bat 316
- skript esbackup.sh 316
- skript escrcm.sh 54
- skript escrcm.vbs 55
- skript esrestore.bat 317
- skript esrestore.sh 317
- skript startccl 317
- skript wp6\_cluster\_uninstall 341
- skripty
  - esbackup.bat 316
  - esbackup.sh 316
  - escrcm.sh 54
  - escrcm.vbs 55
  - escrdb2.sh 46
  - escrdb2.vbs 47
  - escrnote.sh 72
  - escrnote.vbs 74
  - escrvbr.sh 42
  - escrvbr.vbs 43
  - esrestore.bat 317
  - esrestore.sh 317
  - startccl 317
- skripty obnovování
  - popis 315
  - spouštění 317
- skripty zálohování
  - popis 315
  - spouštění 316
- sledování dokumentů
  - konfigurace souboru žurnálu 302
  - popis 302
  - povolení 302
  - sestavy 303
  - soubory žurnálu 304
  - zakázání 302
- sloupcové grafy
  - nejlepší výsledky 231
  - třídy Java pro nejlepší výsledky 231
- slovníky ignorovaných slov
  - opakovaná implementace 189
- slovníky slov pro upřednostnění
  - asociace s kolekcí 200
  - opakovaná implementace 189
  - popis 198
  - přidávání do systému 200
- slovníky synonym
  - asociace s kolekcí 186
  - opakovaná implementace 189
  - popis 184
  - přidávání do systému 186

- slovníky zakázaných slov
  - asociace s kolekcí 188
  - popis 187
  - přidávání do systému 188
- složený výraz, analýza 141
- složitý text, jazyky 157
- souběžná sestavení indexu 166
- soubor ccl.properties 23
- soubor cookies.ini
  - formát 98
  - konfigurování 98
  - popis 97
- soubor crawler\_rdb\_plugin.xml 64
- soubor followindex.rules
  - konfigurování 101
  - popis 101
- soubor nodes.ini 23, 235
- soubor parserdriver.collection.properties 148
- soubor parserTypes.cfg 143, 145, 147
- soubor runtime-generic.properties 191, 300
- soubory archivu
  - formáty identifikátorů URI 113
  - podporované formáty 112
  - prolézání 112
- soubory cookie pro prolézání webu
  - formát 98
  - konfigurování 98
  - popis 97
- soubory tar
  - formáty identifikátorů URI 113
  - prolézání 112
- soubory úložiště klíčů 235
- soubory zip
  - formáty identifikátorů URI 113
  - prolézání 112
- soubory žurnálu
  - cyklické použití 308
  - filtrování 312
  - konfigurace serveru SMTP 309
  - konfigurace velikosti 311
  - maximální velikost 308
  - monitorování 304, 312
  - popis 305
  - pro sledování dokumentů 302
  - protokoly dotazů 311
  - průvodce migrací 346
  - úrovně závažnosti 308
  - volby e-mailu 310
  - výchozí umístění 305
  - zobrazení 312
  - zobrazení odstraněných dokumentů 304
- souhrny
  - dynamické 190
  - přizpůsobení 190, 191
- souhrny dokumentů
  - přizpůsobení 190
  - úprava vlastností 191
- správa identit
  - extrakce skupiny 252
  - konfigurace 257
  - kontext zabezpečení uživatele 252
  - podpora SSO 255
  - popis 252
  - profily uživatelů 254
  - řetězec dotazu XML 252
  - zakázání 252
- spuštění
  - modul pro příjem dat 301
  - moduly analýzy 297
  - prolézací servery 288
  - průvodce migrací 343
  - Přizpůsobení vyhledávacích aplikací 228
  - servery podnikového vyhledávání 279
  - sestavení indexu 298
  - vyhledávací aplikace 234
  - vyhledávací servery 283, 299
- SSL, konfigurace vyhledávacího serveru 235
- startIndexBuild, příkaz 167
- statické třídy důležitosti
  - popis 196
  - povolení pro kolekci 30
  - v migrovaných kolekcích 343
- stav systému
  - kolekce 286
  - moduly analýzy 297
  - prolézací moduly 288
  - sestavení indexu 298
  - vyhledávací servery 299
  - webové prolézací moduly 289
- stavové kódy HTTP
  - přijaté webovými prolézacími moduly 292
  - webový prolézací modul, sestava 292
- stellent.properties, soubor 150
- stellentypes.cfg, soubor 150
- stellentTypes.cfg, soubor 143
- stroj pro analýzu textu
  - asociace s kolekcemi 136
  - mapování obecné struktury analýzy na relační databáze 139
  - mapování prvků XML 137
  - mapování výsledků analýzy na index 138
  - mapování výsledků analýzy na relační databáze 139
  - popis 134
  - přidávání do systému 135
- stromy kategorií
  - migrace z produktu WebSphere Portal 343
  - popis 126
- Systém, popis pohledu 13
- systémové prostředky
  - odhad 285
- T**
  - taxonomie založená na pravidlech, migrace z produktu WebSphere Portal 343
  - taxonomie, migrace z produktu WebSphere Portal 343
  - třídy upřednostnění
    - detekce duplicitních dokumentů 202
    - dotazy s nízkou návratností 202, 205
    - dotazy s vysokou návratností 202, 205
    - konfigurace 204, 205
    - mapování polí 204
    - popis 202
    - výchozí hodnoty 205
  - typ kategorizace
    - výběr 30, 127
    - založené na pravidlech 124
  - typy dokumentů
    - analýza 144
  - typy dokumentů (*pokračování*)
    - detekce 143
    - podporované analyzátoři Stellent 153
    - pro analyzátoři Stellent 150
    - pro službu analyzátoři 145, 147
  - typy prolézacích modulů
    - kombinování v kolekci 35
    - základní hodnoty 35
  - typy zdrojů dat
    - databáze CA-Datacom 44
    - databáze DB2 44, 60, 239
    - databáze DB2 for iSeries 44
    - databáze Domino Document Manager 57
    - databáze IMS 44
    - databáze Informix 44
    - databáze JDBC 60, 62, 64, 239
    - databáze Oracle, 44, 60, 239
    - databáze QuickPlace 78
    - databáze Software AG Adabas 44
    - databáze SQL Server 44, 60
    - databáze Sybase 44
    - databáze VSAM 44
    - DB2 for z/OS 44
    - Exchange Server, veřejné složky 59
    - Lotus Quickr for Domino 78
    - Lotus Quickr for WebSphere Portal 81
    - NNTP, diskusní skupiny 68
    - Notes, databáze 68, 71
    - podpora externích 2, 10
    - podporované podnikovým vyhledáváním 2
    - relační databáze 44
    - souborové systémy UNIX 82
    - souborový systém Windows 107
    - typy položek DB2 Content Manager 53
    - úložiště Content Edition 39, 41, 42
    - webové servery 83
    - Weby produktu Web Content Management 103
    - weby WebSphere Portal 105
- U**
  - UIMA
    - asociace s kolekcemi 136
    - mapování obecné struktury analýzy na index 138
    - mapování obecné struktury analýzy na relační databáze 139
    - mapování prvků XML 137
    - mapování výsledků analýzy na index 138
    - mapování výsledků analýzy na relační databáze 139
    - obecná struktura analýzy 138, 139
    - popis 134
    - přidávání strojů pro analýzu textu do systému 135
    - ukázková vyhledávací aplikace
      - config.properties, soubor 211, 227
      - klonování 230
      - popis 10, 210
      - vyhledávací funkce 209, 210
      - výchozí implementace 227
      - vynucení protokolu HTTPS 235
      - vynucení protokolu SSL 235
      - zakázání zabezpečení 274

- úlohy konzoly pro správu, souhrnné informace 13
- unicode, kódování 159
- úplná shoda - vyhledávací pole, popis 169
- upozornění e-mailem
  - konfigurace serveru SMTP 309
  - pro výstrahy 310
  - pro zprávy 310
- úpravy
  - aplikace modulu pro příjem dat 109
  - kolekce 32
  - prolézané prostory 38
  - vlastnosti prolézacího modulu 37
  - vyhledávací aplikace, vlastnosti 211, 227
- uživatel serveru WebSphere Application Server
  - heslo při konfiguraci jediného serveru 262
  - heslo při konfiguraci více serverů 262
- uživatelští agenti 84

## V

- vbr\_access\_services.jar, soubor 42, 43
- vícebajtové kódování 159
- vlastní analýza textu
  - mapování obecné struktury analýzy na index 138
  - mapování obecné struktury analýzy na relační databázi 139
  - mapování prvků XML 137
  - mapování výsledků analýzy na index 138
  - mapování výsledků analýzy na relační databázi 139
  - popis 134
  - stroj pro analýzu textu 135, 136
- vlastnosti prolézacího modulu
  - popis 4
  - úpravy 37
- volby vyhledávání
  - obsah dokumentu 169
  - parametrické vyhledávání 169
  - pro výsledky vyhledávání 169
  - řaditelná pole 169
  - úplná shoda 169
  - volné vyhledávání textu 169
  - vyhledávání s použitím polí 169
- volné vyhledávání textu, popis 169
- vyhledávací a indexové rozhraní API 10, 209
- vyhledávací aplikace
  - asociace s kolekce 210
  - asociování s externími zdroji 241
  - ID aplikací 248
  - popis 10
  - přístup 234
  - přízpusobení 228
  - spuštění 234
  - ukázka 209, 210
  - vlastní 209
  - zabezpečení na úrovni kolekce 248
- vyhledávací centrum, portál WebSphere
  - popis 323
  - verze 6, integrace s podnikovým vyhledáváním 332
- vyhledávací panel, portál WebSphere
  - verze 5.1, přeměrování na podnikové vyhledávání 327

- vyhledávací panel, portál WebSphere (*pokračování*)
  - verze 6, přeměrování na podnikové vyhledávání 334
- vyhledávací pole HTML
  - Dublin Core, prvky 133
  - mapování prvků 132, 133
  - popis 132
  - vytvoření 133
- vyhledávací pole XML
  - mapování prvků 129, 137
  - popis 129, 137
  - vytvoření 129
- vyhledávací servery
  - asociování slovníků slov pro upřednostnění 200
  - asociování slovníků synonym 186
  - asociování slovníků zakázaných slov 188
  - historie doby odezvy 299
  - konfigurace protokolu HTTPS 235
  - konfigurace protokolu SSL 235
  - konfigurace serveru proxy 236
  - mezipaměť vyhledávání 183, 184
  - monitorování 299, 300
  - oblíbené dotazy 299, 300
  - opakovaná implementace slovníků 189
  - popis 7, 183
  - poslední dotazy 299, 300
  - slovníky slov pro upřednostnění 198
  - slovníky synonym 184, 186
  - slovníky zakázaných slov 187
  - spuštění 279, 283, 299
  - stav systému 299
  - výpočet počtu dotazů 300
  - zastavení 281, 283, 299
- vyhledávání
  - dokumenty HTML 132, 133
  - dokumenty XML 129, 137
  - kategorie 124
  - kolekce 169
  - rychlé odkazy 192
- vyhledávání s použitím polí
  - popis 169
  - řazení řetězců 169
- Vyhledávání, popis stránky 13
- výchozí vyhledávací aplikace 227
- výsledky vyhledávání
  - dodatečně filtrování 249
  - dynamické řazení podle důležitosti 195
  - dynamické shrnutí 190
  - expanze zástupných znaků 176
  - hodnocení na základě textu 195
  - konfigurace tříd upřednostnění 202, 204, 205
  - konfigurace vzoru identifikátoru URI 201
  - popis 195
  - předběžné filtrování 249
  - přízpusobení souhrnů 190, 191
  - řazení podle důležitosti 201
  - sbalení 178, 179, 276
  - seskupování 178, 179
  - souhrny 190, 191
  - statické třídy důležitosti 196
  - zástupné znaky 174
- výsledky vyhledávání, popis polí 169
- výstrahy
  - doba odezvy vyhledávání 306

- výstrahy (*pokračování*)
  - index, limity 285
  - indexované dokumenty 306
  - konfigurace serveru SMTP 309
  - na úrovni kolekce 285, 306
  - na úrovni systému 307
  - popis 305
  - prolezené dokumenty 306
  - příjem e-mailů 310
  - volby e-mailu 306, 307
  - volné místo na serverech 307
- výstrahy týkající se volného místa 307
- vytvoření
  - kategorie založené na pravidlech 127
  - kolekce 29, 30
  - obory 178
  - prolézací moduly 36
  - rychlé odkazy 192
  - vyhledávací pole HTML 133
  - vyhledávací pole XML 129
  - webový prolézací modul, sestavy 292

## W

- Web Content Management
  - integrace s produktem WebSphere Portal 323
- webové prolézací moduly
  - aktivní weby 289, 290
  - cookie, konfigurace souboru 98
  - data prolezených dokumentů 102
  - direktivy no-follow 101
  - direktivy no-index 101
  - formát souborů cookie 98
  - frekvence prolézání 291
  - globální prolézaný prostor 99
  - historie prolézacího modulu 289
  - JavaScript, podpora 87
  - konfigurace 83
  - méně závažné chyby, stránky 93
  - monitorování 289
  - naposledy prolézané adresy URL 289
  - návštěva adresy URL, co nejdříve 93
  - omezení prolézaného prostoru 88
  - opakované prolézání, intervaly 92
  - počáteční adresy URL 88, 93
  - podrobnosti o podprocesech 289, 290
  - podrobnosti webu 289
  - prolézání, pravidla 88
  - proxy, servery 97
  - robots.txt, soubory 84, 85
  - soubor followindex.rules 101
  - soubory cookie 97
  - stav adresy URL 289
  - stav systému 289
  - uživatelští agenti 84
  - vytváření sestav 292
  - weby chráněné heslem 95, 96
- WebSphere II Classic Federation 52
- WebSphere II Event Publisher Edition, konfigurace prolézacího modulu DB2 48
- WebSphere MQ, konfigurace prolézacího modulu DB2 50
- WebSphere MQ, konfigurace prolézacího serveru 46, 47



WebSphere Portal  
 integrace s podnikovým  
 vyhledáváním 323  
 kategorie, migrace stromu 343  
 klastrované systémy 337  
 kolekce, migrace 343  
 konfigurační skripty pro podnikové  
 vyhledávání 324  
 taxonomie, migrace 343  
 verze 5.1, integrační skripty 325  
 verze 5.1, konfigurace vyhledávacího  
 panelu 327  
 verze 5.1, odebrání podnikového  
 vyhledávání 329  
 verze 6, integrační skripty 330  
 verze 6, konfigurace centra  
 vyhledávání 332  
 verze 6, konfigurace vyhledávacího  
 panelu 334  
 verze 6, odebrání podnikového  
 vyhledávání 336  
 vyhledávací centrum, popis 323  
 vyhledávací panel, popis 323  
 výchozí nastavení migrace 345  
 weby chráněné heslem 95  
 ověřování založené na použití  
 formulářů 96  
 základní ověřování HTTP 95  
 Windows  
 podpora protokolu IPv6 25  
 wp5\_install, skript 325  
 wp5\_uninstall, skript 329  
 wp6\_cluster\_install, skript 338  
 wp6\_install, skript 330  
 wp6\_uninstall, skript 336  
 WpsMigratorLog.log, soubor 346

## X

XML, nativní syntaxe dotazů 142  
 XML, nativní vyhledávání fragmentů  
 XML 142  
 XPath, nativní vyhledávání XML 142

## Z

zabezpečení  
 administrativní role 247  
 analýza kotvícího textu 257  
 detekce duplicitních dokumentů 247  
 dokumenty Lotus Domino 267, 268  
 domény systému Windows 271  
 globální zabezpečení produktu  
 WebSphere 274  
 globální, WebSphere Application  
 Server 259, 260  
 instalace na jeden server 262  
 instalace na více serverů 262  
 konfigurace protokolu HTTPS pro  
 vyhledávání 235  
 konfigurace protokolu SSL pro  
 vyhledávání 235  
 moduly plug-in prolézacího modulu 110  
 na úrovni dokumentu 249, 250, 251, 257,  
 275  
 na úrovni kolekce 247, 275

zabezpečení (*pokračování*)  
 ověřování 245, 274  
 podpora SSO 255  
 popis 243  
 povolení podnikového vyhledávání 259  
 povolení pro kolekci 30, 243  
 profily uživatelů 254  
 registr uživatelů LDAP 260  
 řízení přístupu 245  
 sbalené výsledky vyhledávání 276  
 správa identit 252, 257  
 vyhledávací aplikace, ID 248  
 vynechání řízení přístupu na úrovni  
 dokumentu 275  
 zakázání pro podnikovou aplikaci 274  
 zabezpečení lokálních uživatelů, prolézací  
 moduly QuickPlace 269  
 zabezpečení na úrovni dokumentu  
 aktuální pověřovací údaje, ověření 251  
 dodatečné filtrování výsledků 249  
 dokumenty Lotus Domino 268  
 indexované řízení přístupu 250  
 konfigurace prolézacího modulu 35  
 moduly plug-in prolézacího modulu 110  
 ověřování v reálném čase 251  
 podpora SSO 255  
 popis 243, 249  
 pro dokumenty Lotus Domino 267  
 pro souborové systémy Windows 271  
 profily uživatelů 254  
 prvky zabezpečení 250  
 předběžné filtrování výsledků 249  
 správa identit 252, 257  
 zabezpečení na úrovni kolekce  
 analýza kotvícího textu 257  
 detekce duplicitních dokumentů 247  
 ID aplikací 248  
 popis 243, 247  
 povolení 30  
 zabezpečení SSO (single sign-on)  
 konfigurace 257  
 správa identit 255  
 Zabezpečení, popis pohledu 13  
 zákaz časových plánů indexování 165  
 základní ověřování HTTP 95  
 zálohování podnikového vyhledávání 315,  
 316  
 zálohování systému 315, 316  
 zastavení  
 moduly analýzy 297  
 prolézací servery 288  
 servery podnikového vyhledávání 279,  
 281  
 sestavení indexu 298  
 vyhledávací servery 283, 299  
 zástupné znaky  
 expanze dotazu 174, 176  
 expanze indexu 174, 175, 176  
 v dotazech 174  
 zjišťování 4  
 zobrazení  
 podrobnosti identifikátoru URI 286  
 soubory žurnálu 312  
 soubory žurnálů odstraněných  
 dokumentů 304  
 zpracování textu  
 anotační moduly 134

zpracování textu (*pokračování*)  
 obecná struktura analýzy 134  
 stroj pro analýzu textu 134

## Ž

Žurnál, popis stránky 13





**IBM**



**Java**<sup>™</sup>  
**COMPATIBLE**

SC09-3744-04



Spine information:

OmniFind Enterprise Edition

Verze 8.5

Správa podnikového vyhledávání

