

IBM WebSphere Information Integrator
OmniFind Edition



Správa podnikového vyhledávání

Verze 8.3

IBM WebSphere Information Integrator
OmniFind Edition



Správa podnikového vyhledávání

Verze 8.3

Před použitím těchto informací a produktu, k jehož podpoře se vztahují, si přečtěte obecné informace v části Poznámky.

Tento dokument obsahuje proprietární informace společnosti IBM. Je k dispozici na základě licenční smlouvy a je chráněn zákonem o autorských právech. Informace uvedené v této publikaci nezahrnují žádné záruky pro produkt a žádné prohlášení uvedené v této příručce nelze takovým způsobem interpretovat.

Publikace společnosti IBM lze objednat online nebo prostřednictvím místního zástupce společnosti IBM:

- Chcete-li objednat publikace online, přejděte na web IBM Publications Center na adrese www.ibm.com/shop/publications/order.
- Místního zástupce společnosti IBM lze vyhledat prostřednictvím stránky IBM Directory of Worldwide Contacts (celosvětový adresář kontaktů společnosti IBM) na adrese www.ibm.com/planetwide.

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, 2005. Všechna práva vyhrazena.

Obsah

O těchto tématech vii

Komu jsou tato témata určena vii

Přehled správy podnikového vyhledávání 1

Přihlášení ke konzole pro správu 4

Změna hesla administrátora podnikového vyhledávání v konfiguraci s jedním serverem 5

Změna hesla administrátora podnikového vyhledávání v konfiguraci s více servery 7

Změna čísla portu pro systém podnikového vyhledávání 10

Kolekce podnikového vyhledávání 13

Vytvoření kolekce pomocí Průvodce kolekcí 13

Vytvoření kolekce s použitím pohledu Kolekce 14

Úprava kolekce 16

Odstranění kolekce 17

Správa prolézacích modulů podnikového vyhledávání. 19

Vytvoření prolézacího modulu 21

Úprava vlastností prolézacího modulu 22

Úprava prolézaného prostoru 23

Odstranění prolézacího modulu 23

Prolézací moduly Content Edition 24

Přístup k úložištím WebSphere II Content Edition v režimu serveru 24

Přístup k úložištím WebSphere II Content Edition v přímém režimu 25

Konfigurování prolézacího serveru v systému UNIX pro produkt WebSphere II Content Edition 25

Konfigurování prolézacího serveru v systému Windows pro produkt WebSphere II Content Edition 26

Prolézací moduly DB2 28

Konfigurování produktu WebSphere Information Integrator Event Publisher Edition pro prolézací moduly DB2 29

Konfigurování produktu WebSphere MQ pro prolézací moduly DB2 32

Konfigurování prolézacího serveru v systému UNIX pro použití publikování událostí 33

Konfigurování prolézacího serveru v systému Windows pro použití publikování událostí 34

Prolézací moduly DB2 Content Manager 35

Konfigurování prolézacího serveru v systému UNIX pro produkt DB2 Content Manager 36

Konfigurování prolézacího serveru v systému Windows pro produkt DB2 Content Manager 37

Prolézací moduly Domino Document Manager 39

Prolézací moduly Exchange Server 40

Ověřování přístupu k zabezpečeným dokumentům serveru Exchange Server 41

Prolézací moduly NNTP 41

Prolézací moduly Notes 42

Rady pro prolézání databází Lotus Domino 44

Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino 45

Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino 46

Konfigurování serverů používajících protokol DIIOP 48

Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino 49

Prolézací moduly QuickPlace 50

Konfigurování serveru QuickPlace pro používání zabezpečení lokálních uživatelů 52

Konfigurování služby Directory Assistance na serveru QuickPlace 53

Prolézací moduly souborového systému UNIX 53

Webové prolézací moduly 54

Konfigurace uživatelského agenta 55

Podpora skriptů JavaScript 56

Pravidla omezující prolézaný webový prostor. 56

Testy spojení adres URL s webovým prolézacím modulem 60

Nastavení intervalu opakovaného prolézání ve webovém prolézacím modulem 60

Volby pro návštěvy adres URL webovým prolézacím modulem 61

Způsob práce webového prolézacího modulu se stránkami méně závažných chyb 61

Podpora prolézání zabezpečených webových serverů 63

Webové servery obsluhované servery proxy 65

Správa souborů cookie 66

Konfigurace globálního prolézaného webového prostoru 67

Direktivy no-follow a no-index 69

Potlačení direktiv no-follow a no-index na webových stránkách 69

Prolézací moduly WebSphere Portal 70

Implementace podnikové aplikace pro prolézací modul WebSphere Portal 71

Kopírování adresy URL pro prolézání ze systému WebSphere Portal 72

Prolézací moduly souborového systému Windows 73

Konfigurování podpory pro aplikace modulu pro příjem dat 74

Vlastní moduly plug-in pro prolézací moduly 75

Formáty identifikátorů URI v indexu podnikového vyhledávání. 76

Správa analyzátoru podnikového vyhledávání 85

Práce s kategoriemi 86

Kategorie založené na pravidlech 87

Kategorie založené na modelu 89

Stromy kategorií 89

Výběr typu kategorizace 90

Konfigurování kategorií 91

Práce s vyhledávacími poli XML 93

Vyhledávací pole XML 93

Mapování prvků XML na vyhledávací pole	93
Práce s vyhledávacími poli HTML	95
vyhledávací pole HTML	95
Mapování prvků metadat HTML na vyhledávací pole	96
Vlastní zpracování textu	96
Přidávání strojů pro analýzu textu do systému	98
Asociování stroje pro analýzu textu s kolekcí	99
Mapování prvků XML na obecnou strukturu analýzy	99
Mapování obecné struktury analýzy na index	101
Mapování obecné struktury analýzy na tabulky JDBC	101
Konfigurování podprocesů pro službu analyzátoru	102
Povolení rozšířené analýzy složených výrazů	103
Povolení podpory nativního vyhledávání XML	104
Lingvistická analýza dokumentů v čínštině, japonštině a korejštině	104
N-gramová segmentace	105
Odebrání znaků nového řádku z rozsahů znaků mimo rozsah sady ASCII	105
Typy dokumentů přiřazené k analyzátorům kolekcí a k relacím Stellent	106
Přiřazení typů dokumentů k analyzátoru kolekce	106
Výchozí pravidla služby analyzátoru kolekce	108
Přiřazení typů dokumentů k relaci Stellent	109
Výchozí pravidla analýzy pro relace Stellent	111

Správa indexu podnikového vyhledávání 113

Plánování sestavení indexu	114
Změna časového plánu indexu	115
Povolení a zákaz časových plánů indexování	115
Konfigurování souběžných sestavení indexu	116
Volby ovlivňující prohlédatelné zobrazení indexu	117
Volby vytváření indexu pro prohlédávání dokumentů	117
Zástupné znaky v dotazech	119
Konfigurování voleb pro použití zástupných znaků v dotazech	121
Obory	122
Konfigurování oborů	123
Sbalené identifikátory URI	123
Sbalení identifikátorů URI ve výsledcích vyhledávání	124
Odebrání identifikátorů URI z indexu	125

Správa vyhledávacího serveru pro podnikové vyhledávání 127

Mezipaměti pro vyhledávání	128
Konfigurování mezipaměti vyhledávání	128
Vlastní slovníky synonym	128
Přidávání slovníků synonym do systému	130
Asociování slovníku synonym s kolekcí	130
Vlastní slovníky zakázaných slov	131
Přidávání slovníků zakázaných slov do systému	131
Asociování slovníku zakázaných slov s kolekcí	132
Dynamické shrnutí	132
Přizpůsobení shrnutí dokumentů v konzole pro správu	133
Přizpůsobení souhrnů dokumentů provedením úprav vlastností	133
Práce s rychlými odkazy	134
Rychlé odkazy	134
Konfigurování rychlých odkazů	135

Hodnocení důležitosti dokumentů v podnikovém vyhledávání 137

Hodnocení na základě textu	137
Statické řazení	138
Vlastní slovníky slov pro upřednostnění	138
Přidávání slovníků slov pro upřednostnění do systému	139
Asociování slovníku slov pro upřednostnění s kolekcí	140
Hodnocení důležitosti dokumentů na základě vzorů identifikátoru URI	140
Ovlivňování skóre dokumentů shodujících se se vzory identifikátorů URI	141
Hodnocení důležitosti dokumentů na základě tříd upřednostnění	142
Mapování polí na třídy upřednostnění	144
Konfigurování faktorů upřednostnění pro třídy upřednostnění	144
Výchozí hodnoty tříd upřednostnění	145

Vyhledávací aplikace podnikového vyhledávání 149

Asociace vyhledávacích aplikací s kolekcemi	150
Funkce ukázkové vyhledávací aplikace	150
Vlastnosti ukázkové vyhledávací aplikace	151
Úpravy vlastností ukázkové vyhledávací aplikace	160
Přístup k ukázkové vyhledávací aplikaci	161
Povolení zabezpečení pro ukázkovou vyhledávací aplikaci	161

Externí zdroje podnikového vyhledávání 165

Přidávání externích zdrojů do systému	165
Asociování vyhledávacích aplikací s externími zdroji	166

Zabezpečení podnikového vyhledávání 169

Administrativní role	170
Konfigurování administrativních uživatelů	171
Ověřování a řízení přístupu	172
Zákaz zabezpečení pro podnikové aplikace na aplikačním serveru WebSphere	173
Zabezpečení na úrovni kolekce	174
Analýza duplicitních dokumentů	174
Analýza kotvicího textu	175
Indexování kotvicího textu v odkazech na zakázané dokumenty	175
Zabezpečení pomocí ID vyhledávacích aplikací	176
Zabezpečení na úrovni dokumentu	177
Ověřování pomocí uložených prvků zabezpečení	177
Ověřování aktuálního pověření během zpracování dotazu	178
Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows	181
Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino	183
Zakázání zabezpečení na úrovni dokumentu	185

Integrace podnikového vyhledávání s produktem WebSphere Portal 187

Zabezpečení na úrovni dokumentů při použití portálového vyhledávacího stroje	189
--	-----

Implementace vyhledávacího portletu	189
Konfigurování portletu pro vyhledávání a procházení produktu WebSphere Portal pro podnikové vyhledávání	191
Instalace adaptéru podnikového vyhledávání pro centrum vyhledávání	192
Instalace registračního portletu podnikového vyhledávání pro centrum vyhledávání	193

Migrace z produktu WebSphere Portal do podnikového vyhledávání. 195

Migrace taxonomie založené na modelu z produktu WebSphere Portal	195
Migrace kolekce z produktu WebSphere Portal	197
Migrované nastavení kolekce	199
Soubor žurnálu průvodce migrací	200

Spouštění a zastavování serverů podnikového vyhledávání 201

Spuštění serverů podnikového vyhledávání	201
Zastavení serverů podnikového vyhledávání.	203

Monitorování aktivity podnikového vyhledávání 205

Odhad počtu dokumentů v kolekci	205
Kontrola dostupnosti systémových prostředků	206
Monitorování kolekce	207
Zobrazení podrobností o identifikátoru URI	207
Monitorování prolézacích modulů.	209
Zobrazení podrobných informací o aktivitě webového prolézacího modulu	210
Podrobnosti o podprocesech webového prolézacího modulu.	211
Aktivní weby webového prolézacího modulu	212
Frekvence prolézání webového prolézacího modulu	212
Vytváření sestav webového prolézacího modulu	213
Návratové kódy HTTP webového prolézacího modulu	214
Monitorování modulu analýzy	217
Monitorování aktivity indexu pro kolekci	218
Monitorování fronty indexů podnikového vyhledávání	219
Monitorování vyhledávacích serverů	220
Monitorování modulu pro příjem dat	221
Sledování dokumentů	222
Konfigurování souborů žurnálu pro sledování dokumentů	222
Zobrazení sestav odstraněných dokumentů	223

Výstrahy a soubory žurnálu podnikového vyhledávání 225

Výstrahy	225
Konfigurování výstrah na úrovni kolekce	226
Konfigurování výstrah na úrovni systému	227
Konfigurování souborů žurnálu	228
Konfigurování informací o serveru SMTP	229
Příjem e-mailů o protokolovaných zprávách.	230
Zobrazení souborů žurnálu	231

Zálohování a obnovování systému podnikového vyhledávání 233

Zálohování systému podnikového vyhledávání	233
Obnova systému podnikového vyhledávání	234
Obnova souborů systému podnikového vyhledávání na nových serverech.	235

Příkazy podnikového vyhledávání, návratové kódy a ID relací 237

Dokumentace k podnikovému vyhledávání 265

WebSphere II OmniFind Edition - usnadnění přístupu 267

Slovníček pojmů z oblasti podnikového vyhledávání 269

Přístup k informacím o produktech WebSphere Information Integration 279

Odesílání poznámek k dokumentaci 281

Možnosti kontaktování společnosti IBM. 283

Ochranné známky 285

Rejstřík 289

O těchto tématech

Tyto informace jsou určeny uživatelům, kteří provádějí administraci systému IBM WebSphere Information Integrator OmniFind Edition verze 8.3.

Produkt WebSphere Information Integrator OmniFind Edition (WebSphere II OmniFind Edition) nabízí technologii nazývanou *podnikové vyhledávání*. Komponenty podnikového vyhledávání se instalují současně s produktem WebSphere II OmniFind Edition. Pojem *podnikové vyhledávání* se používá v celé dokumentaci k produktu WebSphere II OmniFind Edition s výjimkou odkazů na instalační cesty a popisků na obalech produktů.

Tato dokumentace k administraci podnikového vyhledávání se věnuje následujícím tématům:

- Úvod do koncepce, komponent a funkcí podnikového vyhledávání
- Pokyny pro vytváření, monitorování a administraci kolekcí
- Informace o způsobu analýzy dat a o možnostech přizpůsobení analytických operací s cílem optimalizovat vyhledávání a načítání
- Informace o způsobu indexování analyzovaných dat a o možnostech administrace aktivit souvisejících s indexováním
- Informace o způsobu, kterým vyhledávací servery vyhledávají, řadí a vracejí výsledky vyhledávání
- Informace o asociování vlastních vyhledávacích aplikací s kolekcemi
- Přehled různých úrovní zabezpečení, které jsou k dispozici pro podnikové vyhledávání
- Informace o integraci podnikového vyhledávání s produktem IBM WebSphere Portal
- Pokyny k migraci taxonomií a kolekcí produktu WebSphere Portal do stromů kategorií a kolekcí podnikového vyhledávání
- Pokyny pro zálohování a obnovení systému
- Pokyny pro vytváření a prohlížení souborů žurnálu
- Pokyny pro použití příkazů podnikového vyhledávání a interpretaci vrácených informací

Komu jsou tato témata určena

Tyto informace jsou určeny pro administrátory a operátory systémů zodpovědné za vytváření, monitorování a správu kolekcí podnikového vyhledávání.

Informace, které zde najdete, využijete při vytváření kolekcí, výběru obsahu pro kolekce a konfigurování voleb umožňujících prohledávání obsahu. Dále můžete tyto informace využít při monitorování aktivity kolekcí a systému, definování uživatelů jako administrátorů podnikového vyhledávání a asociování kolekcí a externích prohledávatelných zdrojů s vyhledávacími aplikacemi.

Abyste mohli s těmito informacemi pracovat efektivně, musíte být obeznámeni s webovými aplikacemi a mít zkušenosti se zdroji dat, které chcete prohledávat.

Přehled správy 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.

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: Klepnete-li na tlačítko Zpět nebo Aktualizovat ve webovém prohlížeči, mohou ve výsledcích vzniknout nekonzistence a hrozí i ztráta dat. Potřebujete-li při konfigurování kolekcí přejít k předcházející stránce nebo aktualizovat informace na konzole pro správu, klepněte místo tlačítka Zpět nebo Aktualizovat ve webovém prohlížeči na tlačítko **Předchozí** nebo **Aktualizovat** na konzole pro správu podnikového vyhledávání.

- Na stránce Obecné můžete zadat volby, které budou platit pro celou kolekci:
 - Obecné volby upravte tehdy, chcete-li změnit jméno 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 jména 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 zlepšit možnosti načítání informací zadáním voleb zpracování textu.
 - Můžete asociovat pole s třídami upřednostnění a ovlivňovat tak ohodnocení polí, která odpovídají výrazům dotazu, ve výsledcích vyhledávání.
- Na stránce Index nastavte konfiguraci plánů reorganizace a aktualizace 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 všechny dokumenty pocházející ze stejného zdroje budou ve výsledcích vyhledávání tvořit jednu sbalenou skupinu.
 - Přiřadit faktory upřednostnění a ovlivnit ohodnocení důležitosti dokumentů odpovídajících vzoru identifikátoru URI ve výsledcích vyhledávání.
 - Odebrat z indexu některé identifikátory URI. 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 spuštění komponent je třeba dodržovat stanovené pořadí. Data musí být nejprve zpracována prolézacími moduly, teprve poté je mohou analyzátoři analyzovat, po dokončení analýzy může proběhnout indexování a poté musí být index aktualizován nebo reorganizován. Když jsou provedeny všechny tyto kroky, mohou vyhledávací servery začít zpracovávat vyhledávací požadavky.

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

Pokud jste členem role administrátorů podnikového vyhledávání, můžete klepnout na tlačítko **Systém** na panelu nástrojů a poté provadět následující činnosti (administrátoři kolekcí, operátoři a uživatelé monitorů mají k tomuto pohledu přístup pouze v případě, že jim administrátor podnikového vyhledávání udělí příslušné oprávnění):

- Kontrolovat dostupnost systémových prostředků.
- Konfigurovat klientské aplikace modulu pro příjem dat tak, aby mohly aktualizovat kolekce.
- Přidávat do systému stroje pro analýzu textu.
- Přidávat do systému vlastní slovníky synonym, zakázaných slov a slov pro upřednostnění.
- Určete počet kolekcí, pro které smějí být souběžně sestavovány indexy, a určete, zda lze souběžně zpracovávat požadavky na aktualizaci a reorganizaci pro stejnou kolekci.
- 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 IBM 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.

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

“Správa prolézacích modulů podnikového vyhledávání” na stránce 19

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

“Monitorování aktivity podnikového vyhledávání” na stránce 205

Při monitorování aktivit systému a kolekcí můžete zobrazit stav různých procesů, sledovat příznaky potenciálních problémů nebo zvýšit výkon úpravou nastavení konfigurace.

Související úlohy

“Spuštění serverů podnikového vyhledávání” na stránce 201

Chcete-li uživatelům povolit vyhledávání v kolekci, je nutné spustit systémové procesy a poté spustit servery, které kolekci prolézají, analyzují, indexují a prohledávají.

“Zastavení serverů podnikového vyhledávání” na stránce 203

Server podnikového vyhledávání může být nutné zastavit a znovu spustit, pokud provádíte změny v jeho konfiguraci nebo pokud potřebujete řešit problémy.

Vytvoření kolekce pomocí Průvodce kolekci

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.

“Vytvoření kolekce s použitím pohledu Kolekce” na stránce 14

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

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 WebSphere II OmniFind 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 uživatelské jméno 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 170

Podnikové vyhledávání využívá koncepci rolí pro řízení přístupu k různým funkcím konzoly pro správu.

Související úlohy

“Spuštění serverů podnikového vyhledávání” na stránce 201

Chcete-li uživatelům povolit vyhledávání v kolekci, je nutné spustit systémové procesy a poté spustit servery, které kolekci prolézají, analyzují, indexují a prohledávají.

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`.

Než začnete

Jméno a heslo administrátora podnikového vyhledávání musí být platné v daném operačním systému a musí mu být uděleno oprávnění pro přístup k databázi DB2 Universal Database a pro konfigurování této databáze.

O této úloze

Heslo pro počáteční jméno administrátora podnikového vyhledávání se nastavuje při instalaci produktu WebSphere II OmniFind 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.

Vzhledem k tomu, že skript eschangepw je nainstalován v adresáři ES_INSTALL_ROOT/bin, můžete jej spustit z kteréhokoli místa v systému.

Postup

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

1. Přihlašte se jako administrátor podnikového vyhledávání.
2. Zastavte server zadáním následujícího příkazu: `esadmin stop`
3. Otevřete konzolu pro správu aplikačního serveru WebSphere a zastavte podnikové aplikace server1 a ESSearchServer.
4. Změňte systémové heslo odpovídající uživatelskému jménu administrátora podnikového vyhledávání pomocí příkazů operačního systému UNIX nebo pomocí prostředků operačního systému Microsoft Windows pro změnu hesla.
5. Spusíte následující skript, přičemž parametr *nová_hodnota* označuje heslo zadané v kroku 4:

Operační systém	Příkaz
UNIX	<code>eschangepw.sh nová_hodnota</code>
Windows	<code>eschangepw nová_hodnota</code>

6. Z konzoly pro správu aplikačního serveru WebSphere spusíte aplikace server1 a ESSearchServer.
7. Restartujte společnou komunikační vrstvu (CCL) WebSphere II OmniFind Edition zadáním následujících příkazů:

Operační systém	Příkazy
UNIX	<code>stopccl.sh</code> , poté <code>startccl.sh -bg</code>
Příkazový řádek systému Windows	<code>stopccl</code> , poté <code>startccl</code>
Nástroj pro správu Služby systému Windows	<ol style="list-style-type: none"> 1. Spusíte Služby systému Windows. 2. Klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Zastavit. 3. Znovu klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Vlastnosti. 4. Klepněte na kartu Přihlášení. 5. Změňte heslo zadáním hodnoty <i>nová_hodnota</i> a klepněte na tlačítko OK. 6. Znovu klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Spustit.

8. Restartujte podnikové vyhledávání zadáním následujícího příkazu: `esadmin start`

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

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.

Než začnete

Jméno a heslo administrátora podnikového vyhledávání musí být platné v daném operačním systému a musí mu být uděleno oprávnění pro přístup k databázi DB2 Universal Database a pro konfigurování této databáze.

Heslo administrátora podnikového vyhledávání musí být shodné ve všech počítačích zapojených do dané instalace produktu WebSphere II OmniFind Edition.

O této úloze

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

Chcete-li toto heslo změnit a rozšířit informaci o změně do celého systému podnikového vyhledávání, musíte spustit skript `eschangepw` ve všech počítačích zapojených do konfigurace s více servery. 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.

Vzhledem k tomu, že skript `eschangepw` je nainstalován v adresáři `ES_INSTALL_ROOT/bin`, můžete jej spustit z kteréhokoli místa v systému.

Postup

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

1. Přihlašte se k počítači, kde je nainstalován indexový server, jako administrátor podnikového vyhledávání.
2. Spusíte režim údržby zadáním následujícího příkazu: `esadmin system maintenance`
3. Změňte systémové heslo odpovídající uživatelskému jménu administrátora podnikového vyhledávání pomocí příkazů operačního systému UNIX nebo pomocí prostředků operačního systému Microsoft Windows pro změnu hesla.
4. Spusíte následující skript, přičemž parametr `nová_hodnota` označuje heslo zadané v kroku 3:

Operační systém	Příkaz
UNIX	<code>eschangepw.sh nová_hodnota</code>
Windows	<code>eschangepw nová_hodnota</code>

5. Restartujte službu CCL (Common Communication Layer) produktu WebSphere II OmniFind Edition zadáním následujících příkazů na indexovém serveru:

Operační systém	Příkazy
UNIX	stopccl.sh, poté startccl.sh -bg
Příkazový řádek systému Windows	stopccl, poté startccl
Nástroj pro správu Služby systému Windows	<ol style="list-style-type: none"> 1. Spusíte Služby systému Windows. 2. Klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Zastavit. 3. Znovu klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Vlastnosti. 4. Klepněte na kartu Přihlášení. 5. Změňte heslo zadáním hodnoty <i>nová_hodnota</i> a klepněte na tlačítko OK. 6. Znovu klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Spustit.

6. Přihlašte se k počítači, kde je nainstalován prolézací server, jako administrátor podnikového vyhledávání.
7. Změňte systémové heslo odpovídající uživatelskému jménu administrátora podnikového vyhledávání pomocí příkazů operačního systému UNIX nebo pomocí prostředků operačního systému Windows pro změnu hesla. Toto heslo se musí shodovat s heslem, které jste zadali v kroku 3 na stránce 7.
8. Spusíte následující skript, přičemž parametr *nová_hodnota* označuje heslo zadané v kroku 3 na stránce 7:

Operační systém	Příkaz
UNIX	eschangepw.sh <i>nová_hodnota</i>
Windows	eschangepw <i>nová_hodnota</i>

9. Restartujte vrstvu CCL na prolézacím serveru zadáním následujících příkazů:

Operační systém	Příkazy
UNIX	stopccl.sh, poté startccl.sh -bg
Příkazový řádek systému Windows	stopccl, poté startccl
Nástroj pro správu Služby systému Windows	<ol style="list-style-type: none"> 1. Spusíte Služby systému Windows. 2. Klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Zastavit. 3. Znovu klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Vlastnosti. 4. Klepněte na kartu Přihlášení. 5. Změňte heslo zadáním hodnoty <i>nová_hodnota</i> a klepněte na tlačítko OK. 6. Znovu klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Spustit.

10. Přihlašte se jako administrátor podnikového vyhledávání k jednomu z počítačů, na kterých je nainstalován vyhledávací server.
11. Zastavte webový server IBM HTTP. (Tento krok zabraňuje síťové dispečerské službě směřovat na daný server nové dotazy.)
12. Zastavte službu CCL na vyhledávacím serveru zadáním následujících příkazů:

Operační systém	Příkazy
UNIX	stopccl.sh
Příkazový řádek systému Windows	stopccl
Nástroj pro správu Služby systému Windows	<ol style="list-style-type: none"> 1. Spusíte Služby systému Windows. 2. Klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Zastavit.

13. Otevřete konzolu pro správu aplikačního serveru WebSphere a zastavte aplikace server1 a ESSearchServer.
14. Změňte systémové heslo odpovídající uživatelskému jménu administrátora podnikového vyhledávání pomocí příkazů operačního systému UNIX nebo pomocí prostředků operačního systému Windows pro změnu hesla. Toto heslo se musí shodovat s heslem, které jste zadali v kroku 3 na stránce 7.
15. Spusíte následující skript, přičemž parametr *nová_hodnota* označuje heslo zadané v kroku 3 na stránce 7:

Operační systém	Příkaz
UNIX	eschangepw.sh <i>nová_hodnota</i>
Windows	eschangepw <i>nová_hodnota</i>

16. Znovu spusíte službu CCL na vyhledávacím serveru zadáním následujících příkazů:

Operační systém	Příkazy
UNIX	startccl.sh -bg
Příkazový řádek systému Windows	startccl
Nástroj pro správu Služby systému Windows	<ol style="list-style-type: none"> 1. Spusíte Služby systému Windows. 2. Klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Vlastnosti. 3. Klepněte na kartu Přihlášení. 4. Změňte heslo zadáním hodnoty <i>nová_hodnota</i> a klepněte na tlačítko OK. 5. Znovu klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Spustit.

17. Na druhém vyhledávacím serveru opakujte kroky 10 až 16.
18. Spusíte webový server IBM HTTP.
19. Z konzoly pro správu aplikačního serveru WebSphere spusíte podnikové aplikace server1 a ESSearchServer.
20. Přihlašte se k počítači, kde je nainstalován indexový server, jako administrátor podnikového vyhledávání, a spusíte podnikové vyhledávání zadáním tohoto příkazu:
esadmin start

21. Přihlaste se ke konzole pro správu podnikového vyhledávání, spusťte monitorování kolekce, klepněte na stránku Vyhledávání a poté restartujte vyhledávací servery postupným klepnutím na tlačítka **Zastavit** a **Spustit**. Tento krok opakujte pro všechny kolekce v systému podnikového vyhledávání.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

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 WebSphere II OmniFind 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 Interní varování java.net.BindException: Adresa je již používána: NET_Bind
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. Zadáním následujícího příkazu zastavte systém podnikového vyhledávání:
esadmin stop
2. Přejděte k serveru, pro který je třeba změnit číslo portu, a zastavte službu CCL (Common Communication Layer) zadáním následujících příkazů:

Operační systém	Příkazy
UNIX	stopccl.sh,
Příkazový řádek systému Windows	stopccl
Nástroj pro správu Služby systému Windows	<ol style="list-style-type: none"> 1. Spusťte Služby systému Windows. 2. Klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Zastavit.

3. Upravte soubor \$ES_NODE_ROOT/nodeinfo/es.cfg (v systému UNIX) nebo %ES_NODE_ROOT%\nodeinfo\es.cfg (v systému Windows), vyhledejte následující vlastnost, zadejte novou hodnotu pro číslo portu a poté soubor uložte a zavřete jej.
CCLPort=nové_číslo_portu

4. Znovu spusíte službu CCL zadáním následujících příkazů:

Operační systém	Příkazy
UNIX	startccl.sh -bg
Příkazový řádek systému Windows	startccl
Nástroj pro správu Služby systému Windows	1. Spusíte Služby systému Windows. 2. Klepněte pravým tlačítkem myši na položku WebSphere Information Integrator OmniFind Edition a vyberte volbu Spustit .

5. Přejděte k indexovému serveru a podle pokynů uvedených v kroku 2 na stránce 10 zastavte službu CCL.
6. Upravte soubor `$ES_NODE_ROOT/nodeinfo/es.cfg` (v systému UNIX) nebo `%ES_NODE_ROOT%\nodeinfo\es.cfg` (v systému Windows).
 - a. Vyhleďte následující vlastnost, přičemž parametr *jméno_počítače* určuje jméno serveru, pro který jste upravili číslo portu v kroku 3 na stránce 10. Číslo *N* ve vlastnosti *nodeN* je číslo identifikující příslušný server.
`nodeN.destination=jméno_počítače`
 - b. Vyhleďte následující dílčí vlastnost, zadejte stejné číslo portu, jako jste zadali pro server v kroku 3 na stránce 10, a poté soubor uložte a zavřete:
`nodeN.port=nové_číslo_portu`
7. Podle pokynů v kroku 4 restartujte službu CCL.
8. Zadáním následujícího příkazu restartujte systém podnikového vyhledávání:
`esadmin start`

Po dokončení zpracování tohoto příkazu bude nové číslo portu aktualizováno na všech serverech podnikového vyhledávání.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

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 207

Můžete zobrazit obecné informace o stavu jednotlivých komponent v kolekci nebo vybrat volby pro zobrazení podrobných informací o jednotlivých komponentách a identifikátorech URI.

“Migrace kolekce z produktu WebSphere Portal” na stránce 197

Chcete-li migrovat kolekce z produktu WebSphere Portal do prostředí podnikového vyhledávání, nejprve je připravte v produktu WebSphere Portal a poté proveďte jejich migraci pomocí průvodce migrací.

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.

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 vytvořte 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 jména 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** vytvořte 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 vytvořte 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:
 - **Jméno kolekce.** Zadejte popisné jméno 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 jménu 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.
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 jména 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 zpřístupnit kolekci uživatelům pro vyhledávání prostřednictvím vlastních vyhledávacích aplikací.

Související pojmy

“Přehled správy podnikového vyhledávání” na stránce 1

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

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, reorganizaci a aktualizaci indexu, vyhledávání indexovaného obsahu a protokolování chybových zpráv. V rámci úprav kolekce po jejím vytvoření 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:

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.

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. Můžete také asociovat pole s třídami upřednostnění a ovlivňovat tak ohodnocení důležitosti dokumentů.

Index Určete časové plány pro reorganizaci celého indexu a aktualizaci indexu novým a změněným obsahem. 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.

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 jméno 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**.

Správa prolézacích modulů podnikového vyhledávání

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

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 prolézáného 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.

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 prohlédávání 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 zakódování pravidel ve třídě Java. Při konfigurování vlastností prolézacího modulu můžete určit jméno modulu plug-in, který má prolézací modul při prolézání dokumentů používat. Prvky zabezpečení přidávané 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 prohlédávány. 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.

Plánování činnosti prolézacích modulů

Prolézací moduly, které vytvoříte pro webové zdroje, zdroje NNTP a zdroje WebSphere Portal, pracují nepřetržitě. Spustíte-li takové prohlížeče, obvykle je nemusíte zastavovat, pokud nepotřebujete změnit jejich 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 prolézáném 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 prolézáném prostoru, prolézání zahrnující všechny aktualizace prolézáného prostoru (nové dokumenty, upravené dokumenty a odstraněné dokumenty) nebo prolézání zahrnující pouze nové a upravené dokumenty. Úplné prolézání trvá nejdéle. Prolézání s odebráním odstraněných dokumentů trvá déle než prolézání, při kterém jsou odstraněné dokumenty ignorovány.

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

Vytvoříte-li více plánů prolézání, získáte přesnější kontrolu nad tím, kdy prolézací modul navštívuje cílové zdroje. Při prolézání databází umístěných v různých časových pásmech

můžete například naplánovat spuštění prolézacího modulu na čas po pracovní době většiny uživatelů.

Související pojmy

“Přehled správy podnikového vyhledávání” na stránce 1

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.

“Zabezpečení na úrovni dokumentu” na stránce 177

Pokud je při vytvoření kolekce povoleno její zabezpečení, lze konfigurovat ovládací prvky zabezpečení na úrovni dokumentů. Zabezpečení na úrovni dokumentů zaručuje, že uživatelé prohledávající kolekce budou mít přístup pouze k dokumentům, k jejichž prohlížení mají oprávnění.

Související úlohy

“Monitorování prolézacích modulů” na stránce 209

Můžete zobrazit obecné informace o stavu jednotlivých prolézacích modulů v kolekci nebo vybrat volby pro zobrazení podrobných informací o aktivitě prolézacího modulu.

Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 76

Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

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:

Použit výchozí hodnoty systému pro nový prolézací modul

Naplní počáteční nastavení prolézacího modulu výchozími instalačními hodnotami.

Po výběru této volby zahajte 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** zahajte konfiguraci nového prolézacího modulu.

Otevře se průvodce pro typ vytvářeného prolézacího modulu. Podle pokynů v průvodci vytvořte 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. Vlastnosti, 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 prolezené tímto modulem však zůstávají v indexu, dokud neprovedete jeho reorganizaci.

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 jméno prolézacího modulu již není zobrazeno 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**.

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 lze použít k prolézání úložišť Documentum, FileNet Panagon Content Services, FileNet P8 Content Manager, Hummingbird Document Management (DM), OpenText Livelink a 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.

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 a Portal Document Manager 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 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 používáný prolézacím modulem pro přístup k datům instalován jako podniková aplikace na serveru WebSphere Application Server a prolézací modul přistupuje 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 WebSphere Information Integrator OmniFind Edition a umožňuje prolézacímu modulu Content Edition přistupovat k úložištím na serveru.

Před konfigurováním prolézacího modulu Content Edition pro používání režimu serveru prostřednictvím konzoly pro správu podnikového vyhledávání proveďte úlohu, která odpovídá 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 26.

Přístup k úložištím WebSphere II Content Edition v přímém režimu

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

V přímém režimu prolézací modul používá konektor WebSphere Information Integrator Content Edition instalovaný na prolézací server při instalaci produktu WebSphere II OmniFind Edition.

Postup

Chcete-li konfigurovat systém tak, aby bylo možné přistupovat k úložištím v přímém režimu, postupujte takto:

1. Potvrďte, že proměnné prostředí VBR_HOME a JAVA_HOME v souboru *kořenový_adresář_instalace_produkту_ice/bin/config.sh* (v systému UNIX) nebo *kořenový_adresář_instalace_produkту_ice/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_ice/bin/Admin.bat* (v systému UNIX) nebo *kořenový_adresář_instalace_produkту_ice/bin/Admin.bat* (v systému Windows) systémovou vlastnost Java **vbr.as.operationMode=direct**.
3. Spusťte konzolu pro správu produktu WebSphere Information Integrator Content Edition v přímém režimu a konfiguruje konektor pro prolézací server WebSphere II OmniFind Edition. (Pokyny naleznete v dokumentaci k produktu WebSphere Information Integrator Content Edition.)
4. Vyberte volbu přímého režimu, používáte-li konzolu pro správu produktu WebSphere II OmniFind Edition pro konfigurování prolézacího modulu Content Edition.

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

Pokud jste instalovali produkt WebSphere II OmniFind 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 knihovny Java 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 je schopen pracovat s knihovnami Java, musíte spustit po instalaci produktu WebSphere Application Server spustit instalační skript, který je k dispozici na prolézacím serveru produktu WebSphere II OmniFind Edition.

Produkt WebSphere Information Integrator Content Edition je instalován na prolézací server při instalaci produktu WebSphere II OmniFind 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 WebSphere II OmniFind Edition instalován v konfiguraci s více servery, instalujte a svažte knihovny Java produktu WebSphere Application Server.
2. Spusťte skript pro nastavení prolézacího modulu Content Edition na prolézacím serveru:
 - a. Přihlašte se jako administrátor podnikového vyhledávání.
 - b. Spusťte následující skript instalovaný v adresáři `$ES_INSTALL_ROOT/bin` a odpovězte na výzvy, které se zobrazí:
`escrvbr.sh`
3. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):

```
esadmin stop
stopccl.sh
startccl.sh -bg
esadmin start
```

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 WebSphere II OmniFind Edition do počítače se systémem Microsoft Windows a konfiguruje prolézací modul Content Edition tak, aby při přístupu k

úložištěm používal režim serveru, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožňuje prolézacímu modulu Content Edition přistupovat k úložištěm WebSphere Information Integrator Content Edition.

O této úloze

Prolézací modul Content Edition používá knihovny Java 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 je schopen pracovat s knihovnami Java, musíte spustit po instalaci produktu WebSphere Application Server spustit instalační skript, který je k dispozici na prolézacím serveru produktu WebSphere II OmniFind Edition.

Produkt WebSphere Information Integrator Content Edition je instalován na prolézací server při instalaci produktu WebSphere II OmniFind 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 WebSphere II OmniFind Edition instalován v konfiguraci s více servery, instalujte a svažte knihovny Java produktu WebSphere Application Server.
2. Spusťte skript pro nastavení prolézacího modulu Content Edition na prolézacím serveru:
 - a. Přihlašte se pod jménem administrátora podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind Edition).
 - b. Spusťte následující skript instalovaný v adresáři `%ES_INSTALL_ROOT%\bin` a odpovězte na výzvy, které se zobrazí:
`escrvbr.vbs`
3. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):
 - a. Zadáním následujícího příkazu na příkazový řádek zastavte systém podnikového vyhledávání:
`esadmin stop`
 - b. Vyberte položky **Start** → **Programy** → **Nástroje pro správu** → **Služby** a restartujte službu IBM WebSphere Information Integrator OmniFind Edition.
 - c. Zadáním následujícího příkazu na příkazový řádek spusťte systém podnikového vyhledávání:
`esadmin start`
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 Universal Database. Prolézací modul DB2 lze použít také k zahrnutí přezdívkových tabulek, které vytvoříte pro databáze IBM DB2 Universal Database for z/OS, IBM Informix, Oracle a Microsoft SQL Server.

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 by měly být databázové tabulky, přezdívkové tabulky nebo pohledy. Prolézací modul DB2 nepodporuje spojené tabulky.

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 32.
- “Konfigurování produktu WebSphere Information Integrator Event Publisher Edition pro prolézací moduly DB2” na stránce 29.
- “Konfigurování prolézacího serveru v systému UNIX pro použití publikování událostí” na stránce 33.
- “Konfigurování prolézacího serveru v systému Windows pro použití publikování událostí” na stránce 34.

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.
Jestliže máte v úmyslu prolézat vzdálené databáze, které nejsou katalogizovány na místním databázovém serveru, musíte před použitím prolézacího modulu DB2 k prolézání těchto databází spustit ve vzdáleném serveru server DB2 Administration Server. Při konfigurování prolézacího modulu musíte navíc zadat hostitelské jméno a port vzdáleného databázového serveru.
- Vybrat databázi, kterou 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 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 asociována 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. Spusíte příruční panel replikačního centra:

Operační systém	Příkaz
UNIX	db2rc
Windows	Vyberte položku 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 jméno schématu a klepněte na tlačítko **Další**.
 - d. Zadejte jména správce fronty, fronty administrace a fronty restartu zadaná 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 tlačítko u pole **Mapa fronty publikování** a klepnutím na volbu **Nové** vytvořte mapu fronty publikování.
 - d. Na stránce Obecné zadejte jméno pro mapu fronty.

- e. Na stránce Vlastnosti zadejte jméno fronty odesílání (například jméno datové fronty, které jste zadali při konfigurování produktu WebSphere MQ na tomto serveru), jako typ obsahu zprávy vyberte volbu **Rádková 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 jméno mapy fronty správné, 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 spouštění publikování XML zaškrtnuto, a klepněte na tlačítko **Další**.
 - m. Na stránce Zkontrolovat 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. Spusíte 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 jméno systému, jméno uživatele a heslo pro uživatele produktu DB2, cestu k adresáři, do kterého jsou ukládány žurnály, a jméno 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=jméno_serveru_pro_zachycení LOGSTDOUT=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 UDB, 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 Universal Database (DB2 UDB), uživatelé produktu WebSphere MQ a uživatelé produktu WebSphere II OmniFind 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.

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 jména správce front, jména fronty, jména hostitele serveru, čísla portu serveru a jména 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. Přihlašte 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 opětovném 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:


```
runmqclsr -m QM1 -t TCP -p 1414 &
```
 3. Zadáním následujících příkazů udělte uživateli produktu DB2 UDB 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.DATAQ -p db2inst1 +allmqi
setmqaut -m QM1 -t queue -n ASN.QM1.ADMINQ -p db2inst1 +allmqi
setmqaut -m QM1 -t queue -n ASN.QM1.RESTARTQ -p db2inst1 +allmqi
```
 4. Zadejte následující příkazy pro jméno uživatele použité k vytvoření a spouštění 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.DATAQ -p esuser +allmqi
```

Konfigurování prolézacího serveru v systému UNIX pro použití publikování událostí

Pokud instalujete produkt WebSphere II OmniFind Edition do počítače se systémem IBM AIX či Linux nebo s provozním prostředím Solaris a konfiguruje prolézací modul DB2 tak, aby používal službu publikování událostí, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožňuje prolézacímu modulu přistupovat k frontám a správcům front produktu WebSphere MQ.

O této úloze

Prolézací modul DB2 používá moduly produktu WebSphere MQ 5.3 pro službu Java Messaging pro přístup k frontám a správcům front produktu WebSphere MQ. Tyto moduly musí být instalovány na prolézacím serveru.

Chcete-li ověřit, že prolézací modul DB2 je schopen používat službu publikování událostí, musíte po instalaci modulů produktu WebSphere MQ rovněž spustit instalační skript, který je k dispozici na prolézacím modulu WebSphere II OmniFind Edition.

Postup

Chcete-li konfigurovat prolézací modul pro používání služby publikování událostí, postupujte takto:

1. Instalujte na prolézacím serveru 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. Spusťte skript pro nastavení prolézacího modulu DB2 na prolézacím serveru:
 - a. Přihlaste se jako administrátor podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind Edition).
 - b. Spusťte následující skript instalovaný v adresáři `$ES_INSTALL_ROOT/bin` a odpovězte na výzvy, které se zobrazí:
`escrdb2.sh`
3. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):

```
esadmin stop
stopccl.sh
startccl.sh -bg
esadmin start
```

Konfigurování prolézacího serveru v systému Windows pro použití publikování událostí

Pokud instalujete produkt WebSphere II OmniFind Edition do počítače se systémem Microsoft Windows a konfiguruje prolézací modul DB2 tak, aby používal službu publikování událostí, musíte spustit skript, který nastaví konfiguraci prolézacího serveru. Tento skript umožňuje prolézacímu modulu přistupovat k frontám a správcům front produktu WebSphere MQ.

O této úloze

Prolézací modul DB2 používá moduly produktu WebSphere MQ 5.3 pro službu Java Messaging pro přístup k frontám a správcům front produktu WebSphere MQ. Tyto moduly musí být instalovány na prolézacím serveru.

Chcete-li ověřit, že prolézací modul DB2 je schopen používat službu publikování událostí, musíte po instalaci modulů produktu WebSphere MQ rovněž spustit instalační skript, který je k dispozici na prolézacím modulu WebSphere II OmniFind Edition.

Postup

Chcete-li konfigurovat prolézací modul pro používání služby publikování událostí, postupujte takto:

1. Instalujte na prolézacím serveru 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. Spusťte skript pro nastavení prolézacího modulu DB2 na prolézacím serveru:
 - a. Přihlašte se pod jménem administrátora podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind Edition).
 - b. Spusťte následující skript instalovaný v adresáři %ES_INSTALL_ROOT%\bin a odpovězte na výzvy, které se zobrazí:
escrdb2.vbs
3. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):
 - a. Zadáním následujícího příkazu na příkazový řádek zastavte systém podnikového vyhledávání:
esadmin stop
 - b. Vyberte položky **Start** → **Programy** → **Nástroje pro správu** → **Služby** a restartujte službu IBM WebSphere Information Integrator OmniFind Edition.
 - c. Zadáním následujícího příkazu na příkazový řádek spusťte systém podnikového vyhledávání:
esadmin start

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 WebSphere Information Integrator OmniFind 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 produkt DB2 Content Manager” na stránce 36.
- “Konfigurování prolézacího serveru v systému Windows pro produkt DB2 Content Manager” na stránce 37.

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 prolézáném 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 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 produkt DB2 Content Manager

Pokud jste instalovali produkt WebSphere II OmniFind 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 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 produktu IBM DB2 Information Integrator for Content verze 8.2 nebo novější na prolézací server. 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 WebSphere II OmniFind 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 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 jméno 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 jméno_uzlu remote jméno_hostitele server port
db2 catalog database jméno_databáze as alias at node
jméno_uzlu
```

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

```
. instalační_adresář_Information_Integrator_for_Content/bin/cmbenv81.sh
cd instalační_adresář_Information_Integrator_for_Content/samples/java/icm
javac *.java
java SConnectDisconnect jméno_databáze_ICM ID_administrátora_CM
heslo_administrátora_CM
```

- 5. Spusíte skript pro nastavení prolézacího modulu DB2 Content Manager na prolézacím serveru:
 - a. Změňte adresář ES_INSTALL_ROOT/bin:


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


```
escrcm.sh
```
- 6. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):

```
esadmin stop
stopccl.sh
startccl.sh -bg
esadmin start
```

Konfigurování prolézacího serveru v systému Windows pro produkt DB2 Content Manager

Pokud jste instalovali produkt WebSphere II OmniFind 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 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 produktu IBM DB2 Information Integrator for Content verze 8.2 nebo novější na

prolézací server. 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 WebSphere II OmniFind 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 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**.
 - c. Zadejte jméno 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 jméno_uzlu remote jméno_hostitele server port  
db2 catalog database jméno_databáze as alias at node  
jméno_uzlu
```

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

```
cmbenv81.bat  
cd instalační_adresář_Information_Integrator_for_Content\samples\java\icm  
javac *.java  
java SConnectDisconnect jméno_databáze_ICM ID_administrátora_CM  
heslo_administrátora_CM
```

4. Spusíte skript pro nastavení prolézacího modulu DB2 Content Manager na prolézacím serveru:
 - a. Přejděte do adresáře ES_INSTALL_ROOT\bin:
cd %ES_INSTALL_ROOT%\bin
 - b. Spusíte následující skript a odpovězte na výzvy:
escrcm.vbs
5. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):
 - a. Zadáním následujícího příkazu na příkazový řádek zastavte systém podnikového vyhledávání:
esadmin stop
 - b. Vyberte položky **Start** → **Programy** → **Nástroje pro správu** → **Služby** a restartujte službu IBM WebSphere Information Integrator OmniFind Edition.

- c. Zadáním následujícího příkazu na příkazový řádek spustíte systém podnikového vyhledávání:
esadmin start

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 speciální skript. Tento skript se dodává s produktem WebSphere Information Integrator OmniFind 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.

Je-li produkt WebSphere II OmniFind 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 45.
- “Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 46.
- “Konfigurování serverů používajících protokol DIIOP” na stránce 48.
- “Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 49.

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.
- 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**.

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 vyberete 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 všechny podsložky na všech serverech v prolézaném 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 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**.

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. Spusťte 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.

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 na jednotlivých serverech, které chcete prolézat. Můžete zadat také vzory diskusních skupin, které chcete vyloučit. 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.
- Zadat vzory pro zahrnutí diskusních skupin a vzory, které z prolézaného prostoru některé diskusní skupiny vyloučí.
- 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.

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 speciální skript. Tento skript se dodává s produktem WebSphere Information Integrator OmniFind Edition a umožňuje prolézacímu modulu Notes komunikovat se servery využívajícími protokol NRPC.

Pokud server Lotus Notes, který máte v úmyslu prolézat, používá protokol DIOP (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.

Je-li produkt WebSphere II OmniFind 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 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 45.
- “Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 46.
- “Konfigurování serverů používajících protokol DIOP” na stránce 48.
- “Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 49.

Zabezpečení na úrovni dokumentu

Pokud je povoleno zabezpečení kolekce a server Lotus Notes, 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” na stránce 183.

- Povolení globálního zabezpečení na serveru WebSphere Application Server a konfigurování vyhledávací aplikace pro použití zabezpečení. Tento krok zaručuje, že se při pokusu o použití vyhledávací aplikace uživatelům zobrazí výzva k zadání údajů pověření. Vyhledávací servery poté mohou tyto údaje pověření použít pro ověření oprávnění každého uživatele pro přístup k dokumentům Lotus Notes.

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.
- Vybrat databáze nebo adresáře, které chcete prolézat.
- 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

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino”
na stránce 183

Pokud server Lotus Notes, 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.

Související úlohy

“Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů” na stránce 183

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.

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 registrována jména 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 jména vyhledávacích polí:

Tabulka 1. Výchozí pravidla pro mapování polí

Jméno pole databáze Notes	Jméno 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 **Prolézat Vše**, 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.

Konfigurování prolézacího serveru v systému UNIX pro prolézání zdrojů Lotus Domino

Pokud instalujete produkt WebSphere II OmniFind 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 verze 6.0.2 nebo vyšší 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 WebSphere II OmniFind 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 `notes` a skupinu `notes`:
 - a. Přihlašte se jako uživatel `root`:
`su - root`
 - b. Přidejte uživatele:
`useradd notes`
 - c. Přidejte heslo pro tohoto uživatele:
`passwd notes`
Budete vyzváni ke změně hesla.
2. Instalujte na prolézací server produkt Lotus Domino Server:
 - a. Vložte disk CD serveru Domino verze 6.0.2 nebo vyšší 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.
3. Spusťte instalační skript dodávaný v rámci produktu WebSphere II OmniFind Edition na prolézacím serveru:

- a. Přihlašte se jako administrátor podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind Edition).
- b. Spusťte 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 systém 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. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):

```
esadmin stop
stopccl.sh
startccl.sh -bg
esadmin start
```

Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino

Pokud instalujete produkt WebSphere II OmniFind 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 Notes. Tyto knihovny nainstalujete instalací produktu Lotus Notes verze 6.0.2 nebo vyšší na prolézací server. Chcete-li se ujistit, že prolézací moduly budou moci pracovat s knihovnamí klienta Lotus Notes, spusťte po instalaci knihoven klienta Lotus Notes instalační skript, který je k dispozici na prolézacím serveru produktu WebSphere II OmniFind 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 Lotus Notes verze 6.0.2 nebo vyšší. (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 Notes.
3. Spusťte instalační skript dodávaný v rámci produktu WebSphere II OmniFind Edition na prolézacím serveru:
 - a. Přihlašte se pod jménem administrátora podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind 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. Zastavte a restartujte systém podnikového vyhledávání včetně všech relací CCL (Common Communication Layer):
 - a. Zadáním následujícího příkazu na příkazový řádek zastavte systém podnikového vyhledávání:
esadmin stop
 - b. Vyberte položky **Start** → **Programy** → **Nástroje pro správu** → **Služby** a restartujte službu IBM WebSphere Information Integrator OmniFind Edition.
 - c. Zadáním následujícího příkazu na příkazový řádek spusťte systém podnikového vyhledávání:
esadmin start

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. Konfigurujte 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.

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 \$SES_NODE_ROOT/logs/system_rrrrmmdd.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).

Prolézací moduly QuickPlace

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

Konfigurace prolézacího serveru

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

Pokud server QuickPlace, 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 QuickPlace tak, aby k němu měl prolézací modul QuickPlace přístup.

Pokud server QuickPlace, který chcete prolézat, používá server LDAP (Lightweight Directory Access Protocol), je třeba konfigurovat server QuickPlace 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 server QuickPlace, aby používal server LDAP jako sekundární server Domino.

Je-li produkt WebSphere II OmniFind 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 45.
- “Konfigurování prolézacího serveru v systému Windows pro prolézání zdrojů Lotus Domino” na stránce 46.
- “Konfigurování serverů používajících protokol DIIOP” na stránce 48.
- “Konfigurování serveru QuickPlace pro používání zabezpečení lokálních uživatelů” na stránce 52.
- “Konfigurování služby Directory Assistance na serveru QuickPlace” na stránce 53.
- “Konfigurování modulu I/O Completion Port v systému AIX pro prolézání zdrojů Lotus Domino” na stránce 49.

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 všechny místnosti v místech vybraných pro prolézání nebo zda chcete prolézat konkrétní místnosti.

Chcete-li vytvořit nebo změnit prolézací modul QuickPlace, 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 QuickPlace a komunikační protokol.
- 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 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**.

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/jméno_místa/QP/jméno_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 adreař 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 DIIOP a HTTP.

Postup

Chcete-li konfigurovat službu LDAP Directory Assistance na serveru QuickPlace, postupujte takto:

1. 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** (Jméno databáze služby Directory Assistance) zadejte jméno databáze vytvořené v kroku 1.
 - c. Uložte dokument Server a zavřete jej.

Server QuickPlace nyní může používat server LDAP jako sekundární adresář Domino.

Prolézací moduly souborového systému UNIX

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

Prolézací modul Souborový systém UNIX lze použít k prolézání libovolného počtu souborových systémů UNIX. 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 UNIX (v seznamu dostupných typů prolézacích modulů se neobjeví prolézací modul Souborový systém UNIX).

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 prolézáném 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 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**.

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.

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 prolézáném 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 (prolézání 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

Webový prolézací modul podporuje protokol Robots Exclusion. Chcete-li prolézat webový server využívající tento protokol, 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 jej však podporuje:

- 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 vlastností uživatelského agenta zadáte. Některé aplikační servery, moduly JSP a servlety však podle názvu uživatelského agenta upravují své odpovědi. Tyto různé varianty odpovědí 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.

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, nebude pro webový prolézací modul viditelný a proto jej nebude možné indexovat.

Pravidla omezující prolézání 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 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

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
```

Je-li prvním znakem názvu hostitele hvězdička, bude pravidlo použito pro všechny názvy hostitelů, které končí vzorem zadaným za hvězdičkou. Můžete například stanovit, že nemají být prolézány žádné domény odpovídající následujícímu vzoru:

```
forbid domain *.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 přípojně body 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 (pouze protokol IPv4) a masky podsítě jako cíle. Příklad:

```
allow address 9.0.0.0 255.0.0.0
```

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 pravidlo forbid address 0.0.0.0 0.0.0.0. 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).

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 požadavy 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 podsítí adres IP, omezení způsobí, že téměř všechny identifikátory URL budou klasifikovány s použitím návratového 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.

Seznam pravidel pro adresy obvykle končí následujícím pravidlem, které prolézacímu modulu brání v přístupu k webovým serverům mimo podnikovou síť.

```
forbid 0.0.0.0 0.0.0.0
```

- 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.

Sekce 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 jím 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.

Testy spojení adres URL s webovým prolézacím modulem

Po zadání adres URL, které má prolézací modul prolézat, můžete otestovat konfiguraci pravidel prolézání.

Výsledky testu ukazují, zda má prolézací modul přístup k adresám URL s použitím jména 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 návratový 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 prolézáném 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 nejdříve 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 prolézaná data budou k dispozici pro vyhledávání při příští reorganizaci nebo aktualizaci 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 s návratovým kódem. Návratový 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 návratový 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 irrelevantní 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 návratový 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 (návratový 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 návratový kód 404 (Stránka nebyla nalezena).

Tabulka 2. Příklad stránky méně závažné chyby

Vzor adresy URL	Vzor názvu	Vzor obsahu	Návratový 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 návratové kódy. Pro každý návratový 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:

Položka

*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 návratový 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 návratovým 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 nutnosti zadávat pověřovací údaje s každým požadavkem.

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 prolézáného 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 konfiguruje 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 ověřovací 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 ověřovací údaje s každým požadavkem.

Vzhledem k tomu, že webový prolézací modul nepracuje interaktivně, je nutné zadat ověřovací údaje potřebné pro přístup ke stránkám chráněným heslem ještě předtím, než prolézací modul začne s prolézáním. Při vytváření webových prolézacích modulů a úpravách prolézáného 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 uvedené v prvním vzoru adresy URL v seznamu, který odpovídá kandidátské adrese URL.

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. Před přidáním serveru proxy do prolézaného prostoru si obstarajte následující informace:

Domény serveru proxy

Domény obsluhované daným serverem proxy. Můžete používat hvězdičku (*) jako zástupný znak. Hodnota * například odpovídá všem doménám obsluhovaným tímto serverem proxy a hodnota *.resource.com odpovídá všem doménám, jejichž jméno končí řetězcem resource.com.

Omezení: Pro prolézání serveru proxy nelze určit pravidla pro adresy IP, protože adresa IP serveru proxy je jedinou adresou IP, kterou má prolézací modul k dispozici pro jiného hostitele. Pokud je prolézací modul omezen pomocí pravidel pro adresy IP na podsítí adres IP, omezení způsobí, že téměř všechny identifikátory URL budou klasifikovány s použitím návratového kódu 760 (tento kód určuje, že jsou zakázány webovým prostorem).

Jméno hostitele nebo adresa IP serveru proxy

Jméno DNS hostitele nebo adresu IP v tečkovém formátu pro server proxy.

Číslo portu serveru proxy

Číslo portu TCP/IP, na kterém server proxy přijímá požadavky proxy HTTP.

Po přidání je třeba server proxy vybrat a umístit tak, aby prolézací modul zpracovával servery proxy v požadovaném pořadí. Na začátek seznamu vložte konkrétnější jména domén a na konec seznamu obecnější jména 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=SLNSIDFNLSIDNFLSNL;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 WebSphere II OmniFind 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/col_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 návratových 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 prolézáný 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. Prolézáný 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 jména domén. Toto omezení umožňuje rozdělit prolézáný 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 prolézáný prostor vliv - takové adresy URL jsou i nadále vylučovány.

Globální prolézáný 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 návratový 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 prolézáný prostor, je této adrese URL přiřazen návratový 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 návratový 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 návratový 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 prolézaných 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 zastavíte webový prolézací modul.
2. Přihlašte se k prolézacímu serveru jako administrátor podnikového vyhledávání (toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind 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`
6. Uložte soubor a zavřete jej.
7. Z konzoly pro správu podnikového vyhledávání znovu spusíte webový prolézací modul, který jste zastavili.

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 v produktu WebSphere Portal implementovat podnikovou aplikaci ESPACServer.ear. Tato podniková aplikace je instalována na vyhledávacích serverech při instalaci produktu WebSphere Information Integrator OmniFind Edition. Chcete-li tuto podnikovou aplikaci implementovat, proveďte následující úlohu:

- “Implementace podnikové aplikace pro prolézací modul WebSphere Portal” na stránce 71.

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 poté stáhne portlety, které jsou na určeném portálu k dispozici.

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.
- Zadání adresy 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.

- 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**.

Implementace podnikové aplikace pro prolézací modul WebSphere Portal

Před vytvořením prolézacího modulu WebSphere Portal musíte v produktu WebSphere Portal implementovat podnikovou aplikaci ESPACServer.ear.

O této úloze

Soubor ESPACServer.ear je instalován v adresáři ES_INSTALL_ROOT/bin na vyhledávacích serverech při instalaci produktu WebSphere II OmniFind Edition. Výchozí cesty instalace jsou následující:

Systémy UNIX:

/opt/IBM/es/bin/ESPACServer.ear

Systémy Windows:

C:\Program Files\IBM\es\bin\ESPACServer.ear

Postup

Chcete-li implementovat podnikovou aplikaci, která prolézacímu modulu WebSphere Portal umožňuje prolézat weby WebSphere Portal:

1. Zastavte instanci serveru WebSphere_Portal.
2. Není-li instance serveru server1 produktu WebSphere Application Server již spuštěna, spusťte ji.
3. Na serveru WebSphere Portal spusťte konzolu pro správu produktu WebSphere Application Server. Zobrazí-li se výzva k přihlášení, přihlašte se.

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://jméno_hostitele:port/admin`, kde *jméno_hostitele* je jméno 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://jméno_hostitele:port/ibm/console`, kde *jméno_hostitele* je jméno hostitele vyhledávacího serveru a *port* je číslo portu konzoly pro správu produktu WebSphere Application Server. Konzola pro správu obvykle používá port 9060.

4. Klepněte na volbu **Aplikace** a poté na volbu **Instalovat novou aplikaci**.
5. Klepněte na volbu **Procházet** a vyberte soubor ESPACServer.ear pro příslušný systém.
6. Klepněte dvakrát na tlačítko **Další**. Obdržíte-li varování týkající se souborů zásad, klepněte na tlačítko **Pokračovat**.
7. Klepejte na tlačítko **Další**, dokud se nezobrazí stránka Mapovat moduly na aplikační servery; poté vyberte následující volby a klepněte na tlačítko **Použít**:
 - a. V textovém poli **Klastry a servery** vyberte volbu **WebSphere:cell=jméno_buňky, node=jméno_uzlu, server=WebSphere Portal**.
 - b. Zaškrtněte políčko u modulu **ESPACServer.ear**.
8. Pokračujte v klepání na tlačítko **Další**, dokud se nezobrazí stránka Souhrn; poté klepněte na tlačítko **Dokončit**.
9. Klepněte na odkaz **Uložit do hlavní konfigurace** a poté klepnutím na tlačítko **Uložit** uložte provedené změny do konfigurace serveru WebSphere Application Server.
10. Restartujte server WebSphere Portal.

Související pojmy

“Integrace podnikového vyhledávání s produktem WebSphere Portal” na stránce 187
 Vyhledávací funkce produktu IBM WebSphere Portal můžete rozšířit implementaci portletů podnikového vyhledávání v produktu WebSphere Portal a ve vyhledávacím centru produktu WebSphere Portal.

Kopírování adresy URL pro prolézání ze systému WebSphere Portal

Chcete-li snížit pravděpodobnost zadání nesprávné adresy URL, můžete při konfigurování prolézacího modulu WebSphere Portal zkopírovat adresu URL z webu WebSphere Portal, kterou chcete prolézat, a vložit ji do příslušného pole.

O této úloze

Při vytváření prolézacího modulu WebSphere Portal zadáváte adresu URL portálu na serveru WebSphere Portal, 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 WebSphere Portal prolézat, postupujte takto:

1. Zobrazí-li se v konzole pro správu podnikového vyhledávání stránka Server WebSphere Portal určený k prolézání, ověřte, že je server WebSphere Portal spuštěn, a poté se přihlašte k serveru WebSphere Portal jako administrátor.
2. V pravém horním rohu vyberte **Správa**.
3. V navigační oblasti vlevo vyberte volbu **Nastavení portálu** a poté vyberte volbu **Správa hledání**.
4. Na stránce Správa kolekcí vyhledávání vyberte volbu **PortalCollection** v oblasti Kolekce vyhledávání. (Jsou-li k dispozici další kolekce, můžete vybrat jinou kolekci.)
5. V části Zdroje obsahu v kolekci klepněte na volbu **Přidat zdroj obsahu**.
6. Pro volbu **Typ zdroje prohledávání** vyberte hodnotu **Portálové stránky**. Adresa URL webu je zobrazena v poli **Sbírat dokumenty propojené z této adresy URL**.
7. Zkopírujte adresu URL do schránky. (Můžete například označit adresu URL, držet klávesu Ctrl stisknutou a stisknout klávesu Insert.)

8. Vraťte se ke konzole pro správu podnikového vyhledávání a vložte zkopírovanou adresu URL do pole **Adresa URL serveru WebSphere Portal**.

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

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 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 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 181

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.

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.

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 změnil čísla 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 **Systém** otevřete pohled **Systém**.
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.
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 jmen kolekcí. U každé kolekce, kterou aplikace může aktualizovat, zaškrtněte políčko **Vybrat**.

8. Klepněte na tlačítko **OK**.
9. Pokud jste změnil čísla portu modulu pro příjem dat, 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 pojmy

"Modul pro příjem dat" v příručce "Programming Guide and API Reference for Enterprise Search"

Související úlohy

"Monitorování modulu pro příjem dat" na stránce 221

Při monitorování modulu pro příjem dat zobrazíte jeho stav a podrobnosti o aktivitě klientské aplikace modulu pro příjem dat.

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 a k asociování metadat s dokumenty v indexu podnikového vyhledávání.

Modul plug-in obsahuje třídu Java, která je volána pro každý dokument zpracováváný prolézacím modulem. Třídě Java je předán identifikátor dokumentu (URI) z indexu podnikového vyhledávání, prvky zabezpečení a metadata. Třída vrátí novou nebo upravenou sadu prvků zabezpečení a metadat nebo pokyn k ignorování dokumentu prolézacím modulem.

Po prvním prolezení všech dokumentů v prolézaném prostoru je modul plug-in volán pouze pro nové nebo upravené dokumenty. Chcete-li změnit prvky zabezpečení a metadata pro dokumenty, které jsou obsaženy v indexu podnikového vyhledávání, ale nebyly aktualizovány v původním zdroji dat, spusíte úplné prolézání všech dokumentů v prolézaném prostoru a poté reorganizujete index.

Modul plug-in nelze asociovat s existujícím prolézacím modulem. Jméno třídy modulu plug-in a cestu ke třídě musíte určit při konfigurování vlastností pro nový prolézací modul.

Použití modulu plug-in k 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.

Použití modulu plug-in k přidávání metadat do dokumentů

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 jméno-hodnota. K prohledávání metadat mohou uživatelé použít dotaz pro volný text nebo dotaz s uvedením jména pole metadat.

Modul 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.

Související pojmy

"Moduly plug-in prolézacího modulu" v příručce "Programming Guide and API Reference for Enterprise Search"

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 přidal 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 odebrání 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.

Prolézací moduly Content Edition

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Content Edition v režimu přístupu k serveru je následující:

```
vbr://jméno_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.

jméno_serveru

Jméno serveru WebSphere Information Integrator Content Edition.

systemové_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

Jméno třídy aplikačního serveru používaný k vytvoření manipulátoru EJB.

java.naming.provider.url

Adresa URL služby jmen aplikačního serveru používaného k zadání požadavku na manipulátor EJB.

Příklady

Documentum:

```
vbr://vbrsrv.ibm.com/Documentum/c06b/094e827780000302//CONTENT/?  
java.naming.provider.url=iio%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=iio%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://jméno_databáze/jméno_tabulky  
/jedinečný_identifikátor_jména_sloupce_1/jedinečný_identifikátor_hodnoty_1  
[/jedinečný_identifikátor_jména_sloupce_2/jedinečný_identifikátor_hodnoty_2/...  
/jedinečný_identifikátor_jména_sloupce_N/jedinečný_identifikátor_hodnoty_N]
```

Parametry:

Pro všechna pole je použito kódování adres URL.

jméno_databáze

Interní jméno nebo alias databáze.

jméno_tabulky

Jméno cílové tabulky včetně jména schématu.

jedinečný_identifikátor_jména_sloupce_1

Jméno prvního sloupce jedinečného identifikátoru v tabulce.

jedinečný_identifikátor_hodnoty_1

Hodnota v prvním sloupci jedinečného identifikátoru.

jedinečný_identifikátor_jména_sloupce_N

Jméno *ntého* sloupce jedinečného identifikátoru v tabulce.

jedinečný_identifikátor_hodnoty_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/REMOVEDB/SCHEMA2.TABLE2/NAME/DAVID

Prolézací moduly DB2 Content Manager

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem DB2 Content Manager:

cm://jméno_serveru/jméno_typu_položky/PID

Parametry

jméno_serveru

Jméno serveru knihovny IBM DB2 Content Manager.

jméno_typu_položky

Jméno cílového typu položky.

PID Trvalý identifikátor DB2 Content Manager.

Příklad

cm://cmsrvctg/ITEMTYPE1/92+3+ICM8+icm\lsdb12+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://jméno_serveru:číslo_portu/ID_repliky_databáze/jméno_databáze_s_cestou
/univerzální_ID_pohledu/univerzální_ID_dokumentu
/?AttNo=číslo_přílohy&AttName=jméno_souboru_přílohy

Parametry

Pro všechna pole je použito kódování adres URL.

jméno_serveru

Jméno serveru Domino Document Manager.

číslo_portu

Volitelné: Číslo portu pro server Domino Document Manager.

ID_repliky_databáze

Identifikátor repliky databáze.

jméno_databáze_s_cestou

Cesta a jméno 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.

jméno_souboru_přílohy

Volitelné: Původní jméno 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 Notes

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Notes:

```
domino://jméno_serveru[:číslo_portu]/ID_repliky_databáze/jméno_databáze_s_cestou/[univerzální_ID_pohledu]/univerzální_ID_dokumentu[?AttNo=číslo_přílohy&AttName=jméno_souboru_přílohy]
```

Parametry

Pro všechna pole je použito kódování adres URL.

jméno_serveru

Jméno serveru Lotus Notes.

číslo_portu

Číslo portu serveru Lotus Notes. Číslo portu je nepovinné.

ID_repliky_databáze

Identifikátor repliky databáze.

jméno_databáze_s_cestou

Cesta a jméno 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 neurčujete pohled nebo složku, kterou chcete prolézat (například tehdy, chcete-li prolézat všechny dokumenty v databázi), univerzální ID pohledu se nezadává.

univerzální_ID_dokumentu

Univerzální ID dokumentu definované v dokumentu zpracovávaném prolézacím modulem.

číslo_přílohy

Pořadové číslo přílohy. Přílohy se číslují od nuly. Číslo přílohy je nepovinné.

jméno_souboru_přílohy

Původní jméno souboru přílohy. Jméno souboru přílohy je nepovinné.

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://jméno_serveru:číslo_portu/ID_repliky_databáze/jméno_databáze_s_cestou  
/univerzální_ID_pohledu/univerzální_ID_dokumentu  
/?AttNo=číslo_přílohy&AttName=jméno_souboru_přílohy
```

Parametry

Pro všechna pole je použito kódování adres URL.

jméno_serveru

Jméno serveru Lotus QuickPlace.

číslo_portu

Volitelné: Číslo portu pro server QuickPlace.

ID_repliky_databáze

Identifikátor repliky databáze.

jméno_databáze_s_cestou

Cesta a jméno 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 prolézáném dokumentu.

číslo_přílohy

Volitelné: Pořadové číslo přílohy. Přílohy se číslují od nuly.

jméno_souboru_přílohy

Volitelné: Původní jméno 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 Souborový systém UNIX

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Souborový systém UNIX:

`file:///jméno_adresáře/jméno_souboru`

Parametry

Pro všechna pole je použito kódování adres URL.

jméno_adresáře

Absolutní cesta k adresáři.

jméno_souboru

Jméno souboru.

Příklad

`file:///home/user/test.doc`

Prolézací moduly WebSphere Portal

Formát identifikátorů URI pro dokumenty zpracovávané prolézacím modulem WebSphere Portal:

`wps://cesta_pro_přístup_k_produkту_WPS?portletDefID=ID_definice_portletu&pageID=ID_stránky&useSSL=SSL`

Parametry

Pro všechna pole je použito kódování adres URL.

cesta_pro_přístup_k_produkту_WPS

Cesta k serveru WebSphere Portal bez uvedení protokolu.

ID_definice_portletu

Identifikátor definice portletu pro server WebSphere Portal.

ID_stránky

Identifikátor stránky pro server WebSphere Portal.

SSL Parametr `useSSL=true` je přidáván v případě, že je pro cestu uveden protokol HTTPS. V opačném případě je přidán parametr `useSSL=false`.

Příklady

Tělo dokumentu:

```
wps://wps.server.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+
LKnPy1vM0Y_QjzKCN4g3cQbJgQi0-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8qHRj/de1ta
/base64xml/L01DU1kvd0NrQUORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!
?portletDefID=3_0_3S&pageID=6_0_6J&useSSL=false
```

Příklady

Podpora zabezpečení SSL:

```
wps://wps.server.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+
LKnPy1vM0Y_QjzKCN4g3cQbJgQi0-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8qHRj/de1ta
/base64xml/L01DU1kvd0NrQUORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!
?portletDefID=7_0_A4&pageID=6_0_6J&useSSL=true
```

Prolézací moduly Souborový systém Windows

Formáty identifikátorů URI pro dokumenty zpracovávané prolézacím modulem Souborový systém Windows:

```
file:///jměno_adresáře/jměno_souboru
file:///jměno_síťové_složky/jměno_adresáře/jměno_souboru
```

Parametry

Pro všechna pole je použito kódování adres URL.

jměno_adresáře

Absolutní cesta k adresáři.

jměno_souboru

Jméno souboru.

jměno_síťové_složky

Pouze u dokumentů umístěných na vzdálených serverech: jméno sdílené složky v síti Windows.

Příklady

Lokální souborový systém:

```
file:///d:/adresář/test.doc
```

Síťový souborový systém:

```
file:///server.ibm.com/adresář/soubor.doc
```

Související pojmy

“Správa prolézacích modulů podnikového vyhledávání” na stránce 19

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

Související úlohy

“Konfigurování kategorií” na stránce 91

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.

“Konfigurování oborů” na stránce 123

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.

“Odebrání identifikátorů URI z indexu” na stránce 125

Chcete-li uživatelům zabránit ve vyhledávání dokumentů v kolekci, můžete odebrat identifikátory URI pro tyto dokumenty z indexu.

“Konfigurování rychlých odkazů” na stránce 135

Při vytváření rychlého odkazu pro kolekci podnikového vyhledávání asociujete identifikátor URI dokumentu s klíčovými slovy, která spouštějí jeho zahrnutí do výsledků vyhledávání.

“Zobrazení podrobností o identifikátoru URI” na stránce 207

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.

“Zobrazení sestav odstraněných dokumentů” na stránce 223

Můžete také zobrazit podrobné informace o dokumentech odstraněných ze systému podnikového vyhledávání. Tyto informace jsou k dispozici pouze v případě, že jste pro kolekci povolili sledování dokumentů.

Správa analyzátoru podnikového vyhledávání

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" v dokumentu "Text Analysis Integration"

"Analýza textu integrovaná v podnikovém vyhledávání" v dokumentu "Text Analysis Integration"

"Architektura UIMA (Unstructured Information Management Architecture)" v dokumentu "Text Analysis Integration"

"Aplikace sémantického vyhledávání" v dokumentu "Text Analysis Integration"

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 jméno 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.

Podnikové vyhledávání podporuje dva přístupy k zařazování dokumentů do kategorií:

Na základě pravidel

Tento přístup použijte, chcete-li přiřadit dokumenty ke kategoriím podle pravidel, která zadáte. Můžete například seskupit dokumenty s určitými společnými částmi identifikátoru URI nebo dokumenty s příbuzným obsahem (například dokumenty obsahující nebo neobsahující určitá slova a fráze).

Na základě modelu

Tento přístup použijte, pokud pracujete s produktem IBM WebSphere Portal a kategorie předdefinované v produktu WebSphere Portal chcete používat při prohledávání kolekci podnikového vyhledávání.

Použití změn zařazení do kategorií

K vytváření a administraci kategorií se používá konzola pro správu podnikového vyhledávání:

- Při vytvoření kolekce zvolíte typ kategorizace. Můžete zvolit práci bez kategorií, použití kategorií založených na pravidlech nebo kategorií založených na modelu.
- 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 a analýzu všech dokumentů a reorganizaci indexu.
- Pokud se rozhodnete používat kategorie založené na pravidlech, můžete pomocí konzoly pro správu spravovat strom kategorií, kategorie a pravidla kategorií. Jestliže změníte kategorie nebo pravidla kategorií po zpracování dokumentů prolézacím modulem a sestavení indexu, bude kvalita vyhledávání snížena, dokud neprovedete nové prolezení a analýzu všech dokumentů a reorganizaci indexu.

Pole obsahu dokumentu

Chcete-li vytvořit kolekci s kategoriemi založenými na modelu nebo na pravidlech, které používají pravidla pro obsah dokumentů (a nikoli pravidla určující vzor identifikátorů URI), ujistěte se, že v dokumentech, které mají být rozděleny do kategorií, existují pole obsahu.

Kategorie založené na modelu a pravidla kategorií založená na obsahu dokumentů pracují s poli obsahu dokumentu a neberou v úvahu metadata dokumentu. Tyto volby kategorizace lze proto použít pouze u dokumentů s poli obsahu nebo s poli, která lze při konfigurování prolézacího modulu nastavit jako pole obsahu.

Pokud konfigurujete kategorie založené na modelu nebo vytvoříte pravidla kategorií založená na obsahu dokumentu, budou se pravidla vztahovat pouze na části dokumentu. Pro následující typy prolézacích modulů se pravidla vztahují na soubory (například na obsah souborů souborů v systému UNIX nebo Microsoft Windows) nebo na obsah webových stránek:

- Exchange Server

- NNTP
- Souborový systém UNIX
- webový prolézací modul
- WebSphere Portal
- Souborový systém Windows

Pro následující typy prolézacích modulů můžete určit, která pole jsou pole obsahu dokumentu, při konfigurování voleb pro jednotlivé zdroje dat. Při vytváření prolézacího modulu je tato volba k dispozici na poslední stránce průvodce prolézacím modulem. Pokud prolézací modul již existuje, upravte prolézáný prostor, vyberte zdroj dat, pro který chcete upravit volby, a poté určete, která pole v tomto zdroji dat jsou pole obsahu dokumentu.

- Content Edition
- DB2 Content Manager
- DB2
- Domino Document Manager
- Notes
- QuickPlace

Kategorie založené na pravidlech

Přiřazení dokumentů ke kategoriím v kolekci podnikového vyhledávání lze určovat konfigurováním pravidel.

Kategorie a 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 pravidel 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 kategorií.

Při administraci stromu kategorií (taxonomie) kolekce určujete, na které místo v hierarchii kategorií chcete novou kategorií 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:

Vzor identifikátoru 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/

Obsah 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, které obsahují slovo lz, nikoli však slovo dávky:

+lz -dávky

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. Povoleny jsou 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.

Pravidla pro obsah dokumentů se vztahují pouze na části dokumentu s obsahem. Pro následující typy prolézacích modulů se pravidla vztahují na soubory (například na obsah souborů souborů v systému UNIX nebo Microsoft Windows) nebo na obsah webových stránek:

- Exchange Server
- NNTP
- Souborový systém UNIX
- webový prolézací modul
- WebSphere Portal
- Souborový systém Windows

Pro následující typy prolézacích modulů můžete určit, která pole jsou pole obsahu dokumentu, při konfigurování voleb pro jednotlivé zdroje dat. Při vytváření prolézacího modulu je tato volba k dispozici na poslední stránce průvodce prolézacím modulem. Pokud prolézací modul již existuje, upravte prolézáný prostor, vyberte zdroj dat, pro který chcete upravit volby, a poté určete, která pole v tomto zdroji dat jsou pole obsahu dokumentu.

- Content Edition
- DB2 Content Manager
- DB2
- Domino Document Manager
- Notes
- QuickPlace

Související úlohy

“Migrace kolekce z produktu WebSphere Portal” na stránce 197

Chcete-li migrovat kolekce z produktu WebSphere Portal do prostředí podnikového vyhledávání, nejprve je připravte v produktu WebSphere Portal a poté proveďte jejich migraci pomocí průvodce migrací.

Kategorie založené na modelu

Pokud v systému IBM WebSphere Portal používáte kategorie, můžete stejné kategorie používat i v kolekcích podnikového vyhledávání.

Produkt WebSphere Portal nabízí předdefinovanou taxonomii zahrnující přes 2300 témat. Tato témata jsou rozdělena do hlavních obchodních kategorií, například Počítače, Finance a Doprava. Uživatelé portálu mohou vytvářet aplikace, které automaticky zjišťují, které dokumenty odpovídají těmto tematickým oblastem, a přizpůsobovat kategorie vlastním obchodním potřebám.

Pokud chcete používat kategorie produktu WebSphere Portal pro potřeby podnikového vyhledávání, musíte provést následující kroky:

- Pomocí průvodce migrací importujte soubory taxonomií založených na modelech do prostředí podnikového vyhledávání.
- Při vytvoření kolekce nebo konfigurování pravidel analýzy pro kolekci určete, že chcete používat kategorie založené na modelu.
- Zajistěte instalaci produktu WebSphere Portal na indexový server podnikového vyhledávání.
- Ke správě kategorií použijte kategorizační nástroje dostupné v produktu WebSphere Portal. Kategorie založené na modelu nelze spravovat pomocí konzoly pro správu podnikového vyhledávání.

Související úlohy

“Migrace taxonomie založené na modelu z produktu WebSphere Portal” na stránce 195

Pomocí portletu Taxonomy Management produktu WebSphere Portal můžete určit, kterou taxonomii založenou na modelu chcete používat pro kolekce podnikového vyhledávání. Kolekce, které jste již migrovali do prostředí podnikového vyhledávání, nebudou novou migrací taxonomie nijak ovlivněny.

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 jména podkategorií, takže uživatelé mohou procházet mezi kategoriemi a zobrazovat podmnožiny dokumentů jednu po druhé.

Související úlohy

“Migrace kolekce z produktu WebSphere Portal” na stránce 197

Chcete-li migrovat kolekce z produktu WebSphere Portal do prostředí podnikového vyhledávání, nejprve je připravte v produktu WebSphere Portal a poté proveďte jejich migraci pomocí průvodce migrací.

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, kategorie založené na modelech, které jsou k dispozici v systému portálu IBM WebSphere Portal, 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 všech dokumentů v kolekci a po analýze dokumentů reorganizujte 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 použít taxonomii obsahující pravidla kategorií, která konfiguruje specificky pro tuto kolekci.

- Pokud konfiguruje kolekci, kterou jste vytvořili pro podnikové vyhledávání, vyberte tuto volbu k určení jmen kategorií a pravidel pro kategorizaci dokumentů.
- Jestliže konfiguruje kolekci, kterou jste migrovali z portálu IBM WebSphere, vyberte tuto volbu v případě, že chcete použít nebo změnit kategorie založené na pravidle, které jste importovali.

Založený na modelu

Tuto volbu vyberte, chcete-li dokumenty přiřadit ke kategoriím založeným na modelu, které jsou definovány v systému WebSphere Portal. Při použití této volby musí být na indexovém serveru podnikového vyhledávání nainstalován produkt WebSphere Portal. Ke správě kategorií je třeba použít kategorizační nástroje, které jsou součástí produktu WebSphere Portal.

3. Klepněte na tlačítko **OK**.

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 jména kategorie prohledávat pouze určitou část dané kolekce. Mohou také vybrat kategorii 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 všech dokumentů v kolekci, analyzujte znovu dokumenty a reorganizujte 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 jméno kategorie, bude nová kategorie vnořena pod vybranou kategorii ve stromu kategorií.

Otevře se průvodce umožňující zadat pravidla pro asociování dokumentů s novou kategorií:

- a. Na stránce Vytvořit kategorii zadejte popisné jméno 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 **Jméno pravidla** jedinečné jméno pro pravidlo. Toto jméno musí být jedinečné pro všechny kategorie v kolekci.
- d. Zadejte pravidlo, které chcete použít pro asociování dokumentů s danou kategorií, a potom klepněte na tlačítko **OK**.

- Pokud chcete, aby byl při podnikovém vyhledávání používán identifikátor URI dokumentu k určení, zda dokument patří do dané kategorie, 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

- Pokud chcete, aby byla k určení, zda dokument patří do dané kategorie, používána analýza slov v polích obsahu dokumentu, klepněte na volbu **Obsah dokumentu**, vyberte jazyk dokumentů a zadejte slova, která se musí nebo nesmí vyskytovat v obsahu dokumentu. Vyjádřete pravidlo ve stejném formátu jako dotaz (povoleny jsou pouze operátory dotazu +, - a " ").

Pokud dokument obsahuje nebo neobsahuje zadaná slova, pak je asociován s kategorií.

Příklad: +finance -účetnictví +"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 pojmy

“Migrace z produktu WebSphere Portal do podnikového vyhledávání” na stránce 195 Podnikové vyhledávání nabízí průvodce migrací, kterého můžete použít k migraci taxonomií a kolekci z produktu IBM WebSphere Portal do prostředí podnikového vyhledávání.

Související úlohy

“Migrace taxonomie založené na modelu z produktu WebSphere Portal” na stránce 195 Pomocí portletu Taxonomy Management produktu WebSphere Portal můžete určit, kterou taxonomii založenou na modelu chcete používat pro kolekce podnikového vyhledávání. Kolekce, které jste již migrovali do prostředí podnikového vyhledávání, nebudou novou migrací taxonomie nijak ovlivněny.

Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 76 Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

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. Linux

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 text obsažený v prvcích 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 tyto prvky zadáním mapovaných 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 na tyto konkrétní prvky. Dotaz na text author:Novák nalezne dokumenty XML, u nichž je v prvku author zadána hodnota Novák.

Mapování prvků XML na vyhledávací pole

Při mapování prvků XML na vyhledávací pole můžete určit, které prvky XML mohou uživatelé vyhledávat zadáním jména 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.

O této úloze

Při vytváření mapování polí XML nebo přidávání, změně či odstraňování polí v existujícím mapování polí XML se změny projeví až po následném zastavení a opětném spuštění modulu analýzy. Nová a změněná mapování se týkají pouze dat, která jsou analyzována po restartu modulu analýzy. Nová a změněná mapování nemají vliv na data, která již byla analyzována a zařazena do indexu.

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 **Jméno kořenového prvku XML** zadejte jméno kořenového prvku: **personnel**. Ujistěte se, že zadané jméno přesně odpovídá kořenovému prvku v dokumentech XML, které chcete vyhledat. Při analýze a indexování dokumentů XML podnikové vyhledávání vybere mapování podle jména kořenového prvku.
4. Do pole **Jméno mapování XML** zadejte jméno pro danou sadu pravidel mapování polí XML.
Po vytvoření sady pravidel mapování XML bude toto jméno zobrazeno na stránce Mapování polí XML a po výběru tohoto jména můžete přidávat, odstraňovat nebo měnit pravidla mapování.
5. Mapování prvku XML **jobrole** na vyhledávací pole se jménem **jobrole**:
 - a. Do pole **Jméno pole** zadejte hodnotu **jobrole**.
 - b. Do pole **Jméno prvku XML** zadejte hodnotu **jobrole**.
 - c. Chcete-li uživatelům umožnit zadávání dotazů na pole **jobrole** a jeho zobrazení ve výsledcích vyhledávání, zaškrtněte políčka **Vyhledávání s použitím polí** a **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 **Jméno pole** zadejte hodnotu **jobroles**.
 - c. Do pole **Jméno prvku XML** zadejte hodnotu **jobrole**.

Tip: Jména prvků XML nemusí odpovídat jménům vyhledávacích polí. Na jedno vyhledávací pole lze mapovat více prvků XML.

 - d. Chcete-li uživatelům umožnit zadávání dotazů na pole **jobrole** a jeho zobrazení ve výsledcích vyhledávání, zaškrtněte políčka **Vyhledávání s použitím polí** a **Výsledky vyhledávání**.
7. Mapování prvku XML **section** s atributem **expertise** na vyhledávací pole se jménem **expertise**:
 - a. Klepnutím na volbu **Přidat pole** přidejte do seznamu pravidel mapování polí prázdný řádek.
 - b. Do pole **Jméno pole** zadejte hodnotu **expertise**.

- c. Do pole **Jméno pole** zadejte hodnotu **section**.
 - d. Do pole **Jméno atributu XML** zadejte hodnotu **id**.
 - e. Do pole **Hodnota atributu XML** zadejte hodnotu **expertise**.
 - f. Chcete-li uživatelům umožnit zadávání dotazů na pole **expertise** a zobrazit hodnoty pole **expertise** ve výsledcích vyhledávání, zaškrtněte políčka **Vyhledávání s použitím polí** a **Výsledky vyhledávání**.
8. Klepněte na tlačítko **OK**.

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 v indexu.

K mapování prvků metadat HTML na vyhledávací pole slouží konzola pro správu podnikového vyhledávání.

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**.

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 jména 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 zastavení a opětném spuštění modulu analýzy. Změny se týkají pouze dat, která jsou analyzována po restartu modulu analýzy. Chcete-li aktualizovat dokumenty, které již byly analyzovány a indexovány, je nutné dokumenty znovu prolézt a analyzovat, a pak reorganizovat index.

Postup

Mapování prvků metadat HTML na vyhledávací pole:

1. Upravte kolekci, vyberte stránku Analýza XML klepněte na volbu **Mapovat metadata HTML na pole**.
2. Klepnutím na volbu **Přidat pole** na stránce Mapování polí HTML přidejte do seznamu pravidel mapování polí prázdný řádek.
3. Do pole **Jméno pole** zadejte jméno, které chcete přiřadit k mapovanému prvku metadat HTML. Uživatelé mohou toto jméno pole zadat při dotazování na dokumenty HTML v této kolekci.
4. Do pole **Jméno prvku metadat HTML** zadejte jméno prvku metadat HTML, který chcete namapovat.
5. Chcete-li uživatelům umožnit zadávání dotazů na toto pole a jeho zobrazení ve výsledcích vyhledávání, zaškrtněte políčka **Vyhledávání s použitím polí** a **Výsledky vyhledávání**.
6. Pokud je datový typ tohoto pole DECIMAL, DOUBLE, INTEGER, SHORT, TIME nebo TIMESTAMP a chcete uživatelům umožnit při prohledávání tohoto pole zadávat parametrické dotazy, zaškrtněte políčko **Parametrické vyhledávání**.
7. Klepněte na tlačítko **OK**.

Příklady:

Uživatelé nyní mohou zadáním dotazu na jména 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 o počítačích Thinkpad zadat dotaz podobný následujícímu:

```
abstract:thinkpad
```

Vlastní zpracování textu

Kvalitu a přesnost výsledků vyhledávání lze zvýšit integrací vlastních algoritmů pro zpracování textu do kolekcí podnikového vyhledávání.

Produkt WebSphere Information Integrator OmniFind 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 vytvořen 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 XML 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 datové struktury v rámci obecné struktury 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 datové struktury v rámci obecné struktury analýzy na databázové tabulky s podporou technologie JDBC (Java Database Connectivity). Můžete mapovat data na tabulky databáze IBM DB2 Universal Database (DB2 UDB) 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" v dokumentu "Text Analysis Integration"

"Architektura UIMA (Unstructured Information Management Architecture)" v dokumentu "Text Analysis Integration"

"Workflow pro integraci vlastní analýzy" v dokumentu "Text Analysis Integration"

"Algoritmy analýzy textu" v dokumentu "Text Analysis Integration"

"Aplikace sémantického vyhledávání" v dokumentu "Text Analysis Integration"

"Dotaz sémantického vyhledávání" v dokumentu "Text Analysis Integration"

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 kolekcí 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 zobrazované jméno a vybrat volbu zobrazení zdroje XML. (Prostřednictvím zdroje XML můžete zobrazit informace vytvářené tímto strojem.)

Pokud je stroj pro analýzu textu asociován s kolekcí, 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 popisné jméno pro nový stroj. Systém použije toto zobrazované jméno 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 soubor umístěn ve vašem systému. Je-li soubor větší než 8 MB, musí být umístěn na indexovém serveru.
7. Klepněte na tlačítko **OK**. Stroj pro analýzu textu bude uveden na stránce Stroje pro analýzu textu.

Související pojmy

"Workflow pro integraci vlastní analýzy" v dokumentu "Text Analysis Integration"

"Integrace vlastní analýzy textu" v dokumentu "Text Analysis Integration"

"Architektura UIMA (Unstructured Information Management Architecture)" v dokumentu "Text Analysis Integration"

"Přístupy pro použití značek jazyka XML při analýze a vyhledávání" v dokumentu "Text Analysis Integration"

Související úlohy

"Vytvoření konfiguračního souboru mapování kódu XML na typy architektury UIMA" v dokumentu "Text Analysis Integration"

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 jméno jiného vlastního stroje pro analýzu textu, budou všechna dříve definovaná mapování analýzy textu pro danou kolekci zachována. Pokud například změňte 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 jméno stroje **Výchozí**.
3. Na stránce Vybrat stroj pro analýzu textu pro tuto kolekci vyberte jméno 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.
4. Klepněte na tlačítko **OK**.

Související pojmy

"Workflow pro integraci vlastní analýzy" v dokumentu "Text Analysis Integration"

"Integrace vlastní analýzy textu" v dokumentu "Text Analysis Integration"

"Architektura UIMA (Unstructured Information Management Architecture)" v dokumentu "Text Analysis Integration"

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 JDBC s použitím sémantického vyhledávání, můžete mapovat 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.

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 XML. Soubory mapování XML musí dodržovat specifikaci architektury UIMA pro analýzu textu.

Přidáním souborů mapování XML 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. Výsledky analýzy (v obecné struktuře analýzy) lze poté 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 definovat žádné mapování XML. 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 obecné struktury analýzy** klepněte na volbu **Přidat mapování**.
3. Na stránce Mapovat prvky XML na obecnou strukturu analýzy zadejte popisné zobrazované jméno pro soubor mapování XML.
4. Určete umístění souboru. Je-li soubor mapování XML menší než 8 MB, můžete cestu k souboru zadat nebo ji určit procházením adresářů a výběrem souboru. Je-li soubor mapování XML větší než 8 MB, zkontrolujte, zda se daný soubor nalézá na indexovém serveru, a určete úplnou cestu k souboru.
5. Klepněte na tlačítko **OK**. Nový soubor mapování XML bude přidán do seznamu souborů mapování XML na stránce Volby zpracování textu.

Související pojmy

"Workflow pro integraci vlastní analýzy" v dokumentu "Text Analysis Integration"

"Integrace vlastní analýzy textu" v dokumentu "Text Analysis Integration"

"Architektura UIMA (Unstructured Information Management Architecture)" v dokumentu "Text Analysis Integration"

"Přístupy pro použití značek jazyka XML při analýze a vyhledávání" v dokumentu "Text Analysis Integration"

Související úlohy

"Vytvoření konfiguračního souboru mapování kódu XML na typy architektury UIMA" v dokumentu "Text Analysis Integration"

Mapování obecné struktury analýzy na index

Můžete určit, kterou obecnou strukturu analýzy chcete použít, když uživatelé zadají dotaz na kolekci s použitím 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.

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 jednu obecnou strukturu analýzy na index** klepněte na volbu **Vybrat obecnou strukturu analýzy**.
3. Na stránce Vybrat obecnou strukturu analýzy pro tuto kolekci vyberte mapování, které chcete používat pro index podnikového vyhledávání:
 - Chcete-li používat výchozí systémová pravidla pro mapování datových struktur v rámci obecné struktury analýzy na index podnikového vyhledávání, vyberte volbu **Výchozí**.
 - Chcete-li mapovat konkrétní obecnou strukturu analýzy na index, určete umístění souboru obecné struktury analýzy. Je-li soubor menší než 8 MB, může být soubor umístěn ve vašem systému. Je-li soubor větší než 8 MB, musí být umístěn na indexovém serveru.
4. Klepněte na tlačítko **OK**. Určená obecná struktura analýzy bude zobrazena na stránce Volby zpracování textu.

Související pojmy

"Workflow pro integraci vlastní analýzy" v dokumentu "Text Analysis Integration"

"Integrace vlastní analýzy textu" v dokumentu "Text Analysis Integration"

"Architektura UIMA (Unstructured Information Management Architecture)" v dokumentu "Text Analysis Integration"

"Přístupy pro indexování výsledků vlastní analýzy" v dokumentu "Text Analysis Integration"

Související úlohy

"Vytvoření konfiguračního souboru sestavení indexu" v dokumentu "Text Analysis Integration"

Mapování obecné struktury analýzy na tabulky JDBC

Můžete určit, kterou obecnou strukturu analýzy chcete mapovat na tabulky JDBC pro použití v databázových aplikacích.

Než začnete

Chcete-li mapovat obecnou strukturu analýzy na databázové tabulky JDBC (Java Database Connectivity), musíte mít přiřazenu roli administrátora podnikového vyhledávání nebo administrátora kolekce pro danou kolekci.

O této úloze

Mapováním obecné struktury analýzy na tabulky JDBC umožníte použití dat v databázových aplikacích. Uživatelé například mohou prostřednictvím dotazů SQL zadávaných mimo rámec podnikového vyhledávání prohledávat anotace přidané obecnou strukturou analýzy. Informace můžete rovněž použít pro další zpracování textu, například v aplikacích dolování dat.

Postup

Chcete-li mapovat obecnou strukturu analýzy na tabulky JDBC, postupujte takto:

1. Upravte kolekci, vyberte stránku Analýza a klepněte na volbu **Konfigurovat volby zpracování textu**.
2. V oblasti **Mapovat obecnou strukturu analýzy na tabulky JDBC** klepněte na volbu **Přidat mapování**.
3. Na stránce Mapovat obecnou strukturu analýzy na tabulky JDBC zadejte popisné zobrazované jméno pro obecnou strukturu analýzy, kterou chcete použít pro mapování informací na tabulky JDBC.
4. Určete umístění souboru obecné struktury analýzy. Je-li soubor menší než 8 MB, může být soubor umístěn ve vašem systému. Je-li soubor větší než 8 MB, musí být umístěn na indexovém serveru.
5. Klepněte na tlačítko **OK**. Nový soubor bude přidán do seznamu souborů na stránce Volby zpracování textu. Datum a čas přidání souboru do systému je rovněž zobrazen.

Související pojmy

"Workflow pro integraci vlastní analýzy" v dokumentu "Text Analysis Integration"

"Integrace vlastní analýzy textu" v dokumentu "Text Analysis Integration"

"Architektura UIMA (Unstructured Information Management Architecture)" v dokumentu "Text Analysis Integration"

"Přístupy pro mapování výsledků analýzy na databázi" v dokumentu "Text Analysis Integration"

Související úlohy

"Vytvoření konfiguračního souboru mapování kódu XML" v dokumentu "Text Analysis Integration"

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 podslova v rámci rozsáhlejšího složeného výrazu.

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

"Lingvistická podpora pro sémantické vyhledávání" v dokumentu "Text Analysis Integration"

"Analýza textu integrovaná v podnikovém vyhledávání" v dokumentu "Text Analysis Integration"

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 například 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" v dokumentu "Text Analysis Integration"

"Analýza textu integrovaná v podnikovém vyhledávání" v dokumentu "Text Analysis Integration"

"Aplikace sémantického vyhledávání" v dokumentu "Text Analysis Integration"

"Dotaz sémantického vyhledávání" v dokumentu "Text Analysis Integration"

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.

Odebrání znaků nového řádku z rozsahů znaků mimo rozsah sady ASCII

V jazycích, ve kterých se slova neoddělují mezerami, jako je čínština a japonština, můžete konfigurovat analyzátor tak, aby odebíral některé druhy prázdných znaků oddělujících nové řádky.

Než začnete

Chcete-li dokončit tuto úlohu, musíte se přihlásit jako administrátor podnikového vyhledávání.

O této úloze

Pokud je pro kolekci povolena tato volba a dokument je psán v čínštině nebo japonštině, odebere analyzátor posloupnosti mezerových znaků oddělující dvě písmena. Písmena musí ležet v rozsahu znaků kódu Unicode mimo rozsah sady ASCII. Odebírány jsou následující znaky:

- tabulátor (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 analyzovány a uloženy do indexu, proveďte nové prolezení a analýzu dokumentů a poté index reorganizujte.

Postup

Chcete-li odebrat znaky nového řádku z prázdného prostoru, postupujte takto:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind Edition.
2. Otevřete následující soubor pro úpravy v textovém editoru. *ID_kolekce* je ID kolekce zadané (nebo přiřazené systémem) při vytvoření kolekce:
`ID_kolekce.parserdriver/collection.properties`
3. Změňte hodnotu vlastnosti `removeCjNewlineChars` z `false` na `true`.

Typy dokumentů přiřazené k analyzátorům kolekcí a k relacím 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 relace filtrování 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 relace Stellent omylem odeslán prostý text nebo dokumenty ve formátu HTML. Tato data jsou zasílána zpět ke zpracování některým z vestavěných analyzátorů, což však může negativně ovlivnit výkon systému. 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 relacím 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 relaci 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 spusťte analyzátor.

Přiřazení typů dokumentů k analyzátoru kolekce

Chcete-li k analyzátoru kolekce přiřadit konkrétní typy dokumentů, musíte vytvořit konfigurační soubor. 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 dokumenty mají být odeslány do relace Stellent v závislosti na příponě souboru nebo typu obsahu.
- Jakým způsobem mají být analyzovány soubory, jejichž typ je neznámý vzhledem k neúplným metadatům.

Formát souboru 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 `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 WebSphere II OmniFind 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 relace 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
```

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 jsd 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 jttdc stellent
EXTENSION jttdx 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 html
DEFAULT html
```

Přirazení typů dokumentů k relaci Stellent

Chcete-li určit, které typy dokumentů mají být analyzovány pomocí dokumentových filtrů Stellent, musíte vytvořit konfigurační soubor. 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

Konfigurační soubor určuje následující nastavení:

- Které typy souborů mají být analyzovány pomocí relace Stellent. Typ souboru odpovídá jednomu z typů souborů rozpoznávaných knihovnou Stellent.
- Které typy souborů 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 relace Stellent omylem, v důsledku nesprávných metadat.)
- Které typy souborů mají být odmítnuty, protože je podnikové vyhledávání nepodporuje.

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 relace 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 z následujících pravidel:

```
accept DEFAULT
accept ALL typ_dokumentu
accept typ typ_dokumentu
native DEFAULT
native typ typ_dokumentu
reject typ
```

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 Jedna z hodnot `FI_` v knihovně Stellent. Hodnota *typ_dokumentu* bude použita při aplikaci pravidla v prvku `doctype`.

DEFAULT

Znamená, že seznam přijímaných nebo nativních typů (v závislosti na typu pravidla) zahrnuje výchozí seznam. 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:

- Pokud existuje pravidlo `accept` pro typ *typ*, včetně výchozího seznamu, pokud byla zadána volba `accept DEFAULT`, bude tento typ přijat.
- Není-li splněna předchozí podmínka a existuje-li pro typ *typ* pravidlo `reject`, nebude tento typ přijat.

- Nejsou-li splněny předchozí podmínky a bylo-li zadáno pravidlo `accept ALL`, bude tento typ přijat.
- Není-li splněna žádná z předchozích podmínek, typ nebude přijat.

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. Jestliže typ dokumentu není přijat, probíhají další kroky takto:

- Pokud pro typ `typ` existuje nativní pravidlo (včetně výchozích pravidel analýzy, jestliže bylo zadáno pravidlo `native DEFAULT`), dokument je 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).
- V opačném případě je dokument odmítnut a nebude analyzován.

Postup

Chcete-li přiřadit typy dokumentů k relaci Stellent, postupujte takto:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind 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 kolekcí:

```
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 uložte 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.

Příklad

Při použití následujícího konfiguračního souboru relace Stellent přijímá kromě položek výchozího seznamu také formát Microsoft Visio.

```
accept DEFAULT
accept FI_VISIO3 visio
accept FI_VISIO4 visio
accept FI_VISIO5 visio
accept FI_VISIO6 visio
```

Při použití následujícího konfiguračního souboru budou podporovány dokumenty ve formátu Postscript a bude je možné prohledávat jako dokumenty typu `ps`, soubory formátu `X pixmap` (`XPM`) budou zasilány zpět vestavěnému textovému analyzátoru, formát obrázků `PNG` bude odmítnut 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
accept ALL other
reject FI_PNG
```

Výchozí pravidla analýzy pro relace Stellent

Nevytvoříte-li konfigurační soubor pro mapování typů souborů na filtry dokumentů relace Stellent, budou pro analýzu dokumentů použita výchozí pravidla.

Výchozí pravidla používaná relací Stellent jsou následující:

```
ACCEPT FI_WORD4 doc
ACCEPT FI_WORD5 doc
ACCEPT FI_RTF rtf
ACCEPT FI_WINWORD1 doc
ACCEPT FI_WINWORD1COMPLEX doc
ACCEPT FI_WINWORD2 doc
ACCEPT FI_WORD6 doc
ACCEPT FI_WINWORD6 doc
ACCEPT FI_ICHITAR03 jxw
ACCEPT FI_ICHITAR04 jsw
ACCEPT FI_WINWORD1J doc
ACCEPT FI_WINWORD5J doc
ACCEPT FI_RTFJ rtf
ACCEPT FI_WINWORD7 doc
ACCEPT FI_WORDPRO lwp
ACCEPT FI_WINWORD97 doc
ACCEPT FI_ICHITAR08 jtd
ACCEPT FI_WORDPRO97 lwp
ACCEPT FI_WINWORD2000 doc
ACCEPT FI_WINWORD2002 doc
ACCEPT FI_WINWORD2003 doc
ACCEPT FI_123R1 123
ACCEPT FI_123R2 123
ACCEPT FI_123R3 123
ACCEPT FI_EXCEL xls
ACCEPT FI_EXCEL3 xls
ACCEPT FI_EXCEL4 xls
ACCEPT FI_123R4 123
ACCEPT FI_EXCEL5 xls
ACCEPT FI_123R6 123
ACCEPT FI_EXCEL97 xls
ACCEPT FI_123R9 123
ACCEPT FI_EXCEL2000 xls
ACCEPT FI_EXCEL2002 xls
ACCEPT FI_EXCEL2003 xls
ACCEPT FI_FREELANCE prz
ACCEPT FI_POWERPOINT4 ppt
ACCEPT FI_POWERPOINT3 ppt
ACCEPT FI_POWERPOINT7 ppt
ACCEPT FI_FREELANCE3 prz
ACCEPT FI_POWERPOINTMAC3 ppt
ACCEPT FI_POWERPOINTMAC4 ppt
ACCEPT FI_PDF pdf
ACCEPT FI_EXTPOWERPOINT4 ppt
ACCEPT FI_EXTPOWERPOINTMAC4 ppt
ACCEPT FI_POWERPOINTMACB3 ppt
ACCEPT FI_POWERPOINTMACB4 ppt
ACCEPT FI_POWERPOINT97 ppt
ACCEPT FI_PDFMACBIN pdf
ACCEPT FI_POWERPOINT9597 ppt
ACCEPT FI_POWERPOINT2000 ppt
ACCEPT FI_POWERPOINT2 ppt

NATIVE FI_HTML htm
```

NATIVE FI_HTML_LATIN2	htm
NATIVE FI_HTML_JAPANESESJIS	htm
NATIVE FI_HTML_JAPANESEEUC	htm
NATIVE FI_HTML_CHINESEBIG5	htm
NATIVE FI_HTML_CHINESEEUC	htm
NATIVE FI_HTML_CHINESEGB	htm
NATIVE FI_HTML_KOREANHANGUL	htm
NATIVE FI_HTML_CYRILLIC1251	htm
NATIVE FI_HTML_CYRILLICKO18	htm
NATIVE FI_CYRILLIC1251	txt
NATIVE FI_CYRILLICKO18	txt
NATIVE FI_W2KHTML	htm
NATIVE FI_XL2KHTML	htm
NATIVE FI_PP2KHTML	htm
NATIVE FI_XML	xml
NATIVE FI_WML	xml
NATIVE FI_HTML_JAPANESEJIS	htm
NATIVE FI_WML_CHINESEBIG5	xml
NATIVE FI_WML_CHINESEEUC	xml
NATIVE FI_WML_CHINESEGB	xml
NATIVE FI_WML_CYRILLIC1251	xml
NATIVE FI_WML_CYRILLICKO18	xml
NATIVE FI_WML_JAPANESEJIS	xml
NATIVE FI_WML_JAPANESESJIS	xml
NATIVE FI_WML_JAPANESEEUC	xml
NATIVE FI_WML_KOREANHANGUL	xml
NATIVE FI_WML_LATIN2	xml
NATIVE FI_HTMLUNICODE	htm
NATIVE FI_XML_DOCTYPE_HTML	htm
NATIVE FI_XHTML	htm
NATIVE FI_ASCII	txt
NATIVE FI_ANSI	txt
NATIVE FI_UNICODE	txt
NATIVE FI_ASCII8	txt
NATIVE FI_ANSI8	txt
NATIVE FI_MAC	txt
NATIVE FI_MAC8	txt
NATIVE FI_SHIFTJIS	txt
NATIVE FI_CHINESEGB	txt
NATIVE FI_HANGEUL	txt
NATIVE FI_CHINESEBIG5	txt
NATIVE FI_LATIN2	txt
NATIVE FI_JAPANESE_EUC	txt
NATIVE FI_HEBREW_OLDCODE	txt
NATIVE FI_HEBREW_PC8	txt
NATIVE FI_HEBREW_E0	txt
NATIVE FI_HEBREW_WINDOWS	txt
NATIVE FI_ARABIC_710	txt
NATIVE FI_ARABIC_720	txt
NATIVE FI_ARABIC_WINDOWS	txt
NATIVE FI_7BITTEXT	txt
NATIVE FI_JAPANESE_JIS	txt
NATIVE FI_CENTRALEU_1250	txt
NATIVE FI_UTF8	txt
NATIVE FI_EBCDIC_37	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_500	txt
NATIVE FI_EBCDIC_870	txt
NATIVE FI_EBCDIC_871	txt
NATIVE FI_EBCDIC_1026	txt

Správa indexu podnikového vyhledávání

Aby měli uživatelé zajištěn trvalý přístup k nejaktuálnějším informacím, podnikové vyhledávání vytváří pro každou kolekci index a udržuje jej pravidelnou aktualizací a reorganizováním 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. Když prolézací moduly zpracovávají nové a změněné zdroje dat, podnikové vyhledávání aktualizuje index s použitím nového obsahu. Aktualizovaný obsah je po určité době nutné sloučit se základním indexem. Tento proces sloučení se nazývá reorganizace. Při každé aktualizaci a reorganizaci 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 indexu umožníte uživatelům vyhledávat v nejaktuálnějších datech. Průběžně aktualizovaný index je nutné po určité době reorganizovat. Zvětšující se index spotřebovává stále více systémových prostředků. Chcete-li tedy udržovat optimální výkon, měli byste proto indexy reorganizovat pravidelně.

Frekvence reorganizací 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 reorganizaci indexu),
- četnost, s jakou je prolézací modul spouštěn ručně a nikoli v naplánovanou dobu,
- četnost přidávání a odebírání identifikátorů URI externími prolézacími moduly (tyto typy prolézacích modulů komunikují s podnikovým vyhledávacím prostřednictvím rozhraní API modulu pro příjem dat).

U kolekcí obsahujících několik miliónů dokumentů, sestavovaných zejména z webových dokumentů, by měl být index reorganizován přibližně jednou denně a aktualizován každou hodinu nebo každé dvě hodiny.

Při udržování indexu ve stále aktuálním stavu umožňujícím vyhledávání se používají následující postupy:

- nastavení plánu aktualizace a reorganizace 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 218
Index pro kolekci monitorujete, jestliže potřebujete zobrazit průběh zpracování sestavovaného indexu, povolit nebo zakázat časový plán indexu nebo spustit a zastavit indexování.

“Monitorování fronty indexů podnikového vyhledávání” na stránce 219
Můžete zobrazit stav všech sestavení indexu ve frontě indexů, zastavit sestavování indexu nebo odstranit index z fronty.

Plánování sestavení indexu

Pro reorganizaci 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 pravidelnou reorganizaci a aktualizaci indexu. Při reorganizaci je celý index sestaven znovu. Procesy indexování načtou všechna data shromážděná prolézacími moduly a analyzovaná modulem analýzy. Při aktualizaci indexu jsou pro vyhledávání zpřístupněny informace získané prolézáním od poslední reorganizace indexu.

Při výchozím nastavení je volba plánování sestavení indexu vybrána. Tato volba určuje, že proces plánovače naplánuje úlohy aktualizace a reorganizace indexu při každém 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.

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 novým obsahem, určete na stránce Časový plán sestavení indexu v oblasti **Zadejte plán aktualizace indexu** následující volby:
 - a. V oblasti **Čas spuštění** zadejte do polí **Rok, Měsíc, Den, Hodina** a **Minuta** datum a čas první aktualizace indexu.
 - b. V oblasti **Interval aktualizace** určete prostřednictvím **dní, hodin** a **minut** frekvenci aktualizací indexu.
Obvykle je vhodné index aktualizovat 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 **Zadejte plán reorganizace indexu** následující volby:
 - a. V oblasti **Čas spuštění** zadejte do polí **Rok, Měsíc, Den, Hodina** a **Minuta** datum a čas první reorganizace indexu.
 - b. V oblasti **Interval aktualizace** určete prostřednictvím **dní, hodin** a **minut** frekvenci reorganizací indexu.

Obvykle je vhodné index reorganizovat 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 reorganizace nebo aktualizace 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 s použitím nového obsahu a jak často má být reorganizován.
2. Klepněte na tlačítko **Použít**.

Povolení a zákaz časových plánů indexování

Časové plány aktualizace a reorganizace 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. Chcete-li povolit nebo zakázat plán reorganizace indexu, zaškrtněte políčko **Povolit při spuštění systému** nebo jeho zaškrtnutí zrušte.
 - 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 .
- c. Je-li index naplánován, ale časový plán pro sestavení je zakázán, klepněte na ikonu  **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 pro kolekci souběžně s reorganizací.

Než začnete

Chcete-li pro systém určit volby sestavování indexu, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

O této úloze


Systém 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 reorganizace 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 kolekce mají být zpracovávány souběžně s reorganizací indexu kolekce. Povolíte-li tuto volbu, budou na vyhledávacích serverech aktualizovány údaje týkající se nejnovějších dokumentů (prostřednictvím aktualizace indexu), zatímco bude zpracovávána reorganizace indexu, která 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.

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 zpracování požadavků na aktualizaci a reorganizaci indexu.
6. Klepněte na tlačítko **OK**.

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.

Volby vytváření indexu pro prohledávání dokumentů

Při konfigurování voleb pro prohledávání prolezených dat nebo externích zdrojů nebo při mapování prvků XML a prvků metadat HTML na vyhledávací pole můžete určit způsob prohledávání dokumentů a jejich zobrazování ve výsledcích vyhledávání.

Při úpravách voleb pro prolézací moduly obsahující 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
- Parametrické vyhledávání
- Výsledky vyhledávání
- Obsah dokumentu

Při úpravách voleb pro externí zdroje není volba označení pole jako pole obsahu dokumentu k dispozici.

Při konfigurování analyzátoru a určování, že mají být prvky XML a prvky metadat HTML mapovány na pole v indexu, můžete určit, zda lze příslušná pole používat při vyhledávání s použitím polí, při parametrickém vyhledávání nebo ve výsledcích vyhledávání.

Volby, které určíte, jsou uloženy spolu s dokumenty v indexu. Umožňují vám omezit cíle dotazů uživatelů a položky, které se uživatelům zobrazí ve výsledcích vyhledávání.

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.

Příklad:

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:

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 pole a metadata dokumentu. Můžete například určit, že zadaná slova se musí vyskytnout v názvu dokumentu.

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.

Příklad:

Následující dotaz hledá dokumenty obsahující slovo ibm a frázi podnikové vyhledávání v poli názvu:

title:ibm title:"podnikové vyhledávání"

Chcete-li prohledávat pole podle jejich názvu, musíte pro tato pole při konfigurování prolézacího modulu povolit vyhledávání s použitím polí.

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.

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
```

Chcete-li prohledávat pole pomocí parametrických dotazů, musíte pro tato pole při konfigurování prolézacího modulu povolit parametrické vyhledávání.

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.

Obsah dokumentu

Obsah pole označeného jako pole obsahu dokumentu lze použít k asociování dokumentů s kategorií založenou na modelu a s kategoriemi určujícími pravidla založená na obsahu dokumentů.

Související pojmy

"Syntaxe dotazů" v příručce "Programming Guide and API Reference for Enterprise Search"

"Vyhledávací aplikace podnikového vyhledávání" na stránce 149

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.

Zástupné znaky v dotazech

V případě potřeby můžete uživatelům 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 uživatelům 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 ve jménu 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.

Jakékoli změny nastavení pro práci se zástupnými znaky se projeví po příští reorganizaci 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***.

Při aktualizaci a reorganizaci 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 specifikátoru pole (příklad: `jmenopole:*sphere`). Jméno pole nesmí obsahovat dvojtečku (:).
- Je-li podpora zástupných znaků omezena na sadu polí, musí být jméno 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 jména prvků XML, jména atributů či hodnoty atributů. Výraz sestávající pouze ze zástupného znaku není podporován.

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 příští reorganizaci 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. Zaškrtnutím nebo zrušením zaškrtnutí políčka **Podpora zástupných znaků v dotazech, které nejsou určeny pro prohledávání polí** můžete určit, zda mají být zástupné znaky podporovány v dotazech vyhledávajících text, který není obsažen v polích. 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 povolit podporu dotazů obsahujících zástupné znaky pro všechna pole v dokumentu, vyberte volbu **Všetchna pole**.
 - Chcete-li omezit podporu pro zástupné znaky na některá pole, vyberte volbu **Specifická pole** a zadejte jména 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. Podrobnosti můžete zobrazit klepnutím na volbu **Nápověda** během konfigurování voleb pro zástupné znaky.
6. Klepněte na tlačítko **OK**.

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.

Při konfiguraci oborů bude pravděpodobně nutné dvakrát reorganizovat index, než se provedené změny projeví. Pokud konfigurujete obory před první reorganizací indexu pro kolekci, uživatelé budou moci prohledávat kolekci, ale nebudou moci ve výsledcích vyhledávání zobrazit údaje o oboru. Opakovanou reorganizací indexu zajistíte, že výsledky vyhledávání budou odpovídat rozsahu identifikátorů URI v daném oboru.

Pokud konfigurujete obory a index byl předtím alespoň jednou reorganizován, provedené změny se projeví až po příští reorganizaci 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 náleží.

O této úloze

Pokud vyhledávací aplikace podporuje použití oborů, uživatelé mohou při prohledávání kolekce vyhledávat pouze dokumenty odpovídající identifikátorům URI, které definují hranice oboru.

Při konfiguraci oborů bude pravděpodobně nutné dvakrát reorganizovat index, než se provedené změny projeví. Pokud konfiguruje obory před první reorganizací indexu pro kolekci, uživatelé budou moci prohledávat kolekci, ale nebudou moci ve výsledcích vyhledávání zobrazit údaje o oboru. Opakovanou reorganizací indexu zajistíte, že výsledky vyhledávání budou odpovídat rozsahu identifikátorů URI v daném oboru.

Pokud konfiguruje obory a index byl předtím alespoň jednou reorganizován, provedené změny se projeví až po příští reorganizaci 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 jméno 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 náleží do dané kolekce.

Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 76
Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

Sbalené identifikátory URI

V prostředí podnikového vyhledávání lze výsledky vyhledávání uspořádat tak, že dokumenty z webu a ze zdrojů NNTP se stejnou předponou identifikátoru URI budou ve výsledcích vyhledávání zobrazeny jako skupina a budou 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 ke jménu 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 jméno 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 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 určená předpona identifikátoru URI.

Uspořádání předpon identifikátorů URI a jmen 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 ke jménu 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 jméno skupiny (nebo předponu identifikátoru URI, pokud pravidlo neurčuje jméno skupiny) k výrazům vyhledávaným v dokumentech.
3. Pokud dokument neodpovídá žádné předponě identifikátoru URI:
 - U webových identifikátorů URI použije indexový server jako předponu identifikátoru URI jméno hostitele odpovídající dané adrese URL.
 - U identifikátorů URI protokolu NNTP použije indexový server 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 ke jménu skupiny, můžete tuto předponu identifikátoru URI individuálně vybrat a přesunout v seznamu nahoru nebo dolů.
- Pokud přidanou předponu identifikátoru URI přiřadíte ke jménu 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 z webu a ze zdrojů NNTP, které mají stejnou předponu identifikátoru URI. Můžete rovněž vytvořit jméno skupiny, které 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 provedené za účelem sbalení výsledků vyhledávání se projeví až po příští reorganizaci 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/
4. Můžete zadat také popisné jméno 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 stejné jméno 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 se jménem 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 se jménem 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 jméno skupiny, vyberte předponu identifikátoru URI a klepněte na volbu  **Upravit**.
8. Chcete-li předponu identifikátoru URI odebrat ze seznamu, vyberte předponu 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 po příští aktualizaci nebo reorganizaci indexu.

Zadáte-li vzor pro odebrání více identifikátorů URI, uživatelům se budou identifikátory URI, které odpovídají vzoru, nadále zobrazovat ve výsledcích vyhledávání až do příští aktualizace nebo reorganizace indexu.

Odeberete-li identifikátor URI z indexu, neodeberete jej z prolézáného 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 prolézáného 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*
```

3. Klepněte na tlačítko **OK**.

Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 76
Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidal do kolekce.

Správa vyhledávacího serveru pro podnikové vyhledávání

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 dotazu, můžete určit, že výsledné dokumenty mají být ukládány do mezipaměti, a konfigurovat velikost prostoru, který má být dokumentům 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

“Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

“Vlastní slovníky slov pro upřednostnění” na stránce 138

Chcete-li zvýšit kvalitu výsledků vyhledávání, můžete ovlivnit ohodnocení důležitosti dokumentů ve výsledcích vyhledávání vytvořením vlastního slovníku slov pro upřednostnění.

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 odpovídající výsledné dokumenty, mohou uživateli rychle vrátit výsledky vyhledávání. Pokud vyhledávací servery vhodné výsledné dokumenty nenaleznou, prohledají index.

Po zaplnění mezipaměti vyhledávání jsou odstraňovány nejstarší výsledné dokumenty a výsledné dokumenty odpovídající méně často zadávaným dotazům, aby uvolnily místo novým výsledkům vyhledávání.

Na konzole pro správu podnikového vyhledávání můžete povolit ukládání výsledků vyhledávání do mezipaměti a také zadat kapacitu mezipaměti (počet dotazů, jejichž výsledky lze do mezipaměti současně uložit).

Pokud změníte nastavení mezipaměti pro vyhledávání, začnou změny platit 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 sad výsledků vyhledávání, 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 dotaz 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, seznamům, 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*, *and 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. Výsledný soubor XML je nutné převést do binárního formátu (soubor .dic) pomocí nástroje dodávaného s produktem WebSphere II OmniFind Edition.

Administrátor podnikového vyhledávání odešle binární soubor do systému a přiřadí mu zobrazované jméno. 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.

Související pojmy

"Podpora synonym ve vyhledávacích aplikacích" v dokumentu "Text Analysis Integration"

Související úlohy

"Vytvoření souboru XML se synonymy" v dokumentu "Text Analysis Integration"

"Vytvoření slovníku synonym" v dokumentu "Text Analysis Integration"

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.

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 jedinečné zobrazované jméno pro slovník synonym a případně zadejte popis.
6. Určete umístění souboru .dic. Je-li soubor menší než 8 MB, může být soubor umístěn ve vašem systému. Je-li soubor větší než 8 MB, musí být umístěn na indexovém serveru.
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 **Jméno 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).

Produkt WebSphere II OmniFind Edition při výchozím nastavení provádí rozpoznávání zakázaných slov specifických pro konkrétní jazyk. 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.

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. Výsledný soubor XML je nutné převést do binárního formátu (soubor .dic) pomocí nástroje dodávaného s produktem WebSphere II OmniFind Edition.

Administrátor podnikového vyhledávání odešle binární soubor do systému a přiřadí mu zobrazované jméno. Administrátoři kolekci 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.

Související pojmy

"Vlastní slovníky zakázaných slov" v dokumentu "Text Analysis Integration"

Související úlohy

"Vytvoření souboru XML se zakázanými slovy" v dokumentu "Text Analysis Integration"

"Vytvoření slovníku zakázaných slov" v dokumentu "Text Analysis Integration"

Přidávání slovníků zakázaných slov do systému


Pokud vytvoříte vlastní slovníky zakázaných slov pro odebrání slov z dotazů, je nutné asociovat je se systémem 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.

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 jedinečné zobrazované jméno pro slovník.
6. Určete umístění souboru .dic. Je-li soubor menší než 8 MB, může být soubor umístěn ve vašem systému. Je-li soubor větší než 8 MB, musí být umístěn na indexovém serveru.
7. Klepněte na tlačítko **OK**. Vlastní slovník zakázaných slov bude přidán do systému podnikového vyhledávání a bude k dispozici pro prohledávání kolekcí.

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 **Jméno 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**.

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 souhrných 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 projevíly, 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, které může věta s popisem obsahovat. Výchozí hodnota je 4.

MaxWordsPerSentence

Maximální počet slov, které může věta s popisem obsahovat. Výchozí hodnota je 20.

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. Přihlašte se k vyhledávacím serverům jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind 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 vrácené 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 vrácení 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 76
Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

Hodnocení důležitosti dokumentů v podnikovém vyhledávání

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 142
Mapováním polí na třídy upřednostnění můžete ovlivnit řazení dokumentů ve výsledcích vyhledávání.

“Hodnocení důležitosti dokumentů na základě vzorů identifikátoru URI” na stránce 140
Přiřazením faktorů upřednostnění ke vzorům identifikátorů URI můžete zvýšit nebo snížit důležitost dokumentů.

“Vlastní slovníky slov pro upřednostnění” na stránce 138
Chcete-li zvýšit kvalitu výsledků vyhledávání, můžete ovlivnit ohodnocení důležitosti dokumentů ve výsledcích vyhledávání vytvořením vlastního slovníku slov pro upřednostnění.

Hodnocení na základě textu

Podnikové vyhledávání dynamicky vypočítává hodnotící skóre pro každý dokument odpovídající výrazům v dotazu.

Při určování textového skóre všech dokumentů vyhovujících podmínkám dotazu pracuje podnikové vyhledávání s mnoha faktory, k nimž patří například:

- Četnost výskytu jednotlivých dotazovacích výrazů v celé kolekci. Obecně platí, že dotazovací výrazy, které se vyskytují ve většině dokumentů, přispívají k celkovému hodnocení dokumentu méně než výrazy obsažené pouze v užší skupině dokumentů.
- Počet výskytů jednotlivých dotazovacích výrazů v daném dokumentu. Obecně platí, že vyšší počet výskytů dotazovacího výrazu v dokumentu znamená vyšší hodnocení tohoto dokumentu.
- Vzájemná vzdálenost výskytu jednotlivých dotazovacích výrazů v každém nalezeném dokumentu. Obecně platí, že dotazovací výrazy, které se v dokumentu vyskytují blízko sebe, přispívají k celkovému hodnocení dokumentu více než výrazy, mezi jejichž výskyty v dokumentu je vzdálenost větší.
- Kontext výskytu dotazovacích výrazů v každém nalezeném dokumentu. Dotazovací výraz nalezený například v názvu dokumentu přispívá k celkovému hodnocení tohoto dokumentu více než tentýž výraz nalezený v těle dokumentu.

Dalšími faktory zahrnovanými do hodnocení dokumentu je například délka dokumentu a bohatost použité slovní zásoby.

Statické řazení

K některým dokumentům lze přiřadit statický faktor řazení, který zvyšuje důležitost těchto dokumentů uváděnou ve výsledcích vyhledávání.

Při vytvoření kolekce určíte, zda chcete dokumentům v této kolekci přiřadit statický faktor řazení. V případě webového obsahu může být důležitost dokumentu ve výsledcích vyhledávání zvyšována počtem odkazů na tento dokument z jiných dokumentů a původem těchto odkazů.

U dokumentů zahrnujících kalendářní pole či metadata můžete ke zvýšení významnosti dokumentu použít jeho datum. 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 pocházejících z daného zdroje dat.

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. Kombinováním zdrojů s různými modely řazení do jedné kolekce může být snížena kvalita vyhledávání.

Dále je třeba zkontrolovat, zda dokumenty v kolekci obsahují pole a hodnoty, pro něž má být statické řazení použito. Použijete-li jako faktor řazení dokumentů například datum dokumentu a dokumenty v kolekci přitom neobsahují kalendářní pole či atributy, může to rovněž snížit kvalitu vyhledávání.

Související úlohy

“Migrace kolekce z produktu WebSphere Portal” na stránce 197

Chcete-li migrovat kolekce z produktu WebSphere Portal do prostředí podnikového vyhledávání, nejprve je připravte v produktu WebSphere Portal a poté proveďte jejich migraci pomocí průvodce migrací.

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.

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. (K ohodnocení důležitosti dokumentu ve výsledcích vyhledávání přispívají pouze výrazy dotazu, které jsou upřednostněny, a faktor upřednostnění je pouze jedním faktorem, který ovlivňuje 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 jméno pole. (Chcete-li použít faktory upřednostnění pro výrazy dotazu vyskytující se v polích, můžete mapovat jména 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. Výsledný soubor XML je nutné převést do binárního formátu (soubor `.dic`) pomocí nástroje dodávaného s produktem WebSphere II OmniFind Edition.

Administrátor podnikového vyhledávání odešle binární soubor do systému a přiřadí mu zobrazované jméno. Administrátoři kolekci 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.

Související pojmy

“Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

“Vlastní slovníky slov pro upřednostnění” v dokumentu “Text Analysis Integration”

Související úlohy

“Vytvoření souboru XML se slovy pro upřednostnění” v dokumentu “Text Analysis Integration”

“Vytvoření slovníku slov pro upřednostnění” v dokumentu “Text Analysis Integration”

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.

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 jedinečné zobrazované jméno pro slovník a případně zadejte popis.
6. Určete umístění souboru .dic. Je-li soubor menší než 8 MB, může být soubor umístěn ve vašem systému. Je-li soubor větší než 8 MB, musí být umístěn na indexovém serveru.
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 **Jméno 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 tříd 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.

Související pojmy

“Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

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, reorganizujte index.

Související pojmy

“Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 76

Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

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

Při mapování pole na třídu upřednostnění je nutné určit, zda má být pole použito pro zjišťování duplicitních dokumentů:

- Používá-li se pole k detekci duplicitních dokumentů, je považováno za pole s obsahem; vybírat lze pouze třídy upřednostnění určené pro pole s obsahem.
- 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.

Vysoké a nízké hodnoty návratnosti

Při vyhodnocování dotazu procesy vyhledávání odhadují počet výsledných dokumentů, které budou vráceny. Prahové hodnoty určují, zda je dotaz považován za dotaz s nízkou nebo vysokou hodnotou návratnosti:

Nízká návratnost

Pokud je odhadovaný počet výsledných dokumentů menší než dolní prahová hodnota, je dotaz považován za dotaz s nízkou návratností.

Vysoká návratnost

Pokud je odhadovaný počet výsledných dokumentů větší než horní prahová hodnota, je dotaz považován za dotaz s vysokou návratností.

Kombinace

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 pětinašobně ve srovnání s výrazy dotazu, které se vyskytnou v běžném textu.

Související pojmy

“Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

Mapování polí na třídy upřednostnění

Mapováním jmen 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.

V podnikovém vyhledávání jsou některá jména polí vyhrazena, aby bylo možné počítat skóre pro pole a text bez jakékoli dalších definující charakteristiky (například pro pole, které není polem názvu, a běžný text, který není zvýrazněn). 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 do třídy upřednostnění zadejte jméno pole, které chcete mapovat na třídu upřednostnění.
Můžete zadat jméno pole, které existuje v prolézaném zdroji nebo v externím zdroji, jméno pole, které je mapováno z prvku XML, jméno pole, které je mapováno z prvku metadat HTML, nebo jedno z předdefinovaných jmen 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. Zrušíte-li zaškrtnutí tohoto políčka, bude seznam tříd upřednostnění k dispozici obsahovat třídy vztahující se k polím metadat.
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 projevily změny pro dokumenty indexované dříve, musíte tyto dokumenty prolézat, analyzovat a indexovat znovu.

Související pojmy

“Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

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 projevily, 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

“Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137

Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším hodnocením podle podmínek uvedených v dotazu.

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
```

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 uvedena jména tříd upřednostnění, výchozí faktory upřednostnění pro dotazy s nízkou hodnotou návratnosti, výchozí faktory upřednostnění pro dotazy s vysokou hodnotou návratnosti a jména 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 3. Výchozí hodnoty tříd upřednostnění

Výchozí nízké a vysoké faktory upřednostnění				
Jméno 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	
Třída metadat A	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	Nízký: 1 Vysoký: 1	es_special_field.default_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
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

| “Hodnocení důležitosti dokumentů v podnikovém vyhledávání” na stránce 137
| Když uživatel prohledává určitou kolekci, vyhledávací procesy vrátí výsledky s nejvyšším
| hodnocením podle podmínek uvedených v dotazu.

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.

Související pojmy

“Volby vytváření indexu pro prohledávání dokumentů” na stránce 117

Při konfigurování voleb pro prohledávání prolezených dat nebo externích zdrojů nebo při mapování prvků XML a prvků metadat HTML na vyhledávací pole můžete určit způsob prohledávání dokumentů a jejich zobrazování ve výsledcích vyhledávání.

“Zabezpečení pomocí ID vyhledávacích aplikací” na stránce 176

Chcete-li zajistit zabezpečení na úrovni kolekcí, můžete určit, které vyhledávací aplikace mohou jednotlivé kolekce a externí zdroje prohledávat.

“Zabezpečení na úrovni dokumentu” na stránce 177

Pokud je při vytvoření kolekce povoleno její zabezpečení, lze konfigurovat ovládací prvky zabezpečení na úrovni dokumentů. Zabezpečení na úrovni dokumentů zaručuje, že uživatelé prohledávající kolekce budou mít přístup pouze k dokumentům, k jejichž prohlížení mají oprávnění.

“Zabezpečení na úrovni dokumentů při použití portálového vyhledávacího stroje” na stránce 189

Pomocí vyhledávacího stroje produktu IBM WebSphere Portal můžete při prohledávání kolekcí podnikového vyhledávání uživateli vynutit zabezpečení na úrovni dokumentů.

“Vyhledávací a indexové rozhraní API - přehled” v příručce “Programming Guide and API Reference for Enterprise Search”

“Syntaxe dotazů” v příručce “Programming Guide and API Reference for Enterprise Search”

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 jméno vyhledávací aplikace.
5. Vyberte kolekce, které může aplikace prohledávat:
 - Chcete-li, aby vyhledávací aplikace přistupovala ke všem kolekcím přidaným 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 jmen kolekcí a jmen 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 jednu, několik nebo 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 v dokumentech 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 vraceny 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 jména 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. Zbýlých 98 dokumentů lze zobrazit výběrem volby zobrazení dalších dokumentů ze stejného zdroje.
- Řazení dokumentů podle důležitosti nebo vzestupně či sestupně podle data dokumentu.
- 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 po klepnutí na identifikátor URI příslušného dokumentu. Jsou-li k dispozici klientské aplikace, může být výsledný dokument rovněž zobrazen pomocí nativního prohlížeče. Je-li například instalována klientská aplikace Notes, můžete jejím prostřednictvím zobrazovat dokumenty z databáze Lotus Notes.
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 ukázkové vyhledávací aplikace

Úpravou souboru config.properties pro ukázkovou vyhledávací aplikaci můžete určit volby pro příslušné prostředí, změnit vzhled aplikace a řídit volby, které mají uživatelé k dispozici po spuštění vyhledávací aplikace.

Parametry prostředí

Můžete určit volby, které řídí činnost vyhledávací aplikace.

applicationName

Určuje jméno platné vyhledávací aplikace. Výchozí hodnota je Default.

Chcete-li používat jinou než výchozí vyhledávací aplikaci, změňte výchozí hodnotu.

Tip: Je-li použito jméno aplikace Default, můžete pomocí ukázkové vyhledávací aplikace prohledávat všechny kolekce a externí zdroje jediným dotazem.

hostname

Určuje úplné jméno 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 úplné jméno hostitele, pro jehož použití je konfigurován produkt WebSphere Application Server. Je-li například hostitelský název počítače MyMachine a webový server je konfigurován pro použití názvu hostitele www.ibm.com, zadejte hodnotu www.ibm.com.

port

Určuje číslo portu webového serveru konfigurovaného pro podporu použité instance produktu WebSphere Application Server. Výchozí hodnota je 80. Tuto hodnotu změňte jen tehdy, jestliže jste změnil port u instance webového serveru.

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.

username

Určuje jméno uživatele, které umožňuje systému podnikového vyhledávání vynucovat ověřování uživatelů a ověřovat uživatele na serveru WebSphere Application Server při odeslání požadavků na vyhledávání uživateli. Toto pole je používáno pouze v případě, že jste pro produkt WebSphere Application Server povolili globální zabezpečení.

password

Určuje heslo odpovídající určenému jménu uživatele.

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.

INFO Zaznamenávány jsou informační zprávy.

FINE Zaznamenávány jsou chybové zprávy.

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 WebSphere Information Integrator OmniFind Edition:

**documentSource.vbr.icon**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Content Edition. Výchozí ikona je /images/sourceVBR.gif.

**documentSource.db2.icon**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem DB2. Výchozí ikona je /images/sourceDB2.gif.

**documentSource.cm.icon**

Určuje jméno 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 jméno 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 jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Exchange Server. Výchozí ikona je /images/sourceExchange.gif.

**documentSource.nntp.icon**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem NNTP. Výchozí ikona je /images/sourceNNTP.gif.

**documentSource.notes.icon**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem Notes. Výchozí ikona je /images/sourceNotes.gif.

**documentSource.quickplace.icon**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem QuickPlace. Výchozí ikona je /images/sourceWorkplace.gif.

**documentSource.unixfs.icon**

Určuje jméno 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 jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen webovým prolézacím modulem. Výchozí ikona je /images/sourceWeb.gif.

**documentSource.wps.icon**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že byl dokument prolezen prolézacím modulem WebSphere Portal. Výchozí ikona je /images/sourceWPS.gif.



documentSource.winfs.icon

Určuje jméno 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 jméno 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 jméno 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. Je-li k dispozici klientská aplikace nativní pro typ výsledného dokumentu, lze dokument rovněž zobrazit v prohlížeči klienta. Můžete přizpůsobit obrázky reprezentující typ prohlížeče klienta používaného pro zobrazení dokumentu. V následujícím příkladu je ikona Lotus Notes použita pro indikaci, že příslušný dokument může být zobrazen pomocí klientské aplikace Lotus Notes:

```
client.notes.icon=/images/notes.gif
```

Ve výsledcích vyhledávání jsou ikona a odkaz na klientskou aplikaci 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_jmen_poli_oddělených_mezerou

Význam znaku dvojtečka (:) v předponě identifikátoru URI musíte změnit zadáním znaku obrácené lomítka (\) před tento znak. Chcete-li v seznamu jmen polí pokračovat na dalším řádku, ukončete předchozí řádek znakem obrácené lomítka (\).
Příklad:

```
fields.db2\://=databasename tablename
fields.domino\://=servername databasename databasetitle filename extension \
createddate modifieddate
fields.cm\://=servername itemtypename createddate modifieddate mimetype
fields.file\://=directory filename extension modifieddate filesize title
```

Ikony polí

Pro typy zdrojů dat a dokumenty obsahující pole můžete přizpůsobit obrázky reprezentující pole. V souboru config.properties jsou předdefinovány následující ikony polí.



field.icon.databasetitle

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že pole obsahuje název dokumentu. Výchozí ikona je /images/notesdb.gif.

**field.icon.databasename**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že pole obsahuje jméno databáze, do níž příslušný dokument náleží. Výchozí ikona je /images/db2.gif.

**field.icon.tablename**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že pole obsahuje jméno tabulky, do níž příslušný dokument náleží. Výchozí ikona je /images/table.gif.

**field.icon.directory**

Určuje jméno souboru obrázku (včetně cesty), který indikuje, že pole obsahuje jméno adresáře, do nějž příslušný dokument náleží. Výchozí ikona je /images/closedFolder.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 jméno 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. Jména polí zadaná v této části jsou ve výsledcích vyhledávání formátována jako data určující datum. Formát data odpovídá nastavením lokality ve webovém prohlížeči.

date.fields=seznam_jmen_polí_oddělených_čárkami

Chcete-li v seznamu jmen 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 smysluplnějšími 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 jméno 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.

titles.clean=seznam_názevů_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, IBM Software Group Presentation Template, \
bez názvu, Dokument beze jména, 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 pokud byla prostřednictvím konzoly pro správu podnikového vyhledávání konfigurována pravidla pro sbalení, budou identifikátory URI odpovídající pravidlu pro předpony identifikátorů URI ve výsledcích vyhledávání sbaleny.

preferences.spellCorrections=Yes

Určuje, že mají být zobrazovány navrhované opravy pravopisu, pokud uživatel odešle dotaz obsahující slovo, které pravděpodobně obsahuje chybu.

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

Řídí, zda bude zobrazeno tlačítko **Aktualizovat** a možnost aktualizace vyhledávací aplikace. Nastavíte-li tuto volbu na hodnotu true (zapnuto), budou uživatelé moci aktualizovat seznam kolekcí a externích zdrojů, které jsou k dispozici pro prohledávání (pokud byla například vyhledávací aplikace asociována s dalšími kolekcemi nebo externími zdroji prostřednictvím konzoly pro správu podnikového vyhledávání).

Může být vhodné zobrazit tlačítko **Aktualizovat** po otestování změn provedených v souboru config.properties. Po uložení změn můžete klepnout na tlačítko **Aktualizovat** a zjistit, jak změny ovlivní vyhledávací aplikaci. Bez možnosti použití tlačítka **Aktualizovat** musíte restartovat podnikovou aplikaci ESSearchApplication na serveru WebSphere Application Server, než se změny projeví.

Pokud pro prohledávání nejsou k dispozici žádné kolekce ani externí zdroje (pokud například bylo zadáno nesprávné jméno hostitele, nebyly spuštěny vyhledávací servery nebo nebyla spuštěna podniková aplikace ESSearchServer na serveru WebSphere Application Server), bude tlačítko **Aktualizovat** zobrazeno automaticky, aby vám usnadnilo odstraňování problému.

extraMessages.show=false

Řídí zobrazování oblasti ve spodní části výsledků vyhledávání určené pro zobrazování varování a informačních zpráv. Chybové zprávy se vždy zobrazují v horní části stránky. Chcete-li zobrazovat další zprávy, nastavte tuto volbu na hodnotu true (zapnuto).

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).

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**.

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 vybrat 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.

Vlastní titulek a logo

Můžete přizpůsobit obrázky zobrazované v oblasti titulku v horní části vyhledávací aplikace. Můžete například nahradit výchozí obrázky pro produkt WebSphere II OmniFind 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 produkt WebSphere II OmniFind Edition obrázky odpovídající značce vašeho podniku. 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 jména 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 jména stránek JSP (Java Server Pages) používaných pro stránky s kartami v pohledu Vyhledávání (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 jména 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 Search tabs (Karty vyhledávání) souboru config.properties, aby byla tato volba k dispozici pouze na panelu nástrojů.

Smysluplné popisky typu dokumentů

Přehlednost použití filtru typů dokumentů můžete zvýšit mapováním skutečných jmen 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 jméno zobrazované 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 toto jméno. 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
```

Vlastní filtry

Zobrazení výsledných dokumentů můžete filtrovat pomocí vlastních dotazů, které zadáte.

filterCustom.popisek=výraz_dotazu

Určuje jméno zobrazované 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 náleží 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.

V souboru config.properties je několik vlastních filtrů označeno jako komentář a uvedeno jako příklady.

Úpravy vlastností ukázkové vyhledávací aplikace

Ukázkovou vyhledávací aplikaci pro podnikové vyhledávání lze použít k prohledávání všech aktivních kolekcí a externích zdrojů v systému. Úpravou souboru vlastností 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 serveru podnikového vyhledávání. Tato vyhledávací aplikace se konfiguruje úpravou souboru vlastností `config.properties`, který se instaluje spolu s aplikací.

Chcete-li, aby se změny projevily, musíte zastavit a znovu spustit podnikovou aplikaci `ESSearchApplication` na serveru `WebSphere Application Server`.

Postup

Chcete-li upravit vlastnosti ukázkové vyhledávací aplikace, postupujte takto:

1. 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 `WebSphere II OmniFind Edition` na vyhledávacím serveru:
`ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war/WEB-INF/config.properties`
2. 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.)
3. Zastavte aplikaci `ESSearchApplication` a znovu ji spusťte:
 - a. 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://jméno_hostitele:port/admin`, kde `jméno_hostitele` je jméno 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://jméno_hostitele:port/ibm/console`, kde `jméno_hostitele` je jméno hostitele vyhledávacího serveru a `port` je číslo portu konzoly pro správu produktu `WebSphere Application Server`. Konzola pro správu obvykle používá port 9060.
 - b. Když se zobrazí výzva k zadání jména uživatele a hesla, zadejte jméno uživatele a heslo zaregistrované v registru uživatelů produktu `WebSphere Application Server`. Pokud jako registr uživatelů používáte lokální operační systém, můžete zadat jméno a heslo administrátora podnikového vyhledávání.
 - c. Po přihlášení ke konzole pro správu klepněte na tlačítko **Applications** a poté na položku **Enterprise Applications**.
 - d. Zaškrtněte políčko u položky `ESSearchApplication` a klepněte na tlačítko **Stop**.
 - e. Po zastavení aplikace znovu zaškrtněte políčko u příslušné položky a klepněte na tlačítko **Start**.

Přístup k ukázkové vyhledávací aplikaci

Přístup k ukázkové vyhledávací aplikaci získáte zadáním adresy URL ve webovém prohlížeči.

Než začnete

Ukázkovou 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 ukázkovou vyhledávací aplikaci, postupujte takto:

1. Zadejte adresu URL vyhledávací aplikace ve webovém prohlížeči. Příklad:

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

SearchServer.com je jméno 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. Pokud je na aplikačním serveru WebSphere povoleno 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 pověřených uživatelů během zpracování dotazu, můžete konfigurovat profil uživatele. Na stránce Profil zadejte údaje pověřených potřebné pro přístup k zabezpečeným doménám. Tyto domény pak budete moci prohledávat bez zobrazování výzev k přihlášení.

3. Na stránce Vyhledávání vyberte kolekce a externí zdroje, které chcete prohledávat, a zadejte dotaz.

Povolení zabezpečení pro ukázkovou vyhledávací aplikaci

Povolíte-li globální zabezpečení na aplikačním serveru WebSphere a chcete-li použít ukázkovou vyhledávací aplikaci k prohledávání zabezpečených kolekcí, musíte změnit nastavení konfigurace v ukázkové vyhledávací aplikaci a na serveru WebSphere Application Server.

Než začnete

- Musíte mít přiřazenu roli administrátora podnikového vyhledávání.
- Musíte povolit globální zabezpečení na serveru WebSphere Application Server. Pokud jste produkt WebSphere II OmniFind Edition nainstalovali v konfiguraci s více servery, povolte globální zabezpečení na vyhledávacích serverech. Pokyny pro povolení globálního zabezpečení naleznete v dokumentaci s produktu WebSphere Application Server.
- Pokud jste jako registr uživatelů serveru WebSphere Application Server vybrali volbu LDAP (Lightweight Directory Access Protocol), musíte přidat jméno a heslo administrátora podnikového vyhledávání do registru LDAP (jedná se o jméno a heslo zadané při instalaci produktu WebSphere II OmniFind Edition).

Zní-li jméno administrátora podnikového vyhledávání například adminUser, bude mít položka uživatele v registru LDAP tvar uid=adminUser,ou=Employees,o=IBM,c=US. Pokyny naleznete v dokumentaci k serveru LDAP.

Postup

Chcete-li povolit zabezpečení pro vzorovou aplikaci podnikového vyhledávání, postupujte takto:

1. Aktualizujte vlastnosti vyhledávací aplikace na konzole pro správu:
 - a. Přihlašte se ke konzole správy podnikového vyhledávání jako uživatel s oprávněními administrátora podnikového vyhledávání.
 - b. Klepněte na tlačítko **Zabezpečení** na panelu nástrojů.
 - c. Na stránce Vyhledávací aplikace klepněte na volbu **Konfigurovat vyhledávací aplikace**.
 - d. Klepněte na volbu **Přidat vyhledávací aplikaci** a do pole **Jméno vyhledávací aplikace** zadejte jméno administrátora podnikového vyhledávání, které bylo zadáno při instalaci produktu WebSphere II OmniFind Edition.
 - e. Zkontrolujte, zda je vybrána volba **Všechny kolekce a externí zdroje**, a poté klepněte na tlačítko **OK**.
2. Upravte soubor config.properties:
 - a. Pokud používáte operační systém UNIX, otevřete okno konzoly. Pokud používáte systém Microsoft Windows, otevřete okno příkazového řádku.
 - b. Přejděte do adresáře WEB-INF ukázkové vyhledávací aplikace. Následující příklady jsou pro přehlednost uvedeny na dvou řádcích, ve skutečnosti je však nutné zadávat příkazy na jednom řádku:
UNIX:

```
cd $ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/  
ESSearchApplication.war/WEB-INF
```

Windows:

```
cd %ES_INSTALL_ROOT%\installedApps\ESSearchApplication.ear\  
ESSearchApplication.war\WEB-INF
```
 - c. Otevřete soubor config.properties v textovém editoru.
 - d. Změňte vlastnost username na jméno platného uživatele serveru WebSphere Application Server.
 - e. Změňte vlastnost password na heslo zadaného uživatele.
 - f. Uložte soubor a zavřete jej.
3. Restartujte aplikaci ESSearchApplication na serveru WebSphere Application Server:
 - a. 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://jméno_hostitele:port/admin`, kde *jméno_hostitele* je jméno 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://jméno_hostitele:port/ibm/console`, kde *jméno_hostitele* je jméno hostitele vyhledávacího serveru a *port* je číslo portu konzoly pro správu produktu WebSphere Application Server. Konzola pro správu obvykle používá port 9060.

- b. 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.
- c. Po přihlášení ke konzole pro správu klepněte na tlačítko **Applications** a poté na položku **Enterprise Applications**.
- d. Zaškrtněte políčko u položky ESSearchApplication a klepněte na tlačítko **Stop**.
- e. Po zastavení aplikace znovu zaškrtněte políčko u položky ESSearchApplication a klepněte na tlačítko **Start**.

Externí zdroje podnikového vyhledávání

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:

- Tabulky databáze JDBC (Java Database Connectivity) (pouze databáze IBM DB2 Universal Database (DB2 UDB) a Oracle). Pro každou tabulku v databázi JDBC je vytvořen samostatný externí zdroj.
- 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" v příručce "Programming Guide and API Reference for Enterprise Search"

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.

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.

Postup

Chcete-li do systému přidat externí zdroj, postupujte takto:

1. Klepnutím na volbu **Externí zdroje** otevřete zobrazení Externí zdroje.
2. Klepněte na volbu **Přidat externí zdroj**.
3. Vyberte typ externího zdroje, který chcete přidat (Server LDAP nebo Databáze JDBC).
4. Po klepnutí na tlačítko **Další** bude zahájeno konfigurování externího zdroje.

Otevře se průvodce pro typ vytvářeného zdroje. Podle pokynů v průvodci konfiguruje příslušný externí zdroj. Další informace o volbách, které můžete určit, zobrazíte klepnutím na volbu **Nápověda** na jednotlivých stránkách průvodce.

Následující výchozí jména a umístění ovladačů JDBC vám mohou usnadnit konfigurování informací o připojení pro databáze DB2 Universal Database (DB2 UDB) a Oracle:

DB2: Legacy JDBC Driver

Jméno ovladače: COM.ibm.db2.jdbc.app.DB2Driver

Umístění: *kořenový_adresář_instalace_produkту_db2/java/db2java.zip*

DB2: Universal JDBC Driver

Jméno ovladače: com.ibm.db2.jcc.DB2Driver

Umístění:

kořenový_adresář_instalace_produkту_db2/java/db2jcc.jar

kořenový_adresář_instalace_produkту_db2/java/db2jcc_license_cu.jar

Oracle

Jméno ovladače: oracle.jdbc.driver.OracleDriver

Umístění:

domovský_adresář_produkту_oracle/jdbc/lib/classes12.zip

domovský_adresář_produkту_oracle/jdbc/lib/nls_charset12.zip

5. 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" v příručce "Programming Guide and API Reference for Enterprise Search"

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 jméno 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 jmen kolekcí a jmen 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" v příručce "Programming Guide and API Reference for Enterprise Search"

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.

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 se systém pokusí spravovat. Když se uživatel přihlásí ke konzole pro správu, má k dispozici pouze funkce a kolekce, pro jejichž správu mu bylo uděleno oprávnění.

Vyhledávací aplikace mohou dále využívat podporu ověřování aplikačního serveru WebSphere k ověřování uživatelů, kteří tyto vyhledávací aplikace používají k prohledávání kolekcí.

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.

Pokud je zabezpečení na úrovni kolekce povoleno, procesy globální analýzy používají následující speciální pravidla:

- Aby bylo zajištěno vyhodnocení prvků zabezpečení pro každý dokument, jsou dokumenty s duplicitním (nebo velmi podobným) obsahem indexovány nezávisle a není pro ně vytvářena společná kanonická reprezentace.
- Fáze zpracování kotvícího 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. Cílový dokument je tak možné načíst pomocí dotazů, v nichž je uveden text obsažený ve zdrojovém dokumentu. Tento typ zpracování kotvícího textu představuje bezpečnostní riziko, pokud je uživatel oprávněn zobrazit cílový dokument, nikoli však zdrojový dokument. Pokud je povoleno zabezpečení kolekce, je proto kotvící text odkazů v zakázaných dokumentech z indexu vyloučen. Dokument je ve výsledcích vyhledávání vrácen jen tehdy, odpovídá-li dotazu jeho vlastní obsah nebo metadata.

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 kolekcí je k dispozici také pro vyhledávací aplikace prostřednictvím ID aplikací. Aby bylo možné prohledávat kolekce, musí administrátor podnikového vyhledávání přiřadit k vyhledávací aplikaci konkrétní

kolekce, které tato aplikace smí prohledávat. Poté můžete použít standardní mechanismy řízení přístupu k povolení nebo zákazu přístupu uživatelů k vyhledávací aplikaci.

Zabezpečení na úrovni dokumentu

Při konfigurování prolézacích modulů pro kolekci můžete povolit zabezpečení na úrovni dokumentů. Pokud zvolíte tuto možnost, může prolézací modul jednotlivým prolézaným modulům přiřadit prvky zabezpečení. Prvky zabezpečení se ukládají do indexu spolu s dokumenty. Pro některé prolézací moduly můžete také určit, zda mají být během zpracování dotazu ověřovány pověřovací údaje uživatele s použitím aktuálních dat řízení přístupu (dle konfigurace v nativním zdroji dat).

Vyhledávací aplikace mohou realizovat řízení přístupu pomocí prvků zabezpečení i pomocí pověřovacích údajů uživatelů. Chcete-li uživatelům umožnit prohledávat a načítat pouze dokumenty, pro které je jim uděleno přístupové oprávnění, může vyhledávací aplikace zahrnout pověřovací údaje přihlášeného uživatele do dotazů předávaných vyhledávacím serverům.

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 otevřeny žádné nadbytečné porty. Konfigurujte systém tak, aby přijímal požadavky pouze prostřednictvím portů přiřazených explicitně k aktivitám a aplikacím podnikového vyhledávání.

Administrativní role

Podnikové vyhledávání využívá koncepci rolí pro řízení přístupu k různým funkcím konzoly pro správu.

Během instalace produktu WebSphere Information Integrator OmniFind Edition (WebSphere II OmniFind Edition) instalační program nastaví 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 uživatelské jméno. 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 WebSphere II OmniFind 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 určitých kolekcí nebo všech kolekcí. 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 určitých kolekcí nebo všech kolekcí. Mohou například spouštět a zastavovat činnosti kolekce, 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 určité kolekce nebo všechny kolekce. 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.

Související úlohy

“Přihlášení ke konzole pro správu” na stránce 4

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

“Spuštění serverů podnikového vyhledávání” na stránce 201

Chcete-li uživatelům povolit vyhledávání v kolekci, je nutné spustit systémové procesy a poté spustit servery, které kolekci prolézají, analyzují, indexují a prohledávají.

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 **System**.
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.

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 a řízení přístupu.

Ověřování

Ověřování je proces, při kterém systém ověřuje, zda skutečná totožnost uživatelů odpovídá jejich deklarované totožnosti. Vzhledem k tomu, že přístup k prostředkům je obvykle založen na identitě uživatelů, kteří tyto prostředky požadují, je ověřování základem efektivního zabezpečení.

Při ověřování uživatelů, kteří se pokoušejí o přístup ke konzole pro správu, podnikové vyhledávání využívá podporu zabezpečení poskytovanou aplikačním serverem WebSphere.

K ověřování uživatelů prohledávajících kolekce podnikového vyhledávání mohou vaše vyhledávací aplikace využívat zabezpečení poskytované aplikačním serverem WebSphere a implementovat vlastní upřednostňované metody ověřování pověřovacích údajů uživatele. Pověřovací údaje uživatele obvykle tvoří jméno uživatele a heslo předávané vyhledávací aplikaci v okamžiku, kdy se uživatel přihlásí nebo kdy se pokusí o přístup k vyhledávací aplikaci.

Ověřování uživatelů lze implementovat i jinými způsoby v závislosti na prostředích a protokolech dostupných ve vašem podniku. Uživatelé mohou prokazovat svou totožnost například pomocí karet Smart Card, prostřednictvím infrastruktury digitálních certifikátů a veřejných klíčů nebo přiřazováním tiketů při přihlášení uživatelů a sledování stavu jejich ověření.

Řízení přístupu

Pojem řízení přístupu označuje omezení činností, které mohou uživatelé provádět po ověření. Nejběžnějším nástrojem omezujícím přístup k prostředkům je seznam přístupových práv (ACL). Seznam přístupových práv obsahuje identifikační údaje uživatelů (jména uživatelů, názvy skupiny, role uživatelů apod.). Ke každému takto identifikovanému uživateli je přiřazena sada oprávnění definujících možnosti a práva uživatele.

Prostřednictvím řízení přístupu lze například povolit nebo zakázat přístup k souborům na souborovém serveru a určit, zda uživatel, kterému je udělován přístup, smí soubory na tomto serveru číst, vytvářet, upravovat či odstraňovat.

Všechny aspekty řízení přístupu v prostředí podnikového vyhledávání závisí na tom, zda má uživatel oprávnění ke čtení dat v indexu. V závislosti na tom, jaké varianty řízení přístupu povolíte ve vyhledávacích aplikacích, a na pravidlech, která zadáte pro kolekce a prolézací moduly v rámci administrace podnikového vyhledávání, máte k dispozici následující možnosti:

- Povolit všem uživatelům prohledávání všech dokumentů v kolekci.
- Povolit všem uživatelům prohledávání všech dokumentů zpracovaných určitými prolézacími moduly a omezit přístup k dokumentům zpracovaným jinými prolézacími moduly.
- Povolit konkrétním uživatelům prohledávání konkrétních dokumentů. Určíte-li například databáze, které chcete zpracovávat pomocí prolézacího modulu Notes, můžete nastavit volby povolující některým uživatelům prohledávání dokumentů ve vybraných vybraným pohledech a složkách a bránící ostatním uživatelům v prohledávání dokumentů.

Zákaz zabezpečení pro podnikové aplikace na aplikačním serveru WebSphere

Chcete-li určit, které aktivity produktu WebSphere II OmniFind Edition vyžadují ověření uživatele, můžete na aplikačním serveru WebSphere zakázat globální zabezpečení pro jednotlivé podnikové aplikace.

O této úloze

Instalační program produktu WebSphere II OmniFind 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 produktu WebSphere II OmniFind Edition 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 podnikové aplikace produktu WebSphere II OmniFind Edition. 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 produktu WebSphere II OmniFind Edition.

Postup

Chcete-li zakázat zabezpečení pro některou 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://jméno_hostitele:port/admin, kde *jméno_hostitele* je jméno 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://jméno_hostitele:port/ibm/console`, kde *jméno_hostitele* je jméno 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. Znovu klepněte na tlačítko **Applications** a poté na položku **Enterprise Applications**.
- 12. Zaškrtněte políčko u položky aplikace a klepněte na tlačítko **Stop**.
- 13. Po zastavení aplikace znovu zaškrtněte políčko u příslušné položky a klepněte na tlačítko **Start**.

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 a kotvicího textu v dokumentech.

Po vytvoření vyhledávací aplikace lze pomocí ID vyhledávací aplikace určit, které kolekce může daná vyhledávací aplikace prohledávat a kteří uživatelé budou mít k této vyhledávací aplikaci přístup.

Analýza duplicitních dokumentů

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. Povoláním 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, že uživatelé prohledávají pouze

dokumenty s prvky zabezpečení odpovídajícími jejich pověřovacím údajům. Mohou například existovat dva dokumenty s velmi podobným obsahem, avšak zabezpečené pomocí odlišných seznamů přístupových práv.

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.

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` webového serveru:

```
<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 nepřirazují 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 stránce obsahující kotvicí text.

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, protože kotvicí texty zakázaných dokumentů nejsou přiřazeny k dokumentům, k nimž uživatelé přístup mají.

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 jednotlivých webových prolézacích modulů vybrat volbu určující, zda má být indexován kotvicí text v odkazech na zakázané dokumenty. Jestliže povolíte indexování kotvicího textu, procesy analýzy a indexování zahrnou do indexu kotvicí text obsažený ve všech stránkách načtených webovým prolézacím modulem. Pokud určíte, že kotvicí text nemá být indexován, bude z indexu vyloučen kotvicí text obsažený 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.

Zabezpečení pomocí ID vyhledávacích aplikací

Chcete-li zajistit zabezpečení na úrovni kolekcí, můžete určit, které vyhledávací aplikace mohou jednotlivé kolekce a externí zdroje prohledávat.

Každá vyhledávací aplikace musí rozhraní API podnikového vyhledávání předat identifikátor aplikace. Administrátor podnikového vyhledávání a vaše vyhledávací aplikace mohou tento identifikátor použít k vynucení zabezpečení pro kolekce a externí zdroje.

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.

Chcete-li používat řízení přístupu, můžete k vyhledávací aplikaci přiřadit prvky zabezpečení (například jména uživatelů, ID skupin nebo role uživatelů) a povolit přístup k aplikaci pouze příslušným uživatelům. 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 aplikace podnikového vyhledávání” na stránce 149

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.

“Vyhledávací a indexové rozhraní API - přehled” v příručce “Programming Guide and API Reference for Enterprise Search”

“Zabezpečení vyhledávacího a indexového rozhraní API” v příručce “Programming Guide and API Reference for Enterprise Search”

Zabezpečení na úrovni dokumentu

Pokud je při vytvoření kolekce povoleno její zabezpečení, lze konfigurovat ovládací prvky zabezpečení na úrovni dokumentů. Zabezpečení na úrovni dokumentů zaručuje, že uživatelé prohledávající kolekce budou mít přístup pouze k dokumentům, k jejichž prohlížení mají oprávnění.

K řízení přístupu k dokumentům v kolekci lze použít data zabezpečení získaná prolézacími moduly a uložená do indexu. Pro některé typy zdrojů dat lze také ověřovat aktuální údaje pověření uživatele po odeslání dotazu. Pro účely ověřování aktuálních pověřovacích údajů uživatele lze do vlastních vyhledávacích aplikací vestavět podporu uživatelských profilů. Uložením uživatelských profilů umožníte uživatelům přístup k dokumentům bez opakovaných výzev k zadání pověřovacích údajů.

Související pojmy

“Vyhledávací aplikace podnikového vyhledávání” na stránce 149

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.

“Zabezpečení na úrovni dokumentů při použití portálového vyhledávacího stroje” na stránce 189

Pomocí vyhledávacího stroje produktu IBM WebSphere Portal můžete při prohledávání kolekcí podnikového vyhledávání uživateli vynutit zabezpečení na úrovni dokumentů.

“Zabezpečení vyhledávacího a indexového rozhraní API” v příručce “Programming Guide and API Reference for Enterprise Search”

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.

Každý dokument je ve výchozím stavu považován za veřejný, tzn. mohou jej prohledávat všichni uživatelé. U většiny typů dokumentů lze dosáhnout zabezpečení na úrovni dokumentů přiřazením jednoho nebo více prvků zabezpečení k dokumentům a uložením těchto prvků spolu s dokumenty do indexu. Při konfigurování prolézacího modulu určujete, zda chcete omezit přístup některých uživatelů k dokumentům zpracovaným tímto prolézacím modulem.

Pokud typ zdroje dat obsahuje pole, můžete určit, že chcete k řízení přístupu použít hodnotu v jednom z těchto polí. Pokud zdroj dat pole neobsahuje, pokud nechcete pro účely zabezpečení použít hodnotu v poli nebo pokud zadané pole neobsahuje hodnotu, kterou by bylo možné použít k řízení přístupu, můžete definovat prvky zabezpečení, které prolézací modul přiřadí k dokumentům.

O prvcích zabezpečení, které má prolézací modul přiřadit k dokumentům, rozhodují administrátoři jednotlivých kolekcí. Prvek zabezpečení může reprezentovat například jméno uživatele, ID skupiny, roli uživatele nebo jakoukoli jinou hodnotu, kterou určíte jako platnou pro daný zdroj dat. Když administrátor zdroje dat aktualizuje nativní seznam přístupových práv, jsou aktualizované prvky zabezpečení zpřístupněny při nejbližší aktualizaci nebo reorganizaci indexu.

Prvky zabezpečení doprovázejí dokumenty ve všech fázích analýzy a indexování. Pokud je ve vyhledávacích aplikacích povoleno zabezpečení, můžete pro řízení přístupu k dokumentům použít prvky zabezpečení. Uživatelé prohledávající kolekci uvidí pouze ty dokumenty, které jim dovolují zobrazit jejich pověřovací údaje. Pokud pověřovací údaje uživatele nevyhovují pravidlům zabezpečení, uživatel nemůže prohledávat dokumenty chráněné prvky zabezpečení.

Na základě použití vlastních obchodních pravidel můžete určit hodnotu prvků zabezpečení zakódováním pravidel ve třídě Java. Při konfigurování vlastností prolézacího modulu můžete určit jméno 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.

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.

Při konfigurování následujících typů prolézacích modulů můžete vybrat volbu určující, zda mají být aktuální údaje pověření uživatele ověřovány porovnáním s aktuálními prvky řízení přístupu spravovanými v nativním úložišti:

- Prolézací modul Content Edition (pouze pro typy úložišť Documentum, FileNet Panagon Content Services a Portal Document Manager)
- Prolézací modul Domino Document Manager
- Prolézací modul Notes
- Prolézací modul QuickPlace
- Prolézací modul WebSphere Portal
- Prolézací modul Souborový systém Windows

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.

Tento přístup k zabezpečení na úrovni dokumentů poskytuje vyšší úroveň zabezpečení, protože pověřovací údaje uživatele jsou porovnávány s aktuálními daty zabezpečení a nikoli s daty zabezpečení uloženými v indexu. Kromě toho je přístup v tomto případě řízen

zabezpečovacími mechanismy nativního úložiště bez ohledu na to, jak složité tyto mechanismy jsou. Vzhledem k tomu, že filtrování dokumentů probíhá v reálném čase, výsledky vyhledávání odrážejí nejaktuálnější nastavení řízení přístupu pro každý z dokumentů vyhovujících kritériím hledání.

Další výhodou tohoto přístupu je skutečnost, že nijak neovlivňuje velikost indexu (v indexu není třeba vyhradit místo navíc pro prvky zabezpečení). Jelikož je však během ověřování nutná komunikace s nativními úložišti, může dojít k negativnímu ovlivnění výkonu při zpracování dotazů.

Chcete-li dosáhnout maximální úrovně zabezpečení a minimalizovat pokles výkonu při zpracování dotazů, zkombinujte volbu ukládání prvků zabezpečení do indexu s volbou ověřování aktuálních prvků řízení přístupu. Když pak uživatel odešle dotaz, bude ověřování probíhat ve dvou fázích:

- Vyhledávací servery nejprve na základě dat zabezpečení uložených v indexu rychle zjistí, zda má uživatel oprávnění pro přístup k serveru a databázi, odkud pochází prolezený dokument (index je optimalizován na rychlost a poskytuje dobu odezvy v řádu zlomků sekundy).
- Poté vyhledávací servery vytvoří dočasný seznam dokumentů existujících v doménách na serverech, k nimž má uživatel povolen přístup. Vyhledávací servery se poté na základě tohoto seznamu připojí k nativním úložištím a zjistí, zda má uživatel oprávnění k prohlížení požadovaného dokumentu.

Pokud má uživatel přístup k serveru a doméně, s největší pravděpodobností má i oprávnění k prohlížení dokumentů. Tato závěrečná fáze filtrování však zaručuje, že ve výsledcích vyhledávání budou vráceny pouze dokumenty odpovídající aktuálnímu nastavení oprávnění uživatele.

Související pojmy

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty souborového systému Windows” na stránce 181

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.

“Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino” na stránce 183

Pokud server Lotus Notes, 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.

Související úlohy

“Konfigurování důvěryhodných serverů Lotus Domino Trusted Server pro ověřování pověření uživatelů” na stránce 183

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.

Profily uživatelů a správa identit

Po vytvoření profilů uživatelů pro podnikové vyhledávání budou uživatelé moci uchovávat údaje pověření, které jim umožní prohledávání zabezpečených domén.

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* produktu WebSphere Information Integrator

OmniFind Edition mohou uživatelé uchovávat údaje pověření pro libovolný počet domén v profilu uživatele. Údaje pověření jsou šifrovány a uloženy v zabezpečené formě v systému podnikového 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** (vaše vlastní vyhledávací aplikace může tuto funkci implementovat jiným způsobem).

V profilu uživatele jsou uloženy různé údaje pověření, které musí uživatel zadávat, aby se mohl přihlašovat k doménám, které chce prohledávat. Uživatelé mohou vytvořit uživatelské profily, pokud jsou splněny všechny následující podmínky:

- Na serveru WebSphere Application Server je povoleno globální zabezpečení.
- V konzole pro správu podnikového vyhledávání je povolena správa identit produktu WebSphere II OmniFind Edition.
- Nejméně v jedné z kolekcí, které smí prohledávat vyhledávací aplikace, je povoleno zabezpečení.
- Alespoň jedna zabezpečená kolekce obsahuje dokumenty prolezené prolézacím modulem, který vynucuje řízení přístupu tím, že vyžaduje ověření aktuálních údajů pověření uživatele při odeslání dotazu.
- Při konfigurování zabezpečení na úrovni dokumentu alespoň pro jeden z následujících typů prolézacích modulů byla vybrána volba ověřování aktuálních údajů pověření během zpracování dotazů:
 - Content Edition (pro určité typy úložišť)
 - Domino Document Manager
 - Notes
 - QuickPlace
 - WebSphere Portal
 - Souborový systém Windows

V profilu jsou uvedeny všechny domény dostupné pro vyhledávací aplikaci, které vyžadují ověření údajů pověření uživatele během zpracování dotazu. Uživatelé mohou určit, pro které domény chtějí údaje pověření uložit. Pokud uživatel nezadá údaje pověření pro doménu, budou dokumenty ze zdrojů dat v této doméně vyřazeny z výsledků vyhledávání.

Nepoužíváte-li správu identit produktu WebSphere II OmniFind Edition, musí vyhledávací aplikace předat řetězec kontextu zabezpečení uživatele při dotazech uživatelů na domény vyžadující ověření aktuálních údajů pověření.

Konfigurování správy identit

Prostřednictvím správy identit produktu WebSphere II OmniFind Edition můžete uchovávat profily uživatelů. Profily umožňují uživatelům prohledávat domény vyžadující ověření údajů pověření příslušného uživatele během zpracování dotazu.

Než začnete

Chcete-li konfigurovat volby správy identit, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

O této úloze

Určíte-li při konfigurování voleb zabezpečení na úrovni dokumentu pro prolézací modul, že mají být údaje pověření uživatelů ověřovány během zpracování dotazu, musí uživatelé při

zadávaní dotazu na doménu vyžadující ověření poskytnout příslušné údaje pověření. Pomocí správy identit produktu WebSphere II OmniFind Edition mohou uživatelé vytvářet profily uživatelů a registrovat údaje pověření, které používají pro přihlašování k zabezpečeným doménám. Údaje pověření jsou šifrovány v zabezpečené databázi spravované produktem WebSphere II OmniFind Edition.

Vyhledávací servery používají uložené údaje pověření k ověření uživatele, prohledává-li příslušný uživatel zabezpečenou doménu. Pokud údaje pověření nejsou platné, budou dokumenty ze zabezpečené domény vyloučeny z výsledků vyhledávání.

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, které produktu WebSphere II OmniFind Edition umožňuje spravovat údaje pověření uživatelů v profilech uživatelů.
4. Klepněte na tlačítko **OK**.

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.

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:

JMÉNO_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 úplné jméno 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:

JMÉNO_DOMÉNY\JMÉNO_UŽIVATELE

Při přihlašování uživatelé zadávají tyto informace v následujícím formátu:

JMÉNO_UŽIVATELE@JMÉNO_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 spouštěn prolézací server, musí mít oprávnění **Jednat jako součást operačního systému**. (Toto oprávnění je konfigurováno pro uživatele, která je administrátorem podnikového vyhledávání, při instalaci produktu WebSphere Information Integrator OmniFind 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 \\jméno_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@JMÉNO_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 spouštěn prolézací server, musí mít oprávnění **Jednat jako součást operačního systému**. (Toto oprávnění je konfigurováno pro uživatele, která je administrátorem podnikového vyhledávání, při instalaci produktu WebSphere Information Integrator OmniFind Edition.)
- Uživatelé musí mít přiřazeno uživatelské oprávnění **Přihlásit jako dávkovou úlohu**.

Související pojmy

“Ověřování aktuálního pověření během zpracování dotazu” na stránce 178
Pokud je pro kolekci při jejím vytvoření povoleno zabezpečení, určité typy domén umožňují ověřit aktuální údaje pověření uživatele během odesílání dotazu příslušného uživatele.

“Prolézací moduly souborového systému Windows” na stránce 73
Chcete-li do kolekce podnikového vyhledávání zahrnout dokumenty uložené v souborových systémech Microsoft Windows, musíte konfigurovat prolézací modul Souborový systém Windows.

Prosazování zabezpečení na úrovni dokumentu pro dokumenty Lotus Domino

Pokud server Lotus Notes, 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, které byly prolezeny na serveru Lotus Notes používajícím 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 prohledává zabezpečenou kolekci, oprávněn k zobrazování dokumentů Lotus Notes vyhovujících kritériím vyhledávání. Dokumenty, k jejichž prohlížení nemá uživatel oprávnění, jsou z výsledků vyhledávání před jejich vrácením uživateli.

- “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 zaručuje, že se při pokusu o použití vyhledávací aplikace uživatelům zobrazí výzva k zadání údajů pověření. Vyhledávací servery poté mohou tyto údaje pověření použít pro ověření přístupu každého uživatele k dokumentům Lotus Notes.

Související pojmy

“Ověřování aktuálního pověření během zpracování dotazu” na stránce 178
Pokud je pro kolekci při jejím vytvoření povoleno zabezpečení, určité typy domén umožňují ověřit aktuální údaje pověření uživatele během odesílání dotazu příslušného uživatele.

“Prolézací moduly Notes” na stránce 42
Chcete-li do kolekce podnikového vyhledávání zahrnout databáze IBM Lotus Notes, musíte konfigurovat prolézací modul Notes.

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. Vyberte volbu **File** → **Open server** (Soubor - Otevřít server).
3. Zadejte jméno 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.

jméno_serveru

Zadejte jméno 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 jméno serveru nebo vybrat skupinu **OtherDomainServers**. Musíte rovněž postupovat podle pokynů pro křížovou certifikaci souboru ID serveru Domino produktu WebSphere II OmniFind Edition 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 178
Pokud je pro kolekci při jejím vytvoření povoleno zabezpečení, určité typy domén umožňují ověřit aktuální údaje pověření uživatele během odesílání dotazu příslušného uživatele.

“Prolézací moduly Notes” na stránce 42
Chcete-li do kolekce podnikového vyhledávání zahrnout databáze IBM Lotus Notes, musíte konfigurovat prolézací modul Notes.

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 konfigurovat volby 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

Při konfiguraci prolézacího modulu můžete asociovat prvky zabezpečení s prolézanými dokumenty. 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. Upravte kolekci, vyberte stránku **Obecné** a klepněte na volbu **Konfigurovat 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** 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.**

4. Klepněte na tlačítko **OK**.

Integrace podnikového vyhledávání s produktem WebSphere Portal

Vyhledávací funkce produktu IBM WebSphere Portal můžete rozšířit implementací portletů podnikového vyhledávání v produktu WebSphere Portal a ve vyhledávacím centru produktu WebSphere Portal.

Body integrace

Portlety podnikového vyhledávání jsou s produktem WebSphere Portal integrovány několika způsoby:

WebSphere Portal

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 portlet podnikového vyhledávání do produktu WebSphere Portal, můžete prohledávat kolekce podnikového vyhledávání přímo z rozhraní produktu WebSphere Portal. 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.

portálový vyhledávací stroj

Vyhledávací stroje produktu WebSphere prolézají weby. Administráční portlety umožňují administrátorům sestavovat indexované kolekce a vyhledávací portlety umožňují uživatelům tyto kolekce prohledávat.

Pokud používáte produkt WebSphere Portal ve verzi 5.0.2 nebo vyšší, můžete do prostředí podnikového vyhledávání migrovat informace o konfiguraci kolekcí a taxonomií portálového vyhledávacího stroje.

Pokud používáte produkt WebSphere Portal ve verzi 5.1 nebo vyšší, můžete prohledávat kolekce podnikového vyhledávání pomocí portletu vyhledávání dokumentů portálového vyhledávacího stroje. Konfigurační vlastnosti umožňují administrátorům snadno přepínat mezi těmito dvěma vyhledávacími funkcemi podle potřeby.

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 administrátor v centru vyhledávání zaregistruje.

Rozhraní centra vyhledávání je rozděleno na stránky. Na společné stránce můžete prohledávat všechny dostupné kolekce, na dalších stránkách je pak možné prohledat jednu konkrétní kolekci. Existuje například samostatná stránka pro indexy portálového vyhledávacího stroje a jiná stránka pro knihovny správy dokumentů portálu.

Pokud je třeba povolit prohledávání kolekcí podnikového vyhledávání pomocí centra vyhledávání, poskytuje potřebný adaptér a registrační portlet produkt IBM WebSphere Information Integrator OmniFind Edition (WebSphere II OmniFind

Edition). Adaptér přidá do rozhraní centra vyhledávání kartu podnikového vyhledávání a registrační portlet zaregistruje portlet podnikového vyhledávání v centru vyhledávání.

Před použitím centra vyhledávání k prohledávání kolekcí podnikového vyhledávání je třeba instalovat adaptér a registrační portlet.

Prolézací modul WebSphere Portal

Podniková aplikace ESPACServer.ear je instalována na vyhledávacích serverech při instalaci produktu WebSphere II OmniFind Edition. Po implementaci této podnikové aplikace v produktu WebSphere Portal můžete prostřednictvím konzoly pro správu podnikového vyhledávání konfigurovat prolézací modul WebSphere Portal a přidávat weby vytvořené pomocí produktu WebSphere Portal do kolekcí podnikového vyhledávání.

Výhody integrace

Podnikové vyhledávání obohacuje vyhledávací prostředí produktu WebSphere Portal tím, že je doplňuje o podporu vyhledávání širší škály typů zdrojů dat. Pomocí portletu podnikového vyhledávání můžete prohledávat weby i všechny ostatní typy zdrojů dat podporované produktem WebSphere II OmniFind Edition.

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. Aby bylo možné podporovat kapacity na úrovni velkých podniků, podnikové vyhledávání distribuuje zátěž mezi čtveřici serverů (jeden zajišťuje prolézání dat, jeden jejich analýzu a indexování a dva podporu procesů vyhledávání a načítání).

Přehled implementace portletu

Portlety, které můžete použít pro podnikové vyhledávání, závisí na použité verzi produktu WebSphere Portal:

WebSphere Portal verze 5.0.2

Můžete implementovat portlet podnikového vyhledávání a použít jej k prohledávání kolekcí podnikového vyhledávání. Tento portlet může pracovat souběžně s portlety portálového vyhledávacího stroje.

WebSphere Portal verze 5.1 nebo novější

- Můžete implementovat portlet podnikového vyhledávání a použít jej k prohledávání kolekcí podnikového vyhledávání. Tento portlet může pracovat souběžně s portlety portálového vyhledávacího stroje.
- Po implementaci portletu podnikového vyhledávání můžete konfigurovat portlet vyhledávání dokumentů v portálovém vyhledávacím stroji pro prohledávání kolekcí podnikového vyhledávání.
- Po registraci portletu podnikového vyhledávání v centru vyhledávání produktu WebSphere Portal můžete k prohledávání kolekcí podnikového vyhledávání použít centrum vyhledávání. Můžete vybrat stránku určenou pouze k prohledávání kolekcí podnikového vyhledávání nebo zadat dotaz prohledávající kolekce podnikového vyhledávání a všechny ostatní kolekce dostupné v centru vyhledávání.

Související pojmy

“Migrace z produktu WebSphere Portal do podnikového vyhledávání” na stránce 195
Podnikové vyhledávání nabízí průvodce migrací, kterého můžete použít k migraci taxonomií a kolekcí z produktu IBM WebSphere Portal do prostředí podnikového vyhledávání.

Související úlohy

“Implementace podnikové aplikace pro prolézací modul WebSphere Portal” na stránce 71
Před vytvořením prolézacího modulu WebSphere Portal musíte v produktu WebSphere Portal implementovat podnikovou aplikaci ESPACServer.ear.

Zabezpečení na úrovni dokumentů při použití portálového vyhledávacího stroje

Pomocí vyhledávacího stroje produktu IBM WebSphere Portal můžete při prohledávání kolekcí podnikového vyhledávání uživateli vynutit zabezpečení na úrovni dokumentů.

Pokud prolézací modul podnikového vyhledávání přiřadí k prolézáným dokumentům prvek zabezpečení ID skupiny a pokud konfiguruje portlet prohledávání dokumentů tak, aby umožňoval prohledávání kolekcí podnikového vyhledávání pomocí portálového vyhledávacího stroje, může portálový vyhledávací stroj odvodit ID skupiny pro přihlášeného uživatele a předat prvek zabezpečení odpovídající tomuto ID skupiny spolu s dotazem podnikovému vyhledávání. Prvek zabezpečení zajistí, aby byly ve výsledcích vyhledávání vráceny pouze dokumenty, které je uživatel oprávněn prohlížet.

Pokud prolézací modul přiřadí k dokumentům jiný typ prvků zabezpečení, například jméno uživatele nebo roli uživatele, a chcete-li při prohledávání kolekcí podnikového vyhledávání vynutit zabezpečení na úrovni dokumentů, musíte vytvořit vlastní vyhledávací portlet. Portálový vyhledávací stroj odvozuje prvky zabezpečení pouze pro ID skupin.

Související pojmy

“Vyhledávací aplikace podnikového vyhledávání” na stránce 149

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.

“Zabezpečení na úrovni dokumentu” na stránce 177

Pokud je při vytvoření kolekce povoleno její zabezpečení, lze konfigurovat ovládací prvky zabezpečení na úrovni dokumentů. Zabezpečení na úrovni dokumentů zaručuje, že uživatelé prohledávající kolekce budou mít přístup pouze k dokumentům, k jejichž prohlížení mají oprávnění.

Implementace vyhledávacího portletu

Vyhledávací portlet ESSearchPortlet.war vám umožňuje prohledávat kolekce podnikového vyhledávání pomocí vyhledávacího stroje produktu WebSphere Portal.

Než začnete

Před konfigurováním portálového vyhledávacího stroje pro použití vyhledávacího portletu pro podnikové vyhledávání je nutné implementovat tento portlet v produktu WebSphere Portal.

O této úloze

Po implementaci vyhledávacího portálu můžete i nadále používat portálový vyhledávací stroj k prohledávání indexovaných dat v produktu WebSphere Portal. Aby mohli uživatelé prohledávat kolekce podnikového vyhledávání, musí administrátor produktu WebSphere Portal upravit vlastnosti v konfiguraci portálového vyhledávacího stroje. Můžete přepínat mezi těmito dvěma možnostmi prohledávání a používat vždy tu metodu vyhledávání, která nejlépe vyhovuje vašim potřebám.

Soubor ESSearchPortlet.war je instalován v adresáři ES_INSTALL_ROOT/bin na vyhledávacích serverech při instalaci produktu WebSphere II OmniFind Edition. Výchozí cesty instalace jsou následující:

- Systémy UNIX: /opt/IBM/es/bin/ESSearchPortlet.war
- Systémy Windows: C:\Program Files\IBM\es\bin\ESSearchPortlet.war

Postup

Chcete-li implementovat vyhledávací portlet v produktu WebSphere Portal, postupujte takto:

1. Instalujte portlet:
 - a. Přihlašte se k serveru WebSphere Portal s použitím ID administrátora serveru WebSphere Portal.
 - b. Vyberte z panelu nástrojů volbu **Správa**.
 - c. V navigační oblasti vlevo vyberte volbu **Portlety** a poté z nabídky **Portlety** vyberte volbu **Instalovat**.
 - d. Klepněte na volbu **Procházet**, vyberte v systému soubor ESSearchPortlet.war a klepněte na tlačítko **Další**.
 - e. Na další stránce klepněte na volbu **Instalovat**. Po instalování portletu se zobrazí následující zpráva: APIN0005I: Portlets were successfully installed (Portlety byly úspěšně instalovány).
2. Upravte parametry portletu:
 - a. V navigační oblasti vlevo vyberte volbu **Správa portletů** a poté vyberte volbu **IBM Enterprise Search** (Podnikové vyhledávání IBM).
 - b. Po zobrazení nových ikon vpravo od pole výběru vyberte volbu **Upravit parametry**.
 - c. V seznamu parametrů portletu upravte následující parametry:

port Nastavte tuto hodnotu na číslo portu webového serveru pro vyhledávací server produktu WebSphere II OmniFind Edition. Výchozí hodnota je 80.

applicationName Nastavte tuto hodnotu na jméno platné vyhledávací aplikace pro podnikové vyhledávání. (Jména vyhledávacích aplikací k dispozici jsou uvedena v seznamu na stránce Vyhledávací aplikace v konzole pro správu podnikového vyhledávání.) Výchozí hodnota Default odpovídá vyhledávací aplikaci, která je na vyhledávacích serverech konfigurována vždy.

hostname Nastavte tuto hodnotu na úplné jméno hostitele vyhledávacího serveru produktu WebSphere II OmniFind Edition.
 - d. Klepnutím na tlačítko **Uložit** uložte změny a poté klepnutím na tlačítko **Storno** opusťte stránku Upravit parametry.
3. Vytvořte pro portlet stránku portálu:
 - a. V navigační oblasti vlevo vyberte volbu **Uživatelské rozhraní portálu** a poté vyberte volbu **Správa stránek**.
 - b. Vyberte odkaz **Můj portál**.
 - c. Klepněte na volbu **Nová stránka**, zadejte název pro stránku a klepněte na tlačítko **OK**. Po vytvoření stránky se zobrazí následující zpráva: APPR0010I: *název_stránky* has been created successfully (Stránka *název_stránky* byla úspěšně vytvořena).
 - d. Klepněte na tlačítko **OK**. Vytvořená stránka se zobrazí v seznamu stránek k dispozici.
4. Přidejte vyhledávací portlet na stránku:
 - a. Vyberte ikonu Upravit (symbol tužky) pro vytvořenou stránku.

- b. Klepněte na volbu **Přidat portlety**.
 - c. Do textového pole **Vyhledat** zadejte text **Enterprise** a klepněte na volbu **Hledat**.
 - d. Zaškrtněte políčko u položky **IBM Enterprise Search** (Podnikové vyhledávání IBM) a klepněte na tlačítko **OK**.
 - e. Klepněte na tlačítko **Hotovo**. Tento krok indikuje, že jste úspěšně instalovali vyhledávací portlet a přidali jej na stránku produktu WebSphere Portal.
5. Přistupujte k portletu:
- a. V navigačním pruhu v horní části stránky vyberte volbu **Můj portál**.
 - b. Vyberte stránku odpovídající názvu stránky zadanému při přidávání portletu do systému.
- Při prvním přístupu ke stránce s portletem může být zobrazování stránky pomalejší, protože systém musí pro portlet kompilovat soubory JSP (Java Server Pages).

Konfigurování portletu pro vyhledávání a procházení produktu WebSphere Portal pro podnikové vyhledávání

Produkt WebSphere Portal poskytuje portlet, jehož prostřednictvím můžete prohledávat a procházet nativní kolekce produktu WebSphere Portal. Tento portlet lze konfigurovat tak, aby mohl prohledávat kolekce podnikového vyhledávání.

Postup

Chcete-li konfigurovat portlet pro vyhledávání a procházení tak, aby prohledával kolekce podnikového vyhledávání, postupujte takto:

1. Zastavte instanci serveru WebSphere_Portal.
2. Zkopírujte následující soubory z adresáře ES_INSTALL_ROOT/lib produktu WebSphere II OmniFind Edition do adresáře /WebSphere/PortalServer/shared/app na serveru WebSphere Portal:


```
esapi.jar
siapi.jar
```
3. Spusťte instanci serveru WebSphere_Portal a přihlašte se k serveru WebSphere Portal s použitím ID administrátora serveru WebSphere Portal.
4. Zkopírujte portlet:
 - a. Klepněte na volbu **Správa** v pravém horním rohu a poté rozbalte objekt **Správa portletů** a vyberte volbu **Portlety**.
 - b. Vyhledejte slovo **search**.
 - c. Klepněte na volbu **Zkopírovat portlet** pro portlet pro vyhledávání a procházení, zadejte pro portlet jedinečné jméno a klepněte na tlačítko **OK**.
5. Konfigurujte portlet:
 - a. Klepněte na ikonu **Konfigurovat portlet** u zadaného jména portletu.
 - b. Klepnutím na šipku vpravo přejděte na druhou stránku parametrů.
 - c. Do pole **Nový parametr** zadejte hodnotu **ApplicationInfold**, do pole **Nová hodnota** zadejte jméno platné vyhledávací aplikace produktu WebSphere II OmniFind Edition a poté nový parametr přidejte klepnutím na volbu **Přidat**.
Jména vyhledávacích aplikací k dispozici jsou uvedena v seznamu na stránce Vyhledávací aplikace v konzole pro správu podnikového vyhledávání. Výchozí hodnota Default odpovídá vyhledávací aplikaci, která je na vyhledávacích serverech konfigurována vždy.
 - d. Odstraňte následující parametry:

```
ApplicationInfoId_EXAMPLE
QueryFactoryImp
BrowseFactoryImp
SearchFactoryImp
IndexName
EJB_Example
IIOP_URL_Example
SOAP_URL_Example
EJB
SOAP_URL
IIOP_URL
```

- e. Přidejte následující parametry a zadejte hodnoty uvedené dále. Jméno a hodnotu každého parametru lze přidat po klepnutí na volbu **Přidat**. Ve jménech parametrů se rozlišují velká a malá písmena.

```
IMPLEMENT = ES
ApplicationInfoId = jméno_vyhledávací_aplikace (například Default)
SearchFactoryImp = com.ibm.es.api.search.RemoteSearchFactory
BrowseFactoryImp = com.ibm.es.api.browse.RemoteBrowseFactory
hostname = jméno_hostitele_vyhledávacího_serveru (například omnifind.ibm.com)
port = číslo_portu_vyhledávacího_serveru (například 80)
```

6. Přiřaďte novou kopii portletu pro vyhledávání a procházení ke stránce portálu. Potřebujete-li pomoc, naleznete informace v dokumentaci ke správě produktu WebSphere Portal.
7. Přistupujte k portletu:
- V navigačním pruhu v horní části stránky portálu vyberte volbu **Můj portál**.
 - Vyberte stránku odpovídající jménu zadaném při kopírování portletu v kroku 4 na stránce 191.

Instalace adaptéru podnikového vyhledávání pro centrum vyhledávání

Před použitím centra vyhledávání produktu WebSphere Portal k prohledávání kolekcí podnikového vyhledávání musíte instalovat adaptér pro podnikové vyhledávání.

O této úloze

Adaptér ESSearchAdapter.ear přidá do rozhraní centra vyhledávání stránku podnikového vyhledávání. Po instalování této aplikace a registračního portletu můžete přidávat kolekce podnikového vyhledávání do centra vyhledávání a prohledávat tyto kolekce pomocí univerzálního vyhledávacího portletu.

Soubor ESSearchAdapter.ear je instalován v adresáři ES_INSTALL_ROOT/bin na vyhledávacích serverech při instalaci produktu WebSphere II OmniFind Edition. Výchozí cesty instalace jsou následující:

- Systémy UNIX: /opt/IBM/es/bin/ESSearchAdapter.ear
- Systémy Windows: C:\Program Files\IBM\es\bin\ESSearchAdapter.ear

Postup

Chcete-li instalovat adaptér pro podnikové vyhledávání v centru vyhledávání, postupujte takto:

1. Zastavte instanci serveru WebSphere_Portal.
2. Není-li instance serveru server1 produktu WebSphere Application Server již spuštěna, spusťte ji.
3. Zkopírujte následující soubory z adresáře ES_INSTALL_ROOT/lib produktu WebSphere II OmniFind Edition do adresáře /WebSphere/PortalServer/shared/app na serveru WebSphere Portal:

esapi.jar
siapi.jar

Zobrazí-li se výzva k přepsání souboru siapi.jar, vyberte volbu **Ano**.

4. Na serveru WebSphere Portal spusíte konzolu pro správu produktu WebSphere Application Server. Zobrazí-li se výzva k přihlášení, přihlašte se.
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://jméno_hostitele:port/admin`, kde *jméno_hostitele* je jméno 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://jméno_hostitele:port/ibm/console`, kde *jméno_hostitele* je jméno hostitele vyhledávacího serveru a *port* je číslo portu konzoly pro správu produktu WebSphere Application Server. Konzola pro správu obvykle používá port 9060.
5. Klepněte na volbu **Aplikace** a poté na volbu **Instalovat novou aplikaci**.
6. Klepněte na volbu **Procházet**, vyberte v systému soubor ESSearchAdapter.ear a klepněte dvakrát na tlačítko **Další**. Obdržíte-li varování týkající se souborů zásad, klepněte na tlačítko **Pokračovat**.
7. Pokračujte klepáním na tlačítko **Další**, dokud se nezobrazí stránka Mapovat moduly na aplikační servery; poté vyberte následující volby a klepněte na tlačítko **Použít**:
 - V poli **Klastery a servery** vyberte volbu **WebSphere:cell=jméno_buňky, node=jméno_uzlu, server=server1**.
 - Zaškrtněte políčko u modulu **ESSearchAdapterEJB**.
8. Pokračujte v klepání na tlačítko **Další**, dokud se nezobrazí stránka Souhrn; poté klepněte na tlačítko **Dokončit**.
9. Klepněte na odkaz **Uložit do hlavní konfigurace** a poté klepnutím na tlačítko **Uložit** uložíte provedené změny do konfigurace serveru WebSphere Application Server.
10. Restartujte server WebSphere Portal.

Instalace registračního portletu podnikového vyhledávání pro centrum vyhledávání

Před použitím centra vyhledávání produktu WebSphere Portal k prohledávání kolekcí podnikového vyhledávání musíte instalovat registrační portlet pro podnikové vyhledávání.

O této úloze

Registrační portlet ESSearchAdapterPortlet.war registruje portlet podnikového vyhledávání pro centrum vyhledávání. Po instalování tohoto portletu (a aplikace adaptéru pro podnikové vyhledávání) můžete přidávat kolekce podnikového vyhledávání do centra vyhledávání a prohledávat tyto kolekce pomocí univerzálního vyhledávacího portletu.

Soubor ESSearchAdapterPortlet.war je instalován v adresáři ES_INSTALL_ROOT/bin na vyhledávacích serverech při instalaci produktu WebSphere II OmniFind Edition. Výchozí cesty instalace jsou následující:

- Systémy UNIX: /opt/IBM/es/bin/ESSearchAdapterPortlet.war
- Systémy Windows: C:\Program Files\IBM\es\bin\ESSearchAdapterPortlet.war

Postup

Chcete-li instalovat registrační portlet pro podnikové vyhledávání, postupujte takto:

1. Přihlašte se k serveru WebSphere Portal s použitím ID administrátora serveru WebSphere Portal.
2. Vyberte z panelu nástrojů volbu **Správa**.
3. V navigační oblasti vlevo vyberte volbu **Portlety** a poté z nabídky **Portlety** vyberte volbu **Instalovat**.
4. Klepněte na volbu **Procházet**, vyberte v systému soubor ESSearchAdapterPortlet.war a klepněte na tlačítko **Další**.
5. Na další stránce klepněte na volbu **Instalovat**. Po instalování portletu se zobrazí následující zpráva: APIN00051: Portlets were successfully installed (Portlety byly úspěšně instalovány).
6. Přiřaďte portlet (se jménem IBM WebSphere II OmniFind Edition registration portlet for enterprise search) ke stránce portálu. Potřebujete-li pomoc, naleznete informace v dokumentaci ke správě produktu WebSphere Portal.

Migrace z produktu WebSphere Portal do podnikového vyhledávání

Podnikové vyhledávání nabízí průvodce migrací, kterého můžete použít k migraci taxonomií a kolekcí z produktu IBM WebSphere Portal do prostředí podnikového vyhledávání.

Chcete-li provést migraci taxonomií a kolekcí, spusťte průvodce migrací na indexovém serveru podnikového vyhledávání. Migrovanou taxonomii můžete použít při práci s kolekcemi podnikového vyhledávání. Kromě toho můžete použít podnikové vyhledávání k administraci a prohlédávání kolekcí migrovaných z produktu WebSphere Portal.

V prostředí podnikového vyhledávání se namísto pojmu taxonomie používá pojem strom kategorií. Po migraci taxonomie založené na pravidlech můžete použít konzolu správy podnikového vyhledávání k administraci stromu kategorií. Aby bylo možné používat v prostředí podnikového vyhledávání taxonomie založené na modelu, musí být produkt WebSphere Portal nainstalován na indexovém serveru.

Jestliže chcete provést migraci taxonomií a kolekcí, vždy provádějte nejprve migraci souborů taxonomií založených na modelu a teprve poté migraci kolekcí. V opačném případě nebude kategorizace založená na modelu použitelná pro kolekce migrované z produktu WebSphere Portal.

Související pojmy

“Integrace podnikového vyhledávání s produktem WebSphere Portal” na stránce 187
Vyhledávací funkce produktu IBM WebSphere Portal můžete rozšířit implementací portletů podnikového vyhledávání v produktu WebSphere Portal a ve vyhledávacím centru produktu WebSphere Portal.

Související úlohy

“Konfigurování kategorií” na stránce 91
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.

Migrace taxonomie založené na modelu z produktu WebSphere Portal

Pomocí portletu Taxonomy Management produktu WebSphere Portal můžete určit, kterou taxonomii založenou na modelu chcete používat pro kolekce podnikového vyhledávání. Kolekce, které jste již migrovali do prostředí podnikového vyhledávání, nebudou novou migrací taxonomie nijak ovlivněny.

Než začnete

Před prvním spuštěním průvodce migrací zastavte systém podnikového vyhledávání, aby bylo možné provádět změny konfiguračních souborů. Při prvním spuštěním průvodce migrací musíte určit cestu k instalovaným produktům WebSphere Application Server a WebSphere Portal. Při dalších spuštěních průvodce migrací může být systém podnikového vyhledávání aktivní.

O této úloze

Chcete-li provést migraci taxonomie založené na modelu, musíte ji vybrat a exportovat z produktu WebSphere Portal. Poté provedete vlastní migraci taxonomie do prostředí podnikového vyhledávání pomocí průvodce migrací podnikového vyhledávání.

Postup

Chcete-li provést migraci taxonomie založené na modelu z produktu WebSphere Portal do prostředí podnikového vyhledávání, postupujte takto:

1. Exportujte aktuální taxonomii založenou na modelu z portletu správy taxonomií produktu WebSphere Portal. Taxonomie zahrnuje následující soubory XML:

```
synonyms.xml  
titles.xml  
treenodes.xml
```

2. Zkopírujte tyto soubory na indexový server podnikového vyhledávání.
3. Přihlaste se k indexovému serveru podnikového vyhledávání jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind Edition.
4. Přejděte do instalačního adresáře podnikového vyhledávání:

```
UNIX: cd $ES_INSTALL_ROOT/bin  
Windows: cd %ES_INSTALL_ROOT%\bin
```

5. Při prvním spuštění průvodce migrací zastavte systém podnikového vyhledávání zadáním následujícího příkazu:

```
esadmin stop
```

6. Spusíte průvodce migrací zadáním následujícího příkazu a poté klepněte na tlačítko **Další**.

```
UNIX:  
./eswpsmigrate.sh  
Windows: eswpsmigrate.bat
```

Pokud nejde o první spuštění průvodce migrací a systém podnikového vyhledávání je aktivní, ověřte, že během činnosti průvodce migrací není vytvářena kolekce podnikového vyhledávání.

7. Vyberte volbu **Importovat soubory taxonomie založené na modelu z portálu WebSphere** a poté klepněte na tlačítko **Další**.
8. Při prvním spuštění průvodce migrací zadejte cesty k instalačním adresářům aplikačního serveru WebSphere a produktu WebSphere Portal. Klepněte na tlačítko **Další**.
9. Přejděte do adresáře obsahujícího soubory taxonomie založené na modelu, vyberte soubory XML, které je třeba migrovat, abyste mohli používat taxonomii založenou na modelu (synonyms.xml, titles.xml a treenodes.xml) a poté klepněte na tlačítko **Další**.
10. Jde-li o první spuštění průvodce migrací, po dokončení migrace restartujte systém podnikového vyhledávání zadáním následujícího příkazu:
esadmin start

V případě výskytu chyb vyhledejte bližší informace v souboru MigrationWizard.log umístěném v instalačním adresáři průvodce migrací.

Související pojmy

“Kategorie založené na modelu” na stránce 89

Pokud v systému IBM WebSphere Portal používáte kategorie, můžete stejné kategorie používat i v kolekcích podnikového vyhledávání.

Související úlohy

“Konfigurování kategorií” na stránce 91

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.

Migrace kolekce z produktu WebSphere Portal

Chcete-li migrovat kolekce z produktu WebSphere Portal do prostředí podnikového vyhledávání, nejprve je připravte v produktu WebSphere Portal a poté proveďte jejich migraci pomocí průvodce migrací.

Než začnete

Chcete-li migrovat kolekce a taxonomie založené na modelu, musíte před použitím tohoto postupu pro migraci kolekcí nejprve migrovat soubory taxonomie založené na modelu. V opačném případě nebude kategorizace založená na modelu použitelná pro kolekce migrované z produktu WebSphere Portal.

Postup

Při migraci kolekce 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 WebSphere II OmniFind Edition.
5. Přejděte do instalačního adresáře podnikového vyhledávání:

UNIX: `cd $ES_INSTALL_ROOT/bin`

Windows: `cd %ES_INSTALL_ROOT%\bin`

6. Spusíte průvodce migrací zadáním následujícího příkazu a poté klepněte na tlačítko **Další**.

UNIX:

`./eswpsmigrate.sh`

Windows: `eswpsmigrate.bat`

7. Vyberte volbu **Migrovat nastavení vyhledávání z portálového vyhledávacího stroje portálu WebSphere** a poté klepněte na tlačítko **Další**.
8. 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.
9. Pro každou kolekci zadejte následující informace a poté spusíte migraci kolekci do prostředí podnikového vyhledávání klepnutím na tlačítko **Další**:
 - název, pod kterým chcete kolekci používat v prostředí podnikového vyhledávání,

- kritéria, podle nichž se v kolekci určuje důležitost dokumentu - faktor statického pořadí může být prázdný, založený na datech dokumentů nebo založený na odkazech na webové dokumenty z jiných webových dokumentů,
- typ kategorií, které chcete v této kolekci používat. Můžete vybrat kolekci bez kategorií, kategorie založené na pravidlech nebo kategorie založené na modelu. Pokud vyberete kategorie založené na pravidlech, bude do prostředí podnikového vyhledávání migrována taxonomie spolu s pravidly kolekce produktu WebSphere Portal.

Kategorie založené na modelu můžete vybrat jen tehdy, jestliže jste již provedli migraci taxonomie založené na modelu z produktu WebSphere Portal do prostředí podnikového vyhledávání.

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 zaslány poznámky k prolézacímu modulu, a název uživatelského agenta (pokud potřebujete s konfigurováním vlastností webového prolézacího modulu pomoci, klepněte na tlačítko **Nápověda**).

10. Na konzole správy podnikového vyhledávání spusťte procesy prolézání, analýzy a indexování pro migrovanou kolekci.
11. Jakmile dosáhnete stavu, kdy lze migrovanou kolekci prohledávat v prostředí podnikového vyhledávání, odstraňte původní kolekci z portálového vyhledávacího jádra.
12. 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.

Související pojmy

“Kolekce podnikového vyhledávání” na stránce 13

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.

“Kategorie založené na pravidlech” na stránce 87

Přiřazení dokumentů ke kategoriím v kolekci podnikového vyhledávání lze určovat konfigurováním pravidel.

“Stromy kategorií” na stránce 89

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.

“Statické řazení” na stránce 138

K některým dokumentům lze přiřadit statický faktor řazení, který zvyšuje důležitost těchto dokumentů uváděnou ve výsledcích vyhledávání.

Migrované nastavení kolekce

Při migraci kolekci 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 kolekci 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 adresa URL,
- počet paralelních prolézacích procesů,
- hloubka prolézání,
- časový limit (v sekundách) pro načtení dokumentu,
- výchozí znaková sada,
- pravidla zahrnutí a vyloučení pro prolézání.

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, například kategorizace založená na pravidlech nebo žádná kategorizace,
- zda má být použita mezipaměť vyhledávání a kolik dotazů s výsledky vyhledávání smí mezipaměť vyhledávání 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 aktualizace indexu,
- plán reorganizace 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 prolézáného prostoru a ujistěte se, že byly zadány všechny požadované hodnoty (povinná pole jsou označena červenou hvězdičkou). Zejména nezapomeňte uvést e-mailovou adresu pro zaslání 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 kolekcí podnikového vyhledávání.

Spouštění a zastavování serverů podnikového vyhledávání

Po vytvoření kolekce je nutné spustit servery pro prolézání, analýzu, indexování a prohledávání dat. Provedete-li v kolekci nějaké změny, zastavte tento server a znovu je spusťte.

Většina serverů podnikového vyhledávání může pracovat nepřetržitě nebo podle zadaných plánů. Můžete například nastavit plány reorganizace a aktualizace indexu. Servery pro analýzu dat a prohledávání indexu je po prvním spuštění obvykle třeba zastavit a spustit znovu 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 obvykle vstoupí v platnost až po zastavení a novém spuštění prolézacích modulů. Pokud nezměníte pravidla prolézání, prolézací moduly pracují buď nepřetržitě (v případě webového prolézacího modulu a prolézacího modulu NNTP), nebo podle zadaných plánů.

Spuštění serverů podnikového vyhledávání

Chcete-li uživatelům povolit vyhledávání v kolekci, je nutné spustit systémové procesy a poté spustit servery, které kolekci prolézají, analyzují, indexují a prohledávají.

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. Je například nutné spustit prolézací modul a prolézt data před analýzou a indexováním dokumentů.

Postup

Spouštění serverů podnikového vyhledávání:

1. Použití konzoly pro správu podnikového vyhledávání nebo vyhledávací aplikace:
 - a. Pokud ještě není spuštěn server IBM HTTP Server, spusťte jej.
 - b. Pokud ještě nejsou spuštěny aplikace podnikového vyhledávání ESAdmin a ESSearchApplication, spusťte je prostřednictvím konzoly pro správu aplikačního serveru WebSphere.
2. Pokud není spuštěna služba CCL (Common Communication Layer) podnikového vyhledávání, spusťte ji:
 - a. Přihlaste se k indexovému serveru s použitím jména uživatele administrátora podnikového vyhledávání zadaného při instalaci produktu WebSphere II OmniFind Edition.

Systémy UNIX



Zadejte následující příkaz:


```
startccl.sh -bg
```


Systémy Windows


Vyberte položky **Start** → **Programy** → **Nástroje pro správu** → **Služby** a restartujte službu IBM WebSphere Information Integrator OmniFind Edition.

3. Pokud není spuštěn systém podnikového vyhledávání, spusťte jej:
 - a. Přihlaste se k indexovému serveru s použitím jména uživatele administrátora podnikového vyhledávání zadaného při instalaci produktu WebSphere II OmniFind Edition.
 - b. Zadejte následující příkaz:

```
esadmin start
```
4. 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.
5. V pohledu Kolekce vyhledejte kolekci, kterou chcete spravovat, a klepněte na ikonu  **Monitor**.
6. Na stránce Prolézání klepněte pro každý prolézací modul, který chcete spustit, na volbu  **Spustit**.
 - Pokud spustíte webový prolézací modul nebo prolézací modul NNTP, bude prolézání dat zahájeno ihned. Tyto typy prolézacích modulů souvisle prolézají a opakovaně prolézají dokumenty na webových serverech a v diskusních skupinách NNTP.
 - 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.
7. Po prolezení dat otevřete stránku Analýza a klepnutím na volbu  **Spustit** spusíte 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).
8. Volitelné: Chcete-li vynutit spuštění procesů indexování a nečekat na zahájení indexování v naplánovanou dobu, otevřete stránku Indexovat a v oblasti **Reorganizace** klepněte na volbu  **Spustit**.

Procesy indexování můžete nechat spuštěné bez přerušení. Index bude aktualizován a reorganizován v naplánovaných dnech a časech.
9. Na stránce Vyhledávání klepněte na ikonu  **Spustit**.

Vyhledávací servery můžete nechat spuštěné bez přerušení. Obvykle je není třeba zastavovat, pokud nechcete provést změnu nastavení mezipaměti vyhledávání nebo shrnutí dokumentů.

Související pojmy

“Přehled správy podnikového vyhledávání” na stránce 1

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.

“Administrativní role” na stránce 170

Podnikové vyhledávání využívá koncepci rolí pro řízení přístupu k různým funkcím konzoly pro správu.

Související úlohy

“Přihlášení ke konzole pro správu” na stránce 4

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

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

Zastavení serverů 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 zastavit a znovu spouštět modul analýzy.

Pokud chcete namísto jednotlivých serverů zastavit systém podnikového vyhledávání, můžete se přihlásit k indexovému serveru pomocí ID administrátora podnikového vyhledávání (toto ID bylo určeno během instalace produktu WebSphere II OmniFind Edition). Poté zadejte následující příkaz:

```
esadmin stop
```

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 prolézání prostor nebo vlastnosti prolézacího modulu, zastavte a znovu spusíte prolézací modul, aby se změny projevily. Pokud změníte prolézání prostor a chcete provést změny u dokumentů, které jsou již zahrnuty v kolekci, 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.

Změníte-li pravidla pro analýzu dat, projeví se změny po zastavení a opětném spuštění modulu analýzy. Změny se projeví pouze u nově prolézáných dokumentů. Pokud chcete změny provést u dokumentů, které jsou již v indexu obsaženy, je nutné zahájit úplné prolézání, aby byly všechny tyto dokumenty znovu prolezeny. To umožní jejich následnou opakovanou analýzu a indexování.

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 změnili nastavení mezipaměti vyhledávání nebo souhrnů dokumentů.

Související pojmy

“Přehled správy podnikového vyhledávání” na stránce 1

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.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

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 (prolézání, 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 prolézání nebo sestavování indexu.

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 pojmy

“Přehled správy podnikového vyhledávání” na stránce 1

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.

Související úlohy

“Spuštění serverů podnikového vyhledávání” na stránce 201

Chcete-li uživatelům povolit vyhledávání v kolekci, je nutné spustit systémové procesy a poté spustit servery, které kolekci prolézají, analyzují, indexují a prohlédávají.

“Zastavení serverů podnikového vyhledávání” na stránce 203

Server podnikového vyhledávání může být nutné zastavit a znovu spustit, pokud provádíte změny v jeho konfiguraci nebo pokud potřebujete řešit problémy.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

Odhad počtu dokumentů v kolekci

Při vytvoření nebo úpravě kolekce podnikového vyhledávání nastavujte odhad počtu dokumentů, které bude tato kolekce obsahovat. Správce prostředků vychází z této hodnoty při odhadování paměťových a diskových prostředků, které budou po kolekci zapotřebí, překročení této hodnoty však není nijak omezeno.

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. Správce prostředků vygeneruje varování, pokud váš odhad požadavků na paměťové a diskové prostředky překročí momentální možnosti systému. Varování vám umožňují předejít potenciálním problémům s nedostatkem prostředků.

Pokud pro kolekci konfiguruje výstrahy a povolíte funkci upozornování na překročení limitu počtu dokumentů v indexu, bude výchozí limit shodný se zadaným odhadem počtu dokumentů v kolekci. Funkce monitorování použije tuto hodnotu v kombinaci s procentuálním prahem výstrahy, který zadáte, a odešle e-mail v situaci, kdy se skutečný počet dokumentů v kolekci blíží k nastavenému 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.

Kontrola dostupnosti systémových prostředků

Pokud po vytvoření kolekce nebo prolézacího modulu zkontrolujete ještě před prolézáním dokumentů systémové prostředky, můžete určit, zda jsou dostupné systémové prostředky dostatečné pro spuštění produktu WebSphere II OmniFind Edition s plnou kapacitou, která je založena na aktuálním nastavení konfigurace.

Než začnete

Chcete-li zkontrolovat systémové prostředky, musíte mít přiřazenu roli administrátora podnikového vyhledávání.

O této úloze

Při vytváření kolekce nebo prolézacího modulu systém automaticky zkontroluje dostupnost prostředků. Volbu kontroly dostupnosti systémových prostředků můžete vybrat kdykoli.

Systém porovná odhad množství prostoru, který jste použili při konfiguraci vlastností kolekce a prolézacího modulu, s dostupným prostorem v systému a poté zobrazí zprávu s informacemi o dostupnosti prostředků. Tato zpráva uvádí, u kterých serverů mohou být prostředky nedostatečné, kolik prostoru vyžaduje odhadovaná velikost kolekce a kolik volného prostoru je k dispozici.

V případě nedostatku prostředků zkuste provést některou z následujících opravných akcí:

- Zvětšete velikost souborového systému uvedeného v textu zprávy.
- Upravte obecné volby pro kolekci a zadejte nižší odhadovaný počet dokumentů.
- Upravte vlastnosti prolézacího modulu a zadejte nižší maximální počet prolézáných dokumentů, maximální velikost stránky a maximální počet podprocesů.
- U webového prolézacího modulu upravte vlastnosti prolézacího modulu a zadejte nižší maximální počet aktivních hostitelů, maximální počet nových dokumentů a maximální počet dokumentů v dočasném úložišti.
- Upravte vlastnosti systémového indexu a snižte povolený počet současně probíhajících sestavení indexu.
- Zabraňte souběžné aktivitě příliš mnoha procesů ve více kolekcích. Na prostředky může mít například vliv, jestliže jsou současně spuštěny prolézací moduly a moduly analýzy z několika kolekcí.
- Odstraňte ze systému některé prolézací moduly.
- Odstraňte ze systému některé kolekce.

Pokud nebyla detekována žádná upozornění na nebezpečí nedostatku prostředků, v systému je pravděpodobně dostatek prostředků pro prolézání a indexovaná data.

Postup

Určování, zda jsou pro aktuální konfiguraci systému podnikového vyhledávání k dispozici dostatečné prostředky:

1. Klepnutím na volbu **Systém** otevřete pohled **Systém**.
2. Na stránce **Obecné** klepněte na volbu **Kontrolovat systémové prostředky**.

Monitorování kolekce


Můžete zobrazit obecné informace o stavu jednotlivých komponent v kolekci nebo vybrat volby 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 **Obecné**, 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**.

Související pojmy

“Kolekce podnikového vyhledávání” na stránce 13

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

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

vyhledávání, 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 Obecné, 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 222

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.

Související úlohy

“Zobrazení sestav odstraněných dokumentů” na stránce 223

Můžete také zobrazit podrobné informace o dokumentech odstraněných ze systému podnikového vyhledávání. Tyto informace jsou k dispozici pouze v případě, že jste pro kolekci povolili sledování dokumentů.

Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 76

Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

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 změnilí pravidla pro analýzu dokumentů a chcete tato pravidla použít pro dokumenty, které byly indexovány dříve.

Prolézací moduly NNTP:

Pokud je prolézací modul zastaven, zahájí znovu prolézání a bude prolézat celý prolézaný prostor. Pokud je prolézací modul pozastaven, bude prolézání obnoveno od začátku cíle, ve kterém bylo pozastaveno.

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 podrobnosti pro prolézaný 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.

Související pojmy

“Správa prolézacích modulů podnikového vyhledávání” na stránce 19
Prolézací moduly je třeba konfigurovat pro různé typy dat, které chcete zahrnout do kolekce. Jedna kolekce může obsahovat libovolný počet prolézacích modulů.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237
Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

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


Aktivity 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.
 - a. 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**.
 - b. Klepnutím na tlačítko **Aktivní weby** zobrazíte informace o webových serverech, které prolézací modul aktivně zpracovává.
 - c. 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.
 - d. Klepnutím na tlačítko **Historie prolézacího modulu** zobrazíte zprávy o dřívější činnosti prolézacího modulu.
 - e. V oblasti **Stav adresy URL** zadejte adresu URL webového serveru, o němž chcete zobrazit informace, a poté klepněte na tlačítko **Zobrazit**.

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 prolézání. Dále zde najdete informace o posledním pokusu o prolezení webového serveru.

Po zobrazení podrobných informací o adrese URL můžete klepnutím na tlačítko **Historie webu** zobrazit další statistické informace o aktivitách prolézacího modulu souvisejících s danou adresou URL.

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 prolézání a prolézací modul dostatečně rychle nenalezne další adresy URL určené k prolézání. 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 prolézání 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 prolézáný prostor neobsahuje dostatečný počet webů k tomu, aby bylo možné udržovat podprocesy neustále v činnosti, nebo neexistuje dostatečný počet identifikátorů URL vhodných k prolézání. 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 prolézání,
- 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 prolézacího modulu.

Pokud se počet aktivovaných adres URL blíží nule, znamená to, že prolézací modul nenachází použitelné adresy URL. K přičinám 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 prolézacího modulu pro úpravy a upravte hodnotu časového limitu v poli **Maximální doba, po kterou mohou adresy URL zůstat v paměti** 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 návratových 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 serveru, pro který chcete vytvořit sestavu, klepněte na tlačítko **Zobrazit** a poté na tlačítko **Historie webu**.
5. Chcete-li vytvořit sestavu historie prolézacího modulu nebo historie webu, zaškrtněte políčka odpovídající statistickým údajům, které se mají v sestavě zobrazit, a poté klepněte na tlačítko **Zobrazit sestavu**.

U těchto typů statistických údajů prolézací modul vrátí sestavu konzole pro správu, jakmile načte informace ze své interní databáze.

6. Pokud vytváříte sestavu historie prolézacího modulu, zadejte volby pro vytvoření sestavy webů a poté klepněte 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, zadejte volby pro vytvoření sestavy návratových kódů HTTP a poté klepněte na tlačítko **Spustit sestavu**.

Tato sestava obsahuje informace o počtu návratových 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 návratových 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 návratových kódů HTTP nepřispívají do indexu, můžete zvýšit výkon prolézacího modulu odebráním těchto webů z prolézaného prostoru.

Návratové kódy HTTP webového prolézacího modulu

Při monitorování webového prolézacího modulu můžete zobrazit informace o návratových 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 návratových kódech HTTP vrácených prolézacím 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 návratových kódů HTTP a informují-li tyto návratové 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 návratové kódy HTTP spolu s údaji o tom, jakým způsobem tyto kódy interpretuje webový prolézací modul. Hodnoty v intervalu od 100 do 505 jsou standardní návratové kódy HTTP (další informace viz <http://www.w3.org/Protocols/rfc2616/rfc2616.html>). Ostatní návratové kódy HTTP jsou vlastní kódy podnikového vyhledávání a webového prolézacího modulu.

Tabulka 4. Návratové kódy HTTP pro webový prolézací modul

Kód	Popis	Kód	Popis	Kód	Popis	Kód	Popis
NULL	Neprolezeno	401	Neautorizováno	500	Interní chyba serveru	700	Chyba analýzy (chybí konec hlavičky)
100	Pokračovat	402	Je vyžadována platba	501	Není implementováno	710	Chyba analýzy (hlavička)
200	Úspěch	403	Zakázáno	502	Chybná brána	720	Chyba analýzy (bez kódu HTTP).
201	Vytvořeno	404	Nenalezeno	503	Služba není dostupná	730	Chyba analýzy (tělo)
202	Přijato	405	Nedovolená metoda	504	Časový limit brány	740	Vyloučeno na základě souboru robots.txt.
203	Neautoritativní informace	406	Nepřijatelné	505	Verze HTTP není podporována	741	Roboti jsou dočasně nedostupní
204	Žádný obsah	407	Je vyžadováno ověření na serveru proxy	611	Chyba čtení	760	Vyloučeno dle definice prolézaného prostoru
205	Obnovení obsahu	408	Vypršení časového limitu požadavku	612	Chyba připojení	770	Chybný protokol nebo nestandardní systémový port
206	Neúplný obsah	409	Konflikt	613	Vypršení časového limitu čtení	780	Vyloučeno dle definice vyloučených typů souboru
300	Více voleb	410	Odstraněno	614	Selhání při navazování komunikace SSL	2004	Chybí indexová značka META
301	Trvale přesunuto	411	Požadována délka	615	Jiná chyba čtení	3020	Softwarové přeměrování
302	Nalezeno	412	Selhání nutné podmínky	616	Anomálie FBA	4044	Vyloučeno na základě souboru robots.txt.

Tabulka 4. Návrátové kódy HTTP pro webový prolézací modul (pokračování)

Kód	Popis	Kód	Popis	Kód	Popis	Kód	Popis
303	Viz další	413	Příliš velká entita v požadavku	617	Chyba kódování		
304	Nezměněno	414	Příliš dlouhý identifikátor URI v požadavku	680	Chyba vyhledávání DNS		
305	Použijte server proxy	415	Nepodporovaný typ média				
306	(nepoužito)	417	Požadavek nebyl splněn				
307	Dočasné přeměrování						

Poznámky k tabulce

Návrátové kódy 4xx

Kód 400 (chybný požadavek) se vyskytuje velmi zřídka. Podle standardu návratových kódů HTTP mají kódy 4xx indikovat selhání klienta (prolézacího modulu). Problém je však obvykle na straně serveru nebo v adrese URL, kterou prolézací 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 prolézací 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 návratového kódu HTTP 400, protože požadavek již není považován za platný.

- 615** Informuje o tom, že na prolézacím serveru, který stahuje data z webových serverů, došlo k neočekávané výjimce. Mnohé z návratových kódů tohoto typu mohou indikovat problém na straně prolézacího modulu.

Návrátové kódy 6xx

S výjimkou návratového kódu 615 informují návratové kódy 6xx o problémech, které lze při prolézání očekávat, například o vypršení časových limitů. Následující návratové kódy mohou vyžadovat provedení nápravných kroků:

611, 612 a 613

Indikují pomalé weby nebo nízký výkon sítě.

- 614** Indikuje, že prolézací modul není schopen pracovat se zabezpečenými weby (HTTPS). Jestliže jste přesvědčeni, že tyto weby by měly být dostupné, zkontrolujte správnost konfigurace certifikátů na prolézacím 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 prolézací modul.

Zkontrolujte také konfiguraci certifikátů s vlastním podpisem na serverech, které se pokoušíte prolézat. Prolézací 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 podrízených doménách (např. <http://internal.sales.ibm.com/>). S certifikáty používanými tímto

způsobem nemůže prolézací modul pracovat. Certifikáty s vlastním podpisem přijímá jen tehdy, odpovídá-li název domény subjektu (sales.ibm.com) a podpis na certifikátu názvu domény požadované stránky.

- 616** Informuje o tom, že i po opakovaném ověření se ve stažených datech stále objevuje přihlašovací formulář.
- 617** Informuje o tom, že z bajtového obsahu dokumentu nelze vytvořit řetězec, protože kódovací řetězec (znaková sada) je neplatný nebo dokument obsahuje neplatné bajty.
- 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 návratových kódů tohoto typu výrazně zhoršuje propustnost.

Návratové 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 návratových 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 kotvicí odkazy na vyloučené dokumenty.

4044 Informuje o tom, že kotvicí 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 návratových 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 návratové kódy 7xx se zpravidla vyskytují tehdy, provedete-li změny v prolézaném prostoru po určité době práce prolézacího modulu. Tyto návratové kódy obvykle nenaznačují problémy, které by bylo třeba řešit.

- 3020** Informuje o tom, že dokument s návratovým 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 dokončena reorganizace indexu).

Postup

Monitorování modulu analýzy pro kolekci:

1. V pohledu Kolekce vyhledejte kolekci, kterou chcete monitorovat, a klepněte na volbu  **Monitor**.
2. Otevřete stránku Analýza.

Tip: Pokud upravujete kolekci a nacházíte se již na stránce Analýza, 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 spusíte 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 indexu může vzniknout například v případě, že je nutné vynutit reorganizaci 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.

Související pojmy

“Správa indexu podnikového vyhledávání” na stránce 113

Aby měli uživatelé zajištěn trvalý přístup k nejaktuálnějším informacím, podnikové vyhledávání vytváří pro každou kolekci index a udržuje jej pravidelnou aktualizací a reorganizováním jeho obsahu.

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 (aktualizace nebo reorganizace), 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 aktualizaci indexu zastavit, aby bylo možné vynutit spuštění jeho reorganizace.
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** zahajte aktualizaci nebo reorganizaci indexu.
5. Chcete-li odebrat index z fronty indexů, klepněte na volbu  **Odebrat**.

Související pojmy

“Správa indexu podnikového vyhledávání” na stránce 113

Aby měli uživatelé zajištěn trvalý přístup k nejaktuálnějším informacím, podnikové vyhledávání vytváří pro každou kolekci index a udržuje jej pravidelnou aktualizací a reorganizováním jeho obsahu.

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 Analýza, 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 nejvhodnější 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).

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.

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. Pokud jste změnil čísla portů pro modul pro příjem dat, klepněte na tlačítko  **Restartovat**.

Modul pro příjem dat je spuštěn při spuštění systému podnikového vyhledávání. Pokud nezměníte čísla portů, modul pro příjem dat není nutné restartovat.

Související úlohy

“Konfigurování podpory pro aplikace modulu pro příjem dat” na stránce 74
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.

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 fáze indexace; tyto informace jsou rovněž zaznamenány. Identifikátory URI 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 207

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.

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.
4. Klepněte na tlačítko **OK**.

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 jste pro kolekci povolili 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 Obecné, 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, the 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 207

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.

Související odkazy

“Formáty identifikátorů URI v indexu podnikového vyhledávání” na stránce 76

Identifikátory URI (Uniform Resource Identifier) jednotlivých dokumentů v indexu podnikového vyhledávání určují typ prolézacího modulu, který daný dokument přidá do kolekce.

Výstrahy a soubory žurnálu podnikového vyhledávání

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 zasilání výstrah a pro zasilání 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 zasilání e-mailů vždy, když dojde k určitým 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" v dokumentu "Messages Reference"

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 zasilání 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 části 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 prolézaných 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 části 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. Číslo 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 230

Můžete určit volby pro příjem e-mailů při každém zaznamenání určitých zpráv nebo zpráv určitých typů.

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 230

Můžete určit volby pro příjem e-mailů při každém zaznamenání určitých zpráv nebo zpráv určitých typů.

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.

4. Do pole **Maximální velikost jednotlivých souborů žurnálu** zadejte maximální počet megabajtů pro jednotlivé soubory žurnálu. Výchozí hodnota je 5 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 10.
Chcete-li zajistit, aby byly pro účely kontroly dostupné i starší zprávy žurnálu, zvýšte 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**.

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** úplné jméno 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:
- Klepnutím na volbu **Kolekce** otevřete pohled Kolekce.
  - V seznamu kolekcí vyhledejte kolekci, kterou chcete konfigurovat, a klepněte na volbu  **Upravit**.
  - Na stránce Žurnál klepněte na volbu **Konfigurovat volby e-mailu pro zprávy**.
  - Na stránce Volby e-mailu pro zprávy kolekcí zaškrtněte políčko **Odesílat e-maily se zprávami na úrovni kolekce**.
  - 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.
  - Pokud chcete přijímat e-maily o všech zaznamenaných chybových zprávách, zaškrtněte políčko **Odesílat e-maily se všemi chybovými zprávami**.
  - 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" v dokumentu "Messages Reference"

Související úlohy

"Konfigurování výstrah na úrovni kolekce" na stránce 226

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.

"Konfigurování výstrah na úrovni systému" na stránce 227

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.


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 Žurnál, 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.


Jméno každého souboru žurnálu obsahuje typ souboru žurnálu (například identifikátor systému nebo kolekce), datum vytvoření souboru a číselnou příponu označující pořadí, ve kterém byl k danému datu soubor vytvořen. Příklad:

```
typ_souboru_zurnálu_2005-05-26_1.log  
typ_souboru_zurnálu_2005-05-26_2.log  
typ_souboru_zurnálu_2005-05-25_1.log  
typ_souboru_zurnálu_2005-05-25_2.log  
typ_souboru_zurnálu_2005-05-25_3.log
```

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, jméno 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 zobrazeno jméno hostitele serveru podnikového vyhledávání, na kterém došlo k chybě, jméno souboru, který chybu generoval, jméno funkce a číslo řádku, kde chyba vznikla, ID procesu a ID podprocesu.

Zálohování a obnovování systému podnikového vyhledávání

Skripty zálohování a obnovování vám umožňují zálohovat a obnovovat systém podnikového vyhledávání.

Dojde-li k selhání systému v důsledku nezotavitelné chyby, je nutné znovu instalovat produkt WebSphere Information Integrator OmniFind Edition a poté spustit skript obnovování. Tyto skripty můžete použít také k obnovení základních systémových souborů na jednom nebo více nových serverech.

Skript zálohuje a obnovuje následující soubory:

- konfigurační soubory obsažené v adresáři ES_NODE_ROOT/master_config,
- databázové soubory prolézacích modulů,
- soubory indexu; nejsou-li indexy podnikového vyhledávání umístěny v adresáři ES_NODE_ROOT/data, nebudete moci používat skripty podnikového vyhledávání pro zálohování a obnovování souborů indexu.

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.

Pro zálohování souborů systému podnikového vyhledávání do jiného adresáře potřebujete dostatek místa na disku. Skripty zálohování a obnovení neprovádějí kontrolu souborů. Po spuštění zálohování je většina systémových relací dočasně nedostupná. Vyhledávací procesy budou pracovat i nadále. Zálohování byste měli spustit po reorganizaci indexu, abyste měli k dispozici co nejaktuálnější zálohu indexu.

V případě instalace na více serverů provádějte zálohování a obnovení systému z indexového serveru podnikového vyhledávání. Vzhledem k tomu, že indexový server vytváří databázový katalog, má přístup k databázovým tabulkám produktu na prolézacím serveru a může je zálohovat.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

Zálohování systému podnikového vyhledávání

Systém podnikového vyhledávání lze zálohovat pomocí skriptu esbackup.sh pro systém UNIX nebo pomocí skriptu esbackup.bat pro systém 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.

Jméno uživatele, který je administrátorem podnikového vyhledávání, musí být uděleno oprávnění pro zápis do adresáře určeného při spuštění skriptu zálohování.

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 WebSphere II OmniFind Edition.
2. Zadejte následující příkaz, přičemž parametr *záložní_adresář* určuje adresář, do kterého chcete zálohovat soubory:

UNIX: `esbackup.sh -c -d -i záložní_adresář`

Windows: `esbackup.bat -c -d -i záložní_adresář`

Můžete určit následující volby:

- c Zálohuje konfigurační soubory. Tato volba je výchozí.
- d Zálohuje prolezené dokumenty v úložišti dat prolézacího modulu (tyto dokumenty dosud nebyly analyzovány ani indexovány).
- i Zálohuje indexové soubory. Nejsou-li soubory indexu umístěny v adresáři `ES_NODE_ROOT/data`, nebudete moci použít skript `esbackup` k zálohování souborů indexu.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

Obnovení systému podnikového vyhledávání

Po opětné instalaci produktu WebSphere II OmniFind Edition můžete obnovit konfigurační soubory systému pomocí skriptu `esrestore.sh` pro systém UNIX nebo skriptu `esrestore.bat` pro systém 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.

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 WebSphere II OmniFind Edition.
2. Spusťte službu CCL (Common Communication Layer), pokud již není spuštěna:

UNIX: `startccl.sh -bg`

Windows: `startccl.bat`

Chcete-li službu CCL spustit v systému Windows na pozadí, klepněte na příkaz **Start** → **Programy** → **Nástroje pro správu** → **Služby** a restartujte službu WebSphere Information Integrator OmniFind Edition.

3. Pokud je spuštěn řadič, zastavte jej:
`esadmin stop`
4. Zadejte následující příkaz, přičemž parametr *záložní_adresář* určuje adresář, do kterého jste zálohovali soubory:

UNIX: `esrestore.sh -c -d -i záložní_adresář`

Windows: `esrestore.bat -c -d -i záložní_adresář`

Můžete určit následující volby:

- c Obnoví konfigurační soubory. Tato volba je výchozí.
- d Obnoví prolezené dokumenty v úložišti dat prolézacího modulu. (Tyto dokumenty dosud nebyly analyzovány ani indexovány).
- i Obnoví soubory indexu. Nejsou-li soubory indexu umístěny v adresáři `ES_NODE_ROOT/data`, nebudete moci použít skript `esrestore` k obnovování souborů indexu.

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

Obnovení souborů systému podnikového vyhledávání na nových serverech

Systémové soubory můžete zálohovat na jednom serveru podnikového vyhledávání a poté obnovit nikoli na aktuálních serverech, ale na jednom nebo více nových serverech.

Než začnete

Před spuštěním obnovovacího skriptu je na nové servery nutné instalovat produkt WebSphere II OmniFind Edition.

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.

Jméno uživatele, který je administrátorem podnikového vyhledávání, musí být uděleno oprávnění pro zápis do adresáře určeného při spuštění skriptu zálohování.

O této úloze

Informace o serveru uložené v souboru `ES_NODE_ROOT/master_config/nodes.ini` nejsou do záložních souborů zahrnuty.

Postup

Chcete-li obnovit systém podnikového vyhledávání na jednom nebo více nových serverech, postupujte takto:

1. Přihlašte se k indexovému serveru jako administrátor podnikového vyhledávání. Toto jméno uživatele bylo určeno při instalaci produktu WebSphere II OmniFind Edition.
2. Spusíte zálohovací skript na aktuálním (původním) indexovém serveru, přičemž parametr *záložní_adresář* označuje adresář, do kterého chcete soubory zálohovat:

UNIX: `esbackup.sh záložní_adresář`

Windows: `esbackup.bat záložní_adresář`

3. Pomocí programu FTP odešlete všechny soubory z adresáře *záložní_adresář* na nový indexový server.
4. Na novém indexovém serveru spusíte obnovovací skript:

UNIX: `esrestore.sh záložní_adresář`

Windows: `esrestore.bat záložní_adresář`

Související odkazy

“Příkazy podnikového vyhledávání, návratové kódy a ID relací” na stránce 237

Pomocí příslušných příkazů můžete diagnostikovat problémy, určovat stav řady součástí systému, spouštět a zastavovat relace nebo spouštět a zastavovat systém.

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 jméno_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 5. Příkazy esadmin podnikového vyhledávání

Příkaz	Popis
<code>esadmin ID_relace_prolézacího_modulu start</code>	Spustí relaci prolézacího modulu. Tento příkaz nespouští žádné prolézání. Ukázka použití příkazu: <code>esadmin col1.WEB1.esadmin start</code> Ukázka zpráv a návratového kódu: FFQC5310I Proces WEBCrawler1 (sid: col1.WEB1.esadmin) není spuštěn. FFQC5314I Výsledek: 0
<code>esadmin ID_relace_prolézacího_modulu startCrawl</code>	Zahájí prolézání. Ukázka použití příkazu: <code>esadmin col3.DB21.esadmin startCrawl</code> Ukázka zpráv a návratového kódu: FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650 FFQC5314I Výsledek: 0
<code>esadmin ID_relace_prolézacího_modulu pause</code>	Pozastaví prolézání. Ukázka použití příkazu: <code>esadmin col3.DB21.esadmin pause</code> Ukázka zpráv a návratového kódu: FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650 FFQC5314I Výsledek: 0

Tabulka 5. Příkazy esadmin podnikového vyhledávání (pokračování)

Příkaz	Popis
esadmin <i>ID_relace_prolézacího_modulu</i> resume	<p>Obnoví prolézání.</p> <p>Ukázka použití příkazu: esadmin col3.DB21.esadmin resume</p> <p>Ukázka zpráv a návratového kódu: FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650 FFQC5314I Výsledek: 0</p>
esadmin <i>ID_relace_prolézacího_modulu</i> stopCrawl	<p>Zastaví prolézání.</p> <p>Ukázka použití příkazu: esadmin col3.DB21.esadmin stopCrawl</p> <p>Ukázka zpráv a návratového kódu: FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650 FFQC5314I Výsledek: 0</p>
esadmin <i>ID_relace_prolézacího_modulu</i> stop	<p>Zastaví relaci prolézacího modulu.</p> <p>Ukázka použití příkazu: esadmin col3.DB21.esadmin stop</p> <p>Ukázka zpráv a návratového kódu: FFQC5303I Proces DB2Crawler1 (sid: col3.DB21.esadmin) je již spuštěn. PID: 23650 FFQC5314I Výsledek: 0</p>
esadmin <i>ID_relace_prol_modulu</i> getCrawlerStatus	<p>Zjistí stav prolézacího modulu. Vrácené informace se liší podle toho, zda jde o webový prolézací modul nebo o prolézací modul pro všechny ostatní zdroje dat.</p> <p>Příklad pro webový prolézací modul: esadmin col1.WEB1.esadmin getCrawlerStatus</p> <p>Možné návratové kódy a zprávy pro webový prolézací modul: FFQC5303I Proces WebCrawler1 (sid: col1.WEB1.esadmin) je již spuštěn. PID: 23650</p> <p>Příklad pro jiný než webový prolézací modul: esadmin col3.DB21.esadmin getCrawlerStatus</p> <p>Možné návratové kódy a zprávy pro jiný než webový prolézací modul: FFQC5303I Proces db2crawler (sid: db2col.DB2_96945) je již spuštěn. PID: 5936</p> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech” na stránce 242.</p>

Tabulka 5. Příkazy esadmin podnikového vyhledávání (pokračování)

Příkaz	Popis
esadmin <i>ID_relace_jiného_prolézacího_modulu</i> getCrawlSpaceStatus	Zjistí obecný stav prolézaného prostoru pro každý prolézací modul, který není webovým prolézacím modulem.
esadmin <i>ID_relace_webového_prolézacího_modulu</i> getCrawlStatus -selections <i>hodnota</i>	<p>Ukázka použití příkazu:</p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatus</pre> <p>Ukázka zpráv a návratového kódu:</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>Ukázka použití příkazu:</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 242.</p>
esadmin <i>ID_relace_jiného_prolézacího_modulu</i> getCrawlSpaceStatusDetail -ts <i>ID_cílového_serveru</i>	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.
esadmin <i>ID_relace_webového_prolézacího_modulu</i> getCrawlDetailsPerSite -url <i>řetězec</i> selections <i>počet</i> -threshold <i>počet</i>	<p>Ukázka použití příkazu:</p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatusDetail -ts FOUNTAIN</pre> <p>Ukázka zpráv a návratového kódu:</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>Ukázka použití příkazu:</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 242.</p>
esadmin startParser -cid <i>ID_kolekce</i>	<p>Spustí modul analýzy.</p> <p>Ukázka použití příkazu:</p> <pre>esadmin startParser -cid col1</pre> <p>Ukázka zpráv a návratového kódu:</p> <pre>FFQC5303I Proces Controller (server1) (sid: controller) je již spuštěn. PID: 25917 FFQC5314I Výsledek: 0</pre>

Tabulka 5. Příkazy esadmin podnikového vyhledávání (pokračování)

Příkaz	Popis
esadmin stopParser <i>ID_kolekce</i>	<p>Zastaví modul analýzy.</p> <p>Ukázka použití příkazu: esadmin stopParser -cid coll</p> <p>Ukázka zpráv a návratového kódu: FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 15292 FFQC5312E Chyba při provádění akce: [stopParser] FFQC4823E Proces Session Parser Driver - Collection coll (node1) [coll.parserdriver] není spuštěn.</p>
esadmin monitor getCollectionParserMonitorStatus -cid <i>ID_kolekce</i>	<p>Zjistí stav modulu analýzy.</p> <p>Ukázka použití příkazu: esadmin monitor getCollectionParserMonitorStatus -cid coll</p> <p>Ukázka zpráv a návratového kódu: FFQC5303I Proces Monitor (node1) (sid: monitor) je již spuštěn. PID: 12543</p> <p>Další informace o vrácených stavových zprávách naleznete v části “Podrobné informace o stavových příkazech” na stránce 242.</p>
esadmin startMain -cid <i>ID_kolekce</i>	<p>Zahájí reorganizaci indexu.</p> <p>Ukázka použití příkazu: esadmin startMain -cid coll</p> <p>Ukázka zpráv a návratového kódu: FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 25917 FFQC5314I Výsledek: 1117671147056</p>
esadmin startDelta -cid <i>ID_kolekce</i>	<p>Zahájí aktualizaci indexu.</p> <p>Ukázka použití příkazu: esadmin startDelta -cid coll</p> <p>Ukázka zpráv a návratového kódu: FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 4548 FFQC5314I Výsledek: 1117670603408</p>

Tabulka 5. Příkazy esadmin podnikového vyhledávání (pokračování)

Příkaz	Popis
<pre>esadmin monitor getCollectionIndexMonitorStatus -cid <i>ID_kolekce</i> -buildType [main delta] -numrecords <i>posledních_N_záznamů</i></pre>	<p>Zjistí stav aktualizace nebo reorganizace indexu. (Index <i>main</i> odkazuje na reorganizovaný index; index <i>delta</i> odkazuje na aktualizovaný index.) Volba <i>numrecords</i> zobrazuje posledních <i>N</i> záznamů stavu sestavování indexu. Je-li parametr <i>numrecords</i> vynechán, bude vrácen stav pro posledních 20 sestavení indexu.</p> <p>Ukázka použití příkazu:</p> <pre>esadmin monitor getCollectionIndexMonitorStatus -cid coll -buildType main -numrecords 4</pre> <p>Ukázka zpráv a návratového kódu:</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 242.</p>
<pre>esadmin startSearch -cid <i>ID_kolekce</i></pre>	<p>Spustí procesy vyhledávacího serveru.</p> <p>Ukázka použití příkazu:</p> <pre>esadmin startSearch -cid coll</pre> <p>Ukázka zpráv a návratového kódu:</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>Ukázka použití příkazu:</p> <pre>esadmin stopSearch -cid coll</pre> <p>Ukázka zpráv a návratového kódu:</p> <pre>FFQC5303I Proces Controller (node1) (sid: controller) je již spuštěn. PID: 15292 FFQC5314I Výsledek: 0</pre>

Tabulka 5. Příkazy esadmin podnikového vyhledávání (pokračování)

Příkaz	Popis
esadmin monitor getCollectionSearchMonitorStatus -cid ID_kolekce	Zjistí stav vyhledávacího serveru.
esadmin ID_relace_správce_vyhledávání getStatus -cid ID_kolekce	<p>Ukázka použití příkazu:</p> <pre>esadmin monitor getCollectionSearchMonitorStatus -cid coll</pre> <p>Ukázka zpráv a návratového kódu:</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>Ukázka použití příkazu:</p> <pre>esadmin searchmanager.node1 getStatus -cid coll</pre> <p>Ukázka zpráv a návratového kódu:</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 237 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 6. Informace o stavu webového prolézacího modulu

Prvek	Atributy	Popis
CrawlerStatus	<ul style="list-style-type: none"> CrawlerThreadStateDist ActiveBucketList CrawlRate RecentlyCrawledURLList NumURLsThisSession 	Stav prolézacího modulu.
CrawlerRunLevel hodnota	<ul style="list-style-type: none"> Řetězec (v angličtině) “Not started” (není spuštěn): Relace prolézacího modulu existuje, dosud však neobdržela zprávu s pokynem k zahájení zpracování dokumentů. “Started” (spuštěn): Probíhá spouštění prolézacího modulu. “Running” (běží): Prolézací modul dokončil fázi inicializace a spouštění a provádí aktivní prolézání. “Paused” (pozastaven): Prolézací modul obdržel pokyn k pozastavení aktivního prolézání, nikoli však k ukončení činnosti. “Stopping” (zastavování): Prolézací modul obdržel signál k zastavení a připravuje se na zastavení. “Error” (chyba): Prolézací modul je v nezotavitelném stavu a je třeba zastavit jej a znovu spustit, aby mohlo být prolézání obnoveno. 	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 6. Informace o stavu webového prolézacího modulu (pokračování)

Prvek	Atributy	Popis
ActiveBucket	<ul style="list-style-type: none"> URL: Řetězec (specifikace adresy URL) Protokol, hostitel a port prolézaných adres URL. NumActURLs: Celé číslo (kladné) Počet adres URL ve fondu při jeho zpřístupnění pro prolézání (aktivaci). NumProcURLs: Celé číslo (nezáporné) Počet adres URL ve fondu, které byly dosud zpracovány (prolezeny nebo zamítnuty). TimeRem: Celé číslo Počet sekund zbývajících do vypršení časového limitu pro fond. Duration: Celé číslo (nezáporné) Počet sekund od aktivace fondu. 	Aktuální aktivita určeného webu.
CrawlRate	<p>Hodnota: Celé číslo (nezáporné)</p> <p>Počet prolézaných 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 pro zjištění stavu jiného než webového prolézacího modulu 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 7. Prvky a atributy pro prolézací modul NNTP, prolézací modul DB2 a prolézací modul Notes pro příkaz pro zjištění stavu prolézacího modulu

Jméno prvku a atributu	Prolézací modul NNTP	Prolézací modul DB2	Prolézací modul Notes
Status	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)

Tabulka 7. Prvky a atributy pro prolézací modul NNTP, prolézací modul DB2 a prolézací modul Notes pro příkaz pro zjištění stavu prolézacího modulu (pokračování)

Jméno prvku a atributu	Prolézací modul NNTP	Prolézací modul DB2	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 8. Prvky a atributy pro prolézací modul Exchange Server, prolézací modul DB2 Content Manager a prolézací modul Content Edition pro příkaz pro zjištění stavu prolézacího modulu

Jméno prvku a atributu	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 9. Prvky a atributy pro prolézací modul QuickPlace Server, prolézací modul Domino Document Manager, prolézací modul souborového systému UNIX a prolézací modul souborového systému Windows pro příkaz pro zjištění stavu prolézacího modulu

Jméno prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
Status	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)	Stav (0, 1, 2, -1)

Tabulka 9. Prvky a atributy pro prolézací modul QuickPlace Server, prolézací modul Domino Document Manager, prolézací modul souborového systému UNIX a prolézací modul souborového systému Windows pro příkaz pro zjištění stavu prolézacího modulu (pokračování)

Jméno 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.

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 10. 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.
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 návratového 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"/>
</HTTPCode Code="780" Count="20000">
<?CrawlStatus>

```

Vracená data obsahují některé, všechny nebo žádné z následujících prvků:

Tabulka 11. Vracené informace pro příkaz pro zjištění stavu prolézaného prostoru webového prolézacího modulu

Prvek	Atribut	Popis
CrawlerStatus	<ul style="list-style-type: none"> • NumPagesInRDS • NumSitesDiscovered • NumSitesWithDNS • NumSitesWithoutDNS • NumURLsDiscovered • NumUniquePagesSaved • NumURLsCrawled • NumURLsUncrawled • NumURLsOverdue • HTTPCodeDist 	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 zaplnění úložiště nezpracovaných dat (pouze z příspěvků tohoto prolézacího modulu).
NumSitesDiscovered	<p>Hodnota: Nezáporné celé číslo</p> <p>Počet hostitelů nalezených při prolézání (nebo z počátečních položek).</p>	Míra pokrytí prolézané domény prolézacím modulem (počet hostitelů).
NumSitesWithDNS	<p>Hodnota: Nezáporné celé číslo</p> <p>Počet hostitelů s asociovanými adresami IP (jsou převáděny prolézacím modulem na pozadí).</p>	Míra efektivity prolézacího modulu při získávání adres IP pro hostitele, kteří jsou nalezeni podle jmen DNS v adresách URL.
NumSitesWithoutDNS	<p>Hodnota: Nezáporné celé číslo</p> <p>Počet hostitelů bez asociovaných adres IP (jsou převáděny prolézacím modulem na pozadí).</p>	Míra efektivity prolézacího modulu při získávání adres IP pro hostitele, kteří jsou nalezeni podle jmen DNS v adresách URL.
NumURLsDiscovered	<p>Hodnota: Nezáporné celé číslo</p> <p>Počet jedinečných adres URL navštívených příslušným prolézacím modulem.</p>	Míra pokrytí prolézané domény prolézacím modulem (počet adres URL).
NumUniquePagesSaved	<p>Hodnota: Nezáporné celé číslo</p> <p>Počet jedinečných stránek zapsaných do úložiště nezpracovaných dat k dalšímu zpracování ostatními komponentami podnikového vyhledávání.</p>	Podíl tohoto prolézacího modulu na velikosti indexu.

Tabulka 11. Vrácené informace pro příkaz pro zjištění stavu prolézaného prostoru webového prolézacího modulu (pokračování)

Prvek	Atribut	Popis
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 pro zjištění stavu jiného než webového prolézaného prostoru 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 jméno serveru + jméno databáze. V případě úrovně agregace 1 má položka Server@Name hodnotu jméno serveru + jméno adresáře.

Tabulka 12. Prvky a atributy pro prolézací modul NNTP, prolézací modul DB2 a prolézací modul Notes pro příkaz pro zjištění stavu prolézaného prostoru

Jméno prvku a atributu	Prolézací modul NNTP	Prolézací modul DB2	Prolézací modul Notes
Server@Name	Jméno serveru diskusních skupin	Jméno databáze	Jméno databáze nebo jméno adresáře
Server/Status	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • -1: Chyba 	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • 5: Naplánován • -1: Chyba 	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • 5: Naplánován • -1: Chyba

Tabulka 12. Prvky a atributy pro prolézací modul NNTP, prolézací modul DB2 a prolézací modul Notes pro příkaz pro zjištění stavu prolézaného prostoru (pokračování)

Jméno prvku a atributu	Prolézací modul NNTP	Prolézací modul DB2	Prolézací modul Notes
Server/StatusMessage	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba
Server/NumberOfTargets	Počet diskusních skupin v prolézaném prostoru.	Počet databází v prolézaném prostoru.	Počet pohledů a složek v prolézaném prostoru.
Server/NumberOfCompletedTargets	Počet prolezených diskusních skupin.	Počet prolezených tabulek.	Počet prolezených pohledů a složek.
Server/NumberOfErrors	Nelze použít.	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	Nelze použít.	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno. 	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.
Server/ScheduleTime	Nelze použít.	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	Nelze použít.	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0, 1: <ul style="list-style-type: none"> 0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.) Prolézací modul Notes prolézá dokumenty v režimu adresářů.

Tabulka 13. Prvky a atributy pro prolézací modul Exchange Server, prolézací modul DB2 Content Manager a prolézací modul Content Edition pro příkaz pro zjištění stavu prolézaného prostoru

Jméno prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Server@Name	Jméno serveru Exchange Server.	Servery DB2 Content Manager.	Jméno úložiště.
Server/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba 	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba
Server/StatusMessage	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven 5: Naplánován -1: Chyba
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.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno. 	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno. 	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.
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 13. Prvky a atributy pro prolézací modul Exchange Server, prolézací modul DB2 Content Manager a prolézací modul Content Edition pro příkaz pro zjištění stavu prolézaného prostoru (pokračování)

Jméno prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Server/AggregationLevel	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)

Tabulka 14. Prvky a atributy pro prolézací modul QuickPlace Server, prolézací modul Domino Document Manager, prolézací modul souborového systému UNIX a prolézací modul souborového systému Windows pro příkaz pro zjištění stavu prolézaného prostoru

Jméno 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	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • -1: Chyba 	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • 5: Naplánován • -1: Chyba 	Stav (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • 5: Naplánován • -1: Chyba
Server/StatusMessage	<ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • 5: Naplánován • -1: Chyba 	<ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • 5: Naplánován • -1: Chyba 	<ul style="list-style-type: none"> • 0: Neprolezen • 1: Prolézání • 2: Dokončen (není plánován) • 3: Čekání • 4: Pozastaven • 5: Naplánován • -1: Chyba
Server/NumberOfTargets	Počet databází míst a databází místností v prolézaném prostoru.	Počet kabinetů v prolézaném prostoru.	Počet podadresářů v prolézaném prostoru.
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.

Tabulka 14. Prvky a atributy pro prolézací modul QuickPlace Server, prolézací modul Domino Document Manager, prolézací modul souborového systému UNIX a prolézací modul souborového systému Windows pro příkaz pro zjištění stavu prolézaného prostoru (pokračování)

Jméno prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno. 	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno. 	0, 1, 2 <ul style="list-style-type: none"> 0: Pro prolézací modul není konfigurováno plánování podle konfiguračních souborů prolézacího modulu. 1: Pro prolézací modul je konfigurováno plánování, pro příslušnou relaci je však plánování zakázáno. 2: Pro prolézací modul je konfigurováno plánování a pro příslušnou relaci je plánování povoleno.
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, 1 <ul style="list-style-type: none"> 0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.) Prolézací modul Notes prolézá dokumenty v režimu adresářů. 	0, 1 <ul style="list-style-type: none"> 0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.) Prolézací modul Notes prolézá dokumenty v režimu adresářů. 	0, 1 <ul style="list-style-type: none"> 0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.) Prolézací modul Notes prolézá dokumenty v režimu adresářů.

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 15. 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.
64	Počet prolezených adres URL.
128	Počet adres URL, které nebyly prolezeny.
256	Počet prošlých adres URL.

Tabulka 15. Hodnoty masek výběru pro příkaz pro zjištění podrobného stavu prolézaného prostoru webového prolézacího modulu (pokračování)

Bit masky	Výběr
512	Distribuce návratového 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 vracená pro podrobný stav prolézacího prostoru webového prolézacího modulu:

Tabulka 16. Vrácené informace pro příkaz pro zjištění podrobného stavu prolézaného prostoru webového prolézacího modulu

Prvek	Atributy	Popis
CrawlDetailsPerSite	<ul style="list-style-type: none"> LastActivationTime: LastActivationDuration: IPAddressList: RobotsContent: HTTPCodeDist: 	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 16. Vrácené informace pro příkaz pro zjištění podrobného stavu prolézaného prostoru webového prolézacího modulu (pokračování)

Prvek	Atributy	Popis
NumURLsOverdueBy	Threshold, Value: Celé číslo (kladné nebo záporné) Prahová hodnota pro hledání položek opětovného prohledávání, jejichž časový limit již vypršel (záporná hodnota určuje počet sekund před aktuálním okamžikem) nebo které by měly být zpracovány během určeného časového intervalu (kladná hodnota určuje počet sekund od aktuálního okamžiku).	Počet adres URL, které byly k dispozici pro prolézání alespoň před určitým počtem sekund (prahová hodnota) nebo které budou k dispozici během příštího intervalu určeného určitým počtem sekund (prahová hodnota).
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 Návratový kód HTTP nebo jiný interní kód.	Počet výskytů konkrétního návratového 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 pro zjištění podrobného stavu prolézaného prostoru jiných než webových prolézacích modulů 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 17. Prvky a atributy pro prolézací modul NNTP, prolézací modul DB2 a prolézací modul Notes pro příkaz pro zjištění podrobného stavu prolézaného prostoru

Jméno prvku a atributu	Prolézací modul NNTP	Prolézací modul DB2	Prolézací modul Notes
Target@Name	Jméno diskusní skupiny	Jméno tabulky	Jméno pohledu nebo složky
Target@CrawlType	Nelze použít.	0,1: <ul style="list-style-type: none"> 0: Aktivní prolézání (normální) 1: Pasivní prolézání (publikování událostí produktu DB2) 	0
Target/Status	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	Stav: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba
Target/StatusMessage	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba
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.
Target/AggregationLevel	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0, 1: <ul style="list-style-type: none"> 0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.) Prolézací modul Notes prolézá dokumenty v režimu adresářů.

Tabulka 17. Prvky a atributy pro prolézací modul NNTP, prolézací modul DB2 a prolézací modul Notes pro příkaz pro zjištění podrobného stavu prolézaného prostoru (pokračování)

Jméno prvku a atributu	Prolézací modul NNTP	Prolézací modul DB2	Prolézací modul Notes
Target/LastUpdatedTime	Nelze použít.	Čas poslední aktualizace: <ul style="list-style-type: none"> 0: Aktivní prolézání (normální) 1: Pasivní prolézání (publikování událostí produktu DB2) 	Nelze použít.
Target/LastResetTime	Nelze použít.	Čas posledního vynulování statistiky: <ul style="list-style-type: none"> 0: Aktivní prolézání (normální) 1: Pasivní prolézání (publikování událostí produktu DB2) 	Nelze použít.

Tabulka 18. Prvky a atributy pro prolézací modul Exchange Server, prolézací modul DB2 Content Manager a prolézací modul Content Edition pro příkaz pro zjištění podrobného stavu prolézaného prostoru

Jméno prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
Target@Name	Jméno podsložky	Jméno typu položky	Jméno třídy položek
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba
Target/StatusMessage	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba
Target/NumberOfRecords	Nelze použít.	Nelze použít.	Nelze použít.
Target/NumberOfCompletedRecords	Počet prolezených dokumentů.	Počet prolezených dokumentů.	Počet prolezených dokumentů.
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.

Tabulka 18. Prvky a atributy pro prolézací modul Exchange Server, prolézací modul DB2 Content Manager a prolézací modul Content Edition pro příkaz pro zjištění podrobného stavu prolézaného prostoru (pokračování)

Jméno prvku a atributu	Prolézací modul Exchange Server	Prolézací modul DB2 Content Manager	Prolézací modul Content Edition
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 Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)
Target/LastUpdatedTime	Nelze použít.	Nelze použít.	Nelze použít.
Target/LastResetTime	Nelze použít.	Nelze použít.	Nelze použít.

Tabulka 19. Prvky a atributy pro prolézací modul QuickPlace Server, prolézací modul Domino Document Manager, prolézací modul souborového systému UNIX a prolézací modul souborového systému Windows pro příkaz pro zjištění podrobného stavu prolézaného prostoru

Jméno prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
Target@Name	Jméno databáze míst nebo jméno databáze místností	Jméno databáze kabinetů	Jméno podadresáře
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba
Target/StatusMessage	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba 	<ul style="list-style-type: none"> 0: Neprolezen 1: Prolézání 2: Dokončen (není plánován) 3: Čekání 4: Pozastaven -1: Chyba
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.
Target/NumberOfUpdatedRecords	Počet prolezených databází míst a databází místností.	Počet prolezených kabinetů.	Počet podadresářů v prolézaném prostoru.
Target/NumberOfDeletedRecords	Počet aktualizovaných záznamů.	Počet aktualizovaných záznamů.	Počet aktualizovaných záznamů.

Tabulka 19. Prvky a atributy pro prolézací modul QuickPlace Server, prolézací modul Domino Document Manager, prolézací modul souborového systému UNIX a prolézací modul souborového systému Windows pro příkaz pro zjištění podrobného stavu prolézaného prostoru (pokračování)

Jméno prvku a atributu	Prolézací modul QuickPlace	Prolézací modul Domino Document Manager	Prolézací moduly souborového systému UNIX Windows
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 Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)	0: Prolézací modul Notes prolézá dokumenty v normálním režimu. (Ostatní prolézací moduly mimo prolézací modul Notes vždy vracejí hodnotu 0.)
Target/LastUpdatedTime	Nelze použít.	Nelze použít.	Nelze použít.
Target/LastResetTime	Nelze použít.	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>
  <SnapshotTimeStamp>1124318637564</SnapshotTimeStamp>
  <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>
  <NumberOfDocsToBeParsed>15881</NumberOfDocsToBeParsed>
  <NumberOfDocsAlreadyParsed>29</NumberOfDocsAlreadyParsed>
</CrawlerStatus>
<CrawlerStatus>
  <Name>Data Listener (server1)</Name>
  <Crawlerid>datalistener</Crawlerid>
  <Type>datalistener</Type>
  <ParserStatus>0</ParserStatus>
  <NumberOfDocsToBeParsed>0</NumberOfDocsToBeParsed>
  <NumberOfDocsAlreadyParsed>0</NumberOfDocsAlreadyParsed>
</CrawlerStatus>
</Monitor>
```

V následující tabulce jsou popsány prvky XML pro informace vrácené příkazem pro zjištění stavu modulu analýzy:

Tabulka 20. Prvky pro příkaz pro zjištění stavu modulu analýzy

Prvek	Popis
Status	1, je-li relace modulu analýzy pro tuto kolekci spuštěna, a 0, je-li relace modulu analýzy zastavena.
State	<p>Možné stavy: Idle, Restart, Parsing, Resuming, Stopped, Initializing, Pause, NoParserServiceIsAvailable.</p> <p>Stav Idle (nečinný) indikuje, že je modul analýzy ve stavu spánku na N 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. Pro prostředí 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 Parsing (provádění analýzy) indikuje, že modul analýzy zpracovává dokumenty.</p> <p>Stav Pause (pozastaven) indikuje, že činnost modulu analýzy byla pozastavena relací sestavování indexu pro tuto kolekci. Stav Initializing (inicializace) indikuje, že probíhá spouštění modulu analýzy a inicializace jeho stavu.</p> <p>Stav Resuming (pokračování) indikuje, že činnost modulu analýzy byla převedena 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. To znamená, že všechna prostředí JVM pro analýzu nebo převod na prvky jsou používány jinými kolekcemi.</p>
SnapshotTimeStamp	Příští čas, kdy modul analýzy získá načtené dokumenty z tabulek prolézacích modulů, v sekundách od roku 1970.
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	Jméno 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	Jméno prolézacího modulu.
Type	Typ prolézacího modulu (webový prolézací modul, prolézací modul NNTP, prolézací modul DB2 atd.).
ParserStatus	<ul style="list-style-type: none"> • 0: Dokumenty z tohoto prolézacího modulu dosud nebyly analyzovány. • 1: Dokumenty z tohoto prolézacího modulu jsou aktuálně analyzovány. • 2: Dokumenty z tohoto prolézacího modulu byly analyzovány.
NumberOfDocsToBeParsed	Počet dokumentů z tohoto prolézacího modulu čekajících na analýzu.
NumberOfDocsAlreadyParsed	Počet analyzovaných dokumentů z tohoto prolézacího modulu.

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:

```
FFQC5314I Výsledek: <Monitor Type="Parser">
<Monitor Type="MainIndexHistory" Count="1">
  <IndexStatus Id="1"/>
  <StartTime>1123101789411<StartTime>
  <Progress>0</Progress>
  <CurrentPhase>0</CurrentPhase>
  <TotalPhase>0</TotalPhase>
  <IndexCopyProgress>0</IndexCopyProgress>
  <CurrentServer>0</CurrentServer>
  <TotalServer>0</TotalServer>
  <IndexCopyTime>0</IndexCopyTime>
  <IndexBuildTime>0</IndexBuildTime>
  <Status>0</Status>
  <StopTime>1123101789618</StopTime>
  <NumberOfDocuments>0</NumberOfDocuments>
</IndexStatus>
<CurrentIndexWildcardSupport/>
<NextIndexWildcardSupport/>
<ScheduleStatus>
  <Status>1</Status>
</Monitor>
```

V následující tabulce jsou popsány všechny prvky XML pro informace vrácené příkazem pro zjištění stavu sestavování indexu:

Tabulka 21. Prvky pro příkaz pro zjištění stavu sestavování indexu

Prvek	Popis
IndexStatusId	ID stavu indexu.
StartTime	Doba zahájení tohoto sestavování indexu v sekundách od roku 1970.
Progress	Procentní část dokončení tohoto sestavování indexu.
CurrentPhase	<ul style="list-style-type: none"> • 1: fáze přepisování úložiště • 2: fáze globální analýzy • 3: fáze sestavování indexu
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 všech fází sestavování indexu.
Progress	0 pro úspěšné sestavení a zkopírování indexu a nenulová hodnota pro kód chybové zprávy.
StopTime	Čas dokončení pro sestavování indexu (všechny fáze) a kopírování indexu.
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"> • 0, není-li pro tuto kolekci a typ indexu povolen časový plán. • 1, je-li pro tuto kolekci a typ indexu povolen časový plán.

Tabulka 21. Prvky pro příkaz pro zjištění stavu sestavování indexu (pokračování)

Prvek	Popis
ScheduledTimeEnabled	Příští čas, kdy bude spuštěno sestavování indexu pro tuto kolekci a typ indexu, v sekundách od roku 1970.

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 22. Prvky pro příkaz pro zjištění stavu vyhledávacího serveru

Prvek	Popis
SearchStatusName	Jméno a ID relace správce vyhledávání, která monitoruje a udržuje index vyhledávání pro tuto kolekci.
HostName	Jméno hostitele serveru, na kterém je spuštěn index vyhledávání.
Status	<ul style="list-style-type: none"> 0, není-li index vyhledávání pro tuto kolekci spuštěn. 1, je-li index vyhledávání pro tuto kolekci spuštěn.

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 23. Položky pro příkaz pro zjištění podrobného stavu vyhledávacího serveru

Položka	Popis
CacheHits	Počet výsledků načtených z mezipaměti vyhledávání.
QueryRate	Počet dotazů obdržných 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.

Tabulka 23. Položky pro příkaz pro zjištění podrobného stavu vyhledávacího serveru (pokračování)

Položka	Popis
ResponseTime	Průměrná doba odezvy v milisekundách pro určený časový interval. (Výchozí nastavení je pět minut.)
Status	<ul style="list-style-type: none"> 0, není-li index vyhledávání pro tuto kolekci spuštěn. 1, je-li index vyhledávání pro tuto kolekci spuštěn.
SessionName	Jméno relace pro index vyhledávání této kolekce.

Návratové kódy pro příkazy esadmin

Pro příkazy esadmin mohou být vráceny následující kódy:

Tabulka 24. Návratové kódy pro příkazy esadmin

Kód	Jméno	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 25. Příklady jmen 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
datalistener	indexový server	10582	Spuštěno
discovery	indexový server	10649	Spuštěno
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

Tabulka 25. Příklady jmen relací, serverů původce, ID relací a stavů relací (pokračování)

Relace	Server, na němž je relace spuštěna	ID relace	Stav relace
searchmanager.node1	indexový server	10927	Spuštěno
utilities.node1	indexový server	10384	Spuštěno

Související pojmy

“Zálohování a obnovování systému podnikového vyhledávání” na stránce 233
 Skripty zálohování a obnovování vám umožňují zálohovat a obnovovat systém podnikového vyhledávání.

“Monitorování aktivity podnikového vyhledávání” na stránce 205

Při monitorování aktivit systému a kolekcí můžete zobrazit stav různých procesů, sledovat příznaky potenciálních problémů nebo zvýšit výkon úpravou nastavení konfigurace.

Související úlohy

“Monitorování prolézacích modulů” na stránce 209

Můžete zobrazit obecné informace o stavu jednotlivých prolézacích modulů v kolekci nebo vybrat volby pro zobrazení podrobných informací o aktivitě prolézacího modulu.

“Spuštění serverů podnikového vyhledávání” na stránce 201

Chcete-li uživatelům povolit vyhledávání v kolekci, je nutné spustit systémové procesy a poté spustit servery, které kolekci prolézají, analyzují, indexují a prohledávají.

“Zastavení serverů podnikového vyhledávání” na stránce 203

Server podnikového vyhledávání může být nutné zastavit a znovu spustit, pokud provádíte změny v jeho konfiguraci nebo pokud potřebujete řešit problémy.

Dokumentace k podnikovému vyhledávání

Dokumentaci k produktu WebSphere Information Integrator OmniFind Edition je k dispozici ve formátu PDF a HTML.

Instalační program produktu WebSphere Information Integrator OmniFind Edition může automaticky instalovat Informační centrum. Instalační program instaluje Informační centrum na vyhledávací server. V případě instalace na více serverů 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. Chcete-li zobrazit témata ve formátu HTML týkající se podnikového vyhledávání, spusťte Informační centrum.

Chcete-li zobrazit dokumenty ve formátu PDF, přejděte do adresáře `docs/lokalita/pdf`. Anglická verze dokumentů je například uložena v adresáři `docs/en_US/pdf`. Dokumentaci ve formátu PDF, položky ke stažení, opravy, technické poznámky a Informační centrum lze stáhnout také z webu podpory produktu WebSphere Information Integrator OmniFind Edition.

V následující tabulce je uvedena dostupná dokumentace, jména souborů a příslušná umístění.

Tabulka 26. Dokumentace ve formátu PDF pro podnikové vyhledávání

Záhlaví	Záhlaví	Záhlaví
<i>Instalační příručka podnikového vyhledávání</i> (témata tohoto dokumentu jsou k dispozici také v Informačním centru)	iiysi.pdf	<code>docs/lokalita/pdf/</code>
<i>Správa podnikového vyhledávání</i> (témata tohoto dokumentu jsou k dispozici také v Informačním centru)	iiysa.pdf	<code>docs/lokalita/pdf/</code>
<i>Programming Guide and API Reference for Enterprise Search</i> (témata tohoto dokumentu jsou k dispozici také v Informačním centru)	iiysp.pdf	<code>docs/lokalita/pdf/</code>
<i>Messages Reference</i> (témata tohoto dokumentu jsou k dispozici také v Informačním centru)	iiysm.pdf	<code>docs/lokalita/pdf/</code>
<i>Installation Requirements for Enterprise Search</i> (témata tohoto dokumentu jsou k dispozici také v Informačním centru)	iiysr.txt nebo iiysr.htm	<code>docs/lokalita/</code> (tento soubor lze spustit i z programu První kroky)
<i>Poznámky k verzi</i>	iiysn.pdf	K dispozici pouze na webu s dokumentací k produktu IBM WebSphere Information Integrator OmniFind Edition.
<i>Text Analysis Integration</i>	iiyst.pdf	<code>docs/lokalita/pdf/</code>

WebSphere II OmniFind Edition - usnadnění přístupu

Uživatelská rozhraní a dokumentace produktu IBM WebSphere Information Integrator OmniFind Edition jsou přístupné.

Instalační program

Instalační program produktu WebSphere II OmniFind Edition lze ovládat také pomocí klávesových zkratk. Některé klávesové zkratky jsou uvedeny v následující tabulce.

Tabulka 27. 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

Konzola pro správu podnikového vyhledávání a Informační centrum

Konzola pro správu a Informační centrum jsou rozhraní založená na prohlížeči. Tato rozhraní lze zobrazit v prohlížeči Microsoft Internet Explorer nebo Mozilla FireFox. Seznam klávesových zkratk a dalších funkcí podporujících usnadnění přístupu pro prohlížeč Internet Explorer nebo FireFox naleznete v nápovědě online k příslušnému prohlížeči.

Dokumentace ve formátu PDF

Veškerá dokumentace k podnikovému vyhledávání je k dispozici ve formátu PDF. Dokumenty PDF jsou přístupné prostřednictvím aplikace Adobe Acrobat verze 6.0. Dokumenty PDF jsou strukturované a většina programu pro čtení obrazovky by měla být schopna číst je.

Slovníček pojmů z oblasti podnikového vyhledávání

Tento slovníček definuje pojmy používané v rozhraních a dokumentaci podnikového vyhledávání.

administrativní role

Klasifikace uživatele určující, s kterými funkcemi konzoly pro správu podnikového vyhledávání smí uživatel pracovat. Role rovněž určuje, které kolekce může uživatel spravovat.

administrátor podnikového vyhledávání

Administrativní role, která uživateli umožňuje spravovat celý systém podnikového vyhledávání.

adresa IP

Jedinečná 32bitová adresa, která identifikuje hostitele v síti.

aktualizace indexu

Proces přidávání nových informací do existujícího indexu v systému podnikového vyhledávání. Porovnejte s pojmem reorganizace indexu.

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.

analyzátor

Program, který interpretuje dokumenty přidané do datového skladu podnikového vyhledávání. Analyzátor 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 jméno společnosti. V rámci architektury UIMA 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.

archiv stroje pro zpracování

Archivní soubor PEAR typu ZIP obsahující stroj pro analýzu architektury UIMA 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.

CCL (Common Communication Layer)

Komunikační infrastruktura, která sjednocuje různé komponenty (řadič, analyzátor, prolézací modul, modul analýzy, modul pro indexaci) produktu WebSphere Information Integrator OmniFind Edition.

certifikační úřad

Organizace, která vydává certifikáty a ověřuje entity (jednotlivce nebo organizace), které vystupují v elektronických transakcích. Certifikační úřady zaručují, že obě strany, které si vyměňují informace, jsou skutečně tím, za koho se prohlašují.

certifikát

Digitální dokument, který svazuje veřejný klíč s identitou vlastníka certifikátu, a tím umožňuje ověření vlastníka certifikátu. Certifikát je vydáván certifikačním úřadem.

cesta funkce

Cesta používaná pro přístup k hodnotě funkce v rámci struktury funkcí architektury UIMA.

diakritické znaménko

Znaménko, jehož přidáním nad písmeno se změní výslovnost slova nebo na základě kterého lze rozlišovat mezi podobnými slovy, například čárka nebo přehláska v němčině.

DIIOIP (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 DIIOIP 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.

dotaz v přirozeném jazyce

Typ hledání, který analyzuje napsané výrazy (například "Kdo vede finanční oddělení?") namísto pouhého shromažďování klíčových slov.

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 WebSphere Information Integrator OmniFind 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 jména, termíny a výrazy.

extrakce pojmů

Vyhledávací funkce, která v textových dokumentech hledá důležité slovníkové položky (například jména osob, míst a produktů) a generuje seznam těchto položek. Viz také extrakce témat.

extrakce témat

Typ extrakce pojmů, při níž jsou v textových dokumentech automaticky

rozpoznávají 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 pojmenování, díky kterému může agregovaný systém zpracovávat složená jména platná v rámci systémů pojmenování.

fronta indexů

Seznam požadavků na reorganizaci indexu nebo požadavků na aktualizaci indexu, která má být provedena.

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

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.

jedinečné jméno

Jméno, které jednoznačně identifikuje položku v adresáři. Jedinečné jméno sestává z dvojic atribut: hodnota oddělených čárkami. Může také jít o sadu dvojic jméno-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 uspořádání 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

Skupina dokumentů s podobnými vlastnostmi.

kategorie založená na modelu

Taxonomie předdefinovaných výrazů, která slouží k určení předmětu dokumentu, aby jej bylo možné indexovat a vyhledávat spolu s dokumenty podobného zaměření.

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 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.

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 který nezahrnuje s požadavky na prostředky komplexnějšího protokolu X.500 Directory Access Protocol.

lemma Kanonický tvar slova. Lemmata jsou významná v jazycích s vysokou mírou skloňování, například v češtině.**lemmatizace**

Proces vyhledávání lemmatu pro dané slovo ve slovníku. Lemmatizace se liší od hledání kořene slova tím, že hledání kořene je algoritmický proces, který obecně nepracuje se slovníkem slov jazyka.

lexikální afinita

Vzájemný vztah vyhledávacích slov uvedených v dokumentu blízko sebe. 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

Klientský federátor, který provádí federování v rámci sady prohledávacích objektů.

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

Speciální stránka, která obsahuje podrobné vysvětlení problému, pokud server HTTP nemůže vrátit stránku, kterou si klient vyžádal, a konfiguruje server HTTP tak, aby vracel tyto stránky místo odpovědi obsahující pouze záhlaví s návratovým kódem identifikujícím problém.

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 příslušných stránkách.

no-index, direktiva

Direktiva v rámci webové stránky, která je pokynem pro roboty (například webový prolézací modul), aby obsah příslušných stránek 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

Vrstva architektury produktu Lotus Notes používaná pro veškerou komunikaci mezi produkty Notes.

obecná struktura analýzy

Struktura, v níž je uložen dokument, který je analyzován strojem pro analýzu textu. Informace jsou v obecné struktuře analýzy ukládány ve formě anotací a dalších struktur funkcí.

obor Skupina souvisejících identifikátorů URI používaná k definování rozsahu vyhledávacího požadavku.

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, spuštění a zastavování procesů na úrovni kolekce.

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 rychlému nalezení dokumentů obsahujících dotazovací výrazy.

počáteční adresa URL

Počáteční bod prolézání.

pole Nejmenší identifikovatelná část záznamu.

pořadí popularity

Typ řazení dokumentů, 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.

prolézání prostor

Sada zdrojů odpovídajících určitým vzorům (například názvům databází, cestám v souborovém systému, názvům domén, adresám IP nebo adresám URL), 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 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*).

reorganizace indexu

Proces sestavení indexu v systému podnikového vyhledávání. Porovnejte s pojmem aktualizace indexu.

REP (Robots Exclusion Protocol), protokol

Protokol, který umožňuje administrátorům webu určit, které části příslušného webu nemají být navštěvovány prolézajícími roboty.

rychlý odkaz

Asociace identifikátoru URI a klíčových slov a frázi.

řazení podle důležitosti

Proces přiřazování celočíselných hodnot k jednotlivým dokumentům ve výsledcích vyhledávání, které vrátil dotaz. 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

Proces, při kterém jsou v rámci řízení cest rozdělovány základní jednotky informací na menší jednotky (takzvané segmenty BIU), aby bylo možné využívat vyrovnávací paměť s menší velikostí na sousedních serverech.

segmentace pomocí mezer založená na kódu Unicode

Metoda převodu na prvky, která na základě vlastností znaků kódu Unicode rozlišuje mezi prvky a oddělujícími znaky.

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.

seznam řízení přístupu

Seznam identifikující uživatelé, kteří mají přístup k asociovanému objektu, a určující přístupová oprávnění uživatele k příslušnému objektu.

shrnutí

Proces začlenění vět, které stručně popisují 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 je spojení znaků f a i, které tvoří slitek.

slovník synonym

Slovník, který uživatelům umožňuje vyhledávat při prohledávání kolekce synonyma výrazů uvedených v dotazu.

slovo pro upřednostnění

Slovo, které může ovlivnit relevantní 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.

soubor úložiště klíčů

Soubor s databází klíčů, který obsahuje veřejné klíče uložené jako certifikáty podepisujícího subjektu a soukromé klíče uložené v osobních certifikátech.

soubory vyhledávacího indexu

Sada souborů, do kterých se v prostředí vyhledávacího stroje ukládá index.

správa identit

Možnost šifrovat údaje pověření uživatele v zabezpečeném úložišti.

SSL (Secure Sockets Layer), zabezpečení

Protokol zabezpečení, který poskytuje ochranu soukromí při komunikaci.

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.

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

Hierarchická struktura kategorií zobrazovaná v konzole pro správu podnikového vyhledávání.

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 třídě Java.

taxonomie

Rozdělení objektů do skupin na základě vzájemné podobnosti. V prostředí podnikového vyhledávání taxonomie dělí data do kategorií a podkategorií. Viz také strom kategorií.

třída upřednostnění

Specifikace, která může 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.

UIMA (Unstructured Information Management Architecture), architektura

Architektura společnosti IBM definující strukturu pro implementaci systémů pro analýzu nestrukturovaných dat.

URI (Uniform Resource Identifier)

Souvislý znakový řetězec identifikující abstraktní nebo fyzicky existující prostředek.

URL (Uniform Resource Locator)

Posloupnost znaků reprezentující informační prostředky v počítači nebo v síti, například v síti Internet. Tato posloupnost znaků zahrnuje zkrácené jméno protokolu použitého k přístupu k informačním prostředkům a informace použité protokolem k nalezení tohoto zdroje.

URN (Universal Resource Name), jméno

Prvek internetového protokolu sestávající z krátké posloupnosti znaků odpovídající určité syntaxi. Posloupnost obsahuje jméno nebo adresu, jejímž prostřednictvím lze odkazovat na prostředek.

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.

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.

volné vyhledávání textu

Hledání, u kterého je hledaný výraz vyjádřen volně utvořeným textem.

vyhledávací aplikace

Program, který zpracovává dotazy, prohledává index, vrací výsledky vyhledávání a načítá zdrojové dokumenty kolekci v systému podnikového vyhledávání.

vyhledávací stroj

Program, který přijímá požadavky na vyhledávání a vrací uživateli seznamy dokumentů.

vyhledávání blízkých slov

Typ vyhledávání, při kterém se hledají určitá slova ve stejné větě, odstavci nebo dokumentu.

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 zpracováváné anotačními moduly. Výsledky analýzy, které odpovídají hledaným informacím, jsou zapisovány do datové struktury nazývané obecná struktura analýzy.

výsledky vyhledávání

Seznam dokumentů odpovídajících vyhledávacímu požadavku.

vzdálený federátor

Serverový federátor, který provádí federování sady prohledávatelných objektů.

webový prolézací modul

Třída softwaru robotů, která prozkoumává web načtením webového dokumentu a prolezením odkazů v rámci tohoto dokumentu.

XPath (XML Path Language), jazyk

Jazyk, který jedinečně identifikuje části zdrojového dokumentu XML nebo na ně odkazuje. Jazyk XPath rovněž poskytuje základní prostředky pro práci s řetězci, čísly a logickými operátory.

zakázané slovo

Běžně používané slovo, například *a*, *i*, *nebo*, které aplikace při zpracování dotazu ignoruje.

zařadit do fronty

Umístit položky 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. Některé systémy vyžadují více než jeden znak.

znak změny významu

Znak, který potlačuje nebo aktivuje speciální význam jednoho nebo více následujících znaků.

Přístup k informacím o produktech WebSphere Information Integration

Informace o produktech WebSphere Information Integration jsou k dispozici prostřednictvím telefonu nebo na webu.

Telefonní čísla uvedena v tomto dokumentu jsou platná ve Spojených státech:

- Objednání produktů nebo získání obecných informací: 1-800-IBM-CALL (1-800-426-2255)
- Objednání publikací: 1-800-879-2755

Informace o produktech WebSphere Information Integration naleznete rovněž na webu na adrese www.ibm.com/software/data/integration/db2ii/. Tento web obsahuje nejnovější informace o následujících tématech:

- Dokumentace produktů
- Stahování produktů
- Opravné sady
- Poznámky k verzi a další informace podpory
- Novinky o produktech WebSphere Information Integration
- Odkazy na webové zdroje, například dokumenty white paper a IBM Redbooks
- Odkazy na diskusní skupiny a uživatelské skupiny
- Odkazy na online verze Informačních center pro produkty WebSphere Information Integration
- Objednávání publikací

Přístup k dokumentaci k produktům:

1. Přejděte na web na adrese www.ibm.com/software/data/integration/db2ii/.
2. Z rozevíracího seznamu vyberte požadovaný produkt, například WebSphere Information Integrator OmniFind Edition.
3. Klepněte na odkaz Support na levé straně stránky.
4. V části Learn vyberte požadovaný odkaz. Je-li pro vybraný produkt k dispozici Informační centrum, můžete vybrat odkaz na příslušné Informační centrum. Příklad naleznete na obrázku Obrázek 1 na stránce 280.

Learn

- **Product documentation and manuals** (2 items)
- **Redbooks** (1 item)
- **V8.2 Documentation and release notes**

Information Center

Provides fast, online centralized access to product information.

- 1.0

Obrázek 1. Příklady odkazů na dokumentaci k produktům na webu podpory produktů WebSphere Information Integration

Odesílání poznámek k dokumentaci

Prosíme o zaslání veškerých poznámek, které máte k těmto informacím nebo k jiné dokumentaci k produktu IBM WebSphere Information Integration.

Váš názor pomáhá společnosti IBM poskytovat kvalitní informace. Prosíme o zaslání veškerých poznámek, které máte k těmto informacím nebo k jiné dokumentaci k produktu WebSphere Information Integration. Poznámky lze odesílat kteroukoli z následujících metod:

1. Odešlete poznámky prostřednictvím online formuláře pro poznámky čtenářů na adrese www.ibm.com/software/awdtools/rcf/.
2. Odešlete poznámky e-mailem na adresu comments@us.ibm.com. Uveďte jméno produktu, číslo verze produktu a jméno a výrobní číslo informací (je-li k dispozici). Odesíláte-li poznámky týkající se konkrétního textu, uveďte umístění textu (například nadpis, číslo tabulky nebo číslo stránky).

Možnosti kontaktování společnosti IBM

Chcete-li kontaktovat zákaznický servis společnosti IBM v USA nebo v Kanadě, volejte na telefonní 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

Chcete-li vyhledat pobočku společnosti IBM ve vaší zemi či oblasti, přečtěte si dokument IBM Directory of Worldwide Contacts (Celosvětový adresář kontaktů společnosti IBM) na webu na adrese www.ibm.com/planetwide.

Ochranné známky

V tomto tématu jsou uvedeny ochranné známky společnosti IBM a některé ochranné známky mimo společnost IBM.

Informace o ochranných známkách společnosti IBM naleznete v dokumentu <http://www.ibm.com/legal/copytrade.shtml>.

Následující výrazy jsou ochrannými známkami nebo registrovanými ochrannými známkami jiných společností.

Java a všechny ochranné známky a loga založená na termínu Java jsou ochrannými známkami nebo registrovanými ochrannými známkami společnosti Sun Microsystems, Inc. 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.

Intel, Intel Inside (loga), MMX a Pentium jsou ochranné společnosti Intel 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.

Linux je ochranná známka Linuse Torvaldse ve Spojených státech a případně v dalších jiných zemích.

Názvy dalších společností, produktů nebo služeb mohou být ochrannými známkami nebo značkami služeb ostatních společností.

Poznámky

Tyto informace byly vytvořeny pro produkty a služby poskytované v USA. Společnost IBM nemusí produkty, služby nebo funkce uvedené v tomto dokumentu nabízet ve všech 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 takovéto 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 popsanych 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 dodržáním 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.

Tyto údaje obsahují příklady dat a sestav používaných v běžných obchodních operacích. Aby byla představa úplná, používají se v příkladech jména osob, společností, značek a produktů. Všechna tato jména jsou fiktivní a jejich podobnost se jmény 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. Tyto ukázkové programy můžete bez závazků vůči společnosti IBM v jakékoli formě a jakýmkoli způsobem kopírovat, měnit a distribuovat za účelem vývoje, používání, odbytu či distribuce aplikačních programů odpovídajících rozhraní API společnosti IBM.

Každá kopie nebo část těchto ukázkových programů nebo jakákoli práce z nich odvozená musí obsahovat následující copyrightovou doložku:

Outside In (®) Viewer Technology, ©1992-2005 Stellent, Chicago, IL., Inc. Všechna práva vyhrazena.

IBM XSLT Processor Licencované materiály - vlastnictví společnosti IBM ©Copyright IBM Corp., 1999-2005. Všechna práva vyhrazena.

Rejstřík

A

- adaptér pro podnikové vyhledávání 192
- administrativní role
 - administrátor kolekce 170, 171
 - administrátor podnikového vyhledávání 170, 171
 - konfigurování 171
 - monitor 170, 171
 - operátor 170, 171
 - popis 170
- administrátor kolekce
 - konfigurace role 171
 - popis 170
- administrátor podnikového vyhledávání
 - konfigurace role 171
 - popis 170
 - změna hesla na jednom serveru 5
 - změna hesla na více serverech 7
- administrátor, heslo
 - změna na jednom serveru 5
 - změna na více serverech 7
- adresová pravidla pro webové prolézací moduly 56
- aktivní weby, monitorování 210, 212
- aktualizace indexu
 - plánování 114, 115
 - popis 113
- aktualizace indexů 113
- analýza kotviciho textu
 - globální analýza 175
 - indexování dokumentů 175
 - popis 169
 - zabezpečení kolekce 175
- analýza textu
 - mapování prvků XML 99
 - obecná struktura analýzy 101
 - stroj pro analýzu textu 98, 99
- Analýza, popis stránky 1
- anotační moduly 96
- aplikace ESSearchApplication
 - config.properties, soubor 151, 160
 - povolení zabezpečení 161
 - restart 160, 161
 - zakázání zabezpečení 173

B

- bos.iocp.rte, modul 49

C

- CCLServer_datum.log, soubor 10
- Centrum vyhledávání produktu WebSphere Portal
 - adaptér pro podnikové vyhledávání 187, 192
 - popis 187
 - registrační portlet 187, 193
- config.properties, soubor
 - popisy vlastností 151

- config.properties, soubor (*pokračování*)
 - úpravy 160
- crawl.rules, soubor 67

Č

- čínština
 - analýza znaků nového řádku 105
 - n-gramová segmentace 105
- číslo portu, podnikové vyhledávání 10

D

- DB2 Content Manager, prolézací moduly
 - nastavení v systému Windows 37
- detekce duplicitních dokumentů
 - globální analýza 174
 - konfigurace tříd upřednostnění 142
 - popis 169
 - povolení zabezpečení 174
- DIHOP, protokol, konfigurace prolézacího modulu 48
- direktivy no-follow
 - konfigurování 69
 - popis 69
- direktivy no-index
 - konfigurování 69
 - popis 69
- doba odezvy vyhledávání
 - monitorování 220
 - výstrahy 226
- dokumentace 265
- dokumenty HTML, vyhledávání 95
- dokumenty XML
 - nativní vyhledávání XML 104
 - vyhledávání 93
- doménová pravidla pro webové prolézací moduly 56
- domény produktu Lotus Domino 183
- domény systému Windows 181
- dotaz, ověřování 178
- dotazy s nízkou návratností
 - popis 142
 - výchozí faktory upřednostnění 145
- dotazy s vysokou návratností
 - popis 142
 - výchozí faktory upřednostnění 145
- důležitost dokumentu
 - povolení pro kolekci 14
 - slovníky slov pro upřednostnění 140
 - statické 138
 - třídy upřednostnění 142, 145
 - pro migrovaných kolekcích 197
 - vzory identifikátorů URI 140, 141
- dynamické řazení podle důležitosti 137
- dynamické shrnutí 132

E

- EAR, soubory
 - aplikace ESSearchApplication 173
 - ESAdmin, aplikace 173
 - ESPACServer.ear 71
 - ESSearchAdapter.ear 192
 - ESSearchServer, aplikace 173
- es.cfg, soubor 5, 7
- ES_INSTALL_ROOT, popis 5, 7
- ES_NODE_ROOT, popis 5, 7
- es_special_field.default_field 145
- es_special_field.regular_text, pole 145
- ESAdmin, aplikace
 - přihlášení 4
 - zakázání zabezpečení 173
- esadmin, příkaz 237
- escrdb2.sh, skript 33
- escrdb2.vbs, skript 34
- escrnote.sh, skript 45
- escrnote.vbs, skript 46
- escrvbr.sh, skript 26
- escrvbr.vbs, skript 27
- eschangepw, skript 5, 7
- ESPACServer.ear, soubor 71
- ESSearchAdapter.ear, soubor 192
- ESSearchAdapterPortlet.war, soubor 193
- ESSearchPortlet.war, soubor 189
- ESSearchServer, aplikace
 - zakázání zabezpečení 173
- externí prolézací moduly
 - aplikace modulu pro příjem dat 74
 - konfigurování 74
- externí zdroje
 - asociace s vyhledávacími aplikacemi 166
 - konfigurování 165
 - popis 165
 - vyhledávání 117
 - zabezpečení pomocí ID aplikací 176
- externí zdroje JDBC
 - konfigurování 165
 - odstranění 165
 - ovladače JDBC 165
 - úpravy 165
- externí zdroje LDAP
 - konfigurování 165
 - odstranění 165
 - úpravy 165

F

- faktory upřednostnění
 - konfigurace tříd upřednostnění 142, 145
 - pro slovníky slov pro upřednostnění 138
 - pro třídy upřednostnění 144, 145
 - pro vzory identifikátorů URI 140, 141
- federované kolekce 13
- frekvence prolézání, monitorování 213
- fronta indexů 219

G

- global.rules, soubor 67
- globální analýza
 - analýza kotvícího textu 169, 175
 - detekce duplicitních dokumentů 169, 174
- globální prolézáný webový prostor 67
- globální zabezpečení produktu WebSphere
 - vyhledávací aplikace, vlastnosti 160
 - zakázání 173

H

- heslo administrátora podnikového
 - vyhledávání 5, 7
- historie doby odezvy, monitorování 220
- hledání dokumentace k podnikovému
 - vyhledávání 265
- Hloubka cesty URL 56
- hodnocení na základě textu 137
- HTTP, návratové kódy
 - přijaté webovými prolézacími moduly 213
 - webový prolézací modul, sestava 213
- HTTP, servery proxy 65

CH

- chybové zprávy
 - konfigurace serveru SMTP 229
 - příjem e-mailů 228, 230
 - zobrazení souborů žurnálu 231

I

- I/O Completion Port, modul, konfigurace prolézacího modulu 49
- ID aplikací 176
- ID kolekce, pravidla syntaxe 14
- ID relací, podnikové vyhledávání 237
- identifikátory URI
 - formáty v prostředí podnikového
 - vyhledávání 76
 - obory 122, 123
 - odebrání z indexu 125
 - ovlivňování statického skóre 140, 141
 - pravidla kategorií 87, 91
 - rychlé odkazy 134, 135
 - sbalené ve výsledcích vyhledávání 123, 124
 - zobrazení podrobností 207
- implementace
 - adaptér pro podnikové vyhledávání 192
 - ESPACServer.ear, soubor 71
 - ESSearchAdapter.ear, soubor 192
 - ESSearchAdapterPortlet.war, soubor 193
 - ESSearchPortlet.war, soubor 189
 - portlet pro vyhledávání a procházení 191
 - registrační portlet 193
 - vyhledávací portlet 189
- implementace vyhledávacího portletu 189
- Index, popis stránky 1
- indexové servery
 - spuštění 201
 - zastavení 203

indexy

- formáty identifikátorů URI 76
 - kotvící text 175
 - monitorování 218, 219
 - obory 117, 122
 - odebrání identifikátorů URI 117, 125
 - odstranění z fronty 219
 - paralelní sestavení 116
 - plánování 114
 - popis 113
 - povolení časového plánu 115, 218
 - sbalené identifikátory URI 117, 123, 124
 - souběžná sestavení 116
 - výstrahy 226
 - zakázání časového plánu 115, 218
 - zástupné znaky 117, 119, 121
 - změna plánu 115
- IOCP, konfigurace prolézacího modulu 49
- IP adresy, pravidla pro webové prolézací moduly 56

J

- japonština
 - analýza znaků nového řádku 105
 - n-gramová segmentace 105
- JavaScript, podpora ve webových prolézacích modulech 56

K

- kategorie
 - formáty identifikátorů URI 76
 - migrace z produktu WebSphere Portal 195, 197
 - popis 86
 - stromy kategorií 90
 - typ kategorizace 90
 - vnořování podkategorií 90
 - vyhledávání 86
 - vytvoření 91
 - založené na modelu 89, 90
 - založené na pravidlech 87, 90
- kategorie založené na modelu
 - popis 89
 - výběr typu kategorizace 90
- kategorie založené na pravidlech
 - popis 87
 - výběr typu kategorizace 90
 - vytvoření 91
- klíčová slova v rychlých odkazech 134, 135
- kolekce
 - analýza 85
 - asociace s vyhledávacími aplikacemi 150
 - federování 13
 - koncept 13
 - migrace z produktu WebSphere Portal 195, 197
 - MigrationWizard.log, soubor 197
 - monitorování 207
 - odhad prostředků 205
 - odhad velikosti 14
 - odstranění 17
 - stav systému 207
 - úpravy 16
 - vyhledávací servery 127

kolekce (pokračování)

- vyhledávání 117
 - výchozí nastavení migrace 199
 - vynechání řízení přístupu na úrovni dokumentu 185
 - vytvoření pomocí Průvodce kolekcí 13
 - vytvoření s použitím pohledu Kolekce 14
 - zabezpečení 174
 - zabezpečení duplicitních dokumentů 174
 - zabezpečení kotvícího textu 175
 - zabezpečení pomocí ID aplikací 176
 - způsoby vytvoření 13
- konektor Java pro produkt DB2 Content Manager 36, 37
- konfigurace portletu pro vyhledávání a procházení 191
- konfigurace serveru SMTP 229
- konfigurace serveru Trusted Server 183
- konfigurace služby Directory Assistance 53
- konfigurace uživatelů produktu Domino, prolézací moduly QuickPlace 52
- kontrola systémových prostředků 206
- konzola pro správu
 - přihlášení 4
 - rozhnutí 1
 - úlohy, souhrnné informace 1
- korejština
 - analýza složených výrazů 103
 - n-gramová segmentace 105

L

- lingvistická podpora
 - nativní vyhledávání XML 104
 - sémantické vyhledávání 96, 104
 - slovníky slov pro upřednostnění 138
 - slovníky synonym 128
 - slovníky zakázaných slov 131
 - vlastní analýza textu 96
- Lotus Domino Trusted Server, servery 183

M

- mapování
 - obecná struktura analýzy na index 101
 - obecná struktura analýzy na tabulky JDBC 101
 - pole na třídy upřednostnění 144
 - prvků XML na obecné struktury analýzy 99
 - vyhledávací pole HTML 96
 - vyhledávací pole XML 93
- maximální interval opakovaného prolézání 60
- méně závažné chyby, zpracování stránek webovými prolézacími moduly 61
- mezipaměť vyhledávání
 - konfigurování 128
 - popis 128
- migrace
 - kolekce 197
 - taxonomie založená na modelu 195
 - taxonomie založená na pravidlech 197
- MIME, zahrnutí typů do prolézáných webových prostorů 56
- minimální interval opakovaného prolézání 60

- modul pro příjem dat
 - konfigurování 74
 - monitorování 221
 - restart 74, 221
- moduly analýzy
 - analýza složených výrazů 103
 - analýza typů dokumentů 106
 - monitorování 217
 - n-gramová segmentace 105
 - nativní vyhledávání XML 104
 - nový řádek, znaky 105
 - podprocesy 102
 - popis 85
 - prázdný prostor 105
 - spuštění 217
 - stav systému 217
 - typy dokumentů pro relace Stellent 109, 111
 - typy dokumentů pro služby analyzátoru 106, 108
 - zastavení 217
- moduly plug-in prolézacího modulu 75
- moduly plug-in, pro prolézací moduly 75
- monitor
 - konfigurace role 171
 - popis 170
- monitorování
 - historie doby odezvy 220
 - kolekce 207
 - modul pro příjem dat 221
 - moduly analýzy 217
 - oblíbené dotazy 220
 - odstraněné dokumenty 223
 - podnikové vyhledávání 205
 - podrobnosti identifikátoru URI 207
 - poslední dotazy 220
 - prolézací moduly 209
 - soubory žurnálu 231
 - vyhledávací servery 220
 - webové prolézací moduly 210
 - webový prolézací modul, aktivní weby 212
 - webový prolézací modul, frekvence prolézání 213
 - webový prolézací modul, podrobnosti o podprocesech 211
- Monitorování, popis pohledu 1

N

- n-gramová segmentace 105
- naposledy prolézané adresy URL, monitorování 210
- nativní vyhledávání XML 104
- návratové kódy, podnikové vyhledávání 237
- návštěva adresy URL, co nejdřívější 61
- NNTP, konfigurování prolézacích modulů 41
- NRPC, protokol, konfigurace prolézacího modulu 45, 46

O

- obecná struktura analýzy
 - mapování na index 101
 - mapování na tabulky JDBC 101
 - mapování prvků XML 99

- obecná struktura analýzy (*pokračování*)
 - popis 96
- oblíbené dotazy, monitorování 220
- obnovování
 - systémové soubory na nových serverech 235
 - ze zálohy 234
- obnovování podnikového vyhledávání 233
- obnovování systému 233, 234, 235
- obory
 - formáty identifikátorů URI 76
 - popis 122
 - vyhledávání 122
 - vytvoření 123
 - vzory identifikátorů URI 122, 123
- odebrání identifikátorů URI z indexu 125
- odhad systémových prostředků 205
- odstraněné dokumenty
 - konfigurace souboru žurnálu 222
 - popis 222
 - sestavy 223
- odstranění
 - indexy z fronty 219
 - kolekce 17
 - prolézací moduly 23
- ochranné bariéry firewall, prolézání dokumentů Exchange Server 41
- omezení prolézaného webového prostoru 56
- opakovaná návštěva adresy URL, co nejdřívější 61
- opakované prolézání, intervaly pro webové prolézací moduly 60
- operační systém AIX
 - DB2 Content Manager, prolézací modul, konfigurování 36
 - konfigurace prolézacího modulu Domino Document Manager 45
 - konfigurace prolézacího modulu Notes 45
 - konfigurace prolézacího modulu QuickPlace 45
 - konfigurace publikování událostí 33
 - prolézací modul Content Edition, konfigurace 26
- operační systém Linux
 - DB2 Content Manager, prolézací modul, konfigurování 36
 - konfigurace prolézacího modulu Domino Document Manager 45
 - konfigurace prolézacího modulu Notes 45
 - konfigurace prolézacího modulu QuickPlace 45
 - konfigurace publikování událostí 33
 - prolézací modul Content Edition, konfigurace 26
 - Provozní prostředí Solaris
 - konfigurace publikování událostí 33
- operační systém Windows
 - DB2 Content Manager, prolézací modul, konfigurování 37
 - konfigurace prolézacího modulu 46
 - konfigurace publikování událostí 34
 - prolézací modul Content Edition, konfigurace 27
- operátor
 - konfigurace role 171

- operátor (*pokračování*)
 - popis 170
- ověření aktuálních pověřovacích údajů 178, 181, 183
- ověřování
 - popis 172
 - zakázání pro podnikové aplikace 173
- ověřování založené na použití formulářů 63, 64

P

- paralelní sestavení indexu 116
- parametrické vyhledávání 117
- PDF, dokumentace 265
- plánování
 - prolézací moduly 19
 - sestavení indexu 114, 115
- počáteční adresy URL pro webové prolézací moduly 56, 61
- podnikové aplikace
 - aplikace ESSearchApplication 173
 - ESAdmin, aplikace 173
 - ESPACServer.ear, soubor 71
 - ESSearchAdapter.ear, soubor 192
 - ESSearchServer, aplikace 173
- podnikové vyhledávání
 - administrativní role 170
 - formáty identifikátorů URI 76
 - ID relací 237
 - indexové servery 113
 - integrace s produktem WebSphere Portal 187
 - konfigurace čísla portu 10
 - moduly analýzy 85
 - monitorování 205
 - návratové kódy 237
 - obnovování systémových souborů 235
 - obnovování ze zálohy 234
 - prolézací servery 19
 - příkazy 237
 - skripty obnovování 233
 - skripty zálohování 233
 - soubory žurnálu 225
 - spuštění serverů 201
 - vyhledávací servery 127
 - zabezpečení 169
 - zabezpečení na úrovni dokumentu 177
 - zabezpečení na úrovni kolekce 174
 - zálohování 233
 - zastavování serverů 201, 203
- podprocesy
 - moduly analýzy 102
 - webový prolézací modul 211
- podrobnosti identifikátoru URI monitorování 207
- odstraněné dokumenty 223
- podrobnosti o podprocesech, monitorování 210
- Pohled Kolekce
 - popis 1
 - vytváření kolekcí 14
- pole, mapování na třídy upřednostnění 144
- portálový vyhledávací stroj
 - popis 187
 - předávání prvků zabezpečení 189
 - zabezpečení na úrovni dokumentu 189

- portlet Taxonomy Management 195
- portlety
 - ESSearchAdapterPortlet.war, soubor 193
 - podnikové vyhledávání 187
 - podporované verze produktu WebSphere Portal 187
 - popis 187
 - portlet Taxonomy Management 195
 - registrace 193
 - Vyhledávání 189
 - vyhledávání a procházení 191
- poslední dotazy, monitorování 220
- povolení časových plánů indexování 115
- pravidla kategorií
 - konfigurování 91
 - obsah dokumentu 87, 91
 - vzory identifikátorů URI 87, 91
- profily uživatelů
 - konfigurace 180
 - popis 179
- prolétací moduly
 - aplikace modulu pro příjem dat 74
 - Content Edition 24, 25
 - DB2 28
 - DB2 Content Manager 35
 - Domino Document Manager 39
 - Exchange Server 40, 41
 - formáty identifikátorů URI 76
 - kombinování typů prolétacích modulů 19
 - moduly plug-in 75
 - monitorování 209
 - NNTP 41
 - Notes 42, 44
 - odstranění 23
 - plánování 19
 - počáteční hodnoty 21
 - podpora externích 74
 - přehled konfigurace 19
 - QuickPlace 50
 - souborový systém UNIX 53
 - stav systému 209
 - úprava vlastností prolétacího modulu 22
 - úpravy prolétaného prostoru 23
 - výchozí nastavení migrace 199
 - vytvoření 21
 - webový prolétací modul 54
 - WebSphere Portal 70, 71, 72
 - Windows, souborový systém 73
 - zabezpečení na úrovni dokumentu 19, 177
 - základní hodnoty 19
- prolétací moduly Content Edition
 - konfigurace 24
 - přímý režim 25
 - režim serveru 25
- Prolétací moduly Content Edition
 - formáty identifikátorů URI 76
 - nastavení v operačním systému AIX 26
 - nastavení v operačním systému Linux 26
 - nastavení v provozním prostředí Solaris 26
 - nastavení v systému Windows 27
- prolétací moduly DB2
 - formáty identifikátorů URI 76
 - instalace produktu WebSphere MQ 33
 - instalace produktu WebSphere MQ v systému Windows 34
- prolétací moduly DB2 (*pokračování*)
 - konfigurace 28
 - konfigurace produktu WebSphere MQ 32
 - konfigurace publikování událostí 33, 34
 - publikování událostí 28
 - WebSphere II Event Publisher Edition, konfigurace 29
- Prolétací moduly DB2 Content Manager
 - formáty identifikátorů URI 76
 - konfigurace 35
 - nastavení v operačním systému AIX 36
 - nastavení v operačním systému Linux 36
 - nastavení v provozním prostředí Solaris 36
- prolétací moduly Domino Document Manager
 - IOCP, konfigurace 49
 - nastavení v operačním systému AIX 45
- Prolétací moduly Domino Document Manager
 - DIIOP, konfigurace protokolu 48
 - formáty identifikátorů URI 76
 - konfigurace 39
 - nastavení v operačním systému Linux 45
 - nastavení v provozním prostředí Solaris 45
 - nastavení v systému Windows 46
 - NRPC, protokol 45, 46
- prolétací moduly Exchange Server
 - formáty identifikátorů URI 76
 - konfigurace 40
 - zabezpečené dokumenty 41
- prolétací moduly Notes
 - DIIOP, konfigurace protokolu 48
 - formáty identifikátorů URI 76
 - IOCP, konfigurace 49
 - konfigurace 42
 - Lotus Domino Trusted Server 183
 - nastavení v operačním systému AIX 45
 - nastavení v operačním systému Linux 45
 - nastavení v provozním prostředí Solaris 45
 - nastavení v systému Windows 46
 - NRPC, protokol 45, 46
 - ověření aktuálních pověřovacích údajů 183
 - pravidla pro mapování polí 44
 - rady pro použití 44
 - zabezpečení na úrovni dokumentu 183
- prolétací moduly QuickPlace
 - IOCP, konfigurace 49
 - konfigurace 50
 - konfigurace služby Directory Assistance 53
 - konfigurace uživatelů produktu Domino 52
 - nastavení v operačním systému AIX 45
 - zabezpečení lokálních uživatelů 52
- Prolétací moduly QuickPlace
 - DIIOP, konfigurace protokolu 48
 - formáty identifikátorů URI 76
 - nastavení v operačním systému Linux 45
 - nastavení v provozním prostředí Solaris 45
 - nastavení v systému Windows 46
 - NRPC, protokol 45, 46
- prolétací moduly souborového systému UNIX
 - formáty identifikátorů URI 76
 - konfigurace 53
- prolétací moduly souborového systému Windows
 - formáty identifikátorů URI 76
 - konfigurace 73
 - zabezpečení na úrovni dokumentu 181
- prolétací moduly WebSphere Portal
 - implementace podnikové aplikace 71
 - implementace souboru ESPACServer.ear 71
- Prolétací moduly WebSphere Portal
 - formáty identifikátorů URI 76
 - konfigurace 70
 - kopírování adres URL webů 72
- prolétací servery
 - spuštění 201, 209
 - zastavení 203, 209
- Prolétání, popis stránky 1
- prolétaný prostor
 - úpravy 23
 - výstrahy 226
 - webový prolétací modul, konfigurace 56
- provozní prostředí Solaris
 - DB2 Content Manager, prolétací modul, konfigurování 36
 - konfigurace prolétacího modulu Domino Document Manager 45
 - konfigurace prolétacího modulu Notes 45
 - konfigurace prolétacího modulu QuickPlace 45
 - prolétací modul Content Edition, konfigurace 26
- proxy, servery 65
- Průvodce kolekcí 13
- průvodce migrací
 - kolekce 197
 - popis 195
 - spuštění 195, 197
 - taxonomie založené na modelu 195
 - taxonomie založené na pravidlech 197
 - výchozí nastavení kolekce 199
 - výchozí nastavení prolétacího modulu 199
 - žurnál, soubor 200
- prvky XML
 - mapování na obecnou strukturu analýzy 99
 - mapování na vyhledávací pole 93
 - vyhledávání 93, 99
- prvky zabezpečení
 - konfigurace prolétacího modulu 178
 - portálový vyhledávací stroj, zpracování 189
 - zabezpečení na úrovni dokumentu 178, 189
 - zakázání pro kolekci 185
- předpony, pravidla pro webové prolétací moduly 56
- přihlášení ke konzole pro správu 4
- příkazy, podnikové vyhledávání 237
- přímý režim, úložiště Content Edition 25
- přípony souborů
 - podporované analyzátoři kolekcí 106, 108
 - podporované relacemi Stellent 109, 111
 - vyloučení z prolétaných webových prostorů 56

- publikování událostí
 - konfigurace prolézacího modulu DB2 29, 32
 - nastavení v operačním systému AIX 33
 - nastavení v operačním systému Linux 33
 - nastavení v provozním prostředí Solaris 33
 - nastavení v systému Windows 34
 - popis 28

R

- relace Stellent
 - analýza typů dokumentů 106
 - přřazení typů dokumentů 109
 - výchozí typy dokumentů 111
- reorganizace indexu
 - plánování 114, 115
 - popis 113
- reorganizace indexů 113
- režim serveru, úložiště Content Edition 25
- Robots Exclusion, protokol 55
- robots.txt, soubor 55
- rozhraní API
 - vyhledávání a indexování 149
- rychlé odkazy
 - formáty identifikátorů URI 76
 - popis 134
 - vyhledávání 134
 - vytvoření 135

Ř

- řazení výsledků vyhledávání
 - dynamické 137
 - hodnocení na základě textu 137
 - popis 137
 - slovníky slov pro upřednostnění 140
 - statické 138
 - třídy upřednostnění 142, 144, 145
 - vzory identifikátorů URI 140, 141
- řízení přřstupu
 - aktuální ověření 178
 - popis 172
 - portálový vyhledávací stroj, podpora 189
 - požadavky pro produkt Lotus Domino 183
 - požadavky pro souborové systémy Windows 181
 - profily uživatelů 179
 - zabezpečení na úrovni dokumentu 178, 189
 - zakázání pro kolekci 185

S

- sbalené identifikátory URI
 - konfigurování 124
 - popis 123
- sbalené výsledky vyhledávání
 - konfigurování 124
 - popis 123
- sémantické vyhledávání 96, 99, 104
- servery modulu analýzy
 - konfigurace podprocesů 102
 - spuštění 201

- servery modulu analýzy (*pokračování*)
 - zastavení 203
- sestavení indexu
 - paralelní 116
 - souběžná 116
 - spuštění 218
 - stav systému 219
 - zastavení 218, 219
- sestavy historie prolézacího modulu
 - HTTP, sestava návratového kódu 213
 - popis 210
 - vytvoření 213
 - webová sestava 213
- sestavy historie webu
 - popis 210
 - vytvoření 213
- SIAPI (rozhraní API pro vyhledávání a indexaci) 149
- skript esbackup.bat 233, 235
- skript esbackup.sh 233, 235
- skript escrcm.sh 36
- skript escrcm.vbs 37
- skript esrestore.bat 234, 235
- skript esrestore.sh 234, 235
- skript startccl 234
- skripty
 - esbackup.bat 233, 235
 - esbackup.sh 233, 235
 - escrcm.sh 36
 - escrcm.vbs 37
 - escrdb2.sh 33
 - escrdb2.vbs 34
 - escrnote.sh 45
 - escrnote.vbs 46
 - escrvbr.sh 26
 - escrvbr.vbs 27
 - esrestore.bat 234, 235
 - esrestore.sh 234, 235
 - startccl 234
- skripty obnovování
 - popis 233
 - spouštění 234, 235
- skripty zálohování
 - popis 233
 - spouštění 233, 235
- sledování dokumentů
 - konfigurace souboru žurnálu 222
 - popis 222
 - povolení 222
 - sestavy 223
 - zakázání 222
- slovníky slov pro upřednostnění
 - asociace s kolekcí 140
 - popis 138
 - přřdávání do systému 139
- slovníky synonym
 - asociace s kolekcí 130
 - popis 128
 - přřdávání do systému 130
- slovníky zakázaných slov
 - asociace s kolekcí 132
 - popis 131
 - přřdávání do systému 132
- složený výraz, analýza 103
- souběžná sestavení indexu 116
- soubor cookies.ini
 - formát 66

- soubor cookies.ini (*pokračování*)
 - konfigurování 67
 - popis 66
- soubor followindex.rules
 - konfigurování 69
 - popis 69
- soubor parserTypes.cfg 106
- soubor runtime-generic.properties 133
- soubor synonyms.xml 195
- soubor titles.xml 195
- soubor treenodes.xml 195
- soubory cookie pro prolézací webu
 - formát 66
 - konfigurování 67
 - popis 66
- soubory žurnálu
 - cyklické použití 228
 - filtrování 231
 - konfigurace serveru SMTP 229
 - maximální velikost 228
 - monitorování 231
 - popis 225
 - pro sledování dokumentů 222
 - průvodce migrací 200
 - úrovně závažnosti 228
 - volby e-mailu 230
 - výchozí umístění 225
 - zobrazení 231
- souhrny
 - dynamické 132
 - přřzpůsobení 133
- souhrny dokumentů
 - přřzpůsobení 133
 - úprava vlastností 133
- správa identit
 - konfigurace 180
 - profily uživatelů 179
- spuštění
 - modul pro příjem dat 221
 - moduly analýzy 217
 - prolézací servery 209
 - průvodce migrací 195, 197
 - servery podnikového vyhledávání 201
 - sestavení indexu 218
 - ukázková vyhledávací aplikace 161
 - vyhledávací servery 220
- statické třídy důležitosti
 - popis 138
 - povolení pro kolekci 14
 - v migrovaných kolekcích 197
- stav systému
 - kolekce 207
 - moduly analýzy 217
 - prolézací moduly 209
 - sestavení indexu 219
 - vyhledávací servery 220
 - webové prolézací moduly 210
- stellent.properties, soubor 109
- stellentypes.cfg, soubor 109
- stroj pro analýzu textu
 - asociace s kolekcemi 99
 - mapování prvků XML 99
 - mapování výsledků analýzy 101
 - popis 96
 - přřdávání do systému 98

stromy kategorií
migrace z produktu WebSphere Portal 195, 197
popis 90
Systém, popis pohledu 1
systémové prostředí
kontrola 206
odhad 205

T

taxonomie založená na modelu, migrace z produktu WebSphere Portal 195
taxonomie založená na pravidlech, migrace z produktu WebSphere Portal 195, 197
taxonomie, migrace z produktu WebSphere Portal 195, 197
třídy upřednostnění
detekce duplicitních dokumentů 142
dotazy s nízkou návratností 142, 145
dotazy s vysokou návratností 142, 145
konfigurace 144, 145
mapování polí 144
popis 142
výchozí hodnoty 145
typ kategorizace
výběr 14, 90
založené na modelu 89
založené na pravidlech 87
typy dokumentů
analýza 106
pro relace Stellent 109, 111
pro služby analyzátoru 106, 108
typy prolézacích modulů
kombinování v kolekci 19
základní hodnoty 19
typy zdrojů dat
databáze Domino Document Manager 39
databáze QuickPlace 50
DB2, databáze 28
Exchange Server, veřejné složky 40
NNTP, diskusní skupiny 41
Notes, databáze 42, 44
relační databáze 28
souborové systémy UNIX 53
souborový systém Windows 73
typy položek DB2 Content Manager 35
úložiště Content Edition 24, 25
webové servery 54
weby WebSphere Portal 70

U

UIMA
asociace s kolekcemi 99
mapování prvků XML 99
mapování výsledků analýzy na index 101
mapování výsledků analýzy na tabulky JDBC 101
obecná struktura analýzy 101
popis 96
přidávání strojů pro analýzu textu do systému 98
ukázková vyhledávací aplikace
config.properties, soubor 151, 160

ukázková vyhledávací aplikace (*pokračování*)
globální zabezpečení produktu WebSphere 161
popis 150
povolení zabezpečení 161
přístup 161
spuštění 161
vyhledávací funkce 149, 150
výchozí implementace 160
zakázání zabezpečení 173
úlohy konzoly pro správu, souhrnné informace 1
upozornění e-mailem
konfigurace serveru SMTP 229
pro výstrahy 230
pro zprávy 230
úpravy
aplikace modulu pro příjem dat 74
kolekce 16
prolézání prostory 23
vlastnosti prolézacího modulu 22
vyhledávací aplikace, vlastnosti 151, 160
usnadnění přístupu 267
uživatelští agenti 55

V

vbr_access_services.jar, soubor 26, 27
vlastní analýza textu
mapování obecných struktur analýzy 101
mapování prvků XML 99
mapování výsledků analýzy na index 101
mapování výsledků analýzy na tabulky JDBC 101
popis 96
stroj pro analýzu textu 98, 99
vlastnosti prolézacího modulu
úpravy 22
volby obsahu dokumentu 117
volby vyhledávání
obsah dokumentu 117
parametrické vyhledávání 117
pro výsledky vyhledávání 117
volné vyhledávání textu 117
vyhledávání s použitím polí 117
volné vyhledávání textu 117
vyhledávací a indexové rozhraní API 149
vyhledávací aplikace
asociace s kolekcemi 150
asociování s externími zdroji 166
ID aplikací 176
ukázka 149, 150
vlastní 149
zabezpečení na úrovni kolekce 176
vyhledávací centrum pro produkt WebSphere Portal
adaptér pro podnikové vyhledávání 187, 192
popis 187
registrační portlet 187, 193
vyhledávací pole HTML
mapování prvků 95, 96
popis 95
vytvoření 96
vyhledávací pole XML
mapování prvků 93, 99
popis 93, 99

vyhledávací pole XML (*pokračování*)
vytvoření 93
vyhledávací servery
asociování slovníků slov pro upřednostnění 139, 140
asociování slovníků synonym 130
asociování slovníků zakázaných slov 132
historie doby odezvy 220
mezipaměť vyhledávání 128
monitorování 220
oblíbené dotazy 220
popis 127
poslední dotazy 220
slovníky slov pro upřednostnění 138
slovníky synonym 128, 130
slovníky zakázaných slov 131
spuštění 201, 220
stav systému 220
zastavení 203, 220
vyhledávání
dokumenty HTML 95, 96
dokumenty XML 93, 99
externí zdroje 117
kategorie 86
kolekce 117
rychlé odkazy 134
vyhledávání s použitím polí 117
Vyhledávání, popis stránky 1
výchozí vyhledávací aplikace 160
výsledky vyhledávání
dynamické řazení podle důležitosti 137
dynamické shrnutí 132
expanze zástupných znaků 121
hodnocení na základě textu 137
konfigurace tříd upřednostnění 142, 144, 145
konfigurace vzoru identifikátoru URI 140
popis 137
přízpusobení souhrnů 133
řazení podle důležitosti 141
sbalení 123, 124
seskupování 123, 124
souhrny 133
statické třídy důležitosti 138
zástupné znaky 119
výstrahy
doba odezvy vyhledávání 226
index, limity 205
indexované dokumenty 226
konfigurace serveru SMTP 229
na úrovni kolekce 205, 226
na úrovni systému 227
popis 225
prolezené dokumenty 226
příjem e-mailů 230
volby e-mailu 226, 227
volné místo na serverech 227
výstrahy týkající se volného místa 227
vytvoření
kategorie založené na pravidlech 91
kolekce 13, 14
obory 123
prolézací moduly 21
rychlé odkazy 135
vyhledávací pole HTML 96
vyhledávací pole XML 93
webový prolézací modul, sestavy 213

W

- webové prolézací moduly
 - aktivní weby 210, 212
 - cookie, konfigurace souboru 67
 - direktivy no-follow 69
 - direktivy no-index 69
 - formát souborů cookie 66
 - frekvence prolézání 213
 - globální prolézání prostor 67
 - historie prolézacího modulu 210
 - historie webu 210
 - JavaScript, podpora 56
 - konfigurace 54
 - méně závažné chyby, stránky 61
 - monitorování 210
 - naposledy prolézané adresy URL 210
 - návštěva adresy URL, co nejdříve 61
 - omezení prolézaného prostoru 56
 - opakované prolézání, intervaly 60
 - počáteční adresy URL 56, 61
 - podrobnosti o podprocesech 210, 211
 - prolézání, pravidla 56
 - proxy, servery 65
 - soubor followindex.rules 69
 - soubory cookie 66
 - stav adresy URL 210
 - stav systému 210
 - uživatelští agenti 55
 - vytváření sestav 213
 - weby chráněné heslem 63, 64
- WebSphere II Event Publisher Edition, konfigurace prolézacího modulu DB2 29
- WebSphere II OmniFind Edition 267
 - ID relací 237
 - integrace s produktem WebSphere Portal 187
 - konfigurace čísla portu 10
 - návratové kódy 237
 - příkazy 237
 - správa identit 180
 - usnadnění přístupu 267
 - změna hesla na jednom serveru 5
 - změna hesla na více serverech 7
- WebSphere MQ, konfigurace prolézacího modulu DB2 32
- WebSphere MQ, konfigurace prolézacího serveru 33, 34
- WebSphere Portal
 - implementace vyhledávacího portletu 189
 - integrace s podnikovým vyhledáváním 187
 - kategorie, migrace stromu 195
 - kolekce, migrace 195
 - portlet pro vyhledávání a procházení 191
 - portlet Taxonomy Management 195
 - taxonomie založené na modelu 195
 - taxonomie, migrace 195
 - výchozí nastavení migrace 199
- weby chráněné heslem 63
 - ověřování založené na použití formulářů 64
 - základní ověřování HTTP 63
- WpsMigratorLog.log, soubor 200

X

- XML, nativní syntaxe dotazů 104
- XML, nativní vyhledávání fragmentů XML 104
- XPath, nativní vyhledávání XML 104

Z

- zabezpečení
 - administrativní role 171
 - analýza kotvičích textu 175
 - detekce duplicitních dokumentů 174
 - dokumenty Lotus Domino 183
 - domény systému Windows 181
 - globální zabezpečení produktu WebSphere 173
 - moduly plug-in prolézacího modulu 75
 - na úrovni dokumentu 177, 178, 180, 185
 - na úrovni kolekce 174, 185
 - ověřování 172, 173
 - popis 169
 - povolení pro kolekci 14, 169
 - profily uživatelů 179
 - řízení přístupu 172
 - správa identit 180
 - ukázková vyhledávací aplikace 161
 - vyhledávací aplikace, ID 176
 - vynechání řízení přístupu na úrovni dokumentu 185
 - zakázání pro podnikovou aplikaci 173
- zabezpečení lokálních uživatelů, prolézací moduly QuickPlace 52
- zabezpečení na úrovni dokumentu
 - aktuální pověřovací údaje, ověření 178
 - dokumenty Lotus Domino 183
 - indexované řízení přístupu 178
 - konfigurace prolézacího modulu 19
 - moduly plug-in prolézacího modulu 75
 - ověřování v reálném čase 178
 - popis 169, 177
 - portálový vyhledávací stroj, podpora 189
 - pro dokumenty Lotus Domino 183
 - pro souborové systémy Windows 181
 - profily uživatelů 179
 - prvky zabezpečení 178
 - správa identit 180
- zabezpečení na úrovni kolekce
 - analýza kotvičích textu 175
 - detekce duplicitních dokumentů 174
 - ID aplikací 176
 - popis 169, 174
 - povolení 14
- Zabezpečení, popis pohledu 1
- zákaz časových plánů indexování 115
- základní ověřování HTTP 63
- zálohování podnikového vyhledávání 233
- zálohování systému 233
- zastavení
 - moduly analýzy 217
 - prolézací servery 209
 - servery podnikového vyhledávání 201, 203
 - sestavení indexu 218, 219
 - vyhledávací servery 220
- zástupné znaky
 - expanze dotazu 119, 121

- zástupné znaky (pokračování)
 - expanze indexu 119, 121

- zobrazení
 - podrobnosti identifikátoru URI 207
 - soubory žurnálu 231
- zpracování textu
 - anotační moduly 96
 - obecná struktura analýzy 96
 - stroj pro analýzu textu 96

Ž

- Žurnál, popis stránky 1



Vytištěno v Dánsku společností IBM Danmark A/S.



SC09-3744-02



Spine information:



WebSphere II OmniFind Edition

Správa podnikového vyhledávání

Verze 8.3