

IBM WebSphere Information Integrator
OmniFind Edition



Administrowanie wyszukiwaniem korporacyjnym

Wersja 8.3

IBM WebSphere Information Integrator
OmniFind Edition



Administrowanie wyszukiwaniem korporacyjnym

Wersja 8.3

Przed rozpoczęciem korzystania z niniejszych informacji oraz z produktu, którego one dotyczą, należy zapoznać się z informacjami ogólnymi w rozdziale "Uwagi".

Niniejszy dokument zawiera informacje dotyczące produktów firmy IBM. Są one udostępniane w ramach umowy licencyjnej i chronione prawem autorskim. Informacje zawarte w tej publikacji nie zawierają żadnych gwarancji dotyczących opisywanych produktów i żadnych zapisanych w niej stwierdzeń nie należy interpretować jako takich gwarancji.

Publikacje firmy IBM można zamówić drogą elektroniczną lub za pośrednictwem lokalnego przedstawiciela firmy IBM:

- Aby zamówić książki na stronie WWW, należy skorzystać ze strony IBM Publications Center pod adresem www.ibm.com/shop/publications/order.
- Aby znaleźć najbliższego lokalnego przedstawiciela firmy IBM, należy skorzystać z informacji umieszczonych na stronie IBM Directory of Worldwide Contacts pod adresem: www.ibm.com/planetwide.

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

© Copyright International Business Machines Corporation 2004, 2005. Wszelkie prawa zastrzeżone.

Spis treści

Informacje o przedstawionych tematach vii

Dla kogo przeznaczone są te tematy vii

Administrowanie wyszukiwaniem korporacyjnym - przegląd 1

Logowanie się do konsoli administracyjnej 5

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji jednoserwerowej 6

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji wieloserwerowej 7

Zmianianie numeru portu dla systemu wyszukiwania korporacyjnego 11

Kolekcje wyszukiwania korporacyjnego 13

Tworzenie kolekcji przy użyciu Kreatora kolekcji 13

Tworzenie kolekcji przy użyciu widoku kolekcji 14

Edytowanie kolekcji 16

Usuwanie kolekcji. 17

Administrowanie przeszukiwaczem wyszukiwania korporacyjnego. 19

Tworzenie przeszukiwacza 21

Edycja właściwości przeszukiwacza 22

Edytowanie przestrzeni przeszukiwania 23

Usuwanie przeszukiwacza 23

przeszukiwacze Content Edition 24

Dostęp do repozytoriów WebSphere II Content Edition w trybie serwera 25

Dostęp do repozytoriów WebSphere II Content Edition w trybie bezpośrednim 25

Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu WebSphere II Content Edition . . 26

Konfigurowanie serwera przeszukiwacza w systemie Windows dla programu WebSphere II Content Edition . 27

przeszukiwacze DB2 28

Konfigurowanie programu WebSphere Information Integrator Event Publisher Edition dla przeszukiwaczy DB2 30

Konfigurowanie programu WebSphere MQ dla przeszukiwaczy DB2 32

Konfigurowanie serwera przeszukiwacza w systemie UNIX do korzystania z publikowania zdarzeń 34

Konfigurowanie serwera przeszukiwacza w systemie Windows do korzystania z publikowania zdarzeń . . 35

przeszukiwacze DB2 Content Manager 36

Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu DB2 Content Manager 37

Konfigurowanie serwera przeszukiwacza dla programu DB2 Content Manager w systemie Windows 39

przeszukiwacze Domino Document Manager 40

Przeszukiwacze serwera Exchange Server 41

Sprawdzanie dostępu do zabezpieczonych dokumentów serwera Exchange 42

Przeszukiwacze NNTP 43

przeszukiwacze Notes. 43

Wskazówki dotyczące przeszukiwania baz danych Lotus Domino 45

Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino . . . 46

Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino . . 48

Konfigurowanie serwerów korzystających z protokołu DIIOP 50

Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino . . . 51

Przeszukiwacze QuickPlace 52

Konfigurowanie serwera QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego 54

Konfigurowanie opcji Directory Assistance na serwerze QuickPlace 55

przeszukiwacze systemu plików UNIX 55

przeszukiwacze sieci WWW. 56

Konfiguracja agenta użytkownika 57

Obsługa języka JavaScript 58

Reguły ograniczania przestrzeni przeszukiwania w sieci WWW 58

Testy połączeń URL przy użyciu przeszukiwacza sieci WWW 63

Ustawienia przeszukiwacza sieci WWW określające odstęp czasu przed ponownym przeszukiwaniem. . . 63

Opcje odwiedzania adresów URL przy użyciu przeszukiwacza sieci WWW. 64

Jak przeszukiwacz sieci WWW obsługuje strony błędów programowych 64

Przeszukiwanie zabezpieczonych serwisów WWW . . 66

Serwisy WWW obsługiwane przez serwery proxy . . 68

Administrowanie informacjami cookie. 69

Konfiguracja globalnej przestrzeni przeszukiwania sieci WWW 70

Dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie 72

Ignorowanie dyrektyw blokujących wykorzystywanie odsyłaczy i indeksowanie na stronach WWW. . . . 73

przeszukiwacze WebSphere Portal 73

Wdrażanie aplikacji korporacyjnej dla przeszukiwacza WebSphere Portal 74

Kopiowanie adresu URL do przeszukania z portalu WebSphere Portal 75

przeszukiwacze systemu plików Windows 76

Konfiguracja obsługi aplikacji nasłuchiwania danych . 77

Niestandardowe wtyczki przeszukiwacza 78

Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego 79

Administrowanie analizatorem składni wyszukiwania korporacyjnego. 87

Praca z kategoriami 88

Kategorie oparte na regułach. 89

Kategorie oparte na modelach 91

Drzewa kategorii	92	Konfiguracja pamięci podręcznej wyszukiwania	134
Wybór typu klasyfikacji	93	Niestandardowe słowniki synonimów	134
Konfiguracja kategorii	93	Dodawanie słowników synonimów do systemu	136
Praca z polami wyszukiwania XML	95	Powiązanie słownika synonimów z kolekcją	136
pola wyszukiwania XML	95	Niestandardowe słowniki słów ignorowanych	137
Odwzorowanie elementów XML na pola wyszukiwania	96	Dodawanie do systemu słowników słów	138
Praca z polami wyszukiwania HTML	98	ignorowanych.	138
pola wyszukiwania HTML	98	Powiązanie słownika słów ignorowanych z kolekcją	138
Odwzorowywanie elementów metadanych HTML na		Podsumowywanie dynamiczne.	139
pola wyszukiwania	98	Dostosowywanie podsumowań dokumentów w	
Niestandardowe przetwarzanie tekstu	99	konsoli administracyjnej	139
Dodawanie mechanizmów analizy tekstu do systemu	100	Dostosowywanie podsumowań dokumentów za	
Powiązanie mechanizmu analizy tekstu z kolekcją	102	pomocą edycji właściwości.	140
Odwzorowywanie elementów XML na wspólną		Praca z szybkimi odsyłaczami	141
strukturę analizy	103	Szybkie odsyłacze	141
Odwzorowywanie wspólnej struktury analizy na		Konfiguracja szybkich odsyłaczy	141
indeks	104		
Odwzorowywanie wspólnej struktury analizy na		Ocenianie dokumentów w	
tabele JDBC	105	wyszukiwaniu korporacyjnym	143
Konfigurowanie wątków dla usługi analizatora składni	106	Wartościowanie danych tekstowych	143
Włączanie zaawansowanej analizy terminów złożonych	106	Ocena statyczna	144
Włączanie obsługi rodzimego wyszukiwania XML.	107	Niestandardowe słowniki kryteriów	144
Analiza lingwistyczna dokumentów chińskich, japońskich		Dodawanie do systemu słowników kryteriów	145
i koreańskich	108	Powiązanie słownika kryteriów z kolekcją	146
Podział na ciągi n-znakowe.	108	Ocena dokumentów na podstawie wzorców URI	146
Usuwanie znaków nowego wiersza z zakresów		Wpływanie na wyniki dokumentów zgodnych z	
znaków innych niż ASCII	108	wzorcami URI	147
Typy dokumentów powiązane z analizatorami składni		Ocenianie dokumentów na podstawie klas kryteriów	148
kolekcji i sesjami Stellent	109	Odwzorowywanie pól na klasy kryteriów	150
Powiązanie typów dokumentów z analizatorem		Konfigurowanie współczynników kryteriów dla klas	
składni kolekcji	110	kryteriów	151
Domyślne reguły usługi analizatora składni kolekcji	111	Domyślne wartości klas kryteriów.	152
Powiązanie typów dokumentów z sesją Stellent.	112		
Domyślne reguły analizy składni dla sesji Stellent	114	Aplikacje wyszukujące dla	
		wyszukiwania korporacyjnego	155
Administrowanie indeksem		Przypisywanie aplikacji wyszukujących do kolekcji	156
wyszukiwania korporacyjnego	117	Funkcje przykładowej aplikacji wyszukującej	156
Planowanie operacji budowania indeksu	118	Właściwości przykładowej aplikacji wyszukującej	158
Zmiana harmonogramu operacji na indeksie.	119	Edytowanie właściwości przykładowej aplikacji	
Włączanie i wyłączanie harmonogramów operacji na		wyszukującej	166
indeksach	119	Uzyskiwanie dostępu do przykładowej aplikacji	
Konfiguracja współbieżnych operacji budowania		wyszukującej	167
indeksu	120	Włączanie zabezpieczeń dla przykładowej aplikacji	
Opcje wpływające na podlegający wyszukiwaniu widok		wyszukującej	168
indeksu	121		
Indeksowane opcje wyszukiwania dokumentów	121	Zewnętrzne źródła wyszukiwania	
Znaki zastępcze w zapytaniach.	123	korporacyjnego	171
Konfigurowanie opcji dotyczących znaków		Dodawanie do systemu źródeł zewnętrznych	171
zastępczych w zapytaniach	126	Powiązanie aplikacji wyszukujących ze źródłami	
Zasięgi.	126	zewnętrznymi.	172
Konfiguracja zasięgów	127		
Zwinięte identyfikatory URI	128	Zabezpieczenia wyszukiwania	
Zwijanie identyfikatorów URI w wynikach		korporacyjnego	175
wyszukiwania.	129	Role administracyjne	176
Usuwanie z indeksu identyfikatorów URI	130	Konfiguracja użytkowników z uprawnieniami	
		administracyjnymi	178
Administrowanie serwerem		Uwierzytelnianie a kontrola dostępu	178
wyszukiwania dla wyszukiwania		Wyłączanie zabezpieczenia dla aplikacji korporacyjnej na	
korporacyjnego	133	serwerze WebSphere Application Server	179
Pamięci podręcznej wyszukiwania	134		

Zabezpieczenia na poziomie kolekcji	181	Szybkość przeszukiwania przeszukiwacza sieci	
Analiza duplikatów dokumentów	181	WWW	221
Analiza tekstu zakotwiczenia	181	Tworzenie raportów przeszukiwacza sieci WWW	222
Indeksowanie tekstu zakotwiczenia w odsyłaczach do		Kody powrotu HTTP przeszukiwacza sieci WWW	223
dokumentów zabronionych	182	Monitorowanie analizatora składni	226
Zabezpieczenia przy użyciu identyfikatorów aplikacji		Monitorowanie aktywności indeksu dla kolekcji	228
wyszukującej	183	Monitorowanie kolejki indeksu wyszukiwania	
Zabezpieczenia na poziomie dokumentu	184	korporacyjnego	229
Sprawdzanie poprawności na podstawie zapisanych		Monitorowanie serwerów wyszukiwania	229
tokenów zabezpieczeń	184	Monitorowanie nasłuchiwania danych	231
Sprawdzanie poprawności bieżących referencji		Śledzenie dokumentu	231
podczas przetwarzania zapytań	185	Konfigurowanie plików dziennika do śledzenia	
Wymuszanie zabezpieczeń na poziomie dokumentu		dokumentów	232
dla dokumentów systemu plików Windows	188	Wyświetlanie raportów o usuniętych dokumentach	233
Wymuszanie zabezpieczeń na poziomie dokumentu			
dla dokumentów produktu Lotus Domino	190		
Wyłączanie zabezpieczeń na poziomie dokumentu	192		
		Alerty i pliki dziennika wyszukiwania	
		 korporacyjnego	235
		Alerty	235
		Konfiguracja alertów na poziomie kolekcji	236
		Konfiguracja alertów na poziomie systemu	237
		Konfiguracja plików dziennika	238
		Konfigurowanie informacji o serwerze SMTP	239
		Odbieranie wiadomości poczty elektronicznej	
		dotyczących zarejestrowanych komunikatów	240
		Wyświetlanie plików dziennika	242
		Tworzenie i odtwarzanie kopii	
		 zapasowej systemu wyszukiwania	
		 korporacyjnego	245
		Tworzenie kopii zapasowej wyszukiwania	
		korporacyjnego	246
		Odtwarzanie systemu wyszukiwania korporacyjnego	246
		Odtwarzanie plików systemu wyszukiwania	
		korporacyjnego na nowych serwerach	247
		Komendy, kody powrotu i	
		 identyfikatory sesji wyszukiwania	
		 korporacyjnego	249
		Dokumentacja wyszukiwania	
		 korporacyjnego	277
		Ułatwienia dostępu w produkcie	
		 WebSphere II OmniFind Edition.	279
		Glosariusz terminów związanych z	
		 wyszukiwaniem korporacyjnym.	281
		Uzyskiwanie dostępu do informacji o	
		 produktach WebSphere Information	
		 Integration	291
		Komentarze na temat dokumentacji	293
		Kontakt z firmą IBM.	295
		Znaki towarowe	297
Integracja wyszukiwania			
 korporacyjnego z portalem			
 WebSphere Portal	195		
Zabezpieczenia na poziomie dokumentu przy użyciu			
mechanizmu wyszukiwania środowiska Portal	197		
Wdrażanie portletu Wyszukiwanie	198		
Konfigurowanie portletu Wyszukiwanie i przeglądanie			
produktu WebSphere Portal dla wyszukiwania			
korporacyjnego	199		
Instalowanie adaptera wyszukiwania korporacyjnego dla			
Centrum wyszukiwania	200		
Instalowanie portletu rejestracji wyszukiwania			
korporacyjnego dla Centrum wyszukiwania	202		
Migracja z portalu WebSphere Portal			
 do wyszukiwania korporacyjnego.	203		
Migrowanie systematyki opartej na modelach z portalu			
WebSphere Portal	203		
Migrowanie kolekcji z portalu WebSphere Portal	205		
Ustawienia zmigrowanej kolekcji	207		
Plik dziennika Kreatora migracji	208		
Uruchamianie i zatrzymywanie			
 serwerów wyszukiwania			
 korporacyjnego	209		
Uruchamianie serwerów wyszukiwania korporacyjnego	209		
Zatrzymywanie serwerów wyszukiwania korporacyjnego	211		
Monitorowanie działań związanych z			
 wyszukiwaniem korporacyjnym.	213		
Szacowanie liczby dokumentów w kolekcji	213		
Sprawdzanie dostępności zasobów systemu	214		
Monitorowanie kolekcji.	215		
Wyświetlanie szczegółów dotyczących identyfikatora			
URI.	216		
Monitorowanie przeszukiwaczy	217		
Wyświetlanie szczegółów na temat działalności			
przeszukiwacza sieci WWW	219		
Szczegółowe informacje o wątkach przeszukiwacza			
sieci WWW	220		
Aktywne serwisy przeszukiwacza sieci WWW	221		

Indeks 303

Informacje o przedstawionych tematach

Przedstawione tematy są przeznaczone dla administratorów systemu IBM WebSphere Information Integrator OmniFind Edition, wersja 8.3.

W programie WebSphere Information Integrator OmniFind Edition (WebSphere II OmniFind Edition) zastosowano technologię *wyszukiwania korporacyjnego*. Komponenty wyszukiwania korporacyjnego są instalowane razem z programem WebSphere II OmniFind Edition. Nazwa *wyszukiwanie korporacyjne* jest używana w całej dokumentacji programu WebSphere II OmniFind Edition z wyjątkiem odwołań do ścieżek instalacyjnych i etykiet na opakowaniu produktu.

W dokumentacji wyszukiwania korporacyjnego dla administratorów opisano następujące tematy:

- wprowadzenie do pojęć związanych z wyszukiwaniem korporacyjnym, opis jego komponentów i możliwości
- instrukcje dotyczące tworzenia, monitorowania i administrowania kolekcjami
- informacje na temat sposobu analizowania danych i dostosowania działań analitycznych optymalizujących wyszukiwanie i pobieranie
- informacje na temat sposobu indeksowania analizowanych danych i administrowania operacjami indeksowania
- informacje na temat sposobu wyszukiwania, oceniania oraz zwracania wyników wyszukiwania przez serwery wyszukiwania
- informacje na temat przypisywania niestandardowych aplikacji wyszukujących do kolekcji
- przegląd różnych poziomów zabezpieczeń dostępnych w wyszukiwaniu korporacyjnym
- informacje na temat integracji wyszukiwania korporacyjnego z produktem IBM WebSphere Portal
- instrukcje dotyczące migrowania systematyk i kolekcji z produktu WebSphere Portal do drzew kategorii i kolekcji wyszukiwania korporacyjnego
- instrukcje dotyczące tworzenia kopii zapasowej i odtwarzania systemu
- instrukcje dotyczące tworzenia i wyświetlania plików protokołu
- instrukcje dotyczące używania komend wyszukiwania korporacyjnego i interpretowania zwracanych przez nie informacji.

Dla kogo przeznaczone są te tematy

Te informacje są przeznaczone dla administratorów i operatorów systemu odpowiedzialnych za tworzenie, monitorowanie i administrowanie kolekcjami wyszukiwania korporacyjnego.

Z tych informacji można korzystać przy tworzeniu kolekcji, wybieraniu treści kolekcji i konfigurowaniu opcji przygotowujących treść do wyszukiwania. Są one również przydatne przy monitorowaniu kolekcji i działania systemu, rejestrowaniu użytkowników jako administratorów wyszukiwania korporacyjnego oraz przypisywaniu kolekcji i zewnętrznych źródeł do aplikacji wyszukujących.

Aby efektywnie wykorzystać te informacje, konieczna jest znajomość aplikacji WWW oraz doświadczenie w dziedzinie przeszukiwanych źródeł danych.

Administrowanie wyszukiwaniem korporacyjnym - przegląd


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

Kolekcje, widok

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

Logowanie

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

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


Tworzenie kolekcji

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

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

Konfigurowanie kolekcji

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

Kliknij opcję  **Edytuj** dla nowej kolekcji, a następnie wybierz odpowiednią stronę, aby określić opcje dla tej kolekcji.

Ważne: Kliknięcie przycisku Wstecz lub Odśwież w przeglądarce WWW może spowodować otrzymanie niespójnych wyników lub utratę danych. Jeśli podczas konfigurowania kolekcji trzeba wrócić do poprzedniej strony lub odświeżyć informacje w konsoli administracyjnej, należy kliknąć przycisk **Wstecz** lub **Odśwież** w konsoli administracyjnej wyszukiwania korporacyjnego zamiast tych przycisków w przeglądarce WWW.

- Na stronie Ogólne można określić opcje mające zastosowanie do całej kolekcji:
 - Opcje ogólne można edytować i dzięki temu zmienić nazwę lub opis kolekcji bądź jej szacowaną wielkość.
 - Opcje, których nie można zmienić, można przeglądać. Jest to na przykład identyfikator kolekcji lub metoda oceny statycznej służąca do oceny dokumentów w wynikach wyszukiwania.

– W przypadku kolekcji, dla których włączono zabezpieczenia w czasie tworzenia, można włączyć lub wyłączyć elementy sterujące zabezpieczeniami na poziomie dokumentu.

- Na stronie Przeszukiwanie należy skonfigurować przynajmniej jeden przeszukiwacz.

W pojedynczej kolekcji mogą znajdować się dane pochodzące z różnych źródeł danych. Dla każdego typu źródła danych, które ma zostać uwzględnione w kolekcji, należy skonfigurować co najmniej jeden przeszukiwacz. Po utworzeniu przeszukiwacza można go skonfigurować, korzystając z kreatora przeznaczonego dla określonego typu przeszukiwanych danych.

- Na stronie Analiza składni można skonfigurować opcje dotyczące sposobu analizowania przeszukiwanych danych, aby zapewnić ich wydajne wyszukiwanie:

– Określenie, czy dokumenty XML mają być analizowane, aby umożliwić ich wyszukiwanie za pomocą rodzimego wyszukiwania XML.

– Przypisanie dokumentów do kategorii, dzięki czemu użytkownicy mogą przeszukiwać podzbiór kolekcji lub przeglądać dokumenty będące wynikiem wyszukiwania według kategorii, do których należą.

– Odwzorowanie elementów XML i elementów metadanych HTML na pola wyszukiwania w indeksie, co umożliwia użytkownikom określenie nazw pól w zapytaniach i wyszukiwanie konkretnych części dokumentów.

– Jeśli do systemu wyszukiwania korporacyjnego zostały dodane niestandardowe mechanizmy analizy tekstu, można wybrać jeden z nich do użycia z daną kolekcją, a następnie określić opcje przetwarzania tekstu, aby poprawić efektywność wydobywania informacji.

– Powiązanie pól z klasami kryteriów, zapewniające wpływ na ocenę pól zgodnych z terminami użytymi w zapytaniu w wynikach wyszukiwania.

- Na stronie Indeks należy określić harmonogramy reorganizowania i odświeżania indeksu. Aby zapewnić użytkownikom dostęp do najnowszych informacji, należy zaplanować regularne wykonywanie operacji tworzenia indeksu. Opcjonalnie można także wykonać następujące czynności:

– Umożliwienie użytkownikom stosowania znaków zastępczych w zapytaniach.

– Skonfigurowanie zasięgów, aby użytkownicy przeszukiwali tylko ograniczoną część kolekcji zamiast wszystkich dokumentów w indeksie.

– Włączenie zwijania wyników wyszukiwania, aby wszystkie dokumenty z tego samego źródła były zgrupowane i zwinięte razem w wynikach wyszukiwania.

– Przypisanie współczynników kryteriów w celu oddziaływania na ocenę dokumentów zgodnych ze wzorcem identyfikatora URI w wynikach wyszukiwania.

– Usunięcie identyfikatorów URI z indeksu. Może to mieć na celu uniemożliwienie użytkownikom wyświetlenia określonych dokumentów po utworzeniu kolekcji.

- Na stronie Wyszukiwanie można określić opcje dotyczące wyszukiwania dokumentów w kolekcji:

– Ustawienie wielkości pamięci podręcznej przeznaczonej dla wyników wyszukiwania i zmiana domyślnego języka kolekcji.

– Jeśli do systemu wyszukiwania korporacyjnego zostały dodane niestandardowe słowniki synonimów, słów ignorowanych i słów-kryteriów, można wybrać te słowniki do używania podczas przeszukiwania kolekcji przez użytkowników.

– Określenie długości podsumowania dokumentu wyświetlanego w wynikach wyszukiwania.

- Skonfigurowanie szybkich odsyłaczy w celu automatycznego wyświetlania konkretnych identyfikatorów URI w wynikach wyszukiwania za każdym razem, gdy w zapytaniu wystąpią określone słowa kluczowe lub wyrażenia.
- Na stronie Rejestrowanie można wykonać następujące czynności:
 - Określenie opcji dotyczących typów komunikatów, które mają być umieszczane w dziennikach, oraz częstotliwości wtórnego przetwarzania plików dzienników.
 - Określenie opcji dotyczących otrzymywania alertów dotyczących działań związanych z kolekcją. Na przykład w alercie może zostać przesłana informacja o przekroczeniu średniego czasu odpowiedzi wyszukiwania.
 - Określenie opcji dotyczących otrzymywania wiadomości e-mail po zapisaniu w dzienniku określonych komunikatów lub komunikatów określonego typu.
 - Określenie opcji dotyczących rejestrowania informacji umożliwiających stwierdzenie kiedy, gdzie i z jakiego powodu dokument został usunięty z systemu wyszukiwania korporacyjnego.

Uruchamianie komponentów

Po określeniu źródeł danych, które mają zostać przeszukane, oraz opcji zbierania i wyszukiwania danych można uruchomić procesy tworzenia kolekcji. Bardzo ważna jest kolejność uruchamiania komponentów. Przeszukiwacze muszą najpierw przeszukać dane, aby można było poddać je analizie. Następnie analizatory składni muszą dokonać analizy przeszukanych danych, aby można było wykonać ich indeksowanie. Na koniec indeks musi zostać odświeżony lub zreorganizowany, aby serwery wyszukiwania mogły rozpocząć przetwarzanie żądań wyszukiwania.

Widok Źródła zewnętrzne

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

Widok System

Użytkownik z przypisaną rolą administratora wyszukiwania korporacyjnego może kliknąć ikonę **System** na pasku narzędzi i wykonać następujące czynności (administratorzy i operatorzy kolekcji oraz użytkownicy je monitorujący mogą uzyskać dostęp do tego widoku, tylko jeśli administrator wyszukiwania korporacyjnego nada im takie uprawnienia):

- Sprawdzenie dostępności zasobów systemowych.
- Skonfigurowanie aplikacji klienckich do nasłuchiwania danych w celu umożliwienia aktualizowania kolekcji.
- Dodanie do systemu niestandardowych mechanizmów analizy tekstu.
- Dodanie do systemu niestandardowych słowników synonimów, słów ignorowanych i słów-kryteriów.
- Określenie liczby kolekcji, dla których mogą być równoległe budowane indeksy, a także określenie, czy żądania odświeżania i reorganizacji dla pojedynczej kolekcji mogą przebiegać współbieżnie.

- Skonfigurowanie alertów dla zdarzeń systemowych.
- Określenie opcji rejestrowania komunikatów generowanych przez sesje systemowe.
- Określenie informacji o serwerze poczty w celu otrzymywania wiadomości e-mail dotyczących działań z zakresu wyszukiwania korporacyjnego.


Widok Zabezpieczenia

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

Jeśli na serwerze IBM WebSphere Application Server zostanie włączona funkcja zabezpieczeń, wówczas w widoku Zabezpieczenia można skonfigurować role administracyjne. Konfigurując role administracyjne, można zezwolić dodatkowym użytkownikom na administrowanie systemem i dla każdego z nich określić pewne ograniczenia dostępu do konkretnych funkcji i kolekcji.

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

Widok Monitor

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

Pojęcia pokrewne

“Administrowanie przeszukiwaczem wyszukiwania korporacyjnego” na stronie 19
Konfiguracja przeszukiwaczy dla różnych typów danych, które mają zostać włączone do kolekcji. Pojedyncza kolekcja może zawierać dowolną liczbę przeszukiwaczy.

“Monitorowanie działań związanych z wyszukiwaniem korporacyjnym” na stronie 213
Monitorując działania systemu i działania związane z kolekcjami, można wyświetlić informacje o statusie różnych procesów, obserwować zmienne w celu wykrycia potencjalnych problemów lub dopasować ustawienia konfiguracyjne w celu poprawienia wydajności.

Zadania pokrewne

“Uruchamianie serwerów wyszukiwania korporacyjnego” na stronie 209
Aby umożliwić użytkownikom wyszukiwanie w kolekcji, należy uruchomić procesy systemowe, a następnie uruchomić serwery przeszukiwania, analizy, indeksowania i wyszukiwania w kolekcji.

“Zatrzymywanie serwerów wyszukiwania korporacyjnego” na stronie 211
Może wystąpić konieczność zatrzymania i zrestartowania serwera wyszukiwania korporacyjnego, jeśli dokonano zmian w jego konfiguracji lub w celu rozwiązania problemów.

Tworzenie kolekcji przy użyciu Kreatora kolekcji

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

“Tworzenie kolekcji przy użyciu widoku kolekcji” na stronie 14

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

Logowanie się do konsoli administracyjnej

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

Zagadnienia wstępne

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

- Jeśli nie włączono zabezpieczeń globalnych na serwerze WebSphere Application Server, dostęp do konsoli administracyjnej będzie miał wyłącznie administrator wyszukiwania korporacyjnego wyznaczony podczas instalacji produktu WebSphere II OmniFind Edition.
- Jeśli włączono globalne zabezpieczenia serwera WebSphere Application Server, można użyć konsoli administracyjnej wyszukiwania korporacyjnego do konfigurowania ról administracyjnych. Konfigurowane identyfikatory użytkowników muszą figurować w rejestrze użytkowników serwera WebSphere Application Server. Skonfigurowanie ról administracyjnych umożliwi większej liczbie użytkowników logowanie się w konsoli administracyjnej. Można jednak przy tym zdefiniować funkcje i kolekcje dostępne dla poszczególnych administratorów.

Procedura

Aby zalogować się w konsoli administracyjnej wyszukiwania korporacyjnego:

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

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

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

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

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

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

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

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

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

Pojęcia pokrewne

“Role administracyjne” na stronie 176

Wyszukiwanie korporacyjne wykorzystuje role jako metodę kontroli dostępu do różnych funkcji konsoli administracyjnej.

Zadania pokrewne

“Uruchamianie serwerów wyszukiwania korporacyjnego” na stronie 209
Aby umożliwić użytkownikom wyszukiwanie w kolekcji, należy uruchomić procesy systemowe, a następnie uruchomić serwery przeszukiwania, analizy, indeksowania i wyszukiwania w kolekcji.

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji jednoserwerowej

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

Zagadnienia wstępne

Identyfikator i hasło administratora wyszukiwania korporacyjnego muszą być poprawne w używanym systemie operacyjnym i muszą umożliwiać dostęp do programu DB2 Universal Database oraz jego konfigurowanie.

Informacje o tym zadaniu

Hasło dla pierwszego identyfikatora administratora wyszukiwania korporacyjnego jest określane podczas instalacji produktu WebSphere II OmniFind Edition.

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

ES_INSTALL_ROOT

Katalog instalacyjny wyszukiwania korporacyjnego.

ES_NODE_ROOT

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

Ponieważ skrypt `eschangepw` jest instalowany w katalogu `ES_INSTALL_ROOT/bin`, można go uruchomić z dowolnego miejsca w systemie.

Procedura

Aby zmienić hasło administratora wyszukiwania korporacyjnego w konfiguracji jednoserwerowej:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego.
2. Zatrzymaj serwer, wprowadzając komendę: `esadmin stop`
3. Otwórz konsolę administracyjną serwera WebSphere Application Server i zatrzymaj aplikacje korporacyjne `server1` i `ESSearchServer`.
4. Zmień hasło administratora wyszukiwania korporacyjnego za pomocą komend systemu operacyjnego UNIX lub narzędzia do zmiany haseł dostępnego w systemie Microsoft Windows.
5. Uruchom następujący skrypt, gdzie *nowaWartość* to hasło podane w kroku 4:

System operacyjny	Komenda
UNIX	<code>eschangepw.sh nowaWartość</code>
Windows	<code>eschangepw nowaWartość</code>

6. W konsoli administracyjnej serwera WebSphere Application Server uruchom aplikacje `server1` i `ESSearchServer`.

7. Wykorzystaj ponownie warstwę CCL produktu WebSphere II OmniFind Edition, wprowadzając następujące komendy:

System operacyjny	Komendy
UNIX	stopccl.sh, następnie startccl.sh -bg
Wiersz komend systemu Windows	stopccl, następnie startccl
Narzędzie administracyjne usług systemu Windows	<ol style="list-style-type: none"> 1. Uruchom usługi systemu Windows. 2. Kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Zatrzymaj. 3. Ponownie kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Właściwości. 4. Kliknij kartę Logowanie. 5. Podaj nowe hasło <i>nowaWartość</i> i kliknij przycisk OK. 6. Ponownie kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Uruchom.

8. Uruchom ponownie system wyszukiwania korporacyjnego, wpisując komendę: esadmin start.

Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

Zmiana hasła administratora wyszukiwania korporacyjnego w konfiguracji wieloserwerowej

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

Zagadnienia wstępne

Identyfikator i hasło administratora wyszukiwania korporacyjnego muszą być poprawne w używanym systemie operacyjnym i muszą umożliwiać dostęp do programu DB2 Universal Database oraz jego konfigurowanie.

Hasło administratora wyszukiwania korporacyjnego musi być jednakowe na wszystkich komputerach należących do danej instalacji produktu WebSphere II OmniFind Edition.

Informacje o tym zadaniu

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

Aby zmienić hasło i wprowadzić tę zmianę w całym systemie wyszukiwania korporacyjnego, należy uruchomić skrypt eschangepw na każdym komputerze w danej konfiguracji wieloserwerowej. Program instalacyjny tworzy dwie zmienne środowiskowe, które mogą zostać użyte w skrypcie eschangepw:

ES_INSTALL_ROOT

Katalog instalacyjny wyszukiwania korporacyjnego.

ES_NODE_ROOT

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

Ponieważ skrypt eschangepw jest instalowany w katalogu ES_INSTALL_ROOT/bin, można go uruchomić z dowolnego miejsca w systemie.

Procedura

Aby zmienić hasło administratora wyszukiwania korporacyjnego w konfiguracji wieloserwerowej:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego na komputerze, na którym zainstalowano serwer indeksowania.
2. Uruchom tryb konserwacji systemu, wprowadzając komendę: `esadmin system maintenance`
3. Zmień hasło administratora wyszukiwania korporacyjnego za pomocą komend systemu operacyjnego UNIX lub narzędzia do zmiany haseł dostępnego w systemie Microsoft Windows.
4. Uruchom następujący skrypt, gdzie *nowaWartość* to hasło podane w kroku 3:

System operacyjny	Komenda
UNIX	<code>eschangepw.sh nowaWartość</code>
Windows	<code>eschangepw nowaWartość</code>

5. Wykorzystaj ponownie warstwę CCL produktu WebSphere II OmniFind Edition, wprowadzając następujące komendy na serwerze indeksowania:

System operacyjny	Komendy
UNIX	<code>stopccl.sh</code> , następnie <code>startccl.sh -bg</code>
Wiersz komend systemu Windows	<code>stopccl</code> , następnie <code>startccl</code>
Narzędzie administracyjne usług systemu Windows	<ol style="list-style-type: none"> 1. Uruchom usługi systemu Windows. 2. Kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Zatrzymaj. 3. Ponownie kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Właściwości. 4. Kliknij kartę Logowanie. 5. Podaj nowe hasło <i>nowaWartość</i> i kliknij przycisk OK. 6. Ponownie kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Uruchom.

6. Zaloguj się jako administrator wyszukiwania korporacyjnego na komputerze, na którym zainstalowano serwer przeszukiwacza.
7. Zmień hasło administratora wyszukiwania korporacyjnego za pomocą komend systemu operacyjnego UNIX lub narzędzia do zmiany haseł dostępnego w systemie Windows. Hasło musi być zgodne z hasłem podanym w kroku 3 na stronie 8.
8. Uruchom następujący skrypt, gdzie *nowaWartość* to hasło podane w kroku 3 na stronie 8:

System operacyjny	Komenda
UNIX	eschangepw.sh <i>nowaWartość</i>
Windows	eschangepw <i>nowaWartość</i>

9. Wykorzystaj ponownie warstwę CCL na serwerze przeszukiwacza, wprowadzając następujące komendy:

System operacyjny	Komendy
UNIX	stopccl.sh, następnie startccl.sh -bg
Wiersz komend systemu Windows	stopccl, następnie startccl
Narzędzie administracyjne usług systemu Windows	<ol style="list-style-type: none"> 1. Uruchom usługi systemu Windows. 2. Kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Zatrzymaj. 3. Ponownie kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Właściwości. 4. Kliknij kartę Logowanie. 5. Podaj nowe hasło <i>nowaWartość</i> i kliknij przycisk OK. 6. Ponownie kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Uruchom.

10. Zaloguj się jako administrator wyszukiwania korporacyjnego na jednym z komputerów, na których zainstalowano serwer wyszukiwania.
11. Zatrzymaj serwer IBM HTTP Web Server. Dzięki temu sieciowy program rozsyłający nie będzie kierował zapytań do tego serwera.
12. Zatrzymaj warstwę CCL na serwerze wyszukiwania, wprowadzając następujące komendy:

System operacyjny	Komendy
UNIX	stopccl.sh
Wiersz komend systemu Windows	stopccl
Narzędzie administracyjne usług systemu Windows	<ol style="list-style-type: none"> 1. Uruchom usługi systemu Windows. 2. Kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Zatrzymaj.

13. Otwórz konsolę administracyjną serwera WebSphere Application Server i zatrzymaj aplikacje server1 i ESSearchServer.
14. Zmień hasło administratora wyszukiwania korporacyjnego za pomocą komend systemu operacyjnego UNIX lub narzędzia do zmiany haseł dostępnego w systemie Windows. Hasło musi być zgodne z hasłem podanym w kroku 3 na stronie 8.
15. Uruchom następujący skrypt, gdzie *nowaWartość* to hasło podane w kroku 3 na stronie 8:

System operacyjny	Komenda
UNIX	eschangepw.sh <i>nowaWartość</i>
Windows	eschangepw <i>nowaWartość</i>

16. Zrestartuj warstwę CCL na serwerze wyszukiwania, wprowadzając następujące komendy:

System operacyjny	Komendy
UNIX	startccl.sh -bg
Wiersz komend systemu Windows	startccl
Narzędzie administracyjne usług systemu Windows	<ol style="list-style-type: none"> 1. Uruchom usługi systemu Windows. 2. Kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Właściwości. 3. Kliknij kartę Logowanie. 4. Podaj nowe hasło <i>nowaWartość</i> i kliknij przycisk OK. 5. Ponownie kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Uruchom.

17. Powtórz kroki od 10 na stronie 9 do 16 na drugim serwerze wyszukiwania.
18. Uruchom serwer IBM HTTP Web Server.
19. W konsoli administracyjnej serwera WebSphere Application Server uruchom aplikacje korporacyjne server1 i ESSearchServer.
20. Zaloguj się jako administrator wyszukiwania korporacyjnego na komputerze, na którym zainstalowano serwer indeksowania, i wprowadź następującą komendę, aby uruchomić wyszukiwanie korporacyjne: `esadmin start`
21. Zaloguj się w konsoli administracyjnej wyszukiwania korporacyjnego, wykonaj monitorowanie kolekcji, kliknij stronę Wyszukiwanie, a następnie kliknij kolejno przyciski **Zatrzymaj** i **Uruchom**, aby zrestartować serwery wyszukiwania. Powtórz ten krok dla każdej kolekcji w systemie wyszukiwania korporacyjnego.

Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

Zmianianie numeru portu dla systemu wyszukiwania korporacyjnego

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

Informacje o tym zadaniu

Numer portu dla systemu wyszukiwania korporacyjnego jest określany podczas instalacji produktu WebSphere II OmniFind Edition. (W przypadku konfiguracji wieloserwerowej ten sam numer portu jest określany dla wszystkich serwerów).

Jeśli nie można użyć numeru portu (na przykład numer portu mógł zostać przypisany do innego produktu na tym samym serwerze), występujący konflikt spowoduje zapisanie następującego komunikatu o błędzie w pliku CCLServer_data.log (gdzie zmienna *data* określa datę utworzenia pliku dziennika):

```
FFQ00273W Internal warning java.net.BindException: Address already in use: 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)
```

Procedura

Aby zmienić numer portu używany przez wyszukiwanie korporacyjne:

1. Wprowadź następującą komendę w celu zatrzymania systemu wyszukiwania korporacyjnego:
esadmin stop
2. Przejdź do serwera, na którym należy zmienić numer portu, i zatrzymaj wspólną warstwę komunikacji (CCL, Common Communication Layer), wprowadzając następujące komendy:

System operacyjny	Komendy
UNIX	stopccl.sh
Wiersz komend systemu Windows	stopccl
Narzędzie administracyjne usług systemu Windows	1. Uruchom usługi systemu Windows. 2. Kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Zatrzymaj .

3. Zmodyfikuj plik \$ES_NODE_ROOT/nodeinfo/es.cfg (w systemie Unix) lub plik %ES_NODE_ROOT%\nodeinfo\es.cfg (w systemie Windows), znajdź następującą właściwość, określ nową wartość numeru portu, a następnie zapisz i zamknij plik:

```
CCLPort=nowy_numer_portu
```

4. Zrestartuj warstwę CCL, wprowadzając następujące komendy:

System operacyjny	Komendy
UNIX	startccl.sh -bg
Wiersz komend systemu Windows	startccl

System operacyjny	Komendy
Narzędzie administracyjne usług systemu Windows	<ol style="list-style-type: none"> 1. Uruchom usługi systemu Windows. 2. Kliknij prawym przyciskiem myszy pozycję WebSphere Information Integrator OmniFind Edition i wybierz opcję Uruchom.

5. Przejdź do serwera indeksowania i postępuj zgodnie z instrukcjami podanymi w kroku 2 na stronie 11, aby zatrzymać warstwę CCL.
6. Zmodyfikuj plik \$ES_NODE_ROOT/nodeinfo/es.cfg (w systemie Unix) lub plik %ES_NODE_ROOT%\nodeinfo\es.cfg (w systemie Windows).
 - a. Znajdź następującą właściwość, gdzie zmienna *nazwa_komputera* oznacza nazwę serwera, na którym został zmodyfikowany numer portu w kroku 3 na stronie 11. Zmienna *N* we właściwości *nodeN* jest numerem identyfikującym serwer.

```
nodeN.destination=nazwa_komputera
```
 - b. Znajdź następującą właściwość podrzędną, podaj ten sam numer portu, który został określony dla serwera w kroku 3 na stronie 11, a następnie zapisz i zamknij plik:

```
nodeN.port=nowy_numer_portu
```
7. Postępuj zgodnie z instrukcjami podanymi w kroku 4 na stronie 11, aby zrestartować warstwę CCL.
8. Wprowadź następującą komendę w celu zrestartowania systemu wyszukiwania korporacyjnego:

```
esadmin start
```

Po wykonaniu tej komendy nowy numer portu zostanie zaktualizowany na wszystkich serwerach wyszukiwania korporacyjnego.

Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

Kolekcje wyszukiwania korporacyjnego

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

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

Istnieją dwa sposoby dodawania kolekcji do wyszukiwania korporacyjnego:

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

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

Stowarzyszenie kolekcji

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

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

Zadania pokrewne

“Monitorowanie kolekcji” na stronie 215

Użytkownik może wyświetlić ogólne informacje na temat statusu poszczególnych komponentów kolekcji albo wybrać opcje dotyczące wyświetlania szczegółowych informacji o wybranych komponentach i identyfikatorach URI.

“Migrowanie kolekcji z portalu WebSphere Portal” na stronie 205

Aby przeprowadzić migrację kolekcji z portalu WebSphere Portal do wyszukiwania korporacyjnego, należy przygotować kolekcje w portalu WebSphere Portal, a następnie użyć Kreatora migracji do wykonania migracji.

Tworzenie kolekcji przy użyciu Kreatora kolekcji

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

Zagadnienia wstępne

Aby utworzyć kolekcję, wymagana jest rola administratora wyszukiwania korporacyjnego.


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

Informacje o tym zadaniu

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

Procedura

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

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

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

Tworzenie kolekcji przy użyciu widoku kolekcji

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

Zagadnienia wstępne

Aby utworzyć kolekcję, wymagana jest rola administratora wyszukiwania korporacyjnego.

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


Informacje o tym zadaniu

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

Procedura

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

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

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

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

Następnie należy uruchomić procesy przeszukiwania, analizy składni, indeksowania i wyszukiwania w kolekcji. Do momentu, gdy kolekcja będzie gotowa do wyszukiwania za pomocą niestandardowych aplikacji wyszukujących użytkowników, do wyszukiwania w kolekcji można używać przykładowej aplikacji wyszukującej.

Pojęcia pokrewne

“Administrowanie wyszukiwaniem korporacyjnym - przegląd” na stronie 1 Konsola administracyjna wyszukiwania korporacyjnego służy do tworzenia kolekcji oraz źródeł zewnętrznych i administrowania nimi, uruchamiania i zatrzymywania komponentów, monitorowania działań w systemie, monitorowania plików dziennika, konfigurowania ustawień dotyczących użytkowników z uprawnieniami administracyjnymi, tworzenia powiązań aplikacji wyszukujących z kolekcjami i źródłami zewnętrznymi, a także do określania informacji umożliwiających wymuszanie stosowania zabezpieczeń.

Edytowanie kolekcji

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

Zagadnienia wstępne


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

Informacje o tym zadaniu

Podczas edycji kolekcji określa się opcje dotyczące przeszukiwania źródeł danych, analizy składni dokumentów, reorganizowania i odświeżania indeksu, wyszukiwania zaindeksowanych treści i rejestrowania komunikatów o błędach. Podczas tworzenia kolekcji należy ją poddać edycji w celu dodania do niej treści. Później kolekcję można edytować w celu zaktualizowania treści bądź zmiany sposobu przeszukiwania, analizowania, indeksowania, wyszukiwania i rejestrowania informacji.

Procedura

Aby zmodyfikować kolekcję, należy wykonać poniższe czynności:

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

Ogólne

Służy do podania ogólnych informacji o kolekcji i wyświetlenia ustawień,

których nie można zmieniać. Jeśli podczas tworzenia kolekcji zostało włączone zabezpieczenie, można konfigurować opcje zabezpieczeń na poziomie dokumentu.

Przeszukiwanie

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

Analiza składni

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

Indeks Umożliwia określanie harmonogramów reorganizowania całego indeksu i odświeżania indeksu zawierającego nowe i zmodyfikowane dane. Można tu także skonfigurować opcje związane z: wykorzystywaniem znaków zastępczych w zapytaniach, ograniczaniem widoków kolekcji do zakresu identyfikatorów URI, zwiżaniem wyników wyszukiwania z tej samej witryny sieci WWW i usuwaniem identyfikatorów URI z indeksu.

Wyszukiwanie

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

Dziennik

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

Usuwanie kolekcji

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

Zagadnienia wstępne

Aby usunąć kolekcję, wymagana jest rola administratora wyszukiwania korporacyjnego.

Przed usunięciem kolekcji należy zatrzymać wszystkie powiązane z nią procesy.


Informacje o tym zadaniu

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

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

Procedura

Aby usunąć kolekcję, należy wykonać poniższe czynności:

1. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.
2. Na liście kolekcji znajdź kolekcję, która ma zostać usunięta, i kliknij ikonę  **Usuń**.

Administrowanie przeszukiwaczem wyszukiwania korporacyjnego

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

Konfiguracja przeszukiwaczy

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

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

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

Konfigurowanie nowego przeszukiwacza przy użyciu wartości podstawowych

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

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

Łączenie typów przeszukiwaczy w kolekcji

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

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

Zabezpieczenia na poziomie dokumentu

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

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

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

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

Planowanie działania przeszukiwaczy

Przeszukiwacze tworzone dla źródeł WWW, NNTP oraz produktu WebSphere Portal działają w sposób ciągły. Po uruchomieniu takich przeszukiwaczy zazwyczaj nie ma konieczności zatrzymywania ich, o ile nie zostały zmodyfikowane ustawienia konfiguracji.

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

Konfigurując harmonogram określa się typ przeszukiwania, które ma zostać wykonane. W harmonogramie można uwzględnić pełne przeszukiwanie wszystkich dokumentów w

przestrzeni przeszukiwania, przeszukiwanie obejmujące wszystkie aktualizacje w przestrzeni przeszukiwania (nowe, zmodyfikowane i usunięte dokumenty) lub przeszukiwanie obejmujące wyłącznie nowe i zmodyfikowane dokumenty. Pełne przeszukiwanie zajmuje najwięcej czasu. Przeszukiwanie usuwające dokumenty, które zostały usunięte, trwa dłużej niż przeszukiwanie ignorujące usunięte dokumenty.

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

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

Pojęcia pokrewne

“Administrowanie wyszukiwaniem korporacyjnym - przegląd” na stronie 1
Konsola administracyjna wyszukiwania korporacyjnego służy do tworzenia kolekcji oraz źródeł zewnętrznych i administrowania nimi, uruchamiania i zatrzymywania komponentów, monitorowania działań w systemie, monitorowania plików dziennika, konfigurowania ustawień dotyczących użytkowników z uprawnieniami administracyjnymi, tworzenia powiązań aplikacji wyszukujących z kolekcjami i źródłami zewnętrznymi, a także do określania informacji umożliwiających wymuszanie stosowania zabezpieczeń.

“Zabezpieczenia na poziomie dokumentu” na stronie 184

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentu. Mechanizm zabezpieczeń na poziomie dokumentu jest odpowiedzialny za to, aby użytkownicy wyszukujący w kolekcji mieli dostęp tylko do tych dokumentów, względem których przydzielono im stosowne uprawnienia.

Zadania pokrewne

“Monitorowanie przeszukiwaczy” na stronie 217

Użytkownik może wyświetlić ogólne informacje na temat statusu poszczególnych przeszukiwaczy w kolekcji albo wybrać opcje dotyczące wyświetlania szczegółowych informacji o działaniu przeszukiwacza.

Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79

Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

Tworzenie przeszukiwacza

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

Zagadnienia wstępne

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

Informacje o tym zadaniu

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

Procedura

Aby utworzyć przeszukiwacz, należy wykonać poniższe czynności:

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

Użyj dla nowego przeszukiwacza systemowych wartości domyślnych

Wprowadzane są początkowe ustawienia przeszukiwacza zgodnie z domyślnymi wartościami instalacyjnymi.

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

Klonuj wartości ustawień istniejącego przeszukiwacza dla nowego przeszukiwacza

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

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

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

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

Edycja właściwości przeszukiwacza

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

Zagadnienia wstępne


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

Informacje o tym zadaniu

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

Procedura

Aby zmodyfikować właściwości przeszukiwacza, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Przeszukiwanie, odszukaj przeszukiwacz, który chcesz zmodyfikować, i kliknij ikonę  **Właściwości przeszukiwacza**.
2. Zmodyfikuj właściwości przeszukiwacza i kliknij przycisk **OK**.
3. Aby zmiany zostały uwzględnione, zatrzymaj i ponownie uruchom przeszukiwacz. Jeśli zmiana dotyczyła wyłącznie opisu przeszukiwacza, nie trzeba go restartować.

Edytowanie przestrzeni przeszukiwania

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

Zagadnienia wstępne


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

Informacje o tym zadaniu

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

Procedura

Aby zmodyfikować przestrzeń przeszukiwania, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Przeszukiwanie, znajdź przeszukiwacz, który chcesz zmodyfikować, i kliknij ikonę  **Przestrzeń przeszukiwania**.
2. Zmodyfikuj przestrzeń przeszukiwania, wybierając opcje, które mają zostać zmienione. Dostępne do wybrania opcje zależą od typu przeszukiwacza. W przypadku niektórych operacji, takich jak dodawanie źródeł danych do kolekcji, otwierany jest kreator właściwy dla typu przeszukiwacza umożliwiający zmianę przestrzeni przeszukiwania.
3. Aby zmiany zostały uwzględnione, zatrzymaj i ponownie uruchom przeszukiwacz.

Usuwanie przeszukiwacza

Usunięcie przeszukiwacza powoduje usunięcie wszystkich informacji o przeszukiwaczu z całego systemu wyszukiwania korporacyjnego. Informacje przeszukane wcześniej przez przeszukiwacz pozostają w indeksie do czasu reorganizacji indeksu.

Zagadnienia wstępne

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


Informacje o tym zadaniu

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

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

Procedura

Aby usunąć przeszukiwacz, należy wykonać poniższe czynności:

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

przeszukiwacze Content Edition

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

Przeszukiwacz Content Edition może służyć do przeszukiwania repozytoriów Documentum, FileNet Panagon Content Services, FileNet P8 Content Manager, Hummingbird Document Management (DM), OpenText Livelink oraz Portal Document Manager (PDM).

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich repozytoriów w przestrzeni przeszukiwania.
- Określenie, czy w celu uzyskania dostępu do repozytoriów przeszukiwacz używa trybu bezpośredniego czy też trybu serwera. W przypadku trybu serwera należy również określić informacje umożliwiające przeszukiwaczowi dostęp do serwera aplikacji WWW.
- Wybór repozytoriów do przeszukiwania.
- Określenie identyfikatorów użytkowników i haseł umożliwiających przeszukiwaczowi uzyskanie dostępu do treści wybranych repozytoriów.
- Skonfigurowanie harmonogramu przeszukiwania repozytoriów.
- Wybór klas elementów, które mają zostać przeszukane w każdym z repozytoriów.
- Określenie opcji udostępnienia właściwości klas elementów dla operacji przeszukiwania. Na przykład można wykluczyć określone typy dokumentów z przestrzeni przeszukiwania albo wybrać konkretną wersję repozytorium do przeszukania.

- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie list kontroli dostępu lub tokenów zabezpieczeń.

W przypadku klas elementów Documentum, FileNet Panagon Content Services oraz Portal Document Manager można również wybrać opcję sprawdzania poprawności referencji użytkownika podczas wprowadzania przez niego zapytania. W takim przypadku referencje użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych (ten typ sprawdzania poprawności bieżących referencji nie jest dostępny dla repozytoriów innych typów).

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

Dostęp do repozytoriów WebSphere II Content Edition w trybie serwera

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

W trybie serwera konektor WebSphere Information Integrator Content Edition używany przez przeszukiwacz w celu uzyskiwania dostępu do danych jest zainstalowany jako aplikacja korporacyjna na serwerze WebSphere Application Server, a przeszukiwacz ma dostęp do repozytoriów za pośrednictwem tego serwera. Takie podejście umożliwia korzystanie z zalet środowisk serwerów aplikacji J2EE.

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

Przed użyciem konsoli administracyjnej wyszukiwania korporacyjnego w celu skonfigurowania przeszukiwacza Content Edition do korzystania z trybu serwera należy wykonać zadanie odpowiednie dla używanego środowiska:

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu WebSphere II Content Edition” na stronie 26.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows dla programu WebSphere II Content Edition” na stronie 27.

Dostęp do repozytoriów WebSphere II Content Edition w trybie bezpośrednim

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

W trybie bezpośrednim przeszukiwacz używa konektora WebSphere Information Integrator Content Edition zainstalowanego na serwerze przeszukiwacza podczas instalacji produktu WebSphere II OmniFind Edition.

Procedura

Aby skonfigurować system pod kątem uzyskiwania dostępu do repozytoriów w trybie bezpośrednim:

1. Upewnij się, że zmienne środowiskowe VBR_HOME i JAVA_HOME w pliku *główny_katalog_instalacji_iice/bin/config.sh* (w systemie UNIX) lub w pliku *główny_katalog_instalacji_iice/bin/config.bat* (w systemie Microsoft Windows) określają poprawny katalog.
2. Aby skonfigurować konsolę administracyjną produktu WebSphere Information Integrator Content Edition do działania w trybie bezpośrednim, dodaj właściwość systemową Java **vbr.as.operationMode=direct** do pliku *główny_katalog_instalacji_iice/bin/Admin.bat* (w systemie UNIX) lub do pliku *główny_katalog_instalacji_iice/bin/Admin.bat* (w systemie Windows).
3. Uruchom konsolę administracyjną WebSphere Information Integrator Content Edition w trybie bezpośrednim i skonfiguruj konektor dla serwera przeszukiwacza WebSphere II OmniFind Edition (instrukcje można znaleźć w dokumentacji do produktu WebSphere Information Integrator Content Edition).
4. Wybierz opcję trybu bezpośredniego w przypadku używania konsoli administracyjnej produktu WebSphere II OmniFind Edition do konfigurowania przeszukiwacza Content Edition.

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

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

Informacje o tym zadaniu

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

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

Procedura

Aby skonfigurować serwer przeszukiwacza tak, aby mógł przeszukiwać repozytoria programu WebSphere Information Integrator Content Edition, należy wykonać poniższe czynności:

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

escrvbr.sh

3. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:

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

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

Kopiuj z:

Plik vbr_access_services.jar znajduje się w następującym domyślnym położeniu:

katalog_instalacyjny_was/installedApps/nazwa_serwera/nazwa_aplikacji

katalog_instalacyjny_was jest katalogiem instalacyjnym serwera WebSphere Application Server, *nazwa_serwera* jest nazwą serwera określoną przez użytkownika, a *nazwa_aplikacji* jest określoną przez użytkownika nazwą aplikacji WebSphere Information Integrator Content Edition na serwerze WebSphere Application Server.

Kopiuj do:

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

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

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

Informacje o tym zadaniu

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

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

Procedura

Aby skonfigurować serwer przeszukiwacza tak, aby mógł przeszukiwać repozytoria programu WebSphere Information Integrator Content Edition, należy wykonać poniższe czynności:

1. Jeśli program WebSphere II OmniFind Edition został zainstalowany na wielu serwerach, zainstaluj i powiąż biblioteki Java serwera WebSphere Application Server.
2. Uruchom skrypt konfiguracyjny przeszukiwacza Content Edition na serwerze przeszukiwacza:
 - a. Zaloguj się przy użyciu identyfikatora administratora wyszukiwania korporacyjnego (ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition).
 - b. Uruchom następujący skrypt, który został zainstalowany w katalogu %ES_INSTALL_ROOT%\bin, i odpowiedz na pytania:
escrvbr.vbs
3. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:
 - a. Z poziomu wiersza komend zatrzymaj system wyszukiwania korporacyjnego:
esadmin stop
 - b. Wybierz kolejno opcje: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**, a następnie zrestartuj usługę programu IBM WebSphere Information Integrator OmniFind Edition.
 - c. Z poziomu wiersza komend uruchom system wyszukiwania korporacyjnego:
esadmin start
4. Skopiuj plik vbr_access_services.jar z serwera WebSphere Information Integrator Content Edition na serwer przeszukiwacza.

Kopiuj z:

Plik vbr_access_services.jar znajduje się w następującym domyślnym położeniu:

katalog_instalacyjny_was/installedApps/nazwa_serwera/nazwa_aplikacji

katalog_instalacyjny_was jest katalogiem instalacyjnym serwera WebSphere Application Server, *nazwa_serwera* jest nazwą serwera określoną przez użytkownika, a *nazwa_aplikacji* jest określoną przez użytkownika nazwą aplikacji WebSphere Information Integrator Content Edition na serwerze WebSphere Application Server.

Kopiuj do:

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

przeszukiwacze DB2

Przeszukiwacz DB2 służy do włączenia baz danych IBM DB2 Universal Database do kolekcji. Przeszukiwacz DB2 może także posłużyć do dodania tabel pseudonimów utworzonych dla baz danych IBM DB2 Universal Database for z/OS, IBM Informix, Oracle oraz Microsoft SQL.

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

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

Publikowanie zdarzeń

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

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

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

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

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

- “Konfigurowanie programu WebSphere MQ dla przeszukiwaczy DB2” na stronie 32.
- “Konfigurowanie programu WebSphere Information Integrator Event Publisher Edition dla przeszukiwaczy DB2” na stronie 30.
- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do korzystania z publikowania zdarzeń” na stronie 34.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do korzystania z publikowania zdarzeń” na stronie 35.

Przegląd konfiguracji

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

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

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

Jeśli planowane jest przeszukiwanie zdalnych baz danych, które nie są wpisane do katalogów na lokalnym serwerze baz danych, to aby można było używać przeszukiwacza DB2 do przeszukiwania tych baz danych, na wybranym serwerze zdalnym należy uruchomić produkt DB2 Administration Server. Podczas konfigurowania przeszukiwacza należy podać nazwę hosta i numer portu na zdalnym serwerze bazy danych.

- Wybór bazy danych, która ma zostać przeszukana.
- Określenie identyfikatorów użytkowników i haseł pozwalających przeszukiwaczowi uzyskać dostęp do baz danych, w których używane są funkcje kontroli dostępu.

- Skonfigurowanie harmonogramu przeszukiwania baz danych.
- Wybór tabel, które mają zostać przeszukane w każdej bazie danych.
Ważne: Aby zoptymalizować wydajność procesów wykrywania (i zapobiec przekroczeniu limitu czasu przez proces konfigurowania przeszukiwacza), opcję przeszukiwania wszystkich tabel należy wybierać tylko wtedy, gdy baza danych nie zawiera zbyt wielu tabel lub gdy tabele nie zawierają zbyt wielu kolumn. Tabele do przeszukania można wybrać od razu, natomiast edycja przestrzeni przeszukiwania i dodanie dodatkowych tabel do kolekcji będzie możliwe również w późniejszym czasie.
- Wybór tabel, które mają być przeszukiwane w przypadku opublikowania dla nich aktualizacji w kolejce publikowania zdarzeń, a także określenie informacji umożliwiających przeszukiwaczowi dostęp do kolejki publikowania zdarzeń.
- Określenie opcji udostępnienia kolumn w określonych tabelach dostępnych dla operacji przeszukiwania. Na przykład można uwzględnić pewne kolumny w zapytaniach parametrycznych lub można określić, które kolumny mogą być zwracane w wynikach wyszukiwania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

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

Konfigurowanie programu WebSphere Information Integrator Event Publisher Edition dla przeszukiwaczy DB2

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

Informacje o tym zadaniu

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

- Do opublikowania muszą zostać wybrane zarówno zmienione, jak i niezmienione kolumny w tabelach źródłowych.
- Do opublikowania muszą zostać wybrane usunięte wiersze w tabelach źródłowych.
- Kolejka publikacji zdarzeń nie może być współużytkowana przez wiele baz danych.
- Pojedyncza baza danych może mieć wiele kolejek i map kolejek.
- Tabela powinna mieć jedną publikację XML powiązaną z jedną mapą kolejek publikowania. (Tabela nie powinna mieć więcej niż jednej publikacji XML powiązanej z pojedynczą mapą kolejek publikowania. Może mieć więcej niż jedną publikację XML, jeśli każda z nich jest powiązana z inną mapą kolejek publikowania).

Procedura

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

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

System operacyjny	Komenda
UNIX	db2rc
Windows	Wybierz opcje Start → Centrum replikacji IBM DB2

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

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

Konfigurowanie programu WebSphere MQ dla przeszukiwacza DB2

Przed skonfigurowaniem przeszukiwacza DB2 do użycia funkcji publikowania zdarzeń, upewnij się, że na serwerze, którego będzie nasłuchiwał przeszukiwacz, został skonfigurowany program IBM WebSphere MQ.

Zagadnienia wstępne

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

Ograniczenia

Jeśli docelowy serwer bazy danych został zainstalowany na komputerze z systemem Linux, wszyscy użytkownicy programu DB2 Universal Database (DB2 UDB), użytkownicy programu WebSphere MQ oraz użytkownicy programu WebSphere II OmniFind Edition muszą ustawić następującą zmienną środowiskową:

```
export LD_ASSUME_KERNEL=2.4.19
```

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

Informacje o tym zadaniu

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

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

Procedura

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

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

```
crtmqm QM1
```
 - b. Po wyświetleniu komunikatu wprowadź następującą komendę:

```
strmqm QM1
```
 - c. Po wyświetleniu komunikatu Uruchomiono menedżera kolejek 'QM1' wprowadź następującą komendę:

```
runmqsc QM1
```
 - d. Po wyświetleniu komunikatu Uruchamianie poleceń MQSC dla menedżera kolejek QM1 wprowadź następującą komendę, aby utworzyć kolejkę administracyjną:

```
DEFINE QLOCAL('ASN.QM1.ADMINQ')
```
 - e. Po wyświetleniu komunikatu Utworzono kolejkę komunikatów programu WebSphere wprowadź następującą komendę, aby utworzyć kolejkę restartowania:

```
DEFINE QLOCAL('ASN.QM1.RESTARTQ')
```

- f. Po ponownym wyświetleniu komunikatu **Utworzono kolejkę komunikatów programu WebSphere** wprowadź następującą komendę, aby utworzyć kolejkę danych:

```
DEFINE QLOCAL(' ASN.QM1.DATAQ')
```

- g. Po ponownym wyświetleniu komunikatu **Utworzono kolejkę komunikatów programu WebSphere** wprowadź następującą komendę, aby wyjść:
- ```
end
```

2. Wprowadź następującą komendę, aby uruchomić program MQ Listener na serwerze bazy danych (program MQ Listener musi być uruchomiony podczas tworzenia przeszukiwacza DB2, który korzysta z publikowania zdarzeń). W tym przykładzie 1414 to numer portu serwera. Używany jest domyślny kanał SYSTEM.DEF.SVRCONN:

```
runmqtsr -m QM1 -t TCP -p 1414 &
```

3. Wprowadź następującą komendę, aby autoryzować dostęp użytkownika DB2 UDB do kolejek i menedżera kolejek za pomocą interfejsu MQI (Message Queuing Interface) dla publikowania zdarzeń (w tym przykładzie identyfikator użytkownika to db2inst1):

```
setmqaut -m QM1 -t qmgr -p db2inst1 +allmqi
setmqaut -m QM1 -t queue -n ASN.QM1.DATAQ -p db2inst1 +allmqi
setmqaut -m QM1 -t queue -n ASN.QM1.ADMINQ -p db2inst1 +allmqi
setmqaut -m QM1 -t queue -n ASN.QM1.RESTARTQ -p db2inst1 +allmqi
```

4. Wprowadź następujące komendy dla identyfikatora użytkownika, który został użyty do utworzenia i uruchomienia przeszukiwacza DB2 z publikowaniem zdarzeń. Te komendy autoryzują dostęp ID użytkownika do menedżera kolejek i kolejek za pomocą interfejsu MQI w celu publikowania zdarzeń. W tym przypadku identyfikator użytkownika to esuser:

```
setmqaut -m ASN.QM1.QM2 -t qmgr -p esuser +allmqi
setmqaut -m ASN.QM1.QM2 -t queue -n ASN.QM1.DATAQ -p esuser +allmqi
```

## Konfigurowanie serwera przeszukiwacza w systemie UNIX do korzystania z publikowania zdarzeń

Jeśli program WebSphere II OmniFind Edition zostanie zainstalowany na komputerze z systemem IBM AIX, Linux lub ze Środowiskiem Operacyjnym Solaris, a przeszukiwacz DB2 zostanie skonfigurowany do użycia funkcji publikowania zdarzeń, należy uruchomić skrypt konfiguracyjny serwera przeszukiwacza. Ten skrypt umożliwia przeszukiwaczowi dostęp do menedżerów kolejek i kolejek programu WebSphere MQ.

### Informacje o tym zadaniu

Przeszukiwacz DB2 używa modułów programu WebSphere MQ 5.3 do obsługi usługi Java Messaging w celu uzyskania dostępu do menedżerów kolejek i kolejek programu WebSphere MQ. Moduły te należy zainstalować na serwerze przeszukiwacza.

Aby upewnić się, że przeszukiwacz DB2 może używać publikowania zdarzeń, należy również uruchomić skrypt konfiguracyjny, który jest udostępniany przez program WebSphere II OmniFind Edition na serwerze przeszukiwacza po zainstalowaniu modułów programu WebSphere MQ.

### Procedura

Aby skonfigurować serwer przeszukiwacza do korzystania z publikowania zdarzeń, należy wykonać następujące czynności:

1. Zainstaluj na serwerze przeszukiwacza moduły programu WebSphere MQ 5.3 dla usługi Java Messaging:
  - a. Zaloguj się jako administrator i wprowadź następującą komendę:
 

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

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

```
escrdb2.sh
```
3. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:
 

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

## Konfigurowanie serwera przeszukiwacza w systemie Windows do korzystania z publikowania zdarzeń

W przypadku zainstalowania programu WebSphere II OmniFind Edition na komputerze z systemem Microsoft Windows i skonfigurowania przeszukiwacza DB2 do korzystania z publikowania zdarzeń należy uruchomić skrypt w celu skonfigurowania serwera przeszukiwacza. Ten skrypt umożliwia przeszukiwaczowi dostęp do menedżerów kolejek i kolejek programu WebSphere MQ.

### Informacje o tym zadaniu

Przeszukiwacz DB2 używa modułów programu WebSphere MQ 5.3 do obsługi usługi Java Messaging w celu uzyskania dostępu do menedżerów kolejek i kolejek programu WebSphere MQ. Moduły te należy zainstalować na serwerze przeszukiwacza.

Aby upewnić się, że przeszukiwacz DB2 może używać publikowania zdarzeń, należy również uruchomić skrypt konfiguracyjny, który jest udostępniany przez program WebSphere II OmniFind Edition na serwerze przeszukiwacza po zainstalowaniu modułów programu WebSphere MQ.

### Procedura

Aby skonfigurować serwer przeszukiwacza do korzystania z publikowania zdarzeń, należy wykonać następujące czynności:

1. Zainstaluj na serwerze przeszukiwacza moduły programu WebSphere MQ 5.3 dla usługi Java Messaging:
  - a. Włóż dysk CD z programem WebSphere MQ.
  - b. Uruchom instalator programu WebSphere MQ.
  - c. W oknie wyboru opcji produktu wybierz opcję instalacyjną **Java Messaging**.

2. Uruchom skrypt konfiguracyjny przeszukiwacza DB2 na serwerze przeszukiwacza:
  - a. Zaloguj się przy użyciu identyfikatora administratora wyszukiwania korporacyjnego (ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu %ES\_INSTALL\_ROOT%\bin, i odpowiedz na pytania:  
escrdb2.vbs
3. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:
  - a. Z poziomu wiersza komend zatrzymaj system wyszukiwania korporacyjnego:  
esadmin stop
  - b. Wybierz kolejno opcje: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**, a następnie zrestartuj usługę programu IBM WebSphere Information Integrator OmniFind Edition.
  - c. Z poziomu wiersza komend uruchom system wyszukiwania korporacyjnego:  
esadmin start

---

## przeszukiwacze DB2 Content Manager

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

### Konfiguracja serwera przeszukiwacza

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

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

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu DB2 Content Manager” na stronie 37.
- “Konfigurowanie serwera przeszukiwacza dla programu DB2 Content Manager w systemie Windows” na stronie 39.

### Przegląd konfiguracji

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

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

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

## Konfigurowanie serwera przeszukiwacza w systemie UNIX dla programu DB2 Content Manager

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

### Informacje o tym zadaniu

Przeszukiwacz DB2 Content Manager korzysta z konektora Java dla programu DB2 Content Manager, wersja 8, w celu uzyskiwania dostępu do serwerów DB2 Content Manager. Konektor ten jest instalowany podczas instalowania programu IBM DB2 Information Integrator for Content w wersji 8.2 lub nowszej na serwerze przeszukiwacza. Aby zyskać pewność, że przeszukiwacz DB2 Content Manager może współpracować z produktem DB2 Content Manager, należy uruchomić skrypt konfiguracyjny dostępny w produkcie WebSphere II OmniFind Edition. Skrypt należy uruchomić na serwerze przeszukiwacza po zainstalowaniu konektora.

### Procedura

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

1. Zainstaluj konektor Java dla programu DB2 Content Manager, wersja 8, na serwerze przeszukiwacza:
  - a. Zaloguj się na serwerze przeszukiwacza jako użytkownik root:  
su - root
  - b. Uruchom plik db2profile. Na przykład:

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

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

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

```
. katalog_inst_Information_Integrator_for_Content/bin/cmbenv81.sh
cd katalog_inst_Information_Integrator_for_Content/samples/java/icm
javac *.java
java SConnectDisconnect nazwa_bazy_danych_ICM
ID_administratora_CM hasło_administratora_CM
```

- 5. Uruchom skrypt konfiguracyjny przeszukiwacza DB2 Content Manager na serwerze przeszukiwacza:
  - a. Przejdź do katalogu ES\_INSTALL\_ROOT/bin:  
cd \$ES\_INSTALL\_ROOT/bin
  - b. Uruchom poniższy skrypt i odpowiadaj na komunikaty:  
escrcm.sh
- 6. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:

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



## Konfigurowanie serwera przeszukiwacza dla programu DB2 Content Manager w systemie Windows

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

### Informacje o tym zadaniu

Przeszukiwacz DB2 Content Manager korzysta z konektora Java dla programu DB2 Content Manager, wersja 8, w celu uzyskiwania dostępu do serwerów DB2 Content Manager. Konektor ten jest instalowany podczas instalowania programu IBM DB2 Information Integrator for Content w wersji 8.2 lub nowszej na serwerze przeszukiwacza. Aby zyskać pewność, że przeszukiwacz DB2 Content Manager może współpracować z produktem DB2 Content Manager, należy uruchomić skrypt konfiguracyjny dostępny w produkcie WebSphere II OmniFind Edition. Skrypt należy uruchomić na serwerze przeszukiwacza po zainstalowaniu konektora.

### Procedura

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

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

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

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

```
cmbenv81.bat
cd katalog_inst_Information_Integrator_for_Content\samples\java\icm
javac *.java
java SConnectDisconnect nazwa_bazy_danych_ICM
ID_administratora_CM hasło_administratora_CM
```

4. Uruchom skrypt konfiguracyjny przeszukiwacza DB2 Content Manager na serwerze przeszukiwacza:
  - a. Przejdź do katalogu ES\_INSTALL\_ROOT\bin:  
cd %ES\_INSTALL\_ROOT%\bin
  - b. Uruchom poniższy skrypt i odpowiadaj na komunikaty:  
escrm.vbs
5. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:
  - a. Z poziomu wiersza komend zatrzymaj system wyszukiwania korporacyjnego:  
esadmin stop
  - b. Wybierz kolejno opcje: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**, a następnie zrestartuj usługę programu WebSphere Information Integrator OmniFind Edition.
  - c. Z poziomu wiersza komend uruchom system wyszukiwania korporacyjnego:  
esadmin start

---

## przeszukiwacze Domino Document Manager

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

### Konfiguracja serwera przeszukiwacza

Jeśli na serwerze Domino Document Manager, który ma być przeszukiwany, używany jest protokół NRPC (Notes remote procedure call), na serwerze przeszukiwacza należy uruchomić odpowiedni skrypt. Skrypt ten jest dostarczany razem z produktem WebSphere Information Integrator OmniFind Edition i umożliwi przeszukiwaczowi Domino Document Manager komunikowanie się z serwerami używającymi protokołu NRPC.

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

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

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

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 46.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 48.
- “Konfigurowanie serwerów korzystających z protokołu DIOP” na stronie 50.
- “Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 51.

## Przegląd konfiguracji

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

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

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

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

---

## Przeszukiwacze serwera Exchange Server

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

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają metodę przeszukiwania wszystkich podfolderów na wszystkich serwerach w przestrzeni przeszukiwania.

- Określenie informacji o serwerze folderów publicznych Exchange Server, który ma zostać przeszukany.

Należy podać identyfikator użytkownika i hasło, aby przeszukiwacz mógł uzyskać dostęp do treści na określonym serwerze. Jeśli na serwerze używany jest protokół SSL (Secure Sockets Layer), można określić opcje pozwalające przeszukiwaczowi na uzyskanie dostępu do pliku kluczy na serwerze przeszukiwacza.

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

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

## Sprawdzanie dostępu do zabezpieczonych dokumentów serwera Exchange

Aby użyć przeszukiwacza Exchange Server do przeszukiwania dokumentów chronionych przez firewall, należy sprawdzić, czy serwer przeszukiwacza może uzyskiwać dostęp do serwera folderu publicznego Microsoft Exchange Server.

### Informacje o tym zadaniu

Jeśli serwer przeszukiwacza nie może uzyskać dostępu do zabezpiezonego serwera Exchange Server, z serwera zostanie przesłany kod HTTP 501 (Nie zaimplementowano). Można również zobaczyć komunikaty wskazujące na odebranie nieoczekiwanej odpowiedzi HTTP.

### Procedura

Aby upewnić się, że serwer przeszukiwacza ma dostęp do dokumentów za firewallem, należy wykonać poniższe czynności:

1. Uruchom przeglądarkę WWW na serwerze przeszukiwacza.
2. Przejdź do adresu URL serwera folderu publicznego Exchange Server, który ma zostać przeszukany. Na przykład: <http://exchange.TwojaFirma.com/public/>
3. Sprawdź, że możesz otworzyć stronę serwera Exchange.

Jeśli nie możesz uzyskać dostępu do serwera Exchange Server, skontaktuj się z administratorem serwera w swojej organizacji.

---

## Przeszukiwacze NNTP

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

Przeszukiwacz NNTP może być używany do przeszukiwania dowolnej liczby serwerów NNTP. Podczas konfigurowania przeszukiwacza należy wybrać grupy dyskusyjne na każdym serwerze, który ma zostać przeszukany. Można także określić wzorce dla grup dyskusyjnych, które mają zostać wykluczone. Dzięki temu można zezwolić, aby przeszukiwacz przeszukał większość grup dyskusyjnych dostępnych na serwerze i zabronić przeszukiwania tych grup dyskusyjnych, które nie powinny być udostępnione użytkownikom.

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają metodę przeszukiwania wszystkich grup dyskusyjnych w przestrzeni przeszukiwania.
- Określenie wzorców dla grup dyskusyjnych, które mają zostać włączone, oraz wzorców dla grup dyskusyjnych, które mają zostać wykluczone z przestrzeni przeszukiwania.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

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

---

## przeszukiwacze Notes

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

### Konfiguracja serwera przeszukiwacza

Jeśli na serwerze Lotus Notes, który ma być przeszukiwany, używany jest protokół NRPC (Notes remote procedure call), na serwerze przeszukiwacza należy uruchomić odpowiedni skrypt. Skrypt ten jest dostarczany razem z produktem WebSphere Information Integrator OmniFind Edition i umożliwia przeszukiwaczowi Notes komunikowanie się z serwerami używającymi protokołu NRPC.

Jeśli na serwerze Lotus Notes, który ma zostać przeszukany, używany jest protokół DIIOP (Domino Internet Inter-ORB Protocol), na serwerze przeszukiwacza nie trzeba uruchamiać

skryptu konfiguracyjnego. Konieczne jest jednak skonfigurowanie serwera Lotus Notes w celu zezwolenia przeszukiwaczowi Notes na uzyskanie dostępu do tego serwera.

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

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

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 46.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 48.
- “Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 50.
- “Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 51.

## Zabezpieczenia na poziomie dokumentu

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

- “Konfigurowanie serwerów Lotus Domino Trusted Servers do sprawdzania poprawności referencji użytkownika” na stronie 191.
- Włączenie zabezpieczeń globalnych na serwerze WebSphere Application Server i skonfigurowanie aplikacji wyszukującej do korzystania z zabezpieczeń. Ten krok zapewnia, że użytkownicy będą proszeni o podanie referencji podczas próby użycia aplikacji wyszukującej. Referencje te mogą następnie zostać użyte przez serwery wyszukiwania do zweryfikowania uprawnień dostępu każdego użytkownika do dokumentów Lotus Notes.

## Przegląd konfiguracji

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich dokumentów w przestrzeni przeszukiwania.
- Określenie nazwy hosta, numeru portu i protokołu komunikacyjnego serwera Lotus Notes.

- Wybranie baz danych lub katalogów, które mają zostać przeszukane.
- Skonfigurowanie harmonogramu przeszukiwania baz danych lub katalogów.
- Wybranie dokumentów, które mają zostać przeszukane. Przeszukiwane mogą być wszystkie dokumenty w katalogu, wszystkie dokumenty w bazie danych lub tylko dokumenty w wybranych widokach i folderach bazy danych.
- Określenie opcji udostępniania pól w różnych bazach danych, widokach i folderach dla operacji przeszukiwania. Na przykład można wykluczyć niektóre pola z przestrzeni przeszukiwania i określić opcje wyszukiwania załączników.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

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

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

#### **Pojęcia pokrewne**

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino” na stronie 190

Jeśli na serwerze Lotus Notes, który ma być przeszukiwany, używany jest protokół NRPC (Notes remote procedure call), należy skonfigurować serwer przeszukiwacza w sposób powodujący wymuszenie zabezpieczeń na poziomie dokumentu.

#### **Zadania pokrewne**

“Konfigurowanie serwerów Lotus Domino Trusted Servers do sprawdzania poprawności referencji użytkownika” na stronie 191

Aby wymusić zabezpieczenia dokumentów, które zostały przeszukane przez przeszukiwacz Notes używający protokołu NRPC (Notes remote procedure call), serwery Domino, które mają być przeszukiwane, należy skonfigurować jako zaufane serwery Lotus Domino.

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

Przed skonfigurowaniem przeszukiwacza Notes należy zapoznać się z wytycznymi na temat przeszukiwania baz danych Lotus Domino.

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

- Pola w następującej tabeli są domyślnie odwzorowywane na nazwy pól wyszukiwania:

*Tabela 1. Domyślne reguły odwzorowywania pól*

| Nazwa pola bazy danych Notes | Nazwa pola wyszukiwania |
|------------------------------|-------------------------|
| Title                        | Title                   |
| EventTitle                   | Title                   |
| Subject                      | Title                   |
| Body                         | Body                    |
| Mission                      | Body                    |
| From                         | Creator                 |
| Author                       | Creator                 |
| Keywords                     | Kategorie               |
| Kategorie                    | Kategorie               |
| TeamRoomName                 | Organization            |
| TeamName                     | Organization            |
| Department                   | Organization            |

- Przeszukiwacz Notes może przeszukiwać wszystkie typy pól za wyjątkiem pól obliczanych przed wyświetleniem.
- Statyczny tekst i obrazy umieszczone w formularzach Notes nie są przeszukiwane.
- Zaznaczenie pola wyboru **Przeszukaj wszystkie** podczas konfigurowania przeszukiwacza powoduje przeszukiwanie wszystkich pól i maksymalizowanie danych z pól do przeszukiwania (pole **Przeszukaj wszystkie pola z wyjątkiem** pozwala ograniczyć pola do przeszukiwania).

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

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

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

### Ograniczenia

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

### Informacje o tym zadaniu

Przeszukiwacz używający protokołu NRPC korzysta z bibliotek Domino jako klient. Biblioteki te są instalowane podczas instalacji serwera Lotus Domino, wersja 6.0.2 lub nowsza, na serwerze przeszukiwacza. Aby mieć pewność, że przeszukiwacze mogą



współpracować z bibliotekami Domino, należy uruchomić skrypt konfiguracyjny, który program WebSphere II OmniFind Edition udostępnia na serwerze przeszukiwacza po zainstalowaniu bibliotek Domino.

## Procedura

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

1. Utwórz użytkownika **notes** i grupę **notes** na serwerze przeszukiwacza:
  - a. Zaloguj się jako użytkownik root:  
`su - root`
  - b. Dodaj użytkownika:  
`useradd notes`
  - c. Dodaj hasło tego użytkownika:  
`passwd notes`  
Zostanie wyświetlony komunikat o zmianę tego hasła.
2. Zainstaluj serwer Lotus Domino na serwerze przeszukiwacza:
  - a. Włóż do napędu instalacyjny dysk CD serwera Domino, wersja 6.0.2 lub nowsza, i podłącz go. (W razie braku dysku CD można pobrać jego obraz).
  - b. Przejdź do katalogu właściwego dla systemu operacyjnego.  
  
AIX: `cd /mnt/cdrom/aix`  
System Linux: `cd /mnt/cdrom/linux`  
Środowisko Solaris: `cd /mnt/cdrom/solaris`
  - c. Uruchom program instalacyjny:  
`./install`
  - d. Odpowiedz na komunikaty i zaakceptuj wartości domyślne lub podaj preferowane ustawienia instalacyjne (takie jak ścieżki do katalogu instalacyjnego i katalogu danych).  
  
Więcej informacji na temat instalacji serwera Domino można znaleźć w dokumentacji tego produktu.
3. Uruchom skrypt konfiguracyjny udostępniany przez program WebSphere II OmniFind Edition na serwerze przeszukiwacza:
  - a. Zaloguj się jako administrator wyszukiwania korporacyjnego (z identyfikatorem użytkownika określonym podczas instalowania programu WebSphere II OmniFind Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu `$ES_INSTALL_ROOT/bin`:  
`escrnote.sh`
  - c. Odpowiedz na komunikaty:
    - Na przedstawiony poniżej komunikat udziel odpowiedzi Y, jeśli serwer Domino jest instalowany w katalogu domyślnym, lub N, jeśli nie jest:  
Odnaleziono ścieżkę katalogu programu Lotus Notes  
`/opt/lotus/notes/latest/linux`.  
Czy to jest poprawna ścieżka katalogu programu Lotus Notes?  
  
Domyślna ścieżka dla systemu AIX to `/opt/lotus/notes/latest/ibmpow`.  
Domyślna ścieżka dla systemu Linux to `/opt/lotus/notes/latest/linux`.  
Domyślna ścieżka dla środowiska Solaris to `/opt/lotus/notes/latest/sunspa`.

- Jeśli serwer Domino nie jest instalowany w domyślnym katalogu na serwerze przeszukiwacza, określ właściwą ścieżkę do katalogu Domino w odpowiedzi na poniższy komunikat:

Wprowadź ścieżkę katalogu programu Lotus Notes.

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

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

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

Domyślna ścieżka to /local/notesdata.

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

Wprowadź ścieżkę katalogu danych programu Lotus Notes.

4. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:

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

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

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

### Ograniczenia

Serwer Lotus Domino i klient programu Lotus Notes nie mogą być uruchamiane jednocześnie na tym samym komputerze z przeszukiwaczem Notes, QuickPlace lub Domino Document Manager, który został skonfigurowany do użycia protokołu NRPC. Uruchomienie jednego z tych przeszukiwaczy w trakcie pracy serwera Domino powoduje wystąpienie błędu i zatrzymanie przeszukiwacza.

### Informacje o tym zadaniu

Przeszukiwacze używające protokołu NRPC korzystają z bibliotek klienta programu Lotus Notes. Biblioteki te są instalowane podczas instalacji wydania Lotus Notes, wersja 6.0.2 lub nowsza, na serwerze przeszukiwacza. Aby mieć pewność, że przeszukiwacze mogą współpracować z bibliotekami klienta Lotus Notes, należy uruchomić skrypt konfiguracyjny, który program WebSphere II OmniFind Edition udostępnia na serwerze przeszukiwacza po zainstalowaniu bibliotek klienta Lotus Notes.

### Procedura

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

1. Zaloguj się na serwerze przeszukiwacza przy użyciu identyfikatora użytkownika należącego do grupy Administratorzy. Upewnij się, że użytkownik o podanym identyfikatorze ma uprawnienia pozwalające na instalację programu Lotus Notes.
2. Zainstaluj program Lotus Notes:
  - a. Włóż do napędu dysk instalacyjny CD programu Lotus Notes, wersja 6.0.2 lub nowsza. (W razie braku dysku CD można pobrać jego obraz).
  - b. Uruchom program instalacyjny: `setup.exe`.
  - c. Odpowiedz na komunikaty i zaakceptuj wartości domyślne lub podaj preferowane ustawienia instalacyjne (takie jak ścieżki do katalogu instalacyjnego i katalogu danych).

W razie potrzeby zapoznaj się z dokumentacją programu Lotus Notes.
3. Uruchom skrypt konfiguracyjny udostępniany przez program WebSphere II OmniFind Edition na serwerze przeszukiwacza:
  - a. Zaloguj się przy użyciu identyfikatora administratora wyszukiwania korporacyjnego (ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition).
  - b. Uruchom następujący skrypt, który został zainstalowany w katalogu `%ES_INSTALL_ROOT%\bin`:  
`escrnote.vbs`
  - c. Odpowiedz na komunikaty:
    - Na przedstawiony poniżej komunikat udziel odpowiedzi Y, jeśli program Lotus Notes jest instalowany w katalogu domyślnym, lub N w przeciwnym razie:  
Odnaleziono ścieżkę katalogu programu Lotus Notes `c:\lotus\notes`.  
Czy to jest poprawna ścieżka katalogu programu Lotus Notes?  
  
Domyślna ścieżka instalacji na komputerze z systemem Windows to `c:\lotus\notes` lub `c:\lotus\domino`.
    - Jeśli program Lotus Notes nie jest instalowany w domyślnym katalogu na serwerze przeszukiwacza, określ właściwą ścieżkę do katalogu Lotus Notes w odpowiedzi na poniższy komunikat:  
Wprowadź ścieżkę katalogu programu Lotus Notes.
    - Na przedstawiony poniżej komunikat udziel odpowiedzi Y, jeśli katalog danych programu Lotus Notes został umieszczony w domyślnej lokalizacji, lub N w przeciwnym razie:  
Odnaleziono ścieżkę katalogu danych programu Lotus Notes  
`c:\lotus\notes\data`.  
Czy jest to poprawna ścieżka katalogu danych programu Lotus Notes?  
  
Typowa ścieżka na komputerze z systemem Windows to `c:\lotus\notes\data` lub `c:\lotus\domino\data`.
    - Jeśli katalog danych programu Lotus Notes nie znajduje się w domyślnej lokalizacji na serwerze przeszukiwacza, określ ścieżkę do katalogu danych w odpowiedzi na poniższy komunikat:  
Wprowadź ścieżkę katalogu danych programu Lotus Notes.
4. Zatrzymaj i zrestartuj system wyszukiwania korporacyjnego oraz wszystkie sesje wspólnej warstwy komunikacji (CCL) wyszukiwania korporacyjnego:
  - a. Z poziomu wiersza komend zatrzymaj system wyszukiwania korporacyjnego:  
`esadmin stop`

- b. Wybierz kolejno opcje: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**, a następnie zrestartuj usługę programu IBM WebSphere Information Integrator OmniFind Edition.
- c. Z poziomu wiersza komend uruchom system wyszukiwania korporacyjnego:  
esadmin start

## Konfigurowanie serwerów korzystających z protokołu DIIOP

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

### Zagadnienia wstępne

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

### Procedura

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

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

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

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

### Informacje o tym zadaniu

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

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

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

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

### Procedura

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

1. Zainstaluj moduł IOCP (bos.iocp.rte) z dysku CD produktu AIX na serwerze przeszukiwacza.

Po zainstalowaniu modułu IOCP i przed utworzeniem przeszukiwacza Notes, QuickPlace lub Domino Document Manager należy zastosować do modułu poprawkę oprogramowania. Instrukcje są dostępne po skorzystaniu z następującego odsyłacza:

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

2. Wprowadź następującą komendę, aby upewnić się, że moduł IOCP został zainstalowany na serwerze przeszukiwacza:

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

Dane wyjściowe komendy lspp powinny być analogiczne do poniższego przykładu:

```
Fileset Level State Description

Path: /usr/lib/objrepos
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. Wprowadź następującą komendę, aby upewnić się, że port IOCP ma status **Available** (dostępny):

```
$ lsdev -Cc iocp
```

Dane wyjściowe komendy lsdev powinny być zgodne z poniższym przykładem:

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

4. Jeśli port IOCP ma status **Defined**, należy go zmienić na **Available**:
  - a. Zaloguj się na serwerze przeszukiwacza jako użytkownik root i wprowadź następującą komendę:

```
smit iocp
```
  - b. Wybierz opcję **Zmień/pokaż charakterystykę portów I/O Completion Ports** i zmień wartość opcji **STAN konfigurowany w momencie ładowania systemu z Zdefiniowane** na **Dostępne**.
  - c. Zrestartuj serwer przeszukiwacza.
  - d. Ponownie wprowadź komendę lsdev i potwierdź, że status portu IOCP zmienił się na **Available**.

---

## Przeszukiwacze QuickPlace

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

### Konfiguracja serwera przeszukiwacza

Jeśli na serwerze QuickPlace, który ma być przeszukiwany, używany jest protokół NRPC (Notes remote procedure call), na serwerze przeszukiwacza należy uruchomić odpowiedni skrypt. Skrypt ten jest dostarczany razem z produktem WebSphere Information Integrator OmniFind Edition i umożliwi przeszukiwaczowi QuickPlace komunikowanie się z serwerami używającymi protokołu NRPC.

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

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

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

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

- “Konfigurowanie serwera przeszukiwacza w systemie UNIX do przeszukiwania źródeł Lotus Domino” na stronie 46.
- “Konfigurowanie serwera przeszukiwacza w systemie Windows do przeszukiwania źródeł Lotus Domino” na stronie 48.
- “Konfigurowanie serwerów korzystających z protokołu DIIOP” na stronie 50.
- “Konfigurowanie serwera QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego” na stronie 54.
- “Konfigurowanie opcji Directory Assistance na serwerze QuickPlace” na stronie 55.
- “Konfigurowanie portu zakończenia we/wy w systemie AIX do przeszukiwania źródeł Lotus Domino” na stronie 51.

## Przegląd konfiguracji

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

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

Istnieje także możliwość wybrania opcji sprawdzania poprawności referencji użytkownika w trakcie wprowadzania przez niego zapytania. W takim przypadku referencje

użytkownika nie są porównywane z poindeksowanymi danymi zabezpieczeń, ale z bieżącymi listami kontroli dostępu, które są przechowywane w oryginalnym źródle danych.

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

## Konfigurowanie serwera QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego

Jeśli planujesz skonfigurować przeszukiwacz QuickPlace do korzystania z opcji użytkownika lokalnego w celu wdrożenia zabezpieczeń, przed utworzeniem przeszukiwacza należy skonfigurować katalog Domino na serwerze Lotus QuickPlace.

### Informacje o tym zadaniu

Podczas konfigurowania przeszukiwacza QuickPlace można wybrać tryb zabezpieczeń używany przez przeszukiwacz do wymuszania zabezpieczeń na poziomie dokumentu. W przypadku wybrania trybu użytkownika lokalnego należy się upewnić, że wszystkie lokalne identyfikatory użytkownika i grupy lokalne zostały zarejestrowane w katalogu Domino (hierarchii katalogu Domino musi odpowiadać hierarchii serwera QuickPlace).

Należy się również upewnić, że identyfikator użytkownika i hasło określone dla używanego przeszukiwacza zostały zarejestrowane w katalogu Domino i mają uprawnienia do odczytu przeszukiwanej bazy danych.

Do korzystania z serwera QuickPlace potrzebna jest tylko nazwa użytkownika. Jednak do przeszukania źródeł QuickPlace wymagany jest w pełni rozwinięty identyfikator użytkownika. Rozwinięty identyfikator użytkownika ma następujący format:  
*nazwa użytkownika/nazwa lokalizacji/QP/nazwa domeny*

Ta procedura umożliwi określenia w pełni rozwiniętej wersji identyfikatora użytkownika, upewnienie się, że dany identyfikator użytkownika ma uprawnienia do odczytu z bazy danych QuickPlace, oraz dodanie identyfikatora użytkownika do katalogu Domino. Katalog Domino musi zawierać identyfikator użytkownika, który będzie używany do przeszukiwania baz danych QuickPlace, oraz wszystkich lokalnych użytkowników i grupy programu QuickPlace (hierarchia katalogu Domino musi odpowiadać hierarchii serwera QuickPlace).

### Procedura

Aby skonfigurować serwer QuickPlace do korzystania z zabezpieczeń użytkownika lokalnego, należy wykonać poniższe czynności:

1. Potwierdź uprawnienia identyfikatora użytkownika:
  - a. Otwórz dokument Server na serwerze QuickPlace.
  - b. Otwórz stronę plików, a następnie otwórz listę kontroli dostępu (ACL) dla bazy danych, którą należy przeszukać.
  - c. Potwierdź, że lokalny identyfikator użytkownika, który zostanie uwzględniony w konfiguracji przeszukiwacza, istnieje na liście ACL oraz że ten identyfikator użytkownika ma uprawnienia do odczytu bazy danych.  
W kroku 2 należy określić w pełni rozwiniętą formę tego identyfikatora użytkownika.
2. Dodaj nowego użytkownika do katalogu Domino:
  - a. Otwórz dokument Server na serwerze QuickPlace.
  - b. Na stronie osób i grup w elemencie drzewa osób dodaj w pełni rozwinięty identyfikator użytkownika potwierdzony w kroku 1.



- c. W polu **Hasło dostępu do Internetu** podaj hasło dla tego identyfikatora użytkownika.

## Konfigurowanie opcji Directory Assistance na serwerze QuickPlace

Jeśli planujesz skonfigurować przeszukiwacz QuickPlace do korzystania z katalogu LDAP w celu wdrożenia zabezpieczeń, przed skonfigurowaniem przeszukiwacza należy utworzyć bazę danych Directory Assistance na serwerze Lotus QuickPlace.

### Ograniczenia

Serwer QuickPlace, który ma być przeszukiwany, musi obsługiwać zadania protokołów DIIOP i HTTP.

### Procedura

Aby skonfigurować opcję LDAP Directory Assistance na serwerze QuickPlace, należy wykonać poniższe czynności:

1. 1. Utwórz bazę danych Directory Assistance:
  - a. Otwórz dokument Server na serwerze QuickPlace.
  - b. Utwórz bazę danych za pomocą szablonu **Directory Assistance(6)**. Ten szablon znajduje się na serwerze.
  - c. Kliknij opcję **Dodaj bazę danych Directory Assistance**, aby utworzyć dokument w bazie danych.
  - d. Otwórz kartę Podstawowe i w polu **Typ Domeny** wybierz opcję **LDAP**.
  - e. Otwórz kartę Konteksty nazw i upewnij się, że pole wyboru **Zaufane dla referencji** jest zaznaczone.
  - f. Otwórz kartę LDAP i podaj informacje o serwerze LDAP.
  - g. Zapisz i zamknij dokument Server.
2. Skonfiguruj serwer QuickPlace do korzystania z bazy danych Directory Assistance:
  - a. Otwórz dokument Server na serwerze QuickPlace.
  - b. Otwórz kartę Podstawowe i w polu **Nazwa bazy danych Directory Assistance** podaj nazwę bazy danych utworzonej w kroku 1.
  - c. Zapisz i zamknij dokument Server.

Serwer QuickPlace może teraz używać serwera LDAP jako dodatkowego katalogu Domino.

---

## przeszukiwacze systemu plików UNIX

Aby do kolekcji wyszukiwania korporacyjnego włączyć dokumenty przechowywane w systemie plików UNIX, należy skonfigurować przeszukiwacz system plików UNIX.

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

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

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

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

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

---

## przeszukiwacze sieci WWW

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

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

Przeszukane dane mogą być zapisane w jednym z wielu powszechnie używanych formatów i mogą pochodzić z różnych źródeł w sieci intranet lub Internet. Powszechnie używane formaty to między innymi HTML, PDF, Microsoft Word, Lotus WordPro oraz XML (Extensible Markup Language).

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania wszystkich serwisów WWW w przestrzeni przeszukiwania.

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

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

## Konfiguracja agenta użytkownika

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

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

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

Zgodność z protokołem Robots Exclusion nie jest obowiązkowa, ale przeszukiwacz sieci WWW systemu wyszukiwania korporacyjnego jest zgodny z tym protokołem:

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

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

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

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

## Obsługa języka JavaScript

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

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

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

- Zostaną znalezione pewne adresy URL, które są pomijane przez bardziej rygorystyczne analizatory składni HTML. Przeszukiwacz odrzuci całą treść, która nie stanowi poprawnego składniowo adresu URL, jednak niektóre poprawne adresy URL zwrócone w trakcie skanowania mogą nie być szczególnie przydatne do wyszukiwania.
- Dokumenty generowane przez skrypty JavaScript (np. gdy użytkownik ogląda stronę w przeglądarce, a ta wykonuje skrypt JavaScript), nie będą widoczne dla przeszukiwacza sieci WWW i nie zostaną poindeksowane.

## Reguły ograniczania przestrzeni przeszukiwania w sieci WWW

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

Gdy przeszukiwacz sieci WWW przeszukuje serwis WWW, wykrywa odsyłacze do innych serwisów i umieszcza je w kolejce serwisów, które zostaną przeszukane w dalszej kolejności. Przeszukiwanie i wykrywanie odsyłaczy może być powtarzane, dopóki pozwala na to czas i dostępne zasoby pamięci. Podczas konfigurowania przeszukiwacza sieci WWW należy określić miejsce, w którym przeszukiwacz ma rozpocząć przeszukiwanie. Wychodząc od tych

wstępnych adresów URL (określanych również jako *początkowe adresy URL*) przeszukiwacz sieci WWW może uzyskać dostęp do każdego dokumentu w sieci WWW, który jest powiązany przy użyciu bezpośredniego lub pośredniego odsyłacza.

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

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

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

Format reguł przeszukiwania jest następujący:

`działanie typ element_docelowy`

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

## Reguły dotyczące domen

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

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

Można użyć gwiazdki jako pierwszego znaku w nazwie elementu docelowego. W takim wypadku określona reguła będzie stosowana dla każdej nazwy hosta, w której na końcu występuje pozostała część podanego łańcucha wzorca. Na przykład można nie przeszukiwać wszystkich domen, których nazwy są zgodne z następującym wzorcem:

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

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

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

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

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

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

## Reguły dotyczące przedrostka

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

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

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

Określając regułę dotyczącą przedrostka, można użyć wielu gwiazdek w dowolnym miejscu łańcucha przedrostka, a nie tylko na ostatniej pozycji tego łańcucha. Na przykład poniższa reguła zabrania przeszukiwaczowi przeszukiwania jakichkolwiek dokumentów w katalogu najwyższego poziomu serwisu sales.ibm.com, jeśli nazwa katalogu kończy się na fs. (Na przykład można używać pozycji dołączania systemu plików, w których nie występują żadne informacje, które mogłyby być przydatne w indeksie wyszukiwania).

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

## Reguły dotyczące adresów

Reguła dotycząca adresu umożliwia sterowanie przeszukiwaniem całych hostów lub sieci przez określenie adresu IP (tylko w wersji IPv4) i maski sieci jako elementu docelowego. Na przykład:

```
allow address 9.0.0.0 255.0.0.0
```

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

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

Przydatna reguła, która może być używana jako adres końcowy na liście reguł to forbid address 0.0.0.0 0.0.0.0. Według tej reguły zgodne są wszystkie adresy IP, ponieważ w masce sieci wszystkie bity są określone jako nieistotne (użycie tej reguły powoduje zabronienie dostępu do wszystkich adresów, do których dostęp nie został zezwolony przy użyciu reguł występujących na poprzednich pozycjach na liście reguł).

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

Jeśli przeszukiwacz sieci WWW używa serwera proxy, adres IP tego serwera proxy jest jedynym adresem IP innego hosta dostępnym dla przeszukiwacza. Jeśli reguły adresów IP służą do ograniczania przeszukiwacza do podsieci adresów IP, ograniczenie to powoduje klasyfikowanie prawie wszystkich adresów URL z kodem powrotu 760 (oznaczającym, że są one zabronione przez obszar WWW).

## Porządek reguł przeszukiwania

Przeszukiwacz stosuje reguły przeszukiwania na różnych etapach procesu wykrywania i przeszukiwania adresów URL. Porządek reguł ma istotne znaczenie, lecz dotyczy wyłącznie reguł tego samego typu. Porządek, w jakim występują dwie reguły dotyczące adresu, ma decydujące znaczenie, lecz nie ma znaczenia, czy reguła dotycząca adresu występuje przed, czy po regule dotyczącej przedrostka, ponieważ przeszukiwacz nie stosuje tych reguł w tym samym momencie.

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

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

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

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

```
forbid domain *
```

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

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

Na końcu listy reguł dotyczących adresów występuje zazwyczaj następująca reguła zapobiegająca przeszukiwaniu serwisów WWW znajdujących się poza siecią korporacyjną.

```
forbid 0.0.0.0 0.0.0.0
```

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

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

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

## **Rozszerzenia plików, typy MIME i maksymalna głębokość przeszukiwania**

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

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

## **Początkowe adresy URL**

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

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

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



Następujące adresy URL są poprawnymi początkowymi adresami URL:

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

Następujący adres URL nie jest poprawnym początkowym adresem URL:

```
www.ibm.com
```

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

## Testy połączeń URL przy użyciu przeszukiwacza sieci WWW

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

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

Gdy serwis zostanie przeszukany co najmniej raz, można przetestować adresy URL w celu uzyskania dodatkowych informacji. Na przykład raport z testu może udostępnić ostatni kod powrotu HTTP (wskazujący, czy przeszukiwanie adresu URL zakończyło się pomyślnie), wskazać, kiedy dany adres URL był ostatnio przeszukiwany i na kiedy zaplanowane jest jego kolejne przeszukiwanie, a także wskazać, czy agent użytkownika korzysta z bieżącego pliku robots.txt serwera WWW.

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

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

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

Gdy do utworzenia przeszukiwacza sieci WWW lub edytowania jego właściwości używana jest konsola administracyjna wyszukiwania korporacyjnego, użytkownik może wybrać opcję konfigurowania właściwości zaawansowanych. Na stronie Zaawansowane właściwości przeszukiwacza sieci WWW użytkownik może określić opcje określające minimalny i maksymalny odstęp czasu między kolejnymi przeszukiwaniami. Określone wartości są używane przez przeszukiwacza sieci WWW do obliczenia odstępu czasu między ponownymi operacjami przeszukiwania danych.

Podczas pierwszego przeszukiwania określonej strony przeszukiwacz wykorzystuje datę i godzinę przeszukiwania tej strony oraz wartość średnią obliczoną przy użyciu minimalnego i maksymalnego odstępu czasu przed ponownym przeszukiwaniem do obliczenia daty ponownego przeszukiwania tej strony. Strona ta nie zostanie przeszukana ponownie przed tym terminem. Czas przeszukiwania tej strony po tym terminie zależy od obciążenia przeszukiwacza i bilansu nowych i starych adresów URL w przestrzeni przeszukiwania.

Podczas ponownego przeszukiwania strony wyszukiwacz sprawdza zmiany treści na tej stronie. Jeśli treść uległa zmianie, następny odstęp czasu przed ponownym przeszukiwaniem będzie krótszy od poprzedniego, lecz nigdy nie będzie krótszy niż określony minimalny odstęp czasu przed ponownym przeszukiwaniem. Jeśli treść nie uległa zmianie, wówczas następny odstęp czasu przed ponownym przeszukiwaniem będzie dłuższy od poprzedniego, lecz nigdy nie będzie dłuższy niż określony maksymalny odstęp czasu przed ponownym przeszukiwaniem.

## Opcje odwiedzania adresów URL przy użyciu wyszukiwacza sieci WWW

Istnieje możliwość wymuszenia przeszukania konkretnych adresów URL przez wyszukiwacz sieci WWW w najbliższym możliwym terminie.

Gdy zaistnieje potrzeba odświeżenia przestrzeni przeszukiwania przy użyciu informacji dostępnych w określonych serwisach WWW, użytkownik może użyć funkcji monitorowania wyszukiwacza, wybrać opcję **Adresy URL do natychmiastowego odwiedzenia lub ponownego odwiedzenia**, a następnie określić adresy URL lub wzorce adresów URL stron, które mają zostać przeszukane lub przeszukane ponownie.

Na przykład, jeśli w Dziale Komunikacyjnym w sieci intranet strona WWW została dodana lub zaktualizowana przy użyciu nowych informacji w wyniku istotnych zmian strategii, można określić adres URL tej nowej lub zmodyfikowanej strony. Jeśli wyszukiwacz jest uruchomiony, określony adres URL zostanie umieszczony w kolejce stron do przeszukania i pozostanie w tej kolejce do następnego sprawdzenia listy stron oczekujących na przeszukanie (zazwyczaj kolejka ta przeglądana jest co 10 minut). Jeśli wyszukiwacz nie jest uruchomiony, określony adres URL zostanie umieszczony w kolejce i zostanie przeszukany po następnym uruchomieniu wyszukiwacza.

Należy upewnić się, że wśród reguł przeszukiwania znajduje się reguła zezwalająca wyszukiwaczowi na odwiedzenie określonych adresów URL. Wyszukiwacz będzie mógł przeszukać określone adresy URL wcześniej niż zwykle. Aby jednak adres URL mógł zostać w ogóle przeszukany, musi istnieć odpowiednia reguła przeszukiwania zezwalająca na przeszukanie określonego adresu URL.

Nowo utworzone dane staną się dostępne dla operacji przeszukiwania po wykonaniu następnej reorganizacji lub odświeżeniu indeksu.

## Jak wyszukiwacz sieci WWW obsługuje strony błędów programowych

Wyszukiwacz sieci WWW można skonfigurować w taki sposób, aby obsługiwał niestandardowe strony tworzone przez administratorów serwisów WWW, które są zwracane zamiast standardowych kodów błędów w odpowiedzi na żądania dostępu do pewnych stron.

Gdy serwer HTTP nie może zwrócić strony żądanej przez klienta, zazwyczaj wysyła odpowiedź zawierającą nagłówek z kodem powrotu. Kod powrotu wskazuje rodzaj problem z wykonaniem żądania (np. błąd 404 oznacza, że nie można znaleźć pliku). Niektórzy administratorzy serwisów WWW tworzą specjalne strony, na których problemy są wyjaśnione w bardziej szczegółowy sposób, i konfiguruje serwer HTTP tak, aby zwracał te strony zamiast kodów błędów. Takie niestandardowe strony są nazywane *stronami błędów programowych*.

Strony błędów programowych mogą powodować przekłamanie wyników działania wyszukiwacza sieci WWW. Na przykład zamiast nagłówka z kodem błędu wyszukiwacz otrzymuje stronę błędu programowego i kod powrotu 200, który oznacza pomyślne pobranie poprawnej strony HTML. Jednak pobrana strona błędu programowego nie jest związana z

żądanym adresem URL, jej treść jest niemal identyczna za każdym razem, gdy jest zwracana zamiast żądanej strony. Te nieistotne i prawie identyczne strony zniekształcają indeks i wyniki wyszukiwania.

W celu obsługi tego typu sytuacji można w trakcie konfigurowania przeszukiwacza sieci WWW określić opcje obsługi stron błędów programowych. Przeszukiwacz wymaga podania następujących informacji o każdym serwisie WWW, który zwraca strony błędów programowych:

- Wzorzec adresu URL serwisu, który używa stron błędów programowych. Wzorzec URL składa się z protokołu (HTTP lub HTTPS), nazwy hosta, numeru portu (jeśli jest inny niż standardowy) oraz nazwy ścieżki. Można używać gwiazdki (\*) jako znaku zastępczego dla jednego lub większej liczby znaków, aż do wystąpienia pierwszego znaku, który nie jest znakiem zastępczym. We wzorcach rozróżniana jest wielkość liter.
- Wzorzec tytułu, który odpowiada tekstowi między znacznikami <TITLE> w dokumencie HTML. Aby określić ten wzorzec, można użyć gwiazdki (\*) jako znaku zastępczego. W tym wzorcu rozróżniana jest wielkość liter.
- Wzorzec treści tekstu, który odpowiada treści dokumentu HTML. Treść tekstu to nie tylko zawartość znacznika <BODY> (o ile dokument zawiera taki znacznik). Treść to wszystko, co jest umieszczone za nagłówkiem HTTP w pliku. Aby określić ten wzorzec, można użyć gwiazdki (\*) jako znaku zastępczego. W tym wzorcu rozróżniana jest wielkość liter.
- Liczba całkowita, która reprezentuje kod powrotu dla dokumentów pasujących do określonych wzorców adresu URL, tytułu i treści.

## Przykład

Poniższa konfiguracja powoduje, że przeszukiwacz sieci WWW będzie porównywał wszystkie poprawne strony HTML (o kodzie powrotu 200) z serwisu WWW `http://www.mojserwis.pl/ka/*` z określonymi wzorcami tytułu i treści. Jeśli tekst znacznika <TITLE> strony rozpoczyna się od słów "Sorry, the page...", a treść dokumentu zawiera dowolny tekst (\*), wtedy przeszukiwacz będzie traktował stronę tak, jakby otrzymał kod powrotu 404 (strona nie została znaleziona).

Tabela 2. Przykład strony błędu programowego.

| Wzorzec adresu URL                        | Wzorzec tytułu   | Wzorzec treści | Kod powrotu HTTP |
|-------------------------------------------|------------------|----------------|------------------|
| <code>http://www.mojserwis.pl/ka/*</code> | Sorry, the page* | *              | 404              |

W celu obsługi różnych kodów powrotu można utworzyć wiele wpisów dla tego samego serwisu WWW. Każdy kod powrotu z tego samego serwisu WWW wymaga oddzielnego wpisu w konfiguracji przeszukiwacza sieci WWW.

## Stosowanie znaków zastępczych

Wzorce adresu URL, tytułu oraz treści nie są wyrażeniami regularnymi. Znak gwiazdki zastępuje dowolne znaki, aż do wystąpienia pierwszego znaku niebędącego znakiem zastępczym. Na przykład:

\*404 zastępuje *dowolne znaki*404  
 404: \* zastępuje 404: *dowolne znaki*  
 http://\*.mojserwis.pl/\* zastępuje `http://dowolny host.mojserwis.pl/dowolny plik`  
 \* zastępuje *dowolne znaki*

## Wpływ na wydajność

Skonfigurowanie opcji obsługi stron błędów programowych wydłuża czas przeszukiwania, gdyż wszystkie pomyślnie przeszukane strony muszą zostać sprawdzone. Dłuższy czas przetwarzania wymagany jest do sprawdzenia dopasowania wzorców i określenia, czy zwrócona powinna zostać strona, czy raczej kod powrotu.

## Przeszukiwanie zabezpieczonych serwisów WWW

Określenie referencji w konsoli administracyjnej wyszukiwania korporacyjnego umożliwia przeszukiwaczowi sieci WWW dostęp do zabezpieczonych treści (np. dokumentów wymagających podania hasła dostępu).

Jeśli serwer WWW używa podstawowego uwierzytelniania HTTP lub uwierzytelniania opartego na formularzu HTML w celu ograniczenia dostępu do serwisów WWW, można określić referencje w konfiguracji przeszukiwacza sieci WWW, które umożliwią przeszukiwanie serwisów WWW zabezpieczonych hasłem. Istnieje także możliwość ręcznego skonfigurowania plików z informacjami cookie.

### Serwisy WWW zabezpieczone przez podstawowe uwierzytelnianie HTTP

Jeśli serwer WWW używa podstawowego uwierzytelniania HTTP w celu ograniczenia dostępu do serwisów WWW, można określić referencje uwierzytelniania, które umożliwią przeszukiwaczowi sieci WWW dostęp do stron zabezpieczonych hasłem.

Wiele serwerów WWW używa schematu uwierzytelniania klienta nazywanego podstawowym uwierzytelnianiem HTTP w celu określenia tożsamości użytkownika i ustalenia, czy użytkownik (lub aplikacja klienta) ma wystarczające uprawnienia dostępu do stron w serwisie WWW. Zazwyczaj takie uwierzytelnianie ma charakter interaktywny:

- Kiedy agent HTTP użytkownika (np. przeglądarka WWW) żąda dostępu do strony zabezpieczonej za pomocą podstawowego uwierzytelniania HTTP, serwer WWW zwraca kod powrotu 401, który oznacza, że klient wysyłający żądanie nie ma uprawnień dostępu do żądanej strony.
- Serwer WWW nakazuje także klientowi wysyłającemu żądanie podanie referencji, które pomogą sprawdzić, czy użytkownik może uzyskać dostęp do zabezpieczonych treści.
- W przeglądarce WWW wyświetlane jest okno dialogowe, w którym należy podać nazwę użytkownika, hasło oraz inne wymagane informacje umożliwiające określenie referencji użytkownika.
- Przeglądarka WWW koduje referencje i wykorzystuje je ponownie przy kolejnym żądaniu dostępu do zabezpieczonej strony.
- Jeśli referencje są poprawne, serwer WWW odpowiada kodem powrotu 200 i udostępnia treść żądanej strony.
- Kolejne żądania dostępu do stron wysyłane z tego samego serwera WWW zawierają zazwyczaj te same referencje, co umożliwia autoryzowanym użytkownikom dostęp do zabezpieczonej treści bez konieczności podawania referencji przy każdym żądaniu.

Gdy tożsamość użytkownika zostanie już określona, serwer WWW i agent HTTP użytkownika zazwyczaj wymieniają tokeny nazywane informacjami *cookie*, które dostarczają danych o stanie logowania użytkownika między kolejnymi żądaniami HTTP.

Ponieważ przeszukiwacz sieci WWW nie działa w sposób interaktywny, referencje umożliwiające przeszukiwanie stron zabezpieczonych hasłem należy określić przed rozpoczęciem przeszukiwania. Podczas tworzenia przeszukiwacza sieci WWW lub edycji przestrzeni przeszukiwania należy określić informacje o każdym zabezpieczonym serwisie WWW, który ma zostać przeszukany.

Informacje te należy określić wspólnie z administratorami serwisów WWW lub serwerów WWW używających uwierzytelniania podstawowego HTTP. Administratorzy powinni udostępnić wymagania bezpieczeństwa dla serwisów WWW, które mają zostać przeszukane, w tym wszystkie informacje niezbędne do uwierzytelnienia tożsamości przeszukiwacza sieci WWW i określenia jego uprawnień do przeszukiwania zabezpieczonych stron.

Jeśli w trakcie tworzenia kolekcji zostały włączone zabezpieczenia, to w ramach konfigurowania przeszukiwacza można określić tokeny zabezpieczeń (np. identyfikatory użytkowników, identyfikatory grup lub role użytkowników) w celu kontroli dostępu do dokumentów. Przeszukiwacz sieci WWW powiąże te tokeny zabezpieczeń z każdym przeszukiwanym dokumentem w drzewie systemu plików dla określonego głównego adresu URL. Te tokeny używane są oprócz tokenów zabezpieczeń na poziomie dokumentu, które zostały skonfigurowane dla całej przestrzeni przeszukiwania sieci WWW.

Istotne znaczenia ma kolejność adresów URL. Po dodaniu informacji o serwisie zabezpieczonym hasłem należy określić kolejność, w jakiej przeszukiwacz będzie przetwarzał adresy URL. Szczegółowo określone adresy URL powinny znaleźć się na początku listy, a bardziej ogólne na końcu. Przeszukiwacz sieci WWW podczas oceny danego adresu URL użyje danych uwierzytelnienia określonych dla pierwszego adresu URL, który odpowiada temu adresowi URL.

### **Serwisy WWW zabezpieczone przez uwierzytelnianie oparte na formularzu**

Jeśli serwer WWW używa formularzy HTML w celu ograniczenia dostępu do serwisów WWW, można określić referencje uwierzytelniania, które umożliwią przeszukiwaczowi sieci WWW dostęp do stron zabezpieczonych hasłem.

Wiele serwerów WWW używa formularzy HTML w celu określenia tożsamości użytkownika i ustalenia, czy użytkownik (lub aplikacja klienta) ma wystarczające uprawnienia dostępu do stron w serwisie WWW. Zazwyczaj takie uwierzytelnianie ma charakter interaktywny:

- Kiedy agent HTTP użytkownika (np. przeglądarka WWW) żąda dostępu do strony zabezpieczonej za pomocą uwierzytelniania opartego na formularzu, serwer WWW sprawdza, czy żądanie zawiera informację cookie potwierdzającą tożsamość użytkownika.
- Jeśli nie ma takiej informacji, serwer WWW wyświetla zachętę do wprowadzenia w formularzu danych bezpieczeństwa. Po wysłaniu formularza przez użytkownika serwer WWW zwraca wymagane informacje cookie i zezwala na żądanie dostępu do zabezpieczonej hasłem strony.
- Kolejne żądania, które zawierają wymagane informacje cookie, również mają dostęp do zabezpieczonych stron. Autoryzowany użytkownik uzyskuje dostęp do dodatkowych zabezpieczonych treści bez konieczności wypełniania formularza i podawania referencji przy każdym żądaniu.

Ponieważ przeszukiwacz sieci WWW nie działa w sposób interaktywny, referencje umożliwiające przeszukiwanie stron zabezpieczonych hasłem należy określić przed rozpoczęciem przeszukiwania. Podczas tworzenia przeszukiwacza sieci WWW lub edycji przestrzeni przeszukiwania należy określić informacje o każdym zabezpieczonym serwisie WWW, który ma zostać przeszukany.

Określone pola muszą odpowiadać polom, które użytkownik wypełnia w trybie interaktywnym w przeglądarce WWW. W celu pomyślnego logowania należy określić również wszystkie pola ukryte oraz statyczne.

Informacje te należy określić wspólnie z administratorami serwisów WWW lub serwerów WWW używających uwierzytelniania opartego na formularzu. Administratorzy powinni udostępnić wymagania bezpieczeństwa dla serwisów WWW, które mają zostać przeszukane,

w tym wszystkie informacje niezbędne do uwierzytelnienia tożsamości przeszukiwacza sieci WWW i określenia jego uprawnień do przeszukiwania zabezpieczonych stron.

Istotne znaczenia ma kolejność wzorców adresów URL. Po dodaniu informacji o serwisie zabezpieczonym hasłem należy określić kolejność, w jakiej przeszukiwacz będzie przetwarzał adresy URL. Szczegółowo określone wzorce adresów URL powinny znaleźć się na początku listy, a bardziej ogólne na końcu. Przeszukiwacz sieci WWW podczas oceny danego adresu URL użyje danych formularza określonych dla pierwszego wzorca adresu URL, który odpowiada temu adresowi URL.

## Serwisy WWW obsługiwane przez serwery proxy

Jeśli przeszukiwacz sieci WWW nie ma zezwolenia na bezpośredni dostęp do sieci, można skonfigurować go, aby korzystał z serwera proxy HTTP w celu uzyskania dostępu do treści, która ma zostać przeszukana.

Jeśli na komputerze, na którym ma zostać uruchomiony przeszukiwacz sieci WWW, nie ma dostępu do sieci TCP/IP lub dostęp jest ograniczony do procesów uprzywilejowanych, można skonfigurować przeszukiwacz do korzystania z serwera HTTP proxy. HTTP proxy to proces, który nasłuchuje żądań HTTP na określonym porcie podanego hosta. Serwer proxy przekazuje żądania do serwera WWW i odpowiedzi z tego serwera do klienta wysyłającego żądania (przeszukiwacz sieci WWW). Serwer proxy może działać na tym samym komputerze, na którym działa przeszukiwacz sieci WWW, lub na innym.

W trakcie przeszukiwania bez pośrednictwa serwera proxy żądania adresu URL są wysyłane bezpośrednio do hosta. W trakcie przeszukiwania za pośrednictwem serwera proxy żądania są wysyłane do serwera proxy.

Podczas tworzenia przeszukiwacza sieci WWW lub edycji przestrzeni przeszukiwania należy określić informacje o serwerach proxy, z których ma korzystać przeszukiwacz sieci WWW w trakcie przeszukiwania stron w domenie serwera proxy. Przed dodaniem serwera proxy do przestrzeni przeszukiwania należy uzyskać następujące informacje:

### Domeny serwera proxy

Domeny, które są obsługiwane przez serwer proxy. Jako znaku zastępczego można używać gwiazdki (\*). Na przykład \* odpowiada wszystkim domenom obsługiwanym przez ten serwer proxy, a \*.resource.com odpowiada wszystkim domenom, których nazwa kończy się na resource.com.

**Ograniczenie:** Nie można określać reguł adresów IP do przeszukiwania serwera proxy, ponieważ adres IP serwera proxy jest jedynym adresem IP innego hosta dostępnym dla przeszukiwacza. Jeśli reguły adresów IP służą do ograniczania przeszukiwacza do podsieci adresów IP, ograniczenie to powoduje klasyfikowanie prawie wszystkich adresów URL z kodem powrotu 760 (oznaczającym, że są one zabronione przez obszar WWW).

### Nazwa hosta lub adres IP serwera proxy

Nazwa hosta DNS lub zapisany z kropkami adres IP serwera proxy.

### Numer portu serwera proxy

Numer portu TCP/IP, na którym serwer proxy nasłuchuje żądań proxy HTTP.

Po dodaniu serwera proxy należy go wybrać i określić kolejność, w jakiej przeszukiwacz będzie z niego korzystał. Szczegółowo określone domeny powinny znaleźć się na początku listy, a bardziej ogólne na końcu. Gdy przeszukiwacz sieci WWW ocenia dany adres URL, używa danych serwera proxy, które są określone dla pierwszej domeny z listy, która

odpowiada temu adresowi URL. Zakłada się, że adresy URL, które nie pasują do żadnej reguły proxy, są bezpośrednio dostępne dla przeszukiwacza.

## Administrowanie informacjami cookie

Zazwyczaj administrowanie informacjami cookie odbywa się w sposób automatyczny bez konieczności podejmowania żadnych działań ze strony administratora wyszukiwania korporacyjnego. Jednak w razie potrzeby można ręcznie określić informacje cookie dla sesji przeszukiwania sieci WWW.

Informacje cookie to osłonięte tokeny, które serwer WWW zwraca do agenta użytkownika jako część nagłówka odpowiedzi HTTP. Informacje cookie są ważne jedynie dla tego serwera WWW, które je wystawił, i są używane do utrzymywania stanu między kolejnymi żądaniami HTTP. Na przykład podczas uwierzytelniania klienta serwer WWW może zwracać informacje cookie, które pozwalają serwerowi określić, czy uwierzytelniony użytkownik jest już zalogowany. Obecność informacji cookie umożliwia użytkownikowi wysyłanie dodatkowych żądań dotyczących stron na tym serwerze WWW bez konieczności ponownego logowania.

Przeszukiwacz sieci WWW zachowuje informacje cookie otrzymane z serwerów WWW i używa ich w czasie działania instancji przeszukiwacza. Informacje cookie są przechowywane w pliku cookies.ini, który jest ponownie zapisywany na koniec każdej sesji przeszukiwacza. Po zatrzymaniu działania przeszukiwacza sieci WWW wszystkie informacje cookie, które nie wygasły, są zapisywane i przeladowywane na początku następnej sesji.

Ręcznie określone informacje cookie należy zapisać w oddzielnym pliku, a następnie w razie potrzeby scalić je z informacjami cookie z pliku cookies.ini. Przeszukiwacz nie usuwa informacji cookie, które nie wygasły, ale w razie problemu uniemożliwiającego zapisanie całej kolekcji informacji cookie informacje wprowadzone ręcznie nie zostaną utracone. Przed rozpoczęciem sesji przeszukiwania należy ręcznie scalić własne informacje cookie z informacjami, które są obsługiwane automatycznie przez przeszukiwacz.

### Format pliku cookie

Informacje cookie, które mają być scalone z plikiem cookies.ini systemu wyszukiwania korporacyjnego, muszą mieć ściśle określony format.

- Każda informacja cookie musi się znajdować w osobnym wierszu. Dopuszczalne jest użycie pustych wierszy i komentarzy, które jednak nie zostaną zachowane w pliku cookies.ini.
- Każda informacja cookie musi mieć następujący format:

```
CookieN(długość_cookie,długość_URL)tekst_cookie,adres_URL_sprawdzania_poprawności
```

#### Cookie

Wymagany parametr definiujący początek wpisu informacji cookie.

Parametr Cookie nie może zawierać spacji i musi mieć dołączoną pojedynczą cyfrę: 0, 1 lub 2. Cyfra określa typ informacji cookie: wersja 0 (Netscape), wersja 1 (RFC2109) lub wersja 2 (RFC2965). Informacje cookie typu RFC2965 nie obsługują list portów.

#### długość\_cookie

Długość tekstu cookie, wyrażona liczbą znaków.

#### długość\_URL

Długość adresu URL sprawdzania poprawności, wyrażona liczbą znaków.

#### tekst\_cookie

Zawartość informacji cookie, która ma zostać wysłana do źródłowego serwera WWW. Łańcuch ten (odpowiadający prawej stronie dyrektywy Set-Cookie w nagłówku odpowiedzi HTTP) określa parę nazwa-wartość oraz wszelkie inne dane (na przykład

ścieżkę, ustawienia zabezpieczeń itp.), które będą wysyłane razem z informacją cookie. Po tym łańcuchu następuje separator w postaci przecinka (,).

#### *adres\_URL\_sprawdzania\_poprawności*

Adres URL, pod którym dana informacja cookie została odkryta. Na podstawie tego adresu ustala się, gdzie należy wysłać informację cookie (na przykład przez podanie nazwy domeny i ścieżki). Adres URL sprawdzania poprawności musi spełniać wymagania dotyczące zabezpieczeń i poufności dla informacji cookie obowiązujące na źródłowym serwerze WWW.

W poniższym przykładzie wprowadzono podział na dwa wiersze dla zwiększenia czytelności. Informacje cookie określane przez użytkowników muszą w rzeczywistości mieścić się w jednym wierszu:

```
Cookie0(53,40)ASPSESSIONIDQSQTACSD=SLNSIDFNLSIDNFLSINFLSNL;path=/,
https://www.ibm.com:443/help/solutions/
```

## **Konfigurowanie informacji cookie dla przeszukiwacza sieci WWW**

Informacje cookie można skonfigurować ręcznie na potrzeby sesji przeszukiwania WWW oraz scalić je z informacjami cookie przechowywanymi przez przeszukiwacz WWW.

### **Zagadnienia wstępne**

Aby ręcznie skonfigurować informacje cookie dla przeszukiwacza sieci WWW, trzeba być administratorem wyszukiwania korporacyjnego.

### **Procedura**

Aby ręcznie skonfigurować informacje cookie dla przeszukiwacza sieci WWW:

1. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom monitorowanie kolekcji, dla której chcesz skonfigurować informacje cookie, i zatrzymaj przeszukiwacz WWW.
2. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerze przeszukiwania (ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition).
3. Przejdź do katalogu danych przeszukiwacza, który ma zostać skonfigurowany, gdzie *identyfikator\_sesji\_przeszukiwacza* to identyfikator przypisany sesji przeszukiwacza przez system wyszukiwania korporacyjnego. Na przykład:  
`ES_NODE_ROOT/data/col_56092.WEB_88534`
4. Zmodyfikuj plik `cookies.ini`, dołącz zdefiniowane ręcznie pozycje cookie do pozycji obecnych już w pliku, a następnie zapisz i zamknij plik. Upewnij się, że nowe informacje cookie nie przesłoniły tych, które już były w pliku obecne.
5. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom ponownie zatrzymany wcześniej przeszukiwacz sieci WWW.

## **Konfiguracja globalnej przestrzeni przeszukiwania sieci WWW**

Możliwość konfigurowania globalnej przestrzeni przeszukiwania dla przeszukiwaczy sieci WWW zapewnia lepszą kontrolę usuwania adresów URL z indeksu.

Każdy przeszukiwacz sieci WWW jest skonfigurowany za pomocą przestrzeni przeszukiwania, która określa adresy URL, które mają lub nie mają być przeszukiwane. Adresy URL znajdujące się w przestrzeni przeszukiwania są zachowywane (w bazie danych) w celu późniejszego przeszukania; adresy URL, które nie znajdują się w przestrzeni przeszukiwania, są usuwane. Jeśli przeszukiwacz zostanie uruchomiony z pustą bazą danych, definicja przestrzeni przeszukiwania oraz baza danych pozostają spójne w trakcie działania przeszukiwacza.



Czasami przeszukiwacz zostaje zatrzymany, a jego przestrzeń przeszukiwania zmniejszona (na przykład przez nowe reguły zabraniające przeszukiwania pewnych stron). W takiej sytuacji po zrestartowaniu przeszukiwacza definicja jego przestrzeni przeszukiwania i baza danych stają się niespójne. Baza danych zawiera wówczas adresy URL (zarówno te przeszukane, jak i nieprzeszukane), których nie ma już w nowej, mniejszej przestrzeni przeszukiwania.

Jeśli kolekcja zawiera tylko jeden przeszukiwacz sieci WWW, wtedy przeszukiwacz może przywrócić spójność, zmieniając kody powrotu HTTP dla tych adresów URL na 760 (co określa, że adresy mają zostać wykluczone) i żądając usunięcia nowo wykluczonych stron z indeksu.

Jeśli przestrzeń przeszukiwania jest podzielona między dwa lub więcej przeszukiwaczy sieci WWW (na przykład w celu zapewnienia, że pewne strony będą przeszukiwane częściej niż pozostałe), to każdy przeszukiwacz sieci WWW obsługuje niezależne tabele w bazie danych (początkowo puste) i każdy przeszukuje inną część przestrzeni przeszukiwania sieci WWW. Pierwotna przestrzeń przeszukiwania przeszukiwacza zostaje pomniejszona o wszystkie części, które mają być przeszukiwane przez pozostałe przeszukiwacze. Problemy pojawiają się, kiedy pierwotny przeszukiwacz próbuje przywrócić spójność, usuwając z indeksu przeniesione strony. Ponieważ przeniesione strony są obecnie przeszukiwane przez inne przeszukiwacze, powinny pozostać w indeksie.

Konfiguracja przestrzeni globalnej, która jest przestrzenią przeszukiwania wyższego poziomu, umożliwia określenie adresów URL, które mają nie być przeszukiwane przez pierwotny przeszukiwacz, ale nie powinny być także usunięte z indeksu. Adresy URL, których nie ma w żadnej przestrzeni przeszukiwania, dalej są oznaczone jako wykluczone z procesu wykrywania i zostaną usunięte z indeksu po ponownym przeszukaniu.

Globalna przestrzeń przeszukiwania jest określona za pomocą pliku konfiguracyjnego o nazwie `global.rules`. Plik ten musi istnieć w katalogu konfiguracyjnym przeszukiwacza (istnienie pliku `global.rules` uaktywnia funkcję obsługi globalnej przestrzeni przeszukiwania). Jeśli plik istnieje, jest odczytywany podczas inicjowania przeszukiwacza. Jeśli plik nie istnieje, przeszukiwacz działa w jednopoziomowej przestrzeni przeszukiwania, a dokumenty są usuwane z indeksu w celu zapewnienia spójności między definicją przestrzeni przeszukiwania i bazą danych.

Jeśli istnieje globalna przestrzeń przeszukiwania, obowiązują dalej te same reguły włączania i wyłączania adresów URL, ale żądania usunięcia adresu URL z indeksu będą realizowane tylko, gdy adres URL nie istnieje w żadnej przestrzeni przeszukiwania sieci WWW.

Składnia pliku `global.rules` jest taka sama, jak w przypadku pliku reguł lokalnych `crawl.rules`, jednak plik globalny może zawierać tylko reguły nazw domen. To ograniczenie umożliwia partycjonowanie przestrzeni przeszukiwania między przeszukiwacze wyłącznie na podstawie nazw hostów DNS, a nie na podstawie adresów IP lub wzorców przedrostków HTTP. Globalna przestrzeń przeszukiwania nie ma wpływu na adresy URL, które zostały wykluczone za pomocą reguł dotyczących adresów IP lub przedrostków adresów URL w lokalnej przestrzeni przeszukiwania (określonych w pliku `crawl.rules`); takie adresy URL dalej pozostaną wykluczone.

Globalna przestrzeń przeszukiwania jest używana tylko do zapobiegania usuwaniu z indeksu adresów URL, które zostały wykluczone z przestrzeni przeszukiwania jednego z przeszukiwaczy za pomocą lokalnej reguły domeny. Reguły są stosowane w następującym porządku:

1. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły przedrostków lub reguły adresów, adresowi URL zostaje przypisany kod 760 i adres jest usuwany z indeksu. Adres URL nie będzie ponownie przeszukiwany.

2. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły domeny, a nie istnieje globalna przestrzeń przeszukiwania, wtedy adresowi URL zostaje przypisany kod 760 i adres jest usuwany z indeksu. Adres URL nie będzie ponownie przeszukiwany.
3. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły domeny, ale jest jawnie dozwolony przez regułę w globalnej przestrzeni przeszukiwania, wtedy adresowi URL zostaje przypisany kod 761. Adres URL nie będzie ponownie przeszukiwany, ale nie zostanie usunięty z indeksu (zakłada się, że adres URL znajduje się w lokalnej przestrzeni przeszukiwania innego przeszukiwacza).
4. Jeśli adres URL z bazy danych przeszukiwacza został wykluczony za pomocą lokalnej reguły domeny i nie jest jawnie dozwolony przez regułę w globalnej przestrzeni przeszukiwania, wtedy adresowi URL zostaje przypisany kod 760 i adres jest usuwany z indeksu.

Ponieważ globalna przestrzeń przeszukiwania jest sprawdzana tylko w celu zapobiegania usuwaniu adresów URL wykluczonych już z lokalnej przestrzeni przeszukiwania, to domyślne zabrania ona przeszukiwania adresów URL, o ile dla tych adresów nie zostały zastosowane inne reguły.

Plik `global.rules` musi istnieć w katalogu `master_config` każdego przeszukiwacza, który współużytkuje globalną przestrzeń przeszukiwania. Należy starannie wyedytować wszystkie kopie pliku `global.rules` oraz poszczególnych plików lokalnych `crawl.rules`, aby zapewnić ich wzajemną spójność.

## Dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie

Jakość wyszukiwania można poprawić, określając dyrektywy dla przeszukiwacza sieci WWW, które decydują o tym, czy należy podążać za odsyłaczami znajdującymi się na stronie i czy strona ma być indeksowana.

Pewne strony WWW zawierają dyrektywy blokujące wykorzystywanie odsyłaczy lub indeksowanie, które informują roboty (takie jak przeszukiwacz sieci WWW) o tym, aby nie podążać za odsyłaczami znalezionymi na stronie, nie uwzględniać zawartości strony w indeksie albo nie wykonywać żadnego z tych działań.

Sterowanie tymi ustawieniami może poprawić jakość przeszukiwania. Na przykład pewne strony katalogów mogą zawierać tysiące odsyłaczy, ale żadnej użytecznej treści. Odsyłacze z takich stron powinny zostać przeszukane, ale poindeksowanie samych stron nie przyniesie żadnych korzyści.

Czasami użytkownik chce, aby przeszukiwacz nie przeszukiwał dokumentów poniżej pewnego poziomu w hierarchii, jednak żądane strony-liście zawierają odsyłacze, ale nie zawierają dyrektyw blokujących ich wykorzystywanie. Ponieważ wiele takich stron jest generowanych automatycznie, nie mają one właścicieli, którzy mogliby wstawić wymagane dyrektywy.

W celu określenia reguł przeszukiwania takich stron można utworzyć bądź wyedytować plik konfiguracyjny o nazwie `followindex.rules`. Poniżej przedstawiono wskazówki pomocne przy określaniu reguł w tym pliku:

- Konfigurowane reguły muszą zawierać przedrostki URL (nie można określać serwisów WWW za pomocą adresu IP lub nazwy hosta DNS).
- Przedrostki URL mogą zawierać gwiazdki (\*) jako znaki zastępcze, co umożliwia sformułowanie zezwolenia lub zakazu dla wielu serwisów o podobnych adresach URL.

- Kolejność reguł ma znaczenie (przeszukiwacz stosuje pierwszą regułę, która pasuje do sprawdzanego adresu URL).
- Reguły, które jawnie zezwalają lub zabraniają korzystania z odsyłaczy lub indeksowania, zastępują inne ustawienia (także ustawienia znajdujące się w dokumencie docelowym).

## Ignorowanie dyrektyw blokujących wykorzystywanie odsyłaczy i indeksowanie na stronach WWW

W pliku konfiguracyjnym można określić reguły decydujące o tym, czy przeszukiwacz WWW ma podążać za odsyłaczami do stron lub indeksować strony, które zawierają dyrektywy blokujące takie działania.

### Zagadnienia wstępne

Aby skonfigurować dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie dla przeszukiwacza sieci WWW, trzeba być administratorem wyszukiwania korporacyjnego. Dyrektywy te przesłaniają dyrektywy zawarte w kodzie przeszukiwanych stron.

### Procedura

Aby przesłonić dyrektywy blokujące wykorzystywanie odsyłaczy i indeksowanie:

1. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom monitorowanie kolekcji, dla której chcesz skonfigurować reguły, i zatrzymaj przeszukiwacz sieci WWW.
2. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerze przeszukiwania (ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition).
3. Przejdź do katalogu konfiguracyjnego przeszukiwacza, który ma zostać skonfigurowany, gdzie *identyfikator\_sesji\_przeszukiwacza* to identyfikator przypisany sesji przeszukiwacza przez system wyszukiwania korporacyjnego. Na przykład:  
ES\_NODE\_ROOT/master\_config/col\_56092.WEB\_88534
4. Utwórz lub zmodyfikuj plik o nazwie followindex.rules.
5. Wprowadź reguły dla przeszukiwacza w poniższym formacie, gdzie *przedrostek\_URL* to początkowe znaki w adresach serwisów WWW, których mają dotyczyć reguły zezwalające lub zakazujące wykorzystywania odsyłaczy lub indeksowania:  
forbid follow *przedrostek\_URL*  
allow follow *przedrostek\_URL*  
forbid index *przedrostek\_URL*  
allow index *przedrostek\_URL*
6. Zapisz i zamknij plik.
7. W konsoli administracyjnej wyszukiwania korporacyjnego uruchom ponownie zatrzymany wcześniej przeszukiwacz sieci WWW.

---

## przeszukiwacze WebSphere Portal

Aby do kolekcji wyszukiwania korporacyjnego włączyć strony serwisu IBM WebSphere Portal, należy skonfigurować przeszukiwacz WebSphere Portal.

### Konfiguracja serwera WebSphere Portal

Przed utworzeniem przeszukiwacza WebSphere Portal konieczne jest wdrożenie aplikacji korporacyjnej ESPACServer.ear na serwerze WebSphere Portal. Ta aplikacja korporacyjna jest instalowana na serwerach wyszukiwania podczas instalacji produktu WebSphere Information Integrator OmniFind Edition. Aby wdrożyć tę aplikację korporacyjną, należy wykonać następujące zadanie:

- “Wdrażanie aplikacji korporacyjnej dla przeszukiwacza WebSphere Portal”.

## Przegląd konfiguracji

Przeszukiwacz WebSphere Portal może być używany do przeszukiwania pojedynczego serwisu WebSphere Portal. Podczas konfigurowania przeszukiwacza użytkownik określa adres URL serwisu portalu do przeszukiwania. Przeszukiwacz pobiera następnie portlety dostępne w określonym portalu.

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają sposób przeszukiwania przez niego wszystkich stron serwisu.
- Określenie adresu URL serwisu portalu do przeszukiwania oraz informacji umożliwiających przeszukiwaczowi nawiązanie połączenia z danym serwisem. Ponieważ adresy URL tego typu mogą być długie i zawierać zakodowane znaki spoza zestawu znaków ASCII, taki adres URL wygodnie jest skopiować z serwera WebSphere Portal i wkleić w konsoli administracyjnej wyszukiwania korporacyjnego.
- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

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

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

## Wdrażanie aplikacji korporacyjnej dla przeszukiwacza WebSphere Portal

Przed utworzeniem przeszukiwacza WebSphere Portal konieczne jest wdrożenie aplikacji korporacyjnej ESPACServer.ear na serwerze WebSphere Portal.

### Informacje o tym zadaniu

Plik ESPACServer.ear jest instalowany na serwerach wyszukiwania w katalogu ES\_INSTALL\_ROOT/bin podczas instalacji produktu WebSphere II OmniFind Edition. Domyślne ścieżki instalacyjne to:

#### Systemy UNIX:

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

#### Systemy Windows:

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

### Procedura

Aby wdrożyć aplikację korporacyjną umożliwiającą przeszukiwaczowi WebSphere Portal przeszukiwanie serwisów portalu WebSphere Portal:

1. Zatrzymaj instancję serwera WebSphere\_Portal.
2. Jeśli nie jest jeszcze uruchomiona, uruchom instancję serwera WebSphere Application Server o nazwie server1.
3. Na serwerze WebSphere Portal uruchom konsolę administracyjną serwera WebSphere Application Server. Jeśli pojawi się odpowiednie wezwanie, zaloguj się.  
Konsolę administracyjną można otworzyć na kilka sposobów:
  - W systemie Windows należy użyć menu **Start**, aby wybrać odpowiedni program.
  - W przypadku serwera WebSphere Application Server w wersji 5 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/admin`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9090.
  - W przypadku serwera WebSphere Application Server w wersji 6 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/ibm/console`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9060.
4. Kliknij przycisk **Aplikacje**, a następnie kliknij przycisk **Zainstaluj nowe aplikacje**.
5. Kliknij przycisk **Przełóżaj** i wybierz z systemu plik ESPACServer.ear.
6. Kliknij dwukrotnie przycisk **Dalej**. Jeśli pojawi się ostrzeżenie na temat plików strategii, kliknij przycisk **Kontynuuj**.
7. Klikaj przycisk **Dalej** do czasu wyświetlenia strony Odwzorowanie modułów na serwery aplikacji, dokonaj poniższych wyborów i kliknij przycisk **Zastosuj**:
  - a. W polu tekstowym **Klastry i serwery** wybierz opcję **WebSphere:cell=nazwa\_komórki, node=nazwa\_węzła, server=WebSphere Portal**.
  - b. Zaznacz pole wyboru obok modułu **ESPACServer.ear**.
8. Klikaj przycisk **Dalej** do czasu wyświetlenia strony Podsumowanie, a następnie kliknij przycisk **Zakończ**.
9. Kliknij odsyłacz **Zapisz do głównej konfiguracji**, a następnie kliknij przycisk **Zapisz**, aby zapisać wprowadzone zmiany w konfiguracji serwera WebSphere Application Server.
10. Zrestartuj serwer WebSphere Portal.

#### Pojęcia pokrewne

“Integracja wyszukiwania korporacyjnego z portalem WebSphere Portal” na stronie 195  
Można rozszerzyć możliwości wyszukiwania przy użyciu portalu IBM WebSphere Portal przez wdrożenie portletów wyszukiwania korporacyjnego w portalu WebSphere Portal i w Centrum wyszukiwania portalu WebSphere Portal.

## Kopiowanie adresu URL do przeszukania z portalu WebSphere Portal

Aby zmniejszyć prawdopodobieństwo wpisania niepoprawnego adresu URL, podczas konfigurowania przeszukiwacza WebSphere Portal adres URL serwisu WebSphere Portal do przeszukiwania można skopiować i wkleić w odpowiednim polu.

#### Informacje o tym zadaniu

Podczas tworzenia przeszukiwacza WebSphere Portal użytkownik określa adres URL portalu na serwerze WebSphere Portal, który ma być przeszukiwany. Ponieważ takie adresy URL są

długie i zawierają zwykle zakodowane znaki spoza zestawu znaków ASCII, wygodne może okazać się użycie tej procedury w celu skopiowania adresu URL z serwera WebSphere Portal i wklejenia go w konsoli administracyjnej wyszukiwania korporacyjnego.

### Procedura

Aby określić adres URL do przeszukiwania przez przeszukiwacz WebSphere Portal:

1. Gdy w konsoli administracyjnej wyszukiwania korporacyjnego wyświetlana jest strona Serwer WebSphere Portal do przeszukiwania, upewnij się, że serwer WebSphere Portal jest uruchomiony, a następnie zaloguj się na nim jako administrator.
2. W prawym górnym rogu wybierz opcję **Administrowanie**.
3. Wybierz opcję **Ustawienia portalu** w obszarze nawigacyjnym po lewej stronie, a następnie wybierz opcję **Administrowanie wyszukiwaniem**.
4. Na stronie Zarządzanie kolekcjami wyszukiwania wybierz opcję **Kolekcja portalu** w obszarze Kolekcje wyszukiwania. (Jeśli dostępne są inne kolekcje, również można je wybrać).
5. W obszarze Źródła treści w kolekcji kliknij opcję **Dodaj źródło treści**.
6. W polu **Typ źródła przeszukiwania** wybierz opcję **Serwis portalu**. Adres URL serwisu zostanie wyświetlony w polu **Pobieraj powiązane dokumenty, zaczynając od tego adresu URL**.
7. Skopiuj ten adres URL do schowka (na przykład podświetl adres URL, a następnie przytrzymaj klawisz Ctrl, naciskając klawisz Insert).
8. Wróć do konsoli administracyjnej wyszukiwania korporacyjnego i wklej skopiowany adres URL w polu **Adres URL serwisu WebSphere Portal**.

---

## przeszukiwacze systemu plików Windows

Aby do kolekcji wyszukiwania korporacyjnego włączyć dokumenty zapisane w systemie plików Microsoft Windows, należy skonfigurować przeszukiwacz system plików Windows.

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

Jeśli serwer przeszukiwacza jest zainstalowany na komputerze z systemem UNIX, nie można go używać do przeszukiwania źródeł w systemie plików Windows (przeszukiwacz system plików Windows nie jest dostępny na liście dostępnych typów przeszukiwaczy).

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

Podczas tworzenia przeszukiwacza kreator pomaga wykonać następujące zadania:

- Zdefiniowanie właściwości określających działanie przeszukiwacza i używanej przez niego metody korzystania z zasobów systemu. Właściwości przeszukiwacza określają metodę przeszukiwania wszystkich podkatalogów w przestrzeni przeszukiwania.
- Skonfigurowanie harmonogramu przeszukiwania systemów plików.
- Wybór podkatalogów do przeszukania.

Można zdefiniować liczbę poziomów katalogów przeszukiwanych przez określony przeszukiwacz. Aby umożliwić przeszukanie zdalnych systemów plików, należy podać także identyfikator użytkownika i hasło umożliwiające przeszukiwaczowi uzyskanie dostępu do danych.

- Określenie opcji udostępniania dokumentów w podkatalogach dla operacji przeszukiwania. Na przykład można wykluczyć z przestrzeni przeszukiwania określone typy dokumentów albo określić identyfikator użytkownika i hasło, które umożliwi przeszukiwaczowi dostęp do plików w konkretnym podkatalogu.

- Skonfigurowanie opcji zabezpieczeń na poziomie dokumentu. Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, przeszukiwacz może powiązać dane dotyczące zabezpieczeń z dokumentami w indeksie. Dzięki nim aplikacje wyszukujące mogą wymuszać kontrolę dostępu na podstawie przechowywanych list kontroli dostępu lub tokenów zabezpieczeń.

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

Do wymuszenia zabezpieczeń na poziomie dokumentu konieczne jest poprawne skonfigurowanie informacji na temat konta użytkownika i domeny na serwerze przeszukiwacza.

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

#### **Pojęcia pokrewne**

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows” na stronie 188

Aby włączyć opcję sprawdzania bieżących referencji podczas wyszukiwania przez użytkownika dokumentów przeszukanych przez przeszukiwacz system plików Windows, należy skonfigurować konto domeny zarówno na serwerze przeszukiwacza, jak i na serwerze Microsoft Windows.

---

## **Konfiguracja obsługi aplikacji nasłuchiwania danych**

Istnieje możliwość rozszerzenia wyszukiwania korporacyjnego przez wykorzystanie funkcji API nasłuchiwania danych do utworzenia przeszukiwacza zewnętrznego. Niestandardowe aplikacje nasłuchiwania danych mogą dodawać dane do kolekcji lub wskazywać sposób, w jaki przeszukiwacz sieci WWW ma odwiedzać adresy URL.

#### **Zagadnienia wstępne**

Aby konfigurować aplikacje nasłuchiwania danych, wymagana jest rola administratora wyszukiwania korporacyjnego.

#### **Informacje o tym zadaniu**


Aplikacja kliencka nasłuchiwania danych umożliwia przeszukiwanie typów danych, które nie mogą być przeszukiwane przez domyślne przeszukiwacze w wyszukiwaniu korporacyjnym. Przed rozpoczęciem korzystania z aplikacji nasłuchiwania danych należy skonfigurować referencje, aby aplikacja mogła uzyskać dostęp do kolekcji i mogła ją aktualizować.



Gdy aplikacja kliencka nasłuchiwania danych łączy się z aplikacją nasłuchiwania danych, musi przekazać swój identyfikator i hasło oraz identyfikator aktualizowanej kolekcji. Informacje te muszą być zgodne z informacjami skonfigurowanymi dla aplikacji w konsoli administracyjnej.

Proces nasłuchiwania danych jest uruchamiany automatycznie przy uruchamianiu systemu wyszukiwania korporacyjnego. Jeżeli po skonfigurowaniu aplikacji za pomocą konsoli administracyjnej zostanie zmieniony numer portu, należy zatrzymać i ponownie uruchomić aplikację nasłuchiwania danych.

### Procedura

Aby skonfigurować aplikację nasłuchiwania danych, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Nasłuchiwanie danych kliknij opcję **Konfiguruj aplikację nasłuchiwania danych**.
4. Na stronie Aplikacje nasłuchiwania danych określ liczbę wątków, które aplikacja nasłuchiwania danych może utworzyć podczas przetwarzania żądań z aplikacji klienckich, i numer portu, na którym aplikacja nasłuchiwania danych będzie nasłuchiwać żądań.
5. Kliknij opcję **Dodaj aplikację nasłuchiwania danych**, aby dodać informacje o aplikacji klienckiej.
6. Na stronie Dodaj aplikację nasłuchiwania danych określ informacje uwierzytelniające, które pozwolą aplikacjom klienckim nasłuchiwania danych na dostęp do kolekcji wyszukiwania korporacyjnego. Identyfikatory klientów nasłuchiwania danych muszą być unikalne w obrębie systemu wyszukiwania korporacyjnego.
7. Wybierz kolekcje, które aplikacja nasłuchiwania danych będzie mogła aktualizować:
  - Kliknij opcję **Wszystkie kolekcje**, jeśli aplikacja ma aktualizować wszystkie kolekcje.
  - Kliknij opcję **Konkretne kolekcje**, jeśli aplikacja ma aktualizować tylko wskazane kolekcje.

Po wybraniu tej opcji wyświetlona zostaje lista nazw kolekcji. Zaznacz pole wyboru **Wybierz** dla każdej kolekcji, którą aplikacja może aktualizować.
8. Kliknij przycisk **OK**.
9. Jeśli został zmieniony numer portu nasłuchiwania danych, należy zrestartować nasłuchiwanie danych:
  - a. Kliknij opcję  **Monitoruj**, aby przejść do widoku monitorowania systemu.
  - b. Na stronie Nasłuchiwanie danych kliknij opcję  **Restartuj**.

#### Pojęcia pokrewne

"Nasłuchiwanie danych" w podręczniku "Programming Guide and API Reference for Enterprise Search"

#### Zadania pokrewne

"Monitorowanie nasłuchiwania danych" na stronie 231

Monitorowanie nasłuchiwania danych służy do wyświetlania statusu i szczegółów działania klienta aplikacji nasłuchiwania danych.

---

## Niestandardowe wtyczki przeszukiwacza

Podczas konfigurowania właściwości dla przeszukiwacza można określić klasę Java używaną do wymuszenia kontroli dostępu na poziomie dokumentu oraz do powiązania metadanych z dokumentami w indeksie wyszukiwania korporacyjnego.

Wtyczka zawiera klasę Java wywoływaną dla każdego dokumentu przeszukiwanego przez przeszukiwacz. Do klasy Java jest przekazywany identyfikator dokumentu (URI) z indeksu wyszukiwania korporacyjnego, tokeny zabezpieczeń i metadane. Klasa zwraca nowy lub zmodyfikowany zbiór tokenów zabezpieczeń i metadanych albo może wskazać, że dokument ma zostać zignorowany przez przeszukiwacz.



Po jednokrotnym przeszukaniu wszystkich dokumentów w przestrzeni przeszukiwania wtyczka jest wywoływana tylko dla nowych lub zmodyfikowanych dokumentów. Aby zmienić tokeny zabezpieczeń i metadane dla dokumentów znajdujących się w indeksie wyszukiwania korporacyjnego, które nie zostały zaktualizowane w oryginalnym źródle danych, należy uruchomić pełne przeszukiwanie wszystkich dokumentów w przestrzeni przeszukiwania, a następnie zreorganizować indeks.

Wtyczki nie można powiązać z istniejącym przeszukiwaczem. Podczas konfigurowania właściwości dla nowego przeszukiwacza należy określić nazwę klasy wtyczki i ścieżkę klasy.

## **Korzystanie z wtyczki do wymuszania zabezpieczeń**

Zabezpieczenia na poziomie dokumentu są wymuszane poprzez powiązanie jednego lub więcej tokenów zabezpieczeń (łańcuchów rozdzielanych przecinkami) z każdym dokumentem przeszukiwanym przez przeszukiwacz. Jako tokeny zabezpieczeń są często używane identyfikatory grupy.

Domyślnie każdy dokument ma przypisany publiczny token, który powoduje udostępnienie dokumentu wszystkim użytkownikom. Publiczny token można zastąpić wartością podaną przez administratora lub wartością wyodrębnioną z pola przeszukiwanego dokumentu.

Wtyczka umożliwia zastosowanie własnych reguł biznesowych w celu określenia wartości tokenów zabezpieczeń dla przeszukiwanych dokumentów. Tokeny zabezpieczeń powiązane z każdym dokumentem są zapisywane w indeksie. Są one używane do filtrowania dokumentów zgodnych z tokenami zabezpieczeń i zapewniają, że w wynikach wyszukiwania są zwracane tylko dokumenty, do których przeglądania użytkownik ma uprawnienia.

## **Korzystanie z wtyczki do dodawania metadanych do dokumentów**

Metadane dokumentu, takie jak data ostatniej modyfikacji dokumentu, są tworzone dla wszystkich przeszukiwanych dokumentów. Wtyczka przeszukiwacza umożliwia zastosowanie własnych reguł biznesowych w celu określenia wartości metadanych, które mają zostać poindeksowane dla każdego dokumentu.

Metadane są tworzone jako para nazwa-wartość. Użytkownicy mogą przeszukiwać metadane za pomocą zapytań pełnotekstowych lub zapytań określających nazwę pola metadanych.

## **Wtyczka przeszukiwacza sieci WWW**

Za pomocą aplikacyjnych interfejsów programistycznych przeszukiwacza sieci WWW można sterować sposobem przeszukiwania dokumentów i sposobem ich przygotowywania do analizy. Na przykład można dodać pola do nagłówka żądania HTTP, który zostanie użyty, gdy przeszukiwacz zażąda dokumentu. Po przeszukaniu dokumentu, a przed jego przeanalizowaniem i podzieleniu na leksemy, można zmienić jego treść, tokeny zabezpieczeń i metadane. Można również wyłączyć wysyłanie dokumentu do analizatora składni.

### **Pojęcia pokrewne**

"Crawler plug-ins" (Wtyczki przeszukiwaczy) w podręczniku "Programming Guide and API Reference for Enterprise Search"

---

## **Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego**

Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

Identyfikatory URI lub wzorce URI można określić podczas konfigurowania kategorii, zasięgów oraz szybkich odsyłaczy do kolekcji. Można również określić identyfikator URI w razie konieczności usunięcia dokumentów z indeksu lub wyświetlenia szczegółowych informacji o statusie konkretnego identyfikatora URI.

Przeszukaj kolekcję w celu określenia identyfikatorów URI lub ich wzorców dla dokumentu. Kliknięcie identyfikatorów URI w wynikach wyszukiwania powoduje pobranie interesujących użytkownika dokumentów. Identyfikator URI można skopiować z wyników wyszukiwania i korzystać z niego na konsoli administracyjnej wyszukiwania korporacyjnego. Na przykład można określić wzorzec identyfikatora URI do automatycznego powiązania dokumentów, które odpowiadają temu wzorcowi identyfikatora URI, z kategorią wyszukiwania korporacyjnego.

## Przeszukiwacze Content Edition

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Content Edition działający w trybie dostępu do serwera to:

```
vbr://nazwa_serwera/systemowy_id_repozytorium/trwały_id_repozytorium
/id_elementu/id_wersji
/typ_elementu/[?Page=nr_strony&] właściwości_JNDI
```

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Content Edition działający w trybie dostępu bezpośredniego to:

```
vbr:///systemowy_id_repozytorium/trwały_id_repozytorium
/id_elementu/id_wersji
/typ_elementu/[?Page=nr_strony]
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_serwera*

Nazwa serwera WebSphere Information Integrator Content Edition.

*systemowy\_id\_repozytorium*

Systemowy identyfikator repozytorium.

*trwały\_id\_repozytorium*

Trwały identyfikator repozytorium.

*id\_elementu*

Identyfikator elementu.

*id\_wersji*

Identyfikator wersji. Jeśli identyfikator wersji jest pusty, ta wartość wskazuje najnowszą wersję dokumentu.

*typ\_elementu*

Typ elementu (CONTENT lub FOLDER).

*nr\_strony*

Numer strony.

*właściwości\_JNDI*

Właściwości JNDI klienta aplikacji J2EE. Istnieją dwa typy właściwości:

**java.naming.factory.initial**

Nazwa klasy serwera aplikacji użytej do utworzenia uchwytu EJB.

**java.naming.provider.url**

Adres URL do usługi nazw serwera aplikacji użytej do żądania uchwytu EJB.

## Przykłady

Documentum:

```
vbr://vbrsrv.ibm.com/Documentum/c06b/094e827780000302//CONTENT/?
java.naming.provider.url=iiop%3A%2F%2Fmyvbr.ibm.com%3A2809&
java.naming.factory.initial=com.ibm.websphere.naming.WsnInitContextFactory
```

FileNet PanagonCS:

```
vbr://vbrsrv.ibm.com/PanagonCS/4a4c/003671066//CONTENT/?Page=1&
java.naming.provider.url=iiop%3A%2F%2Fmyvbr.ibm.com%3A2809&
java.naming.factory.initial=com.ibm.websphere.naming.WsnInitContextFactory
```

## Przeszukiwacze DB2

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz DB2 to:

```
db2://nazwa_bazy_danych/nazwa_tabeli
/nazwa_kolumny1_unikalnego_id/wartość1_unikalnego_id
[/nazwa_kolumny2_unikalnego_id/wartość2_unikalnego_id/...
/nazwa_kolumnyN_unikalnego_id/wartośćN_unikalnego_id]
```

### Parametry:

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_bazy\_danych*

Wewnętrzna nazwa bazy danych lub alias bazy danych.

*nazwa\_tabeli*

Nazwa tabeli docelowej, w tym nazwa schematu.

*nazwa\_kolumny1\_unikalnego\_id*

Nazwa pierwszej kolumny unikalnego identyfikatora w tabeli.

*wartość1\_unikalnego\_id*

Wartość pierwszej kolumny unikalnego identyfikatora.

*nazwa\_kolumnyN\_unikalnego\_id*

Nazwa *n*-tej kolumny unikalnego identyfikatora w tabeli.

*wartośćN\_unikalnego\_id*

Wartość *n*-tej kolumny unikalnego identyfikatora.

### Przykłady

Lokalna baza danych wpisana do katalogu:

```
db2://LOCALDB/SCHEMAT1.TABELA1/MODEL/ThinkPadA20
```

Zdalna baza danych nie wpisana do katalogu:

```
db2://mojserwer.mojafirma.com:50001/REMOVEDB/SCHEMAT2.TABELA2/NAZWA/DAWID
```

## Przeszukiwacze DB2 Content Manager

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz DB2 Content Manager to:

```
cm://nazwa_serwera/nazwa_typu_elementu/PID
```

### Parametry

*nazwa\_serwera*

Nazwa serwera bibliotecznego IBM DB2 Content Manager.

*nazwa\_typu\_elementu*

Nazwa typu docelowego elementu.

*PID*

Trwały identyfikator programu DB2 Content Manager.

### Przykład

```
cm://cmsrvctg/ITEMTYPE1/92+3+ICM8+icmn1sdb12+ITEMTYPE159+26+A1001001A
03F27B94411D1831718+A03F27B+94411D183171+14+1018
```

## Przeszukiwacze Domino Document Manager

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Domino Document Manager to:

```
dominodoc://nazwa_serwera:nr_portu/id_repliki_bazy_danych/ścieżka_i_nazwa_bazy_danych
/uniwersalny_id_widoku/uniwersalny_id_dokumentu
/?AttNo=nr_zalącznika&AttName=nazwa_pliku_zalącznika
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_serwera*

Nazwa serwera Domino Document Manager.

*nr\_portu*

Opcjonalnie: Numer portu serwera Domino Document Manager.

*id\_repliki\_bazy\_danych*

Identyfikator repliki bazy danych.

*ścieżka\_i\_nazwa\_bazy\_danych*

Ścieżka i nazwa pliku bazy danych NSF dokumentów na docelowym serwerze Domino Document Manager.

*uniwersalny\_id\_widoku*

Uniwersalny identyfikator widoku używany przy przeszukiwaniu dokumentów Domino Document Manager.

*uniwersalny\_id\_dokumentu*

Uniwersalny identyfikator dokumentu zdefiniowany w przeszukiwanym dokumencie.

*nr\_zalącznika*

Opcjonalnie: Kolejny numer, zaczynając od zera, dla każdego załącznika.

*nazwa\_pliku\_zalącznika*

Opcjonalnie: Pierwotna nazwa pliku załącznika.

### Przykłady

Dokument Domino Document Manager:

```
dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf
/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
```

Załącznik do dokumentu:

```
dominodoc://dominodocsvr.ibm.com/49256D3A000A20DE/domdoc%2FADMN-6FAJXL.nsf
/8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226
?AttNo=0&AttName=PlikZalącznika.doc
```

## Przeszukiwacze Exchange Server

Format identyfikatorów dla dokumentów przeszukiwanych przez przeszukiwacz Exchange Server to:

```
exchange://ścieżka_OWA[?useSSL=true]
```

### Parametry

*ścieżka\_OWA*

Ścieżka programu Outlook Web Access (OWA) bez protokołu.

### **useSSL=true**

Dodawany, gdy protokół pierwotnej ścieżki OWA to HTTPS.

### **Przykłady**

Treść dokumentu:

```
exchange://exchangesvr.ibm.com/public/FolderGłówny1/Folder1/Dokument.EML
```

Załącznik do dokumentu:

```
exchange://exchangesvr.ibm.com/public/FolderGłówny1/Folder1/Dokument.EML/
DolaczonyPlik.doc
```

Włączone dla SSL:

```
exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/Document.EML
?useSSL=true
```

## **Przeszukiwacze Notes**

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz Notes to:

```
domino://nazwa_serwera[:nr_portu]/id_repliki_bazy_danych/ścieżka_i_nazwa_bazy_danych
/[uniwersalny_id_widoku]/uniwersalny_id_dokumentu
[?AttNo=nr_zalącznika&AttName=nazwa_pliku_zalącznika]
```

### **Parametry**

Kodowanie URL ma zastosowanie do wszystkich pól.

#### *nazwa\_serwera*

Nazwa serwera Lotus Notes.

#### *nr\_portu*

Numer portu serwera Lotus Notes. Numer portu jest opcjonalny.

#### *id\_repliki\_bazy\_danych*

Identyfikator repliki bazy danych.

#### *ścieżka\_i\_nazwa\_bazy\_danych*

Ścieżka i nazwa pliku bazy danych NSF na serwerze docelowym Lotus Notes.

#### *uniwersalny\_id\_widoku*

Uniwersalny identyfikator widoku zdefiniowany w docelowej bazie danych. Ten identyfikator jest określony tylko wtedy, gdy dokument został wybrany z widoku lub foldera. Jeśli widok lub folder do przeszukiwania nie zostanie wskazany (na przykład, jeśli określono, że przeszukiwane mają być wszystkie dokumenty w bazie danych), Uniwersalny identyfikator widoku nie jest określony.

#### *uniwersalny\_id\_dokumentu*

Uniwersalny identyfikator dokumentu zdefiniowany w dokumencie, który jest przeszukiwany przez przeszukiwacz.

#### *nr\_zalącznika*

Kolejny numer, zaczynając od zera, dla każdego załącznika. Numer załącznika jest opcjonalny.

#### *nazwa\_pliku\_zalącznika*

Pierwotna nazwa pliku załącznika. Nazwa pliku załącznika jest opcjonalna.

### **Przykłady**

Dokument wybrany do przeszukania według widoku lub folderu:

domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf/  
8178B1C14B1E9B6B8525624F0062FE9F/0205F44FA3F45A9049256DB20042D226

Dokument, który nie został wybrany do przeszukania według widoku lub folderu:

domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf//  
0205F44FA3F45A9049256DB20042D226

Załącznik do dokumentu:

domino://dominosvr.ibm.com/49256D3A000A20DE/Database.nsf//  
0205F44FA3F45A9049256DB20042D226?AttNo=0&AttName=PlikZalacznika.doc

## Przeszukiwacze QuickPlace

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz QuickPlace to:

*quickplace://nazwa\_serwera:nr\_portu/id\_repliki\_bazy\_danych/ścieżka\_i\_nazwa\_bazy\_danych  
/uniwersalny\_id\_widoku/uniwersalny\_id\_dokumentu  
/?AttNo=nr\_zalacznika&AttName=nazwa\_pliku\_zalacznika*

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_serwera*

Nazwa serwera Lotus QuickPlace.

*nr\_portu*

Opcjonalnie: Numer portu serwera QuickPlace.

*id\_repliki\_bazy\_danych*

Identyfikator repliki bazy danych.

*ścieżka\_i\_nazwa\_bazy\_danych*

Ścieżka i nazwa pliku bazy danych NSF dokumentów na docelowym serwerze QuickPlace.

*uniwersalny\_id\_widoku*

Uniwersalny identyfikator widoku używany przy przeszukiwaniu dokumentów QuickPlace.

*uniwersalny\_id\_dokumentu*

Uniwersalny identyfikator dokumentu zdefiniowany w przeszukiwanym dokumencie.

*nr\_zalacznika*

Opcjonalnie: Kolejny numer, zaczynając od zera, dla każdego załącznika.

*nazwa\_pliku\_zalacznika*

Opcjonalnie: Pierwotna nazwa pliku załącznika.

### Przykłady

Dokument:

*quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace  
%5CPageLibrary4925704300021490.nsf  
/A7986FD2A9CD47090525670800167225  
/2B02B1DE3A82B2CE49257043001C2498*

Załącznik do strony:

*quickplace://1twsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace  
%5CPageLibrary4925704300021490.nsf  
/A7986FD2A9CD47090525670800167225  
/2B02B1DE3A82B2CE49257043001C2498  
?AttNo=0&AttName==QPCons3.ppt*

## Przeszukiwacze system plików UNIX

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz system plików UNIX to:

```
file:///nazwa_katalogu/nazwa_pliku
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_katalogu*

Bezwzględna nazwa ścieżki do katalogu.

*nazwa\_pliku*

Nazwa pliku.

### Przykład

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

## Przeszukiwacze WebSphere Portal

Format identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz WebSphere Portal to:

```
wps://ścieżka_WPS?portletDefID=ID_def_portletu&pageID=id_strony&useSSL=SSL
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*ścieżka\_WPS*

Ścieżka serwera WebSphere Portal, bez nazwy protokołu.

*ID\_def\_portletu*

Identyfikator definicji portletu dla serwera WebSphere Portal.

*id\_strony*

Identyfikator strony dla serwera WebSphere Portal.

*SSL* Jeśli protokołem w ścieżce jest HTTPS, dodawana jest opcja `useSSL=true`.  
W przeciwnym wypadku dodawana jest opcja `useSSL=false`.

### Przykłady

Treść dokumentu:

```
wps://wpsserver.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+
LKnPy1vM0Y_QjzKCN4g3cQbJgQi0-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8qHRj/delta
/base64xml/L01DU1kvd0NrQUpORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!
?portletDefID=3_0_3S&pageID=6_0_6J&useSSL=false
```

### Przykłady

Włączone dla SSL:

```
wps://wpsserver.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+
LKnPy1vM0Y_QjzKCN4g3cQbJgQi0-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8qHRj/delta
/base64xml/L01DU1kvd0NrQUpORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!
?portletDefID=7_0_A4&pageID=6_0_6J&useSSL=true
```

## Przeszukiwacze system plików Windows

Formaty identyfikatorów URI dla dokumentów przeszukiwanych przez przeszukiwacz system plików Windows to:

```
file:///nazwa_katalogu/nazwa_pliku
```

```
file:///nazwa_folderu_sieciowego/nazwa_katalogu/nazwa_pliku
```

### Parametry

Kodowanie URL ma zastosowanie do wszystkich pól.

*nazwa\_katalogu*

Bezwzględna nazwa ścieżki do katalogu.

*nazwa\_pliku*

Nazwa pliku.

*nazwa\_folderu\_sieciowego*

Tylko dla dokumentów znajdujących się na serwerach zdalnych, nazwa folderu współużytkowanego w sieci Windows.

### **Przykłady**

Lokalny system plików:

file:///d:/katalog/test.doc

Sieciowy system plików:

file:///filesvr.ibm.com/katalog/plik.doc

### **Pojęcia pokrewne**

“Administrowanie przeszukiwaczem wyszukiwania korporacyjnego” na stronie 19  
Konfiguracja przeszukiwaczy dla różnych typów danych, które mają zostać włączone do kolekcji. Pojedyncza kolekcja może zawierać dowolną liczbę przeszukiwaczy.

### **Zadania pokrewne**

“Konfiguracja kategorii” na stronie 93

Dla kolekcji można utworzyć dowolną liczbę kategorii, a każda kategoria może zawierać dowolną liczbę reguł. Reguły określają, które dokumenty będą automatycznie powiązane z daną kategorią.

“Konfiguracja zasięgów” na stronie 127

Przy konfigurowaniu zasięgu dla kolekcji wyszukiwania korporacyjnego należy określić identyfikatory URI lub wzorce URI dla zakresu dokumentów w indeksie, który użytkownicy mogą wyszukiwać.

“Usuwanie z indeksu identyfikatorów URI” na stronie 130

Aby zapobiec wyszukiwaniu dokumentów w kolekcji przez użytkowników, należy usunąć z indeksu identyfikatory URI tych dokumentów.

“Konfiguracja szybkich odsyłaczy” na stronie 141

Aby utworzyć szybki odsyłacz do kolekcji wyszukiwania korporacyjnego, należy powiązać identyfikator URI dokumentu ze słowami kluczowymi, które spowodują jej włączenie do wyników wyszukiwania.

“Wyświetlanie szczegółów dotyczących identyfikatora URI” na stronie 216

Użytkownik może wyświetlić szczegółowe informacje na temat identyfikatora URI. Wyświetlić można aktualne i historyczne informacje odnośnie przeszukiwania, indeksowania i wyszukiwania dokumentu reprezentowanego przez dany identyfikator URI.

“Wyświetlanie raportów o usuniętych dokumentach” na stronie 233

Użytkownik może wyświetlić szczegółowe informacje na temat dokumentów, które zostały usunięte z systemu wyszukiwania korporacyjnego. Informacje te dostępne są tylko wtedy, gdy włączona jest opcja śledzenia dokumentów danej kolekcji.



---

## Administrowanie analizatorem składni wyszukiwania korporacyjnego

Aby poprawić wydajność pobierania dokumentów, można określić opcje definiujące sposób analizowania i klasyfikowania dokumentów i metadanych przed dodaniem ich do indeksu wyszukiwania korporacyjnego.

Analizę składniową zawartości dokumentów i optymalizację pobierania informacji można skonfigurować przy użyciu następujących operacji:

### **Konfigurowanie opcji analizy dokumentów chińskich, japońskich i koreańskich**

Na potrzeby analizy składni dokumentów w językach chińskim, japońskim i koreańskim można określić opcje podziału na ciągi n-znakowe. Ponadto można usunąć znaki nowego wiersza z dokumentów chińskich i japońskich.

### **Włączanie rodzimego wyszukiwania XML**

Jeśli do kolekcji należą dokumenty XML, można włączyć przeszukiwanie ich przy użyciu rodzimej składni zapytań XML (np. XPath i fragmenty XML). Rodzime wyszukiwanie XML pozwala formułować zapytania w oparciu o związki między różnymi elementami XML.

### **Konfiguracja kategorii**

Dokumenty, których identyfikator URI jest zgodny z określonym wzorcem, lub dokumenty zawierające określone słowa mogą zostać zgrupowane w kategorii. Gdy użytkownicy przeszukują kolekcję, mogą ograniczyć wyniki wyszukiwania do dokumentów należących do określonych kategorii.

### **Konfigurowanie pól wyszukiwania**

Elementy w dokumentach XML można odwzorowywać na pola wyszukiwania w indeksie. Ponadto do pól wyszukiwania można przypisywać elementy metadanych ze źródeł HTML. Tworząc pola wyszukiwania w indeksie wyszukiwania korporacyjnego, daje się użytkownikom możliwość wprowadzania zapytań dotyczących określonych części dokumentów XML i HTML oraz poprawiania precyzji wyników wyszukiwania.

### **Konfigurowanie opcji przetwarzania tekstu**

Jeśli do systemu wyszukiwania korporacyjnego dodano niestandardowe mechanizmy analizy tekstu, można wybrać jeden z nich do użytku przez daną kolekcję. Po powiązaniu mechanizmu analizy z kolekcją można określić opcje odwzorowywania treści sterujące analizą lingwistyczną i adnotowaniem. Można również określić sposób odwzorowywania wyników analizy na indeks wyszukiwania korporacyjnego lub tabelę bazy danych JDBC.

### **Odwzorowywanie pól na klasy kryteriów**

Można określić, że dokumenty mające pola zgodne z terminami użytymi w zapytaniu będą sklasyfikowane w wynikach wyszukiwania wyżej niż inne dokumenty zgodne z terminami użytymi w zapytaniu. Podczas odwzorowywania pól na klasy kryteriów należy określić, do jakich pól treści i metadanych mają być stosowane kryteria. Można również skonfigurować wyniki, które są używane do oceny dokumentów przez klasę kryteriów.

### **Pojęcia pokrewne**

"Linguistic support for semantic search" (Obsługa lingwistyczna w wyszukiwaniu semantycznym) w podręczniku "Text Analysis Integration"

"Text analysis included in enterprise search" (Analiza tekstu w wyszukiwaniu korporacyjnym) w podręczniku "Text Analysis Integration"

"Unstructured information management architecture (UIMA)" (Architektura zarządzania informacjami bez określonej struktury - UIMA) w podręczniku "Text Analysis Integration"

"Semantic search applications" (Aplikacje wyszukiwania semantycznego) w podręczniku "Text Analysis Integration"

"Semantic search query" (Zapytanie wyszukiwania semantycznego) w podręczniku "Text Analysis Integration"

---

## Praca z kategoriami

Użycie kategorii pozwala na grupowanie dokumentów o wspólnych charakterystykach, a także wyszukiwanie i pobieranie tylko tych dokumentów, które spełniają kryteria przynależności do określonej grupy.

Jeśli dokumenty są przypisane do kategorii, a używane aplikacje wyszukiwujące obsługują tę opcję, użytkownicy mogą przeszukiwać tylko podzbiór kolekcji, określając nazwę kategorii. Jeśli przeszukiwana jest cała kolekcja, użytkownicy mogą ograniczyć zbiór wyników wyszukiwania i przeglądać tylko dokumenty należące do określonej kategorii, do której należy jeden z dokumentów wynikowych.

W wyszukiwaniu korporacyjnym obsługiwane są dwie metody klasyfikowania dokumentów:

### Oparta na regułach

Należy wybrać tę metodę, aby przypisywać dokumenty do kategorii na podstawie zdefiniowanych reguł. Na przykład można zgrupować dokumenty o wspólnym wzorcu URI lub dokumenty, w których występują lub nie występują konkretne słowa i frazy.

### Oparta na modelach

Tę metodę należy stosować, jeśli używany jest system IBM WebSphere Portal i do przeszukiwania kolekcji wyszukiwania korporacyjnego mają być używane predefiniowane kategorie zdefiniowane w systemie WebSphere Portal.

## Wprowadzanie zmian klasyfikowania

Do tworzenia kategorii i zarządzania nimi służy konsola administracyjna wyszukiwania korporacyjnego:

- Typ klasyfikowania wybierany jest podczas tworzenia kolekcji. Istnieje możliwość wyłączenia funkcji klasyfikacji, użycia kategorii opartych na regułach lub kategorii opartych na modelach.
- W razie potrzeby można zmienić typ klasyfikowania, konfigurując reguły analizowania dla kolekcji. Jeśli typ klasyfikowania zostanie zmieniony po przeszukaniu dokumentów i poindeksowaniu ich, jakość wyników wyszukiwania pozostanie niska do czasu ponownego przeszukania wszystkich dokumentów, przeanalizowania ich i zreorganizowania indeksu.
- Jeśli zostanie wybrana opcja kategorii opartych na regułach, wówczas do administrowania drzewem kategorii, kategoriami i regułami kategorii używa się konsoli administracyjnej. Jeśli kategorie lub reguły kategorii zostaną zmienione po przeszukaniu dokumentów i poindeksowaniu ich, jakość wyników wyszukiwania pozostanie niska do czasu ponownego przeszukania wszystkich dokumentów, przeanalizowania ich i zreorganizowania indeksu.

## Pola treści dokumentu

Aby można było utworzyć kolekcję z kategoriami opartymi na modelach lub kategoriami opartymi na regułach, które wykorzystują reguły treści dokumentu (a nie reguły oparte na wzorcach URI), dokumenty mające zostać sklasyfikowane muszą zawierać pola treści.

Kategorie oparte na modelach oraz reguły kategorii oparte na treści dokumentu wykorzystują pola treści dokumentów i nie uwzględniają metadanych dokumentu. Przy użyciu tych opcji mogą być klasyfikowane jedynie te dokumenty, które zawierają pola treści lub pola, które mogą zostać skonfigurowane jako pola treści podczas konfigurowania przeszukiwacza.

W przypadku skonfigurowania kategorii opartych na modelach lub utworzenia reguł kategorii opartych na treści dokumentu reguły te odnoszą się wyłącznie do tych części dokumentu, które stanowią treść. Dla następujących typów przeszukiwaczy reguły te odnoszą się do plików (takich jak treść plików systemu UNIX lub Microsoft Windows) lub do treści stron WWW:

- Exchange Server
- NNTP
- system plików UNIX
- sieć WWW
- WebSphere Portal
- system plików Windows

W przypadku następujących typów przeszukiwaczy pola będące polami treści dokumentu określa się podczas konfigurowania opcji poszczególnych źródeł danych. Podczas tworzenia przeszukiwacza ta opcja jest dostępna na ostatniej stronie kreatora przeszukiwacza. Jeśli przeszukiwacz już istnieje, należy dokonać edycji przestrzeni przeszukiwania, wybrać źródło danych, którego opcje mają zostać zmienione, a następnie określić, które pola w tym źródle danych są polami treści dokumentu.

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

## Kategorie oparte na regułach

Użytkownik konfiguruje reguły określa dokumenty przypisywane do kategorii w kolekcji wyszukiwania korporacyjnego.

Kategorie i reguły kategorii można tworzyć dla kolekcji utworzonych w wyszukiwaniu korporacyjnym i dla kategorii opartych na regułach migrowanych z kolekcji produktu IBM WebSphere Portal. Aby skonfigurować reguły klasyfikowania dokumentów, podczas tworzenia kolekcji lub definiowania reguł analizowania dla kolekcji należy wybrać opcję użycia kategorii opartych na regułach.

W analizatorze składni podczas przypisywania dokumentów do jednej lub wielu kategorii użyte zostaną reguły określone przez użytkownika:

- Jeśli dokument jest zgodny z przynajmniej jedną regułą w kategorii, analizator składni przypisze go do tej kategorii.
- Jeśli dokument jest zgodny z przynajmniej jedną regułą w kilku kategoriach, analizator składni przypisze go do wszystkich tych kategorii.
- Jeśli dokument nie jest zgodny z żadną regułą dla kategorii, analizator składni nie przypisze go do tej kategorii. Podczas przeszukiwania kolekcji użytkownicy mogą wyszukać ten dokument i pobrać go, lecz nie znajdą go, jeśli przeszukiwać będą poszczególne kategorie.

Administrator używanego dla kolekcji drzewa kategorii (lub systematyki) decyduje, gdzie w hierarchii kategorii ma zostać dodana nowa kategoria. W drzewie kategorii można także wybrać kategorię, która ma zostać poddana edycji, a następnie dodać lub usunąć reguły klasyfikowania dokumentów albo zmienić treść poszczególnych reguł.

Konfigurując regułę klasyfikowania dokumentów należy określić, czy w wyszukiwaniu korporacyjnym przynależność dokumentu do określonej kategorii jest ustalana na podstawie identyfikatora URI tego dokumentu, czy na podstawie jego treści:

### **Wzorzec URI**

Reguła oparta na identyfikatorze URI jest stosowana względem identyfikatora URI dokumentu. Użytkownik określa częściowy identyfikator URI (wzorzec) i z regułą zgodne są te dokumenty, których identyfikatory URI zawierają ten wzorzec.

Na przykład jeśli określony jest łańcuch reguły `/ka/`, wówczas pierwszy pokazany poniżej identyfikator URI jest zgodny z tą regułą, a drugi identyfikator URI jest niezgodny:

```
file:///korporacja/ka/formularz.doc
http://firma.com/kadry/formularz.htm
```

Ponieważ wszystkie identyfikatory URI są traktowane jak wzorce, system ignoruje wszystkie gwiazdki wprowadzone jako znaki zastępcze na początku lub na końcu wzorca. Na przykład wzorce `*/ka/*` i `/ka/` odpowiadają temu samemu zestawowi identyfikatorów URI.

W regułach wzorców URI nie jest rozróżniana wielkość liter. W przypadku identyfikatorów URI zawierających spacje wzorzec URI musi być zgodny z regułami wyszukiwania korporacyjnego dotyczącymi kodowania identyfikatorów URI. W poniższym przykładzie przedstawiono poprawny i niepoprawny sposób określania identyfikatorów URI dla ścieżki systemu plików Windows:

Niepoprawny identyfikator URI: `file:///c:/program files/`  
Poprawny identyfikator URI: `file:///c:/program+files/`

### **Treść dokumentu**

Reguły treści dokumentu należy podawać w takim samym formacie jak zapytania. Jeśli dokument jest poprawny dla określonego zapytania, wówczas jest on zgodny z daną regułą. Podczas konfigurowania reguły należy określić słowa i frazy, które muszą występować lub nie mogą występować w dokumencie, oraz wybrać język reguł tematyzacji słów (określenia słów na podstawie rdzenia).

Na przykład w następującej regule określono, że jeśli w dokumencie występuje słowo `ka` lub `kadry`, wówczas dokument jest zgodny z tą regułą:

`ka "kadry"`

Natomiast w następującej regule określono, że jeśli w dokumencie występuje słowo `ka` lecz nie występuje słowo `zyski`, wówczas dokument jest zgodny z tą regułą:

`+ka -zyski`

Reguły dotyczące treści podlegają tym samym normom językowym, co zapytania tworzone za pomocą interfejsu SI-API (Search and Index API). Jednak w składni reguł dotyczących treści obsługiwany jest tylko podzbiór operacji dostępnych w składni zapytań tworzonych za pomocą interfejsu SI-API. Dozwolone są tylko następujące operatory zapytania:

**+** Wstawienie znaku plus przed terminem oznacza, że dokument musi zawierać dany termin.

- Wstawienie znaku minus przed terminem oznacza, że dokument nie może zawierać danego terminu.
- " Ujęcie dwóch lub większej liczby terminów w cudzysłów oznacza, że dokument musi zawierać dokładnie taką frazę.

Reguły dotyczące treści dokumentu odnoszą się wyłącznie do tych części dokumentu, które stanowią treść. Dla następujących typów przeszukiwaczy reguły te odnoszą się do plików (takich jak treść plików systemu UNIX lub Microsoft Windows) lub do treści stron WWW:

- Exchange Server
- NNTP
- system plików UNIX
- sieć WWW
- WebSphere Portal
- system plików Windows

W przypadku następujących typów przeszukiwaczy pola będące polami treści dokumentu określa się podczas konfigurowania opcji poszczególnych źródeł danych. Podczas tworzenia przeszukiwacza ta opcja jest dostępna na ostatniej stronie kreatora przeszukiwacza. Jeśli przeszukiwacz już istnieje, należy dokonać edycji przestrzeni przeszukiwania, wybrać źródło danych, którego opcje mają zostać zmienione, a następnie określić, które pola w tym źródle danych są polami treści dokumentu.

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

#### **Zadania pokrewne**

“Migrowanie kolekcji z portalu WebSphere Portal” na stronie 205

Aby przeprowadzić migrację kolekcji z portalu WebSphere Portal do wyszukiwania korporacyjnego, należy przygotować kolekcje w portalu WebSphere Portal, a następnie użyć Kreatora migracji do wykonania migracji.

## **Kategorie oparte na modelach**

Jeśli używane są kategorie oparte na modelach zdefiniowanych w systemie IBM WebSphere Portal, kategorie te mogą być używane również w wyszukiwaniu korporacyjnym.

W systemie WebSphere Portal dostępna jest predefiniowana systematyka, w której wykorzystywanych jest ponad 2300 tematów. Tematy te są zgrupowane według głównych kategorii biznesowych, takich jak Komputery, Finanse lub Transport. Użytkownicy portalu mogą tworzyć aplikacje, w których automatycznie określone są dokumenty zgodne z tymi obszarami tematycznymi, oraz dostosowywać istniejące kategorie do swoich własnych potrzeb.

Aby podczas wyszukiwania korporacyjnego używać kategorii zdefiniowanych w systemie WebSphere Portal, należy wykonać następujące czynności:

- Użyj kreatora migracji, aby zaimportować pliki systematyki opartej na modelach do wyszukiwania korporacyjnego.
- Podczas tworzenia kolekcji lub konfigurowania reguł analizowania kolekcji określ, że mają być używane kategorie oparte na modelach.

- Upewnij się, że na serwerze indeksowania wyszukiwania korporacyjnego jest zainstalowany system WebSphere Portal.
- Do administrowania kategoriami użyj narzędzi klasyfikacji dostępnych w systemie WebSphere Portal. Konsola administracyjna wyszukiwania korporacyjnego nie umożliwia administrowania kategoriami opartymi na modelach.

#### **Zadania pokrewne**

“Migrowanie systematyki opartej na modelach z portalu WebSphere Portal” na stronie 203

Za pomocą portletu zarządzania systematykami WebSphere Portal można wybrać systematykę opartą na modelach, która ma być używana względem kolekcji wyszukiwania korporacyjnego. Migracja nowej systematyki nie wpływa na kolekcje, które migrowano już wcześniej do wyszukiwania korporacyjnego.

## **Drzewa kategorii**

Drzewo kategorii umożliwia wyświetlenie wszystkich - używanych w kolekcji - kategorii opartych na regułach. Drzewo kategorii służy do tworzenia i usuwania kategorii oraz do edytowania reguł przypisywania dokumentów do kategorii.

Drzewo kategorii, zwane także systematyką, jest uporządkowane hierarchicznie. Początkowym punktem tej hierarchii jest kategoria podstawowa; wszystkie pozostałe kategorie są podkategoriami kategorii podstawowej. Można utworzyć dowolną liczbę zagnieżdżonych kategorii i podkategorii w celu udostępnienia użytkownikom odpowiednich opcji wyboru podczas przeglądania i pobierania dokumentów.

Na przykład jeśli dokument jest zgodny z regułami związanymi z kilkoma kategoriami, jest przypisywany do wszystkich tych kategorii. Gdy podczas przeglądania wyników wyszukiwania użytkownicy przeszukują kategorię lub przeglądają dokumenty należące do określonej kategorii, fakt przynależności określonego dokumentu do wielu kategorii zapewnia większe prawdopodobieństwo, że dokument ten zostanie znaleziony.

Administrator drzewa kategorii może określić, czy dokument należy do jednej czy do wielu kategorii, zagnieżdżając nowe kategorie w istniejących kategoriach. Podczas tworzenia kategorii należy określić, czy ma ona zostać utworzona jako kategoria główna czy jako podkategoria innej kategorii. Drzewo kategorii służy także do usuwania kategorii z kolekcji oraz do modyfikowania reguł przypisywania dokumentów do kategorii. Podczas edycji kategorii można zmienić jej nazwę, dodać lub usunąć reguły klasyfikowania albo zmodyfikować treść poszczególnych reguł.

Podczas administrowania drzewem kategorii warto korzystać z następujących opisów działania funkcji wyszukiwania i przeglądania:

- Jeśli przeszukiwana jest kategoria wysokiego poziomu, dokument spełniający kryteria wyszukiwania jest poszukiwany w całej tej kategorii i we wszystkich jej podkategoriach. Jeśli kategoria nie zawiera żadnych podkategorii, wówczas przeszukiwana jest tylko ta określona kategoria.
- Jeśli podczas przeglądania wyników wyszukiwania użytkownik wybierze opcję przeglądania dokumentów należących do konkretnej kategorii, wówczas wyświetlone zostaną tylko dokumenty należące do tej kategorii. W wynikach wyszukiwania wyświetlane są także nazwy wszystkich podkategorii, aby użytkownik mógł jednocześnie nawigować wśród kategorii i wyświetlać podzbiory dokumentów.

#### **Zadania pokrewne**

“Migrowanie kolekcji z portalu WebSphere Portal” na stronie 205

Aby przeprowadzić migrację kolekcji z portalu WebSphere Portal do wyszukiwania korporacyjnego, należy przygotować kolekcje w portalu WebSphere Portal, a następnie użyć Kreatora migracji do wykonania migracji.

## Wybór typu klasyfikacji

Po wybraniu typu klasyfikacji należy określić metodę, przy użyciu której dokumenty będą kojarzone z kategoriami w kolekcji.

### Zagadnienia wstępne

Aby zmienić typ klasyfikacji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora zmienianej kolekcji.

### Informacje o tym zadaniu

Typ klasyfikacji jest określany w momencie tworzenia kolekcji. W razie potrzeby istnieje możliwość zmiany sposobu klasyfikowania dokumentów w kolekcji. Można zastosować kategorie oparte na regułach skonfigurowane dla kolekcji, kategorie oparte na modelach istniejące w systemie IBM WebSphere Portal bądź w ogóle nie korzystać z kategorii.

**Ważne:** Jeśli typ klasyfikacji zostanie zmieniony po przeszukaniu danych i utworzeniu indeksu dla kolekcji, indeks będzie niespójny. Aby zapewnić dokładność wyników wyszukiwania, należy ponownie przeszukać wszystkie dokumenty w kolekcji i po ich przeanalizowaniu zreorganizować indeks.

### Procedura

Aby wybrać typ klasyfikacji, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Wybór typu klasyfikowania**.
2. Na stronie Wybór typu klasyfikowania wybierz jedną z następujących opcji:

**Brak** Wybierz tę opcję, jeśli dokumenty w tej kolekcji nie mają być klasyfikowane.

#### Oparty na regułach

Należy wybrać tę opcję, jeśli ma być używana systematyka klasyfikacji zawierająca reguły skonfigurowane specjalnie dla tej kolekcji.

- W przypadku konfigurowania kolekcji utworzonej na potrzeby wyszukiwania korporacyjnego tę opcję można wybrać, aby określić nazwy kategorii i reguły klasyfikacji dokumentów.
- W przypadku konfigurowania kolekcji zmigrowanej z portalu WebSphere Portal wybranie tej opcji umożliwia zastosowanie lub zmianę zaimportowanych kategorii opartych na regułach.

#### Oparty na modelach

Wybierz tę opcję, jeśli dokumenty mają być powiązane z kategoriami opartymi na modelach istniejącymi w systemie WebSphere Portal. Aby można było korzystać z tej opcji, na serwerze indeksowania wyszukiwania korporacyjnego musi być zainstalowane oprogramowanie WebSphere Portal. Do administrowania kategoriami należy używać narzędzi klasyfikacji dostępnych w portalu WebSphere Portal.

3. Kliknij przycisk **OK**.

## Konfiguracja kategorii

Dla kolekcji można utworzyć dowolną liczbę kategorii, a każda kategoria może zawierać dowolną liczbę reguł. Reguły określają, które dokumenty będą automatycznie powiązane z daną kategorią.

### Zagadnienia wstępne

Aby skonfigurować kategorie, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą kategorie.

Jako typ klasyfikowania należy wybrać kategorie oparte na regułach.

Na przykład w celu określenia reguł dotyczących powiązania dokumentów z kategoriami należy kliknąć opcję **Pomoc** w trakcie tworzenia lub edycji kategorii.

### Informacje o tym zadaniu

Jeśli używane aplikacje wyszukujące zapewniają obsługę kategorii, użytkownicy mogą wykonywać wyszukiwanie w podzbiorze kolekcji, podając nazwę kategorii. Użytkownicy mogą także wybrać kategorię w wynikach wyszukiwania i przeglądać wyłącznie dokumenty należące do wybranej kategorii.

**Ważne:** Jeśli po przeszukaniu dokumentów i utworzeniu indeksu dla kolekcji zmienione zostaną kategorie lub reguły kategorii, indeks stanie się niespójny. Aby zapewnić dokładność wyników wyszukiwania, należy ponownie przeszukać wszystkie dokumenty w kolekcji, przeanalizować je i zreorganizować indeks.

### Procedura

Aby skonfigurować kategorię, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguracja drzewa kategorii**.
2. Na stronie Drzewo kategorii wybierz w drzewie lokalizację, w której chcesz dodać kategorię, i kliknij opcję **Tworzenie kategorii**.

Jeżeli wybrany został poziom główny, nowa kategoria zostanie utworzona na tym poziomie. Jeśli wybrana zostanie nazwa kategorii, nowa kategoria zostanie utworzona poniżej wybranej kategorii w drzewie kategorii.

Otwarty zostanie kreator umożliwiający określanie reguł dotyczących powiązania dokumentów z nową kategorią:

- a. Na stronie Tworzenie kategorii wprowadź opisową nazwę kategorii i kliknij przycisk **Dalej**.
- b. Na stronie Tworzenie reguł kategorii kliknij przycisk **Dodaj regułę**.
- c. Na stronie Tworzenie reguły kategorii wprowadź unikalną nazwę reguły w polu **Nazwa reguły**. Nazwa ta musi być niepowtarzalna w obrębie wszystkich kategorii w danej kolekcji.
- d. Określ regułę, która ma być stosowana do powiązania dokumentów z daną kategorią, a następnie kliknij przycisk **OK**.

- Jeśli określenie, czy dokument należy do kategorii, ma odbywać się na podstawie jego identyfikatora URI, kliknij opcję **Wzorzec URI**, a następnie określ wzorzec URI.

Jeśli identyfikator URI będzie zawierać podany tekst, dokument zostanie powiązany z kategorią.

Na przykład: file:///c:/program+files/finance

- Jeśli podczas określania przynależności dokumentu do kategorii wyszukiwanie korporacyjne ma analizować wyrazy w polach zawartości dokumentu, kliknij opcję **Treść dokumentu**, wybierz język dokumentu i podaj wyrazy, które muszą występować lub nie mogą występować w polach treści dokumentu. Format definicji reguły jest taki sam, jak format zapytania (dopuszczalne są tylko operatory zapytania +, - i " ").



Dokument zawierający podane słowa lub dokument, który ich nie zawiera, jest kojarzony z kategorią.

Na przykład: +finanse -księgowość +”rok podatkowy”

e. Kliknij przycisk **Zakończ**.

Nowa kategoria jest wyświetlana na stronie Drzewo kategorii razem z innymi kategoriami należącymi do tej kolekcji.

#### **Pojęcia pokrewne**

“Migracja z portalu WebSphere Portal do wyszukiwania korporacyjnego” na stronie 203  
W wyszukiwaniu korporacyjnym dostępny jest kreator migracji, za pomocą którego można dokonać migracji systematyk i kolekcji z portalu IBM WebSphere Portal do wyszukiwania korporacyjnego.

#### **Zadania pokrewne**

“Migrowanie systematyki opartej na modelach z portalu WebSphere Portal” na stronie 203

Za pomocą portletu zarządzania systematykami WebSphere Portal można wybrać systematykę opartą na modelach, która ma być używana względem kolekcji wyszukiwania korporacyjnego. Migracja nowej systematyki nie wpływa na kolekcje, które migrowano już wcześniej do wyszukiwania korporacyjnego.

#### **Odsyłacze pokrewne**

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79  
Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

---

## **Praca z polami wyszukiwania XML**

W celu umożliwienia użytkownikom wyszukiwania określonych części dokumentów XML należy odwzorować elementy XML na pola wyszukiwania.

Do odwzorowania elementów XML na pola wyszukiwania należy użyć konsoli administracyjnej wyszukiwania korporacyjnego. Linux

### **pola wyszukiwania XML**

Pola wyszukiwania XML umożliwiają wykonywanie zapytań dotyczących określonych części dokumentów XML.

Dokumenty XML stają się coraz bardziej powszechne, ponieważ zawierają zarówno tekst z częściową strukturą, jak i tekst bez określonej struktury. Struktura XML jest hermetyczna i korzysta z kontekstu zdefiniowanego jawnie przez występujące wokół tekstu elementy XML. Na przykład imię i nazwisko autora może występować w następujący sposób:

```
<author>John Smith</author>
```

W tym kontekście tekst John Smith identyfikuje autora dokumentu XML.

Wyszukiwanie korporacyjne może powiązać lub odwzorować tekst wewnątrz elementów XML na nazwy pól wyszukiwania. Przy konfigurowaniu opcji analizowania składni dla kolekcji określa się, które elementy XML mają być odwzorowane na które nazwy pól wyszukiwania. Dzięki odwzorowaniu elementów XML na pola wyszukiwania użytkownicy będą mogli wyszukiwać te elementy, podając w zapytaniach odwzorowane nazwy pól. Zapytania, w których użyto konkretnych pól, umożliwiają uzyskanie bardziej precyzyjnych wyników wyszukiwania niż zapytania pełnotekstowe, które wyszukują cały dokument.

Na przykład, jeśli kolekcja zawiera dokumenty XML, a elementy tytuł i autor zostały oznaczone w indeksie jako pola wyszukiwania, użytkownicy mogą używać w zapytaniach tych konkretnych elementów. W wyniku wyszukiwania frazy `author:Smith` zostaną znalezione dokumenty XML zawierające tekst `Smith` w elemencie `author`.

## Odwzorowanie elementów XML na pola wyszukiwania

Przy odwzorowaniu elementu XML na pole wyszukiwania użytkownik określa, którzy użytkownicy elementów XML mogą wyszukiwać, określając nazwę pola w zapytaniu.

### Zagadnienia wstępne

Aby odwzorować elementy XML na pola wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienie administratora kolekcji, do której należą dane dokumenty XML.

### Informacje o tym zadaniu

Po utworzeniu odwzorowania pól XML lub dodaniu, zmianie bądź skasowaniu pól z istniejącego odwzorowania pól XML zmiany stają się widoczne po zatrzymaniu i zrestartowaniu analizatora składni. Nowe i zmienione odwzorowania mają zastosowanie do danych analizowanych po zrestartowaniu analizatora składni. Nowe i zmienione odwzorowania nie mają wpływu na dane, które są już analizowane i indeksowane.

To zadanie korzysta z poniższego przykładowego dokumentu XML w celu przedstawienia sposobu odwzorowania rekordów dotyczących pracowników i umożliwienia użytkownikom bezpośredniego zapytania o niektóre elementy.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<personnel>
 <personnelrecord>
 <phone>5555</phone>
 <email>joe@us.ibm.com</email>
 <jobroles>Manager, architect
 <jobrole>Managing Search Development Group</jobrole>
 <jobrole>Architecting Search Technology</jobrole>
 </jobroles>
 <location>New York</location>
 <section id="expertise">
 <text>Linguistics</text>
 </section>
 </personnelrecord>
</personnel>
```

### Procedura

Aby odwzorować użyte w przykładzie elementy XML na pola wyszukiwania, należy wykonać następujące czynności:

1. Przejdź do edycji kolekcji, wybierz stronę **Analiza składni**, a następnie kliknij opcję **Odwzorowanie elementów XML na pola**.
2. Na stronie **Odwzorowania pól XML** kliknij przycisk **Utwórz odwzorowanie XML**. Zostanie otwarta strona **Tworzenie odwzorowania pola XML**.
3. W polu **Nazwa elementu głównego XML** wpisz nazwę elementu głównego: `personnel`.  
Upewnij się, że określona w tym miejscu nazwa dokładnie odpowiada elementowi głównemu w dokumentach XML, które mają być wyszukiwane. Podczas analizowania i indeksowania dokumentów XML wyszukiwanie korporacyjne wybiera odwzorowanie do użycia zgodnie z nazwą elementu głównego.

4. W polu **Nazwa odwzorowania XML** wpisz nazwę tego zestawu reguł odwzorowywania pól XML.  
Po utworzeniu zestawu reguł odwzorowania XML nazwa ta będzie wyświetlana na stronie Odwzorowania pól XML i będzie można ją wybrać w celu dodania, usunięcia lub zmiany reguł odwzorowania.
5. Odwzoruj element XML jobrole na pole wyszukiwania o nazwie jobrole:
  - a. W polu **Nazwa pola** wpisz jobrole.
  - b. W polu **Nazwa elementu XML** wpisz jobrole.
  - c. Aby umożliwić użytkownikom używanie w zapytaniu pola jobrole i wyświetlanie ról zadań w wynikach wyszukiwania, zaznacz pola wyboru **Wyszukiwanie dla pól** i **Wyniki wyszukiwania**.
6. Odwzoruj element XML jobroles na to samo pole wyszukiwania:
  - a. Kliknij opcję **Dodaj pole**, aby dodać pusty wiersz do listy reguł odwzorowywania pól.
  - b. W polu **Nazwa pola** wpisz jobroles.
  - c. W polu **Nazwa elementu XML** wpisz jobrole.  
  
**Wskazówka:** Nazwy elementów XML nie muszą odpowiadać nazwom pól wyszukiwania, a wiele elementów XML może być odwzorowanych na to samo pole wyszukiwania.
  - d. Aby umożliwić użytkownikom używanie w zapytaniu pola jobrole i wyświetlanie ról zadań w wynikach wyszukiwania, zaznacz pola wyboru **Wyszukiwanie dla pól** i **Wyniki wyszukiwania**.
7. Odwzoruj element XML section z atrybutem expertise na pole wyszukiwania o nazwie expertise:
  - a. Kliknij opcję **Dodaj pole**, aby dodać pusty wiersz do listy reguł odwzorowywania pól.
  - b. W polu **Nazwa pola** wpisz expertise.
  - c. W polu **Nazwa pola** wpisz section.
  - d. W polu **Nazwa atrybutu XML** wpisz id.
  - e. W polu **Wartość atrybutu XML** wpisz expertise.
  - f. Aby umożliwić użytkownikom używanie w zapytaniu pola expertise i wyświetlanie wartości atrybutu expertise w wynikach wyszukiwania, zaznacz pola wyboru **Wyszukiwanie dla pól** i **Wyniki wyszukiwania**.
8. Kliknij przycisk **OK**.

### Przykłady:

Aby znaleźć w organizacji wszystkie osoby pracujące na produktach wyszukiwania, podaj następujące zapytanie:

```
jobrole:search
```

Aby znaleźć w organizacji wszystkie osoby znajdujące się na lingwistyce, podaj następujące zapytanie:

```
expertise:linguistics
```

---

## Praca z polami wyszukiwania HTML

Aby umożliwić użytkownikom wyszukiwania określonych sekcji metadanych dokumentów HTML, należy odwzorować elementy metadanych HTML na pola wyszukiwania w indeksie.

Do odwzorowania elementów metadanych HTML na pola wyszukiwania należy użyć konsoli administracyjnej wyszukiwania korporacyjnego.

### polo wyszukiwania HTML

Pola wyszukiwania HTML umożliwiają wykonywanie zapytań dotyczących atrybutów dokumentów HTML.

Elementy metadanych w dokumentach HTML przypominają atrybuty dokumentu, ponieważ również udostępniają informacje o dokumencie, o jego formacie oraz o metodzie dostępu do niego w sieci WWW. Na przykład:

```
<meta http-equiv="Content-Type" content="text/html; charset=utf-8" />]
<meta name="copyright" content="(C) Copyright IBM Corporation 2005" />
<meta name="content.owner" content="(C) Copyright IBM Corporation 2005" />
<meta name="security" content="public" />
<meta name="abstract" content="Opis produktu firmy IBM." />
<meta name="format" content="XHTML" />
```

Wyszukiwanie korporacyjne może powiązać (odwzorować) nazwy elementów metadanych HTML z nazwami pól wyszukiwania. Podczas konfigurowaniu opcji analizowania składni dla kolekcji określa się, jakie elementy metadanych HTML mają być odwzorowane na jakie nazwy pól wyszukiwania. Dzięki odwzorowaniu elementów metadanych HTML na pola wyszukiwania użytkownicy mogą wyszukiwać dokumenty zawierające te elementy, podając w zapytaniach nazwy pól wyszukiwania. Zapytania, w których użyto konkretnych pól, umożliwiają uzyskanie bardziej precyzyjnych wyników wyszukiwania niż zapytania pełnotekstowe, które wyszukują cały dokument.

Na przykład, jeśli kolekcja zawiera dokumenty HTML, a elementy metadanych "copyright" i "abstract" mają być poindeksowane jako pola wyszukiwania, użytkownicy mogą używać tych elementów w zapytaniach. W wyniku zapytania copyright:IBM zostaną znalezione dokumenty HTML, które zawierają tekst IBM w elemencie metadanych copyright.

### Odwzorowywanie elementów metadanych HTML na pola wyszukiwania

Podczas odwzorowywania elementów metadanych HTML na pola wyszukiwania określa się elementy metadanych HTML, które użytkownik może wyszukiwać, podając w zapytaniu nazwę pola.

#### Zagadnienia wstępne

Aby odwzorować elementy metadanych HTML na pola wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą dokumenty HTML.

#### Informacje o tym zadaniu

Po utworzeniu odwzorowania pól HTML, dodaniu, zmianie lub usunięciu istniejącego odwzorowania pól HTML, zmiany zostaną uwzględnione po zatrzymaniu i ponownym uruchomieniu analizatora składni. Zmiany są stosowane tylko do danych analizowanych po

ponownym uruchomieniu analizatora. Aby zaktualizować dokumenty, które zostały już przeanalizowane i poindeksowane, należy powtórnie przeszukać i przeanalizować dokumenty, po czym zreorganizować indeks.

### Procedura

Aby odwzorować elementy metadanych HTML na pola wyszukiwania:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Odwzorowanie metadanych HTML na pola**.
2. Na stronie Odwzorowania pól HTML kliknij opcję **Dodaj pole**, aby dodać pusty wiersz do listy reguł odwzorowania pól.
3. W polu **Nazwa pola** wpisz nazwę, którą chcesz skojarzyć z odwzorowywanym elementem metadanych HTML. Użytkownicy będą mogli podawać tę nazwę pola podczas wyszukiwania dokumentów HTML w danej kolekcji.
4. W polu **Nazwa elementu metadanych HTML** wpisz nazwę elementu metadanych HTML, który ma zostać odwzorowany.
5. Aby umożliwić użytkownikom wprowadzanie zapytań odnoszących do tego pola i wyświetlanie go w wynikach wyszukiwania, zaznacz pola wyboru **Wyszukiwanie dla pól** i **Wyniki wyszukiwania**.
6. Jeśli pole jest typu DECIMAL, DOUBLE, INTEGER, SHORT, TIME lub TIMESTAMP i użytkownicy mają mieć możliwość podawania zapytań parametrycznych podczas wyszukiwania tego pola, zaznacz pole wyboru **Wyszukiwanie parametryczne**.
7. Kliknij przycisk **OK**.

### Przykłady:

Użytkownicy mogą teraz używać w zapytaniach odwzorowanych nazw pól, aby znaleźć dokumenty HTML z określonymi metadanymi. Jeśli na przykład element metadanych HTML o nazwie `description` został odwzorowany na pole wyszukiwania o nazwie `abstract`, użytkownik może uruchomić zapytanie podobne do poniższego, aby znaleźć dokumenty o komputerach Thinkpad:

```
abstract:thinkpad
```

---

## Niestandardowe przetwarzanie tekstu

Dzięki zintegrowaniu niestandardowych algorytmów przetwarzania tekstu z kolekcjami wyszukiwania korporacyjnego można poprawić jakość i precyzję wyników wyszukiwania.

Program WebSphere Information Integrator OmniFind Edition obsługuje architekturę IBM Unstructured Information Management Architecture (UIMA), środowisko do tworzenia, wykrywania, składania i wdrażania funkcji analizy tekstu. Twórcy aplikacji mogą tworzyć i testować algorytmy analizy dla przeszukiwanej treści, a następnie tworzyć archiwum mechanizmu przetwarzania (plik `.pear`), które zawiera wszystkie zasoby niezbędne do wykorzystania tego archiwum podczas wyszukiwania korporacyjnego. Aby mieć możliwość przeszukiwania kolekcji za pomocą niestandardowych algorytmów analizy, użytkownik musi dodać do systemu wyszukiwania korporacyjnego archiwum zawierające mechanizm analizy tekstu.

Komponent logiki analizy w mechanizmie analizy tekstu jest nazywany *adnotatorem*. Każdy adnotator wykonuje konkretne zadania analizy lingwistycznej. Mechanizm przetwarzania tekstu może zawierać wiele adnotatorów; może się również składać z kilku mechanizmów analizy tekstu, z których każdy zawiera własne niestandardowe adnotatory.

Informacje tworzone przez adnotatory są nazywane *wynikami analizy*. Wyniki analizy, które odpowiadają wyszukiwanym informacjom, są zapisywane w strukturze danych nazywanej *wspólną strukturą analizy*.

Podczas konfigurowania opcji przetwarzania tekstu dla kolekcji użytkownik wykonuje następujące zadania:

- Wybiera mechanizm analizy tekstu, który ma być używany do tworzenia adnotacji do dokumentów w kolekcji.
- Jeśli kolekcja zawiera dokumenty XML, w których elementy języka znaczników są normalnymi słowami, i ten język znaczników ma zostać użyty do niestandardowej analizy tekstu, może powiązać pliki odwzorowania XML z kolekcją oraz odwzorować dane wyjściowe odwzorowania XML na wspólną strukturę analizy.

Na przykład można odwzorować zawartość elementów <adresat> i <klient> na adnotacje Osoba we wspólnej strukturze analizy. Do tych adnotacji mogą później uzyskać dostęp niestandardowe adnotatory, wykrywające dodatkowe informacje (na przykład płeć osoby). Można również odwzorować adnotacje osoby w indeksie wyszukiwania korporacyjnego, co umożliwi użytkownikom wyszukiwanie osób bez znajomości oryginalnych elementów XML.

Jeśli użytkownicy mają mieć możliwość określania w zapytaniach oryginalnych elementów XML, nie trzeba definiować żadnych odwzorowań XML. Zamiast tego można skonfigurować opcje analizy składni i umożliwić rodzime wyszukiwanie XML w kolekcji.

- Odwzorowuje struktury danych znajdujące się we wspólnej strukturze analizy na indeks wyszukiwania korporacyjnego, który umożliwia wyszukiwanie dokumentów z adnotacjami za pomocą wyszukiwania semantycznego.

Na przykład w zależności od encji i związków wykrytych przez adnotatory użytkownicy mogą wyszukiwać pojęcia, które występują w tym samym zdaniu (takie jak nazwisko konkretnej osoby i nazwa dowolnego konkurenta), lub słowo kluczowe i pojęcie (takie jak termin Jan i numer telefonu).

- Odwzorowuje struktury danych znajdujące się we wspólnej strukturze analizy na tabelę bazy danych obsługującą interfejs JDBC. Można odwzorować dane na tabelę bazy danych IBM DB2 Universal Database (DB2 UDB) lub Oracle. Ten typ odwzorowania umożliwia użycie wyników analizy na przykład w aplikacjach bazodanowych służących do eksploracji danych. Pozwala też na wykorzystanie zapytań SQL do wyszukiwania danych znajdujących się poza systemem wyszukiwania korporacyjnego.

#### **Pojęcia pokrewne**

"Custom text analysis integration" (Integracja niestandardowej analizy tekstu) w podręczniku "Text Analysis Integration"

"Unstructured information management architecture (UIMA)" (Architektura zarządzania informacjami bez określonej struktury - UIMA) w podręczniku "Text Analysis Integration"

"Workflow for custom analysis integration" (Przepływ pracy dla integracji analizy niestandardowej) w podręczniku "Text Analysis Integration"

"Text analysis algorithms" (Algorytmy analizy tekstu) w podręczniku "Text Analysis Integration"

"Semantic search applications" (Aplikacje wyszukiwania semantycznego) w podręczniku "Text Analysis Integration"

"Semantic search query" (Zapytanie wyszukiwania semantycznego) w podręczniku "Text Analysis Integration"

## **Dodawanie mechanizmów analizy tekstu do systemu**

Jeśli użytkownik utworzy niestandardowy mechanizm analizy tekstu, należy go dodać do systemu, aby można było z niego korzystać do wyszukiwania korporacyjnego. Używając

mechanizmu analizy w kolekcjach, można analizować dokumenty i opatrywać je adnotacjami oraz zwiększać dokładność wyników wyszukiwania.

### **Zagadnienia wstępne**

Aby dodać mechanizmy analizy tekstu do systemu, wymagana jest rola administratora wyszukiwania korporacyjnego.

### **Informacje o tym zadaniu**


Twórcy aplikacji mogą utworzyć archiwum mechanizmu przetwarzania (.pear), które jest zgodne ze strukturą UIMA do analizy tekstu. Archiwum zawiera wszystkie zasoby niezbędne do wyszukiwania w kolekcjach wyszukiwania korporacyjnego. Aby mieć możliwość przeszukiwania kolekcji za pomocą niestandardowych algorytmów analizy, użytkownik musi dodać do systemu wyszukiwania korporacyjnego archiwum zawierające mechanizm analizy tekstu.

Po dodaniu do systemu mechanizmu analizy tekstu można zmienić jego nazwę wyświetlaną i wybrać opcję przeglądania źródła XML. (Na podstawie źródła XML można ustalić, jakie informacje są wytwarzane przez dany mechanizm).

Jeśli mechanizm analizy tekstu jest powiązany z kolekcją, nie można go usunąć z systemu.

### **Procedura**

Aby dodać niestandardowy mechanizm analizy tekstu do systemu wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Analiza składni kliknij opcję **Konfiguruj mechanizmy analizy tekstu**.
4. Na stronie Mechanizmy analizy tekstu kliknij opcję **Dodaj mechanizm analizy tekstu**.
5. Na stronie Dodawanie mechanizmu analizy tekstu wpisz opisową nazwę nowego mechanizmu. W systemie nazwa wyświetlana służy do identyfikowania tego mechanizmu analizy tekstu w konsoli administracyjnej.
6. Określ położenie pliku .pear. Jeśli plik ten jest mniejszy niż 8 MB, może znajdować się w systemie użytkownika. Jeśli plik jest większy niż 8 MB, musi znajdować się na serwerze indeksowania.
7. Kliknij przycisk **OK**. Mechanizm analizy tekstu będzie wyświetlany na stronie Mechanizmy analizy tekstu.

### **Pojęcia pokrewne**

"Workflow for custom analysis integration" (Przeływ pracy dla integracji analizy niestandardowej) w podręczniku "Text Analysis Integration"

"Custom text analysis integration" (Integracja niestandardowej analizy tekstu) w podręczniku "Text Analysis Integration"

"Unstructured information management architecture (UIMA)" (Architektura zarządzania informacjami bez określonej struktury - UIMA) w podręczniku "Text Analysis Integration"

"Approaches for using XML markup in analysis and search" (Metody stosowania języka znaczników XML w analizie i wyszukiwaniu) w podręczniku "Text Analysis Integration"

### **Zadania pokrewne**

| "Creating an XML to UIMA types mapping configuration file" (Tworzenie pliku  
| konfiguracyjnego odwzorowania typów XML na UIMA) w podręczniku "Text Analysis  
| Integration"

## Powiązanie mechanizmu analizy tekstu z kolekcją

Jeśli z systemem wyszukiwania korporacyjnego są powiązane niestandardowe mechanizmy analizy tekstu, można wybrać jeden z nich do użycia z określoną kolekcją. Użytkownicy, wyszukujący dane w tej kolekcji, będą mogli tworzyć zapytania semantyczne, a przy tym podnieść jakość i dokładność wyników wyszukiwania.

### Zagadnienia wstępne

Aby powiązać mechanizm analizy tekstu z kolekcją, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji.

### Informacje o tym zadaniu

Jeśli mechanizm analizy tekstu już jest powiązany z tą kolekcją, podczas próby powiązania innego mechanizmu zostaną wykonane następujące działania:

- Jeśli zostanie wybrana opcja **Brak analizy niestandardowej**, wszystkie odwzorowania analizy tekstu zdefiniowane wcześniej dla tej kolekcji zostaną zresetowane. Odtąd dla kolekcji będą używane systemowe wartości domyślne.
- Jeśli zostanie wybrana nazwa innego niestandardowego mechanizmu analizy tekstu, wszystkie odwzorowania analizy tekstu zdefiniowane wcześniej dla tej kolekcji zostaną zachowane. Na przykład, jeśli mechanizm\_1 zostanie zmieniony na mechanizm\_2, to mechanizm\_2 odziedziczy pliki odwzorowań XML, które skonfigurowano dla mechanizmu\_1.

### Procedura

Aby powiązać mechanizm analizy tekstu z kolekcją, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. Kliknij opcję **Wybierz mechanizm analizy tekstu**. Jeśli żaden niestandardowy mechanizm analizy tekstu nie został dodany do systemu wyszukiwania korporacyjnego lub jeśli dla kolekcji używane są domyślne algorytmy analizy, mechanizm analizy ma nazwę **Domyślnie**.
3. Na stronie Wybór mechanizmu analizy tekstu dla tej kolekcji wybierz nazwę mechanizmu, który ma być używany z tą kolekcją. Jeśli nie są dostępne żadne mechanizmy analizy tekstu lub jeśli wybrano opcję **Brak analizy niestandardowej**, analizator składni zastosuje domyślne reguły analizy tekstu podczas dodawania do dokumentów adnotacji i przygotowywania dokumentów pod kątem indeksu.
4. Kliknij przycisk **OK**.

### Pojęcia pokrewne

"Workflow for custom analysis integration" (Przeływ pracy dla integracji analizy niestandardowej) w podręczniku "Text Analysis Integration"

"Custom text analysis integration" (Integracja niestandardowej analizy tekstu) w podręczniku "Text Analysis Integration"

"Unstructured information management architecture (UIMA)" (Architektura zarządzania informacjami bez określonej struktury - UIMA) w podręczniku "Text Analysis Integration"



## Odwzorowywanie elementów XML na wspólną strukturę analizy

Jeśli kolekcja zawiera dokumenty XML, w których elementy języka znaczników są normalnymi słowami, i ten język znaczników ma zostać użyty do umożliwienia użytkownikom wyszukiwania w indeksie wyszukiwania korporacyjnego lub w tabelach JDBC przy użyciu wyszukiwania semantycznego, można odwzorować elementy XML na wspólną strukturę analizy.

### Zagadnienia wstępne

Aby odwzorować elementy XML na wspólną strukturę analizy, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji.

Aby umożliwić procesom niestandardowej analizy tekstu dostęp do konkretnych elementów w dokumentach XML lub odwzorować kilka elementów XML na wspólny typ używany podczas wyszukiwania semantycznego, można utworzyć niestandardowe pliki odwzorowania XML. Pliki odwzorowania XML muszą być zgodne ze strukturą UIMA do analizy tekstu.

Po dodaniu plików odwzorowania XML do kolekcji, która używa niestandardowego mechanizmu analizy tekstu, zostanie włączona możliwość odwzorowania elementów XML w dokumentach źródłowych na adnotacje we wspólnej strukturze analizy. Adnotacje mogą następnie zostać wykorzystane przez niestandardowy mechanizm analizy tekstu. Następnie można odwzorować wyniki analizy (we wspólnej strukturze analizy) na indeks i umożliwić użytkownikom wysyłanie zapytań do adnotacji podczas wyszukiwania semantycznego w kolekcjach.

Na przykład można odwzorować zawartość elementów adresat i klient na adnotacje Osoba we wspólnej strukturze analizy. Do tych adnotacji mogą później uzyskać dostęp niestandardowe adnotatory, wykrywające dodatkowe informacje (na przykład płeć osoby). Można również odwzorować adnotacje osoby w indeksie wyszukiwania korporacyjnego, co umożliwi użytkownikom wyszukiwanie osób bez znajomości oryginalnych elementów XML.

Jeśli użytkownicy mają mieć możliwość określania w zapytaniach oryginalnych elementów XML, nie trzeba definiować żadnych odwzorowań XML. Zamiast tego można skonfigurować opcje analizy składni i umożliwić rodzime wyszukiwanie XML w kolekcji.

### Procedura

Aby odwzorować elementy XML na wspólną strukturę analizy, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. W obszarze **Odwzoruj elementy XML na wspólne struktury analizy** kliknij opcję **Dodaj odwzorowanie**.
3. Na stronie Odwzorowanie elementów XML na wspólną strukturę analizy wpisz opisową nazwę wyświetlaną pliku odwzorowania XML.
4. Określ położenie pliku. Jeśli plik odwzorowania XML jest mniejszy niż 8 MB, można wpisać ścieżkę lub wyszukać plik. Jeśli plik odwzorowania XML jest większy niż 8 MB, należy się upewnić, że plik znajduje się na serwerze indeksowania, i podać pełną ścieżkę do pliku.
5. Kliknij przycisk **OK**. Nowy plik odwzorowania XML zostanie dodany do listy plików na stronie Opcje przetwarzania tekstu.

### Pojęcia pokrewne

"Workflow for custom analysis integration" (Przeływ pracy dla integracji analizy niestandardowej) w podręczniku "Text Analysis Integration"

"Custom text analysis integration" (Integracja niestandardowej analizy tekstu) w podręczniku "Text Analysis Integration"

"Unstructured information management architecture (UIMA)" (Architektura zarządzania informacjami bez określonej struktury - UIMA) w podręczniku "Text Analysis Integration"

"Approaches for using XML markup in analysis and search" (Metody stosowania języka znaczników XML w analizie i wyszukiwaniu) w podręczniku "Text Analysis Integration"

#### **Zadania pokrewne**

"Creating an XML to UIMA types mapping configuration file" (Tworzenie pliku konfiguracyjnego odwzorowania typów XML na UIMA) w podręczniku "Text Analysis Integration"

## **Odwzorowywanie wspólnej struktury analizy na indeks**

Można określić wspólną strukturę analizy używaną, gdy użytkownicy wysyłają zapytania do kolekcji za pomocą wyszukiwania semantycznego.

### **Zagadnienia wstępne**

Aby odwzorować wspólną strukturę analizy na indeks, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji.

### **Informacje o tym zadaniu**

Odwzorowanie wspólnej struktury analizy na indeks wyszukiwania korporacyjnego umożliwia użytkownikom określanie zapytań dokładnych pod względem semantycznym i poprawienie jakości wyników wyszukiwania.

Na przykład w zależności od encji i związków wykrytych przez adnotatory użytkownicy mogą wyszukiwać pojęcia, które występują w tym samym zdaniu (takie jak nazwisko konkretnej osoby i nazwa dowolnego konkurenta), lub słowo kluczowe i pojęcie (takie jak termin Jan i numer telefonu).

### **Procedura**

Aby odwzorować wspólną strukturę analizy na indeks, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. W obszarze **Odwzoruj jedną wspólną strukturę analizy na indeks** kliknij opcję **Wybierz wspólną strukturę analizy**.
3. Na stronie Wybór wspólnej struktury analizy dla tej kolekcji wybierz odwzorowanie, które ma być używane dla indeksu wyszukiwania korporacyjnego:
  - Aby użyć domyślnych reguł systemowych dotyczących odwzorowywania struktur danych we wspólnej strukturze analizy na indeks wyszukiwania korporacyjnego, wybierz opcję **Domyślnie**.
  - Aby odwzorować konkretną wspólną strukturę analizy na indeks, określ położenie pliku wspólnej struktury analizy. Jeśli plik ten jest mniejszy niż 8 MB, może znajdować się w systemie użytkownika. Jeśli plik jest większy niż 8 MB, musi znajdować się na serwerze indeksowania.
4. Kliknij przycisk **OK**. Określona wspólna struktura analizy jest wyświetlana na stronie Opcje przetwarzania tekstu.

### **Pojęcia pokrewne**

"Workflow for custom analysis integration" (Przeptyw pracy dla integracji analizy niestandardowej) w podręczniku "Text Analysis Integration"

"Custom text analysis integration" (Integracja niestandardowej analizy tekstu) w podręczniku "Text Analysis Integration"

"Unstructured information management architecture (UIMA)" (Architektura zarządzania informacjami bez określonej struktury - UIMA) w podręczniku "Text Analysis Integration"

"Approaches for indexing custom analysis results" (Metody indeksowania wyników analizy niestandardowej) w podręczniku "Text Analysis Integration"

#### **Zadania pokrewne**

"Creating the index build configuration file" (Tworzenie pliku konfiguracyjnego budowania indeksu) w podręczniku "Text Analysis Integration"

## **Odwzorowywanie wspólnej struktury analizy na tabele JDBC**

Można określić wspólną strukturę analizy, która ma być odwzorowywana na tabele JDBC do użycia w aplikacjach baz danych.

#### **Zagadnienia wstępne**

Aby odwzorować wspólną strukturę analizy na tabele bazy danych JDBC, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji.

#### **Informacje o tym zadaniu**

Odwzorowanie wspólnej struktury analizy na tabele JDBC umożliwia użycie danych przez aplikacje baz danych. Na przykład użytkownicy mogą określać zapytania SQL poza wyszukiwaniem korporacyjnym w celu wyszukania adnotacji dodanych przez wspólną strukturę analizy. Tych informacji można również użyć do dalszego przetwarzania tekstu, np. korzystając z nich w aplikacjach służących do eksploracji danych.

#### **Procedura**

Aby odwzorować wspólną strukturę analizy na tabele JDBC, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Konfiguruj opcje przetwarzania tekstu**.
2. W obszarze **Odwzoruj wspólne struktury analizy na tabele JDBC** kliknij opcję **Dodaj odwzorowanie**.
3. Na stronie Odwzorowywanie wspólnej struktury analizy na tabele JDBC wpisz opisową nazwę wyświetlaną dla wspólnej struktury analizy, która ma zostać użyta do odwzorowania informacji na tabele JDBC.
4. Określ położenie pliku wspólnej struktury analizy. Jeśli plik ten jest mniejszy niż 8 MB, może znajdować się w systemie użytkownika. Jeśli plik jest większy niż 8 MB, musi znajdować się na serwerze indeksowania.
5. Kliknij przycisk **OK**. Nowy plik zostanie dodany do listy plików na stronie Opcje przetwarzania tekstu. Zostanie również wyświetlona data i godzina dodania pliku do systemu.

#### **Pojęcia pokrewne**

"Workflow for custom analysis integration" (Przeptyw pracy dla integracji analizy niestandardowej) w podręczniku "Text Analysis Integration"

"Custom text analysis integration" (Integracja niestandardowej analizy tekstu) w podręczniku "Text Analysis Integration"

"Unstructured information management architecture (UIMA)" (Architektura zarządzania informacjami bez określonej struktury - UIMA) w podręczniku "Text Analysis Integration"

"Approaches for mapping analysis results to a database" (Metody odwzorowywania wyników analizy na bazę danych) w podręczniku "Text Analysis Integration"

#### **Zadania pokrewne**

"Creating the XML mapping configuration file" (Tworzenie pliku konfiguracyjnego odwzorowania XML) w podręczniku "Text Analysis Integration"

---

## **Konfigurowanie wątków dla usługi analizatora składni**

Jeśli istnieją wystarczające zasoby pamięci, można zwiększyć liczbę wątków dostępnych dla analizatora składni podczas analizowania dokumentów.

### **Zagadnienia wstępne**

W przypadku dużej liczby kolekcji można zwiększyć liczbę wątków analizatora składni. Należy się upewnić, że system ma wystarczającą ilość pamięci do obsługi dodatkowych wątków. Analizator składni z jednym wątkiem wymaga 200 MB pamięci. Każdy dodatkowy wątek wymaga kolejnych 50 MB pamięci.

Aby skonfigurować liczbę wątków uruchamianych na potrzeby analizatora składni, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

### **Procedura**

Aby skonfigurować liczbę wątków analizatora składni:

1. Przejdź do edycji kolekcji, wybierz stronę Analiza składni, a następnie kliknij opcję **Skonfiguruj opcje analizowania**.
2. Określ maksymalną liczbę wątków analizatora składni, które mają być uruchamiane podczas uruchamiania analizatora składni, i kliknij przycisk **OK**.
3. Zrestartuj analizator składni.

---

## **Włączanie zaawansowanej analizy terminów złożonych**

Jakość wyszukiwania można poprawić, umożliwiając analizatorowi składni użycie zaawansowanych metod analizy dla terminów złożonych. Analiza zaawansowana umożliwia rozbiór terminów złożonych, dzięki czemu każdy z jego członów może być traktowany jak pojedynczy termin.

### **Zagadnienia wstępne**

Aby określić opcje analizowania terminów złożonych, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

### **Informacje o tym zadaniu**

Niektóre języki łączą terminy zawierające kilka słów w pojedyncze słowo bez spacji (terminy *złożone*). Funkcja zaawansowanej analizy i rozbioru terminów złożonych jest przydatna podczas wyszukiwania w językach takich jak niemiecki i ma niewrażliwe znaczenie w przypadku wyszukiwania w językach takich jak koreański.

Jeśli zaawansowana analiza terminów złożonych zostanie włączona, użytkownicy mogą wyszukiwać terminy bez konieczności używania znaków zastępczych w celu odnalezienia złożonych form terminów użytych w zapytaniu. Na przykład wyszukanie terminu **Organ** (organ) może zwrócić dokumenty zawierające termin **Organspender** (dawca organu), ale nie zwróci dokumentów zawierających termin **Organisation** (organizacja). W odróżnieniu od zapytania ze znakiem zastępczym **Organ\***, które może zwrócić dowolny łańcuch zawierający ciąg **Organ**, wyszukiwanie dopasowuje tylko pełne słowa podrzędne w ramach większego terminu złożonego.

Definiowane przez użytkownika terminy słownikowe, takie jak synonimy i słowa-kryteria, są również stosowane do członów terminów złożonych, które są używane jako oddzielne słowa w zapytaniu.

### Procedura

Aby włączyć zaawansowaną analizę terminów złożonych, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij opcję **Skonfiguruj opcje analizowania**.
2. Zaznacz pole wyboru **Włącz zaawansowaną analizę terminów złożonych** i kliknij przycisk **OK**.

#### Pojęcia pokrewne

"Linguistic support for semantic search" (Obsługa lingwistyczna w wyszukiwaniu semantycznym) w podręczniku "Text Analysis Integration"

"Text analysis included in enterprise search" (Analiza tekstu w wyszukiwaniu korporacyjnym) w podręczniku "Text Analysis Integration"

---

## Włączanie obsługi rodzimego wyszukiwania XML

Jeśli kolekcja zawiera dokumenty XML, można umożliwić użytkownikom korzystanie z języka znaczników XML podczas wyszukiwania dokumentów, włączając rodzime wyszukiwanie XML dla kolekcji.

### Zagadnienia wstępne

Aby włączyć obsługę wyszukiwania dokumentów XML za pomocą rodzimego wyszukiwania XML, użytkownik musi mieć uprawnienia administracyjne wyszukiwania korporacyjnego lub być administratorem danej kolekcji.

### Informacje o tym zadaniu

Rodzime wyszukiwanie XML, takie jak XPath lub fragmenty XML, może zapewnić dokładniejsze wyniki wyszukiwania dzięki wykorzystaniu języka znaczników XML dokumentów w zapytaniu. Na przykład użytkownicy mogą określić, że termin zapytania musi występować w pewnym elemencie lub atrybucie XML.

Na przykład faktury od dostawcy komputerów, które są w formacie XML, mogą zawierać wpisy <zamówienie>, które zawierają elementy <firma> i <typkomputera>. Podczas pobierania faktur, które zawierają zamówienia na notebooki firmy IBM, wyszukiwanie według słowa kluczowego może spowodować pobranie dokumentów, które zawierają notebooki firmy Dell i modele desktop firmy IBM. Dzięki użyciu wyszukiwania XML można określić, że w elemencie <firma> musi wystąpić IBM, w elemencie <typkomputera> musi wystąpić notebook oraz że obydwa elementy muszą wystąpić w tym samym elemencie <zamówienie>. W ten sposób można pobrać faktury, które dotyczą notebooków firmy IBM.

## Procedura

Aby umożliwić użytkownikom przeszukiwanie kolekcji za pomocą rodzimego wyszukiwania XML, należy wykonać poniższe czynności:

1. Przejdź do edycji kolekcji, wybierz stronę Analiza składni, a następnie kliknij opcję **Skonfiguruj opcje analizowania**.
2. Wybierz opcję **Zezwól użytkownikom na wyszukiwanie w dokumentach XML za pomocą rodzimego wyszukiwania XML**.
3. Kliknij przycisk **OK**.

### Pojęcia pokrewne

"Linguistic support for semantic search" (Obsługa lingwistyczna w wyszukiwaniu semantycznym) w podręczniku "Text Analysis Integration"

"Text analysis included in enterprise search" (Analiza tekstu w wyszukiwaniu korporacyjnym) w podręczniku "Text Analysis Integration"

"Semantic search applications" (Aplikacje wyszukiwania semantycznego) w podręczniku "Text Analysis Integration"

"Semantic search query" (Zapytanie wyszukiwania semantycznego) w podręczniku "Text Analysis Integration"

---

## Analiza lingwistyczna dokumentów chińskich, japońskich i koreańskich

Aby zwiększyć wydajność pobierania dokumentów w języku chińskim, japońskim i koreańskim, można określić opcje analizy lingwistycznej.

Dla dokumentów w języku chińskim, japońskim i koreańskim można skonfigurować analizator składni, aby używał do analizy leksykalnej podziału na ciągi n-znakowe. Dla dokumentów w języku chińskim i japońskim można ponadto skonfigurować analizator składni w taki sposób, aby ze znaków niedrukowanych usuwał znaki nowego wiersza.

## Podział na ciągi n-znakowe

Podczas tworzenia kolekcji wybierany jest typ analizy leksykalnej, która ma być używana do analizowania składni dokumentów napisanych w językach, w których nie stosuje się spacji do oddzielania słów.

Podział według znaków spacji w kodzie Unicode wykorzystuje znaki spacji jako separatory wyrazów. Według reguł podziału na ciągi n-znakowe nakładające się sekwencje dowolnej liczby znaków są uznawane za pojedynczy wyraz. W przypadku takich języków, jak japoński, chiński i koreański, w których nie używa się odstępów jako znaków oddzielających wyrazy, podział na ciągi n-znakowe może dać lepsze wyniki niż podział według znaków spacji w kodzie Unicode.

Metodę podziału, która ma zostać użyta do analizowania składni dokumentów, wybiera się podczas tworzenia kolekcji. Po utworzeniu kolekcji można przeglądać te ustawienia wśród opcji analizowania składni, ale nie można ich zmienić.

## Usuwanie znaków nowego wiersza z zakresów znaków innych niż ASCII

Dla języków, w których biały znak nie jest używany do oddzielania granic słów, takich jak chiński i japoński, można skonfigurować analizator składni tak, aby usuwał niektóre białe znaki powodujące wstawienie znaków nowego wiersza.

### Zagadnienia wstępne

Aby wykonać to zadanie, należy się zalogować jako administrator wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Jeśli ta opcja jest włączona dla kolekcji, a dokument został napisany po chińsku lub japońsku, analizator składni usunie ciągi znaków tabulacji i nowego wiersza oddzielające dwa oddzielne znaki liter. Znak litery musi pochodzić z zakresu znaków Unicode innych niż ASCII. Zostaną usunięte następujące znaki:

- Tabulacja (0x09)
- LF lub znak nowego wiersza (0x0A)
- Znak CR lub znak powrotu karetki (0x0D)

Aby zmiany zostały uwzględnione, należy zatrzymać i restartować analizator składni. Aby zastosować zmianę do dokumentów, które zostały już przeanalizowane i zapisane w indeksie, dokumenty należy ponownie przeszukać i wprowadzić do indeksu, a następnie należy przeprowadzić reorganizację indeksu.

### Procedura

Aby usunąć znaki nowego wiersza z białych znaków, należy wykonać następujące czynności:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
2. Za pomocą edytora tekstu zmień następujący plik, gdzie *identyfikator\_kolekcji* jest identyfikatorem, który został określony (lub przypisany przez system) podczas tworzenia kolekcji:  
`identyfikator_kolekcji.parserdriver/collection.properties`
3. Zmień wartość właściwości `removeCjNewlineChars` z `false` na `true`.

---

## Typy dokumentów powiązane z analizatorami składni kolekcji i sesjami Stellent

Aby zagwarantować dokładne i efektywne analizowanie dokumentów w przestrzeni przeszukiwania, można utworzyć pliki konfiguracyjne określające, które typy dokumentów mają być analizowane przez analizator składni kolekcji, a które w ramach sesji filtrowania dokumentów Stellent.

W kolekcji wyszukiwania korporacyjnego dokumenty w większości formatów przetwarzane są przy użyciu wbudowanych analizatorów składni HTML lub XML. Niektóre typy dokumentów z reguły nie podlegają analizie (np. dokumenty w języku Postscript), a inne są obsługiwane przez funkcje analizujące Stellent (np. dokumenty programów Microsoft Word, Microsoft Excel, Microsoft PowerPoint, Lotus Freelance, Lotus 123, PDF, RT i Ichitaro).

Ponieważ metadane być mylące, dokumenty w formacie zwykłego tekstu i HTML mogą zostać błędnie skierowane do sesji Stellent, po czym odesłane z powrotem do jednego z wbudowanych analizatorów składni. Sytuacja taka niekorzystnie odbije się na wydajności systemu. Aby tego uniknąć, można utworzyć pliki konfiguracyjne definiujące miejsce i sposób analizowania różnych typów dokumentów.

Kojarzenie typów dokumentów z analizatorem składni kolekcji oraz z sesją Stellent wymaga wykonania następujących czynności:

1. Skonfigurowanie typów dokumentów dla analizatora składni kolekcji. W tym kroku należy utworzyć plik konfiguracyjny, który przypisuje typy dokumentów do analizatora składni używanego przez kolekcję. Można utworzyć po jednym pliku konfiguracyjnym na kolekcję.
2. Skonfigurowanie typów dokumentów dla sesji Stellent. W tym kroku należy utworzyć plik konfiguracyjny, który przypisuje typy dokumentów do filtrów dokumentów Stellent używanych przez kolekcję. Można utworzyć po jednym pliku konfiguracyjnym na kolekcję.
3. Zatrzymanie i ponowne uruchomienie analizatora składni. Aby zmiany odniosły skutek, należy użyć konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania kolekcji, dla której skonfigurowano typy dokumentów, a następnie zatrzymać i ponownie uruchomić analizator składni.

## Powiązanie typów dokumentów z analizatorem składni kolekcji

Aby powiązać konkretne typy dokumentów z analizatorem składni kolekcji, należy utworzyć plik konfiguracyjny. Konsola administracyjna nie umożliwia wykonania tego zadania.

### Zagadnienia wstępne

Aby wykonać to zadanie, należy się zalogować jako administrator wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Jeśli plik konfiguracyjny nie istnieje, analizator składni kolekcji używa domyślnych reguł usługi analizatora składni. Jeśli plik konfiguracyjny istnieje, reguły w pliku określają:

- Które dokumenty mają zostać wysłane do sesji Stellent, w zależności od rozszerzenia pliku lub typu treści.
- W jaki sposób analizować składnię dokumentów, których typ jest nieznan ze względu na niekompletne metadane.

Formatem pliku jest sekwencja wierszy, w której każdy wiersz odpowiada jednej z następujących reguł:

#### **EXTENSION** *rozszerzenie analizator składni*

Wszystkie dokumenty, których adres URL kończy się podanym rozszerzeniem, zostaną przetworzone przez podany analizator składni. W rozszerzeniu nie należy wstawiać kropki, a podczas porównywania nie jest rozróżniana wielkość liter.

#### **CONTENTTYPE** *typ/podtyp analizator składni*

Wszystkie dokumenty, których typ treści odpowiada podanemu typowi/podtypowi, zostaną przetworzone przez podany analizator składni. Jeśli typ treści dokumentu zostanie podany w postaci t/p, zgodność występuje wtedy, gdy typ równa się t oraz podtyp równa się p albo podtypem jest znak zastępczy (gwiazdka, \*).

#### **UNKNOWN** *analizator składni*

Wszystkie dokumenty, których rozszerzenie i typ treści są nieznanne (czyli nie zostały udostępnione przez przeszukiwacz), zostaną przetworzone przez podany analizator składni.

#### **DEFAULT** *analizator składni*

Wszystkie dokumenty, które nie zostały uwzględnione w innych regułach, zostaną przetworzone przez podany analizator składni.

We wszystkich przypadkach *analizator składni* musi być równy html, xml, stellent lub none, gdzie wartość none oznacza, że dokumenty tego typu nie zostaną przeanalizowane.



Jeśli dokument odpowiada więcej niż jednej regule, pierwszeństwo ma reguła, która jest bardziej szczegółowa, bez względu na kolejność występowania reguł:

- Reguła EXTENSION jest bardziej szczegółowa niż reguła CONTENTTYPE.
- Reguła CONTENTTYPE, która obejmuje podtyp, jest bardziej szczegółowa niż ta sama reguła ze znakiem zastępczym. Reguła dla typu treści `application/postscript` ma pierwszeństwo przed regułą `application/*`.
- Nie powinno być dwóch reguł dotyczących tego samego rozszerzenia lub typu treści. W takim przypadku kolejność reguł zależy od implementacji.

### Procedura

Aby powiązać typy dokumentów z analizatorem składni kolekcji, należy wykonać poniższe czynności:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
2. W następujący sposób utwórz plik konfiguracyjny, gdzie *identyfikator\_kolekcji* identyfikuje kolekcję, która ma zostać skonfigurowana:  
`ES_NODE_ROOT/master_config/identyfikator_kolekcji.parserdriver/parserTypes.cfg`
3. Za pomocą edytora tekstu zmodyfikuj zawartość pliku i określ reguły usługi analizatora składni, a następnie zapisz i zamknij plik.
4. Aby zmiany zostały uwzględnione, użyj konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania analizatora składni kolekcji, a następnie zatrzymaj i ponownie uruchom analizator składni.

### Przykład

W tym przykładzie wbudowany analizator składni HTML przetwarza wszystkie dokumenty z rozszerzeniem `txt`, `htm` lub `html` oraz z typem treści, który zaczyna się od `text/`, albo z nieznanym rozszerzeniem i typem treści. Wbudowany analizator składni XML przetwarza wszystkie dokumenty z rozszerzeniem `xml` lub z typem treści `text/xml`. Wszystkie pozostałe dokumenty, włącznie z mającymi typ treści rozpoczynający się od `application/`, są wysyłane do sesji 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
```

## Domyślne reguły usługi analizatora składni kolekcji

Jeśli nie zostanie utworzony plik konfiguracyjny, który odwzorowuje typy plików i typy treści na analizator składni dla kolekcji, zostaną użyte domyślne reguły analizy składni dokumentów.

Domyślne reguły używane przez analizatora składni kolekcji:

```
EXTENSION pdf stellent
EXTENSION ppt stellent
EXTENSION prz stellent
EXTENSION lwp stellent
EXTENSION doc stellent
```

```
EXTENSION rtf stellent
EXTENSION xls stellent
EXTENSION 123 stellent
EXTENSION vsd stellent
EXTENSION vdx stellent
EXTENSION jxw stellent
EXTENSION jsw stellent
EXTENSION jtw stellent
EXTENSION jaw stellent
EXTENSION juw stellent
EXTENSION jbw stellent
EXTENSION jvw stellent
EXTENSION jfw stellent
EXTENSION jtt stellent
EXTENSION jtd stellent
EXTENSION jttc stellent
EXTENSION jtdc stellent
EXTENSION jtdx stellent
EXTENSION ps none
EXTENSION xml xml
EXTENSION txt text
EXTENSION htm html
EXTENSION html html
EXTENSION shtml html
EXTENSION xhtml html
EXTENSION asp html
```

```
CONTENTTYPE application/postscript none
CONTENTTYPE application/* stellent
CONTENTTYPE text/rtf stellent
CONTENTTYPE text/richtext stellent
CONTENTTYPE text/xml xml
CONTENTTYPE text/html html
CONTENTTYPE text/plain text
```

```
UNKNOWN html
DEFAULT html
```

## Powiązanie typów dokumentów z sesją Stellent

Aby określić, jakie typy dokumentów mają zostać przeanalizowane przez filtry dokumentów Stellent, należy utworzyć plik konfiguracyjny. Konsola administracyjna nie umożliwia wykonania tego zadania.

### Zagadnienia wstępne

Aby wykonać to zadanie, należy się zalogować jako administrator wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Plik konfiguracyjny określa:

- Jakie typy plików mają zostać przeanalizowane przez sesję Stellent. Typ pliku odpowiada jednemu z typów plików rozpoznawanych przez bibliotekę Stellent.
- Jakie typy plików mają zostać zwrócone do analizatora składni w celu przetworzenia przez jeden z wbudowanych analizatorów składni. (To działanie jest niezbędne, ponieważ analizator składni kolekcji mógł błędnie wysłać dokument do sesji Stellent ze względu na mylące metadane).
- Jakie typy plików mają zostać odrzucone, ponieważ nie są obsługiwane przez wyszukiwanie korporacyjne.

Jeśli plik konfiguracyjny został określony, ale nie istnieje, nie można będzie uruchomić analizatora składni. Jeśli dla właściwości `OutsideInSupportedTypes` w pliku `stellent.properties` nie został określony żaden plik konfiguracyjny, zostaną użyte domyślne reguły analizy sesji Stellent.

W pliku konfiguracyjnym są wymienione typy dokumentów i sposoby ich obsługi. Formatem pliku jest sekwencja wierszy, w której każdy wiersz odpowiada jednej z następujących reguł:

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

Gdzie:

*typ\_dokumentu*

Jest wartością używaną dla tokenu zapytania `doctype`. Dokumenty można wyszukiwać według typu dokumentu. Na przykład użytkownik mógł podać `$doctype::pdf` w celu wyszukania dokumentów PDF.

*typ* Jest jedną z wartości `FI_` w bibliotece Stellent, a *typ\_dokumentu* jest wartością używaną przez token `doctype`, jeśli stosowana jest reguła.

#### **DEFAULT**

Oznacza, że lista akceptowanych lub rodzimych typów, w zależności od typu reguły, uwzględnia listę domyślną. Ta opcja umożliwia rozszerzenie domyślnej konfiguracji bez konieczności jej zastępowania.

**All** Oznacza, że dla podanego tokenu `doctype` będą akceptowane wszystkie typy, które nie zostały jawnie wymienione.

Reguły w pliku konfiguracyjnym są przetwarzane w następujący sposób:

- Jeśli istnieje reguła `accept` dla elementu *typ*, która obejmuje domyślną listę, o ile został podany element `accept DEFAULT`, zostanie ona zaakceptowana.
- W przeciwnym przypadku, jeśli istnieje reguła `reject` dla elementu *typ*, nie zostanie ona zaakceptowana.
- W przeciwnym przypadku, jeśli została określona reguła `accept ALL`, zostanie ona zaakceptowana.
- W przeciwnym przypadku nie zostanie ona zaakceptowana.

Jeśli typ dokumentu zostanie zaakceptowany, zostanie użyta wartość *typ\_dokumentu*, która została określona w zastosowanej regule. Ta wartość jest zwracana do analizatora składni kolekcji wraz z przeanalizowaną treścią. Jeśli typ dokumentu nie został zaakceptowany, wystąpi następujące zachowanie:

- Jeśli istnieje rodzima reguła dla elementu *type* (włącznie z domyślnymi regułami analizy w przypadku określenia reguły `native DEFAULT`), dokument zostanie zwrócony do wbudowanego analizatora składni wraz z wartością elementu `doctype`, która jest określana przez tę regułę. Wartość elementu `doctype` musi być równa `txt`, `htm` lub `xml`, co wskazuje odpowiednio na zwykły tekst, HTML lub XML.
- W przeciwnym przypadku dokument zostanie odrzucony i nie zostanie przeanalizowany.

#### **Procedura**

Aby powiązać typy dokumentów z sesją Stellent, należy wykonać następujące czynności:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
2. Zmień zawartość pliku  
ES\_NODE\_ROOT/master\_config/identyfikator\_kolekcji.stellent/stellent.properties, gdzie *identyfikator\_kolekcji* identyfikuje kolekcję, która ma zostać skonfigurowana.
3. We właściwości `OutsideInSupportedTypes` podaj bezwzględną ścieżkę do tworzonego pliku konfiguracyjnego.  
Na przykład można utworzyć następujący plik konfiguracyjny dla pojedynczej kolekcji i zapisać go wraz z innymi plikami danej kolekcji:  
ES\_NODE\_ROOT/master\_config/identyfikator\_kolekcji.stellent/stellenttypes.cfg  
Innym przykładem może być utworzenie następującego pliku konfiguracyjnego w celu wykorzystania tych samych ustawień dla wszystkich kolekcji oraz zapisanie go z innymi plikami na poziomie systemu. (W przypadku wykorzystania tego podejścia należy się upewnić, że ta ścieżka została podana w pliku `stellent.properties` dla każdej kolekcji, co zostało określone w kroku 2).  
ES\_INSTALL\_ROOT/default\_config/stellent/stellenttypes.cfg
4. Za pomocą edytora tekstu utwórz plik konfiguracyjny i określ reguły analizy Stellent, a następnie zapisz i zamknij plik.
5. Aby zmiany zostały uwzględnione, użyj konsoli administracyjnej wyszukiwania korporacyjnego do monitorowania analizatora składni kolekcji, a następnie zatrzymaj i ponownie uruchom analizator składni.

### Przykład

W następującym pliku konfiguracyjnym sesja Stellent akceptuje, oprócz domyślnej listy, format programu Microsoft Visio.

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

W następującym pliku konfiguracyjnym dokumenty Postscript będą obsługiwane z możliwością wyszukiwania dla typu dokumentu `ps`, format mapy pikselowej X (XPM) zostanie zwrócony do wbudowanego analizatora składni tekstu, format PNG zostanie odrzucony, a wszystkie inne typy plików zostaną zaakceptowane z możliwością wyszukiwania dla typu dokumentu `other`.

```
accept DEFAULT
accept FI_POSTSCRIPT ps
native FI_XPIXMAP txt
accept ALL other
reject FI_PNG
```

## Domyślne reguły analizy składni dla sesji Stellent

Jeśli nie zostanie utworzony plik konfiguracyjny, który przypisuje typy plików do filtrów dokumentów sesji Stellent, zostaną użyte domyślne reguły analizy składni dokumentów.

Domyślne reguły używane przez sesję Stellent:

```
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_JAPANESEJIS	xml
NATIVE FI_WML_JAPANESEJIS	xml
NATIVE FI_WML_JAPANESEJIS	xml
NATIVE FI_WML_JAPANESEJIS	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_OLDPCODE	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

---

## Administrowanie indeksem wyszukiwania korporacyjnego

Aby użytkownicy mieli zapewniony dostęp do najnowszych informacji, w wyszukiwaniu korporacyjnym dla każdej kolekcji tworzony jest indeks, a jego treść jest regularnie odświeżana i reorganizowana.

Aby udostępnić zgromadzone przez przeszukiwacze dane dla operacji przeszukiwania, należy utworzyć indeksy. Podczas tworzenia kolekcji w wyszukiwaniu korporacyjnym tworzony jest indeks wszystkich danych, które zostały wstępnie przeszukane. Gdy przeszukiwacze przeszukują nowe i zmodyfikowane źródła danych, w wyszukiwaniu korporacyjnym indeks jest odświeżany o nową treść. Następnie odświeżona treść jest scalana z indeksem podstawowym. Ten proces scalania jest określany jako reorganizacja. Za każdym razem, gdy wykonywane jest odświeżenie lub reorganizacja indeksu, nowa treść jest kopiowana na serwerze przeszukiwania i udostępniana dla operacji przeszukiwania.

Przeszukiwacze zbierają dane ciągle lub regularnie według harmonogramu. Częste odświeżanie indeksu zapewnia, że użytkownicy mogą przeszukiwać najbardziej aktualne dane. Ciągłe odświeżanie indeksu musi być na zakończenie zreorganizowany. Rozbudowywanie odświeżanego indeksu wiąże się z coraz większym zapotrzebowaniem na zasoby systemu. Dlatego w celu zapewnienia optymalnej wydajności należy regularnie reorganizować indeksy.

Częstotliwość reorganizowania indeksu zależy od następujących czynników:

- zasoby systemu (obszar systemu plików, szybkość procesora, pamięć)
- liczba dokumentów przeznaczonych do przeszukania i ponownego przeszukania
- typ danych do przeszukania
- częstotliwość modyfikowania reguł kategorii (zmiany są uwzględniane dopiero po zreorganizowaniu indeksu)
- częstotliwość wymuszonych uruchomień przeszukiwacza (w przeciwieństwie do uruchomień zaplanowanych w harmonogramie)
- częstotliwość usuwania lub dodawania identyfikatorów URI przez przeszukiwacze zewnętrzne (te typy przeszukiwaczy są powiązane z wyszukiwaniem korporacyjnym za pośrednictwem interfejsu API nasłuchiwanie danych)

W przypadku kolekcji, w których znajduje się kilka milionów dokumentów (przy czym większość z nich to dokumenty sieci WWW), indeks powinien być reorganizowany codziennie i odświeżany co jedną lub dwie godziny.

Aby zapewnić użytkownikom dostęp do indeksu, który jest aktualny i przygotowany dla operacji przeszukiwania, należy wykonać następujące zadania:

- określenie harmonogramów odświeżania i reorganizowania indeksu
- zmodyfikowanie harmonogramu operacji na indeksie
- włączenie i wyłączenie harmonogramu operacji na indeksie
- skonfigurowanie współbieżnych operacji tworzenia indeksów

Aby określić istotne dla użytkowników opcje mające wpływ na indeks, można dodatkowo wykonać następujące zadania:

- skonfigurowanie obsługi znaków zastępczych w zapytaniach
- skonfigurowanie zasięgów w celu ograniczenia zakresu dokumentów dostępnych do przeglądania przez użytkowników

- zwijanie w wynikach wyszukiwania dokumentów pochodzących z tego samego źródła
- usunięcie identyfikatorów URI z indeksu

#### Zadania pokrewne

“Monitorowanie aktywności indeksu dla kolekcji” na stronie 228

Monitorowanie operacji na indeksie związanym z kolekcją pozwala wyświetlić postęp operacji budowania indeksu, włączyć lub wyłączyć harmonogram operacji na indeksie bądź uruchomić lub zatrzymać operację indeksowania.

“Monitorowanie kolejki indeksu wyszukiwania korporacyjnego” na stronie 229

Użytkownik może wyświetlać status wszystkich operacji budowania indeksów w kolejce indeksów, zatrzymać operację budowania indeksu, a także usunąć indeks z kolejki.

---

## Planowanie operacji budowania indeksu

Istnieje możliwość określenia harmonogramów reorganizowania indeksu oraz odświeżania indeksu zawierającego nową treść.

### Zagadnienia wstępne

Do planowania operacji budowania indeksu wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla danej kolekcji.

### Informacje o tym zadaniu

Aby upewnić się, że użytkownicy zawsze mają dostęp do najnowszych informacji w przeszukiwanych źródłach, należy zaplanować reorganizowanie i odświeżanie indeksu w regularnych odstępach czasu. Reorganizacja polega na przebudowaniu całego indeksu. Procesy indeksowania odczytują wszystkie dane zebrane przez przeszukiwacze i przeanalizowane przez analizator składni. Podczas odświeżania indeksu do wyszukiwania udostępnione zostają informacje przeszukane od czasu ostatniej reorganizacji indeksu.

Domyślnie włączona jest opcja planowania operacji budowania indeksu. Opcja ta nakazuje procesowi planowania utworzenia harmonogramu odświeżania i reorganizacji indeksu przy każdym uruchomieniu systemu wyszukiwania korporacyjnego. Aby zapobiec uruchamianiu zaplanowanych operacji budowania indeksu, w dowolnym momencie można usunąć zaznaczenie opcji **Włączanie podczas uruchamiania systemu**. Na przykład konieczne może być wyłączenie harmonogramu w celu rozwiązania problemów.

### Procedura

Aby utworzyć harmonogram operacji budowania indeksu, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Tworzenie harmonogramu budowania indeksów**.
2. Aby określić, jak często indeks ma być odświeżany nową zawartością, należy na stronie Tworzenie harmonogramu budowania indeksów, w obszarze **Określ harmonogram odświeżania indeksu** podać następujące opcje:
  - a. W obszarze **Termin uruchomienia** w polach **Rok**, **Miesiąc**, **Dzień**, **Godzina** i **Minuta** określ, kiedy indeks ma zostać odświeżony po raz pierwszy.
  - b. W obszarze **Okres aktualizacji** w polach **dni**, **godziny** i **minuty** określ częstotliwość odświeżania indeksu.

Zwykle wskazane jest stosunkowo częste odświeżanie indeksu, na przykład co godzinę lub dwie. W zależności od tego, jak często zmienia się treść źródłowa, należy podać większy lub mniejszy odstęp czasu. Na przykład może to być proces wykonywany co godzinę (0 dni 1 godzina) albo co 12 godzin (0 dni i 12 godzin).



3. Aby określić, na ile często indeks ma być całkowicie przebudowywany, określ następujące opcje w obszarze **Określ harmonogram reorganizowania indeksu**:
  - a. W obszarze **Termin uruchomienia** w polach **Rok, Miesiąc, Dzień, Godzina i Minuta** określ, kiedy indeks ma zostać zreorganizowany po raz pierwszy.
  - b. W obszarze **Okres aktualizacji** w polach **dni, godziny i minuty** określ częstotliwość reorganizacji indeksu.

Zazwyczaj zalecane jest regularne reorganizowanie indeksu, na przykład co 24 godziny. W zależności od tego, jak często zmienia się treść źródłowa, należy podać większy lub mniejszy odstęp czasu. Na przykład może to być proces wykonywany co 12 godzin (0 dni 12 godzina) albo co dwie i pół doby (2 dni i 12 godzin).
4. Kliknij przycisk **OK**.

## Zmiana harmonogramu operacji na indeksie

Można zmodyfikować harmonogram reorganizacji i odświeżania indeksu.

### Zagadnienia wstępne

Aby użytkownik mógł zmodyfikować harmonogram operacji na indeksie, musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub administratora określonej kolekcji.

### Procedura

Aby zmienić harmonogram operacji na indeksie:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i zmień odpowiednie wartości w polach **Miesiąc, Dzień, Rok i Godzina**. Określ, jak często indeks ma być odświeżany z uwzględnieniem nowej treści i jak często ma być reorganizowany.
2. Kliknij przycisk **Zastosuj**.

## Włączanie i wyłączanie harmonogramów operacji na indeksach

Istnieje możliwość włączania i wyłączania harmonogramów odświeżania i reorganizowania indeksu.

### Zagadnienia wstępne

Aby użytkownik mógł włączyć lub wyłączyć harmonogram operacji na indeksie, musi mieć przypisaną rolę administratora wyszukiwania korporacyjnego lub administratora określonej kolekcji.




### Informacje o tym zadaniu

Istnieje możliwość wyłączenia harmonogramu w celu uniknięcia zaplanowanego procesu tworzenia indeksu. Na przykład może zaistnieć potrzeba wyłączenia harmonogramu w celu uniknięcia utworzenia indeksu w zaplanowanym czasie w celu rozwiązania zaistniałych problemów.

Użytkownik może włączyć lub wyłączyć harmonogram podczas edytowania lub monitorowania kolekcji.

### Procedura

1. Aby podczas edytowania kolekcji włączyć lub wyłączyć harmonogram operacji na indeksie, wykonaj następujące kroki:
  - a. Rozpocznij edycję kolekcji, w której ma nastąpić zmiana.

- b. Na stronie Indeks odpowiednio zaznacz lub usuń zaznaczenie pola wyboru **Włączanie podczas uruchamiania systemu**, aby włączyć albo wyłączyć harmonogram odświeżania indeksu.
  - c. Aby włączyć lub wyłączyć harmonogram reorganizowania indeksu, zaznacz lub usuń zaznaczenie pola wyboru **Włączanie podczas uruchamiania systemu**.
  - d. Kliknij przycisk **Zastosuj**.
2. Aby podczas monitorowania kolekcji włączyć lub wyłączyć harmonogram operacji na indeksie, wykonaj następujące kroki:
- a. Rozpocznij monitorowanie kolekcji, w której ma nastąpić zmiana.
  - b. Jeśli istnieje harmonogram operacji na indeksie i indeks ten ma zostać utworzony w określonym czasie, kliknij na stronie Indeks opcję  **Wyłącz harmonogram**. Indeks nie zostanie utworzony, dopóki użytkownik nie włączy harmonogramu lub nie kliknie ikony  **Uruchom** w celu uruchomienia procesu tworzenia indeksu.
  - c. Jeśli harmonogram operacji na indeksie istnieje, ale jest wyłączony, kliknij ikonę  **Włącz harmonogram**.  
Proces tworzenia indeksu zostanie umieszczony w kolejce z datą i godziną określoną w harmonogramie operacji na indeksie.

## Konfiguracja współbieżnych operacji budowania indeksu

Użytkownik steruje korzystaniem z zasobów indeksowania, określając liczbę kolekcji, które mogą jednocześnie przetwarzać własne żądania budowy indeksu. Jeśli w systemie są wystarczające zasoby, można poprawić jakość wyszukiwania, umożliwiając jednoczesne odświeżanie indeksu kolekcji i jego reorganizację.

### Zagadnienia wstępne

Aby określić opcje budowania indeksu dla systemu, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Informacje o tym zadaniu


Wyszukiwanie korporacyjne umożliwia współbieżne budowanie wielu indeksów dzięki współużytkowaniu zasobów przez kolekcje, co umożliwia równoległe przetwarzanie żądań budowy indeksu w wielu kolekcjach. Współużytkując procesy, można się upewnić, że reorganizacja bardzo dużego indeksu nie spowoduje zablokowania dostępności innych indeksów oczekujących w kolejce na zbudowanie.

Gdy operacja budowania indeksu zostanie zainicjowana na żądanie lub zgodnie z harmonogramem, indeks zostaje umieszczony w kolejce indeksów, gdzie oczekuje na przetworzenie. Ponieważ każda kolekcja ma własny indeks, w kolejce indeksów jednocześnie może być kilka żądań budowy indeksu z różnych kolekcji. Podczas konfigurowania w systemie opcji indeksowania można określić liczbę kolekcji, które mogą równoległe współużytkować zasoby indeksowania i przetwarzać żądania.

Można również określić, że żądania odświeżenia indeksu kolekcji mają być przetwarzane równoległe z reorganizacją indeksu kolekcji. W przypadku włączenia tej opcji serwery wyszukiwania będą odświeżane najnowszymi dokumentami (poprzez odświeżenie indeksu) w trakcie przetwarzania wolniejszej reorganizacji indeksu. Jednak budowanie indeksu jest procesem w znacznym stopniu obciążającym zasoby systemu. Podczas operacji budowania indeksu wykorzystywana jest duża ilość pamięci systemowej i znaczny obszar dysku. W przypadku włączenia tej opcji i posiadania niewystarczającej ilości miejsca na dysku lub pamięci ogólna wydajność systemu może ulec pogorszeniu.

### Procedura

Aby określić w systemie opcje budowania indeksu, należy wykonać następujące czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Indeksowanie kliknij opcję **Konfiguruj opcje indeksu**.
4. Na stronie Opcje indeksowania na poziomie systemu wpisz liczbę kolekcji, które mogą współużytkować zasoby systemu i dla których żądania budowy indeksu mogą być przetwarzane równolegle.  
Liczba kolekcji współużytkujących zasoby indeksowania nie może przekroczyć liczby kolekcji w systemie wyszukiwania korporacyjnego. Na przykład, jeśli w systemie istnieje pięć kolekcji, należy wprowadzić liczbę nie większą niż pięć.
5. Jeśli w systemie są wystarczające zasoby do obsługi wielu współbieżnych operacji budowy indeksu, można zaznaczyć opcję umożliwiającą współbieżne odświeżanie i reorganizację indeksu.
6. Kliknij przycisk **OK**.

---

## Opcje wpływające na podlegający wyszukiwaniu widok indeksu

Po poindeksowaniu dokumentów można określić opcje decydujące o sposobach wyszukiwania dokumentów i przeglądania ich w wynikach wyszukiwania przez użytkowników.

Aby określić opcje wpływające na sposób wyświetlania indeksu dla użytkownika, można wykonać następujące działania:

- Skonfigurować obsługę znaków zastępczych w terminach użytych w zapytaniu. W indeks można wbudować obsługę zapytań ze znakami zastępczymi albo można określić opcje rozwijające kryteria podczas przetwarzania zapytania.
- Skonfigurowanie zasięgów, aby ograniczyć zakres dokumentów dostępnych do przeszukiwania dla użytkownika. Przy wyszukiwaniu w kolekcji uwzględniane są tylko dokumenty znajdujące się w zdefiniowanym zasięgu, a nie wszystkie dokumenty z indeksu.
- Zwinięcie w wynikach wyszukiwania dokumentów pochodzące z jednego źródła. Dokumenty pasujące do identyfikatora URI lub jego wzorca w indeksie można grupować i w wynikach wyszukiwania wyświetlać tylko dokumenty położone najwyżej. Użytkownik w takiej sytuacji może sam zdecydować o wyświetleniu zwiniętych dokumentów wynikowych.
- Usunięcie identyfikatorów URI z indeksu. Można tymczasowo uniemożliwić użytkownikom wyszukiwanie określonych dokumentów w indeksie.

## Indeksowane opcje wyszukiwania dokumentów

Podczas konfigurowania opcji wyszukiwania przeszukanych danych lub źródeł zewnętrznych (albo w przypadku odwzorowania elementów metadanych XML i HTML na pola wyszukiwania) należy określić sposób wyszukiwania dokumentów i ich prezentacji w wynikach wyszukiwania.

Podczas edycji opcji dla przeszukiwaczy, które zawierają pola, można określić następujące opcje określające, czy i w jaki sposób można wyszukiwać dane pole oraz czy może ono zostać zwrócone w wynikach wyszukiwania:

- Wyszukiwanie pełnotekstowe
- Wyszukiwanie dla pól
- Wyszukiwanie parametryczne
- Wyniki wyszukiwania

- Treść dokumentu

Podczas edycji opcji dla źródeł zewnętrznych nie jest dostępna opcja oznaczenia pola jako pola treści dokumentu.

Jeśli podczas konfiguracji analizatora składni użytkownik wskaże, że elementy XML i elementy metadanych HTML mają być odwzorowywane na pola w indeksie, należy określić, czy tych pól można używać w wyszukiwaniu opartym na polach, w wyszukiwaniu parametrycznym lub w wynikach wyszukiwania.

Określone opcje są zapisywane w indeksie wraz z dokumentami. Umożliwiają one ograniczenie zakresu zapytań wprowadzanych przez użytkowników oraz ograniczenie danych zwracanych w wynikach wyszukiwania.

## Wyszukiwanie pełnotekstowe

Indeks wyszukiwania korporacyjnego jest indeksem pełnotekstowym zawierającym treści z różnych źródeł danych. Można przeszukiwać treść przy użyciu prostego zapytania sformułowanego w języku naturalnym. Procesy wyszukiwania przeglądają pola i treść dokumentów, wyszukując dokumenty odpowiadające zapytaniu.

### Przykład:

Wyszukiwanie pełnotekstowe można zainicjować nawet tak prostym zapytaniem:

łańcuch rowerowy

W celu zaznaczenia słów, które muszą albo nie mogą wystąpić w dokumencie, można użyć specjalnych notacji. Na przykład poprzedzenie słowa znakiem plusa (+) oznacza, że dokument musi zawierać dane słowo, aby spełnił kryteria. Słowo poprzedzone znakiem minusa (-) powoduje wykluczenie dokumentów zawierających to słowo spośród wyników wyszukiwania. Ujęcie dwóch lub więcej słów w cudzysłów ("") powoduje wyszukanie identycznej frazy.

### Przykład:

W przypadku następującego zapytania pełnotekstowego dokument będzie zgodny tylko pod warunkiem, że będzie on zawierać frazę *fantastyka naukowa*, przy tym nie będzie zawierać słowa *robot*:

+*"fantastyka naukowa"* -robot

## Wyszukiwanie dla pól

Wyszukiwanie dla pól umożliwia ograniczenie obiektu zapytania do określonych pól i metadanych dokumentu. Przykładem może być wyszukiwanie określonych słów tylko w tytułach dokumentów.

Aby użyć wyszukiwania dla pól w wyszukiwaniu korporacyjnym, należy podać nazwę pola oraz słowo lub frazę, które powinny się znajdować w polu w zapytaniu.

### Przykład:

Poniższe zapytanie powoduje wyszukanie dokumentów zawierających słowo *ibm* oraz frazę *wyszukiwanie korporacyjne* w polu tytułu:

title:ibm title:"wyszukiwanie korporacyjne"

Aby możliwe było wyszukiwanie pola po nazwie pola, pole to musi zostać włączone do wyszukiwania dla pól przy konfigurowaniu przeszukiwacza.

## Wyszukiwanie parametryczne

Wyszukiwanie parametryczne jest odmianą wyszukiwania dla pól umożliwiającą użycie zapytań porównujących lub obliczających, które operują na polach liczbowych i daty oraz na metadanych. Przykładem może być wyszukanie dokumentów mających określoną wielkość lub utworzonych po określonej dacie. Można również wyszukiwać dokumenty o atrybutach większych, mniejszych lub równych określonym wartościom.

### Przykład 1:

Poniższe zapytanie pozwala wyszukać elementy kosztujące dokładnie 50 złotych (lub innej waluty przypisanej wartościom w polu cena):

```
#cena::=50
```

### Przykład 2:

Poniższe zapytanie pozwala wyszukać dokumenty mające rozmiar pliku powyżej 1024 i nie większy niż 2048:

```
#rozmiar_pliku::>1024<=2048
```

Aby możliwe było wyszukiwanie pola za pomocą zapytania parametrycznego, pole to musi zostać włączone do wyszukiwania parametrycznego przy konfigurowaniu przeszukiwacza.

## Wyniki wyszukiwania

Można wyszukać pewne pola i nie pokazać ich w wynikach wyszukiwania. Można też pokazać w wynikach wyszukiwania pole, mimo że nie było ono uwzględnione w zapytaniu. Użytkownik może na przykład wprowadzić zapytanie dotyczące danych finansowych w celu utworzenia odpowiedniego raportu, ale może jednocześnie wyświetlić pensje pracowników w wynikach, które zawierają również nazwiska pracowników.

## Treść dokumentu

Treść pola oznaczonego jako pole treści dokumentu może służyć do powiązania dokumentów z kategorią opartą na modelu i z kategoriami określającymi reguły oparte na treści dokumentu.

### Pojęcia pokrewne

"Składnia zapytania" w podręczniku "Programming Guide and API Reference for Enterprise Search"

"Aplikacje wyszukujące dla wyszukiwania korporacyjnego" na stronie 155

Aplikacja wyszukująca umożliwia wyszukiwanie w kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Liczba możliwych do utworzenia aplikacji wyszukujących nie jest ograniczona, a pojedyncza aplikacja wyszukująca pozwala przeszukać dowolną liczbę kolekcji i źródeł zewnętrznych.

## Znaki zastępcze w zapytaniach

Można umożliwić użytkownikom stosowanie znaków zastępczych w terminach zapytania oraz wyszukiwanie słów zgodnych z określonym wzorcem.

Termin zapytania ze znakiem zastępczym jest to termin, który zawiera znak gwiazdki (\*). Jeśli użytkownik wprowadzi zapytanie, które zawiera znak zastępczy, w wynikach wyszukiwania zostaną uwzględnione wszystkie dokumenty w indeksie, które są zgodne z terminem zapytania oraz wszystkie dokumenty w indeksie, które są zgodne ze wzorcem reprezentowanym przez znak zastępczy. Na przykład znak zastępczy kończący w wyszukiwanym terminie SZU\* może odpowiadać terminom szukaj, szufla i szuler.

Podczas konfigurowania opcji znaków zastępczych dla indeksu należy wskazać, czy użytkownicy będą mieli możliwość podawania znaków zastępczych w zapytaniach, a jeśli tak, to trzeba też określić sposób ich obsługi:

- Możliwość wyszukiwania słów zgodnych ze wzorcem zawierającym znak zastępczy można włączyć dla wszystkich części dokumentu lub ograniczyć tylko pól.
- Obsługę zapytań zawierających znaki zastępcze można włączyć dla wszystkich pól lub ograniczyć wyłącznie do określonych pól.
- Można ograniczyć dozwoloną pozycję znaku zastępczego do ostatniego znaku terminu użytego w zapytaniu (znak zastępczy kończący) lub można zezwolić na występowanie znaku zastępczego w każdym miejscu tego terminu. (Nie można stosować znaków zastępczych w nazwach pól).
- W zależności od miejsc, w których dozwolone jest występowanie znaku zastępczego, można wybrać sposób rozwijania terminów użytych w zapytaniu (terminy zawierające znaki zastępcze rozwijane są do wszystkich zgodnych z nimi terminów indeksu). W indeksie mogą być zapisane wszystkie możliwe rozwinięcia terminów lub procesy wyszukiwania mogą rozwijać terminy podczas przetwarzania zapytania.

Wszystkie zmiany wprowadzone w ustawieniach znaków zastępczych zaczynają obowiązywać od następnej reorganizacji indeksu.

## Rozwinięcie indeksu

Aby uwzględnić rozwijanie terminów znajdujących się w indeksie, należy określić, ile początkowych znaków słowa musi być zgodnych ze wzorcem zawierającym znaki zastępcze umieszczonym w zapytaniu, żeby nastąpiło uzgodnienie. Wyniki zwracane są tylko przez te terminy, które mają przynajmniej taką liczbę znaków (z wyłączeniem znaku \*). Na przykład jeśli podana zostanie wartość 4, to uzyskanie zgodności możliwe będzie tylko wtedy, gdy termin użyty w zapytaniu będzie zawierać przynajmniej cztery znaki.

Gdy podana zostanie wartość 4, wtedy słowo *technologia* będzie zgodne z terminem *tech\** oraz z terminem *techno\**, ale nie będzie zgodne z terminem *te\**.

Po odświeżeniu lub reorganizacji indeksu zostaną w nim uwzględnione, oprócz oryginalnych terminów, wszystkie możliwe rozszerzenia każdego terminu w dokumencie. Zaletą tego podejścia jest oszczędność czasu potrzebnego na rozwijanie terminów podczas przetwarzania zapytania. Jednak to podejście powoduje zwiększenie wielkości indeksu, co oznacza konieczność posiadania wystarczających zasobów systemowych do obsługi większego indeksu.

To podejście jest najbardziej użyteczne, gdy wielkość kolekcji jest stosunkowo nieduża lub gdy czas i miejsce potrzebne do zbudowania indeksu są mniej ważne niż czas odpowiedzi na zapytanie. Na przykład to podejście można wybrać w celu przeszukania katalogu lub katalogu pracowników.

Takie podejście dostępne jest tylko w przypadku włączonej obsługi znaków zastępczych kończących. Włączenie obsługi znaków zastępczych występujących w dowolnym miejscu terminu użytego w zapytaniu powoduje, że nie można wybrać opcji pozwalającej na włączenie rozwinięć terminów do indeksu.

## Rozwinięcie zapytania

Aby rozwijać zapytania i stosować reguły dopasowywania wzorca podczas wprowadzania przez użytkowników zapytań zawierających znaki zastępcze, należy podać liczbę zgodnych wariantów terminu zapytania. Jeśli na przykład zostanie podana wartość 50, do 50 wariantów terminu zapytania może zostać zakwalifikowane jako zgodne z terminem zapytania.

W tym przykładzie termin zapytania `tech*` jest zgodny ze słowami `techniczny`, `technika`, `technologia` oraz z innymi, ale co najwyżej 50 słowami, które rozpoczynają się od ciągu znaków `tech`.

Chociaż rozwinięcie zapytania ma niewielki wpływ na wielkość indeksu, może ono pogorszyć wydajność zapytania. Procesy wyszukiwania muszą iteracyjnie szukać wszystkich możliwych rozwinięć terminu zapytania ze znakiem zastępczym aż do osiągnięcia limitu określonego w ustawieniach znaku zastępczego.

To podejście jest najbardziej użyteczne, gdy wielkość kolekcji jest stosunkowo duża i należy zminimalizować miejsce i czas budowy indeksu. Na przykład to podejście można wybrać dla repozytoriów poczty elektronicznej, gdzie indeks musi nadążać za szybko zmieniającymi się dokumentami, a czas odpowiedzi na zapytanie jest mniej ważny.

Takie podejście dostępne jest bez względu na to, czy włączona jest obsługa znaków zastępczych kończących czy też obsługa znaków zastępczych występujących w dowolnym miejscu terminu użytego w zapytaniu.

## Obsługa znaków zastępczych w zapytaniach

Zbiór rozwinięć terminu zapytania ze znakiem zastępczym zawiera wszystkie terminy indeksu, które można uzyskać, zastępując znak zastępczy arbitralnymi sekwencjami znaków. Zbiór ten jest określony w następujący sposób:

- Jeśli w kolekcji obsługiwane są znaki zastępcze mogące wystąpić w dowolnym miejscu terminu użytego w zapytaniu, to każdy termin zapytania zawierający gwiazdkę jest interpretowany jako termin ze znakiem zastępczym.
- Zbiór zawiera najwyżej maksymalną liczbę rozwinięć, którą skonfigurował administrator wyszukiwania korporacyjnego. Jeśli indeks zawiera więcej rozwinięć, są one ignorowane. (W wynikach wyszukiwania zawarta jest informacja, czy niektóre rozwinięcia znaków zastępczych zostały zignorowane).
- Jeśli obsługa znaków zastępczych jest ograniczona do pewnego zbioru pól, to zbiór zawiera tylko te terminy, które występują w jednym z podanych pól. Wystarczy, jeśli termin występuje tylko w jednym z pól w przynajmniej jednym dokumencie w indeksie.
- Jeśli termin użyty w zapytaniu dotyczy określonego pola, to znak zastępczy musi wystąpić po specyfikatorze pola (na przykład `nazwa_pola:*sphere`). Nazwa pola nie może zawierać dwukropka (:).
- Jeśli obsługa znaków zastępczych jest ograniczona do pewnego zbioru pól, to nazwa pola występująca w zapytaniu, które zawiera znaki zastępcze, musi być nazwą jednego z pól określonych w konsoli administracyjnej wyszukiwania korporacyjnego. W przeciwnym razie nie zostaną znalezione żadne rozwinięcia danego terminu.
- Znaki zastępcze są obsługiwane tylko w terminach określonych zwykłym tekstem, natomiast nie można ich używać w nazwach elementów XML, nazwach atrybutów ani ich wartościach. Terminy składające się wyłącznie ze znaków zastępczych nie są obsługiwane.

## Konfigurowanie opcji dotyczących znaków zastępczych w zapytaniach

Podczas konfigurowania opcji indeksowania dla kolekcji wyszukiwania korporacyjnego można określić, czy użytkownicy mają mieć możliwość stosowania znaków zastępczych w terminach używanych w zapytaniach.

### Zagadnienia wstępne

Do skonfigurowania opcji znaków zastępczych wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należy dany indeks.

### Informacje o tym zadaniu

Zmienione opcje znaków zastępczych zaczynają obowiązywać po najbliższej reorganizacji indeksu.

### Procedura

Aby skonfigurować obsługę znaków zastępczych w zapytaniach, należy wykonać poniższe czynności:

1. Przejdź do edycji kolekcji, wybierz stronę Indeksowanie, a następnie kliknij opcję **Konfiguruj opcje dotyczące znaków zastępczych**.
2. Na stronie Opcje znaków zastępczych zaznacz pole wyboru **Obsługa znaków zastępczych w zapytaniach**.
3. Za pomocą pola wyboru **Obsługa znaków zastępczych w zapytaniach, które nie wyszukują pól** wskaż, czy znaki zastępcze są obsługiwane w zapytaniach dotyczących tekstu, który nie znajduje się w polach. Na przykład zapytanie pełnotekstowe `tech*`, które nie dotyczy pola o zadanej nazwie, zwraca wyniki rozwinięte (takie jak technologia lub technika) tylko wtedy, gdy to pole wyboru jest zaznaczone.
4. Określ pola, w których obsługiwane są znaki zastępcze:
  - Aby włączyć obsługę znaków zastępczych w zapytaniach dla wszystkich pól w dokumencie, wybierz opcję **Wszystkie pola**.
  - Aby ograniczyć obsługę znaków zastępczych do niektórych pól, wybierz opcję **Konkretne pola**, a następnie wpisz nazwy pól. Wyniki rozwinięte zwracane są tylko dla podanych pól. Na przykład zapytanie `author:jan*` zwraca wyniki rozwinięte tylko wtedy, gdy dla pola `author` włączona jest obsługa znaków zastępczych.
5. Wskaż, czy znak zastępczy musi wystąpić na ostatniej pozycji terminu użytego w zapytaniu (znak zastępczy kończący) czy też jego pozycja jest dowolna i może on wystąpić w każdym miejscu terminu użytego w zapytaniu.

Po określeniu pozycji i typu znaku zastępczego trzeba też podać sposób włączenia obsługi znaków zastępczych. Kliknij przycisk **Pomoc** podczas konfigurowania opcji, aby uzyskać szczegółowe informacje.
6. Kliknij przycisk **OK**.

## Zasięgi

Skonfigurowanie zasięgu pozwala przedstawić użytkownikom ograniczony widok kolekcji.

Zasięg jest grupą pokrewnych identyfikatorów URI w indeksie. Skonfigurowanie zasięgu ogranicza zbiór dokumentów wyświetlanych w kolekcji. Przy wyszukiwaniu w kolekcji uwzględniane są tylko dokumenty znajdujące się w zdefiniowanym zasięgu, a nie wszystkie



dokumenty z indeksu. Aby można było korzystać z tej funkcji, aplikacje wyszukujące muszą zapewniać obsługę wyszukiwania w zakresach.

Tworzenie zasięgu polega na wydzieleniu zakresu identyfikatorów URI w indeksie, które będą dostępne do wyszukiwania dla użytkowników. Ograniczenie zakresu dokumentów dostępnych do wyszukiwania umożliwia lepsze dopasowanie wyników wyszukiwania do oczekiwań użytkowników.

Na przykład można utworzyć jeden zasięg obejmujący identyfikatory URI dla pracowników działu obsługi technicznej oraz drugi zasięg obejmujący identyfikatory URI dla pracowników działu kadr. Jeśli używana aplikacja wyszukująca obsługuje zasięgi, pracownicy działu obsługi technicznej i pracownicy działu kadr będą pobierali dokumenty z odpowiadających tym działom zasięgów.

Istnieje możliwość tworzenia dowolnej liczby zasięgów, jednak utworzenie zbyt wielu zasięgów może mieć wpływ na wydajność. Zasięgi należy konfigurować w taki sposób, aby realizacja większości zapytań wymagała filtrowania tylko przez jeden lub dwa zasięgi. Ponieważ zasięgi mogą zawierać pełne identyfikatory URI lub ich zwróce, jeden dokument może należeć do więcej niż jednego zasięgu.

Podczas konfigurowania zasięgów konieczne może być dwukrotne zreorganizowanie indeksu, aby wprowadzone zmiany zostały uwzględnione. Jeśli zasięgi zostaną skonfigurowane przed wykonaniem pierwszej reorganizacji indeksu dla kolekcji, użytkownicy będą mogli wyszukiwać dokumenty w kolekcji, lecz w wynikach wyszukiwania nie będą mogli wyświetlić danych o zasięgu. Aby mieć pewność, że wyniki wyszukiwania odpowiadają zakresowi identyfikatorów URI w zasięgu, należy ponownie zreorganizować indeks.

Jeśli zasięgi zostaną skonfigurowane po przynajmniej jednej reorganizacji indeksu, zamiany zaczną obowiązywać po kolejnej reorganizacji.

## Konfiguracja zasięgów

Przy konfigurowaniu zasięgu dla kolekcji wyszukiwania korporacyjnego należy określić identyfikatory URI lub wzorce URI dla zakresu dokumentów w indeksie, który użytkownicy mogą wyszukiwać.

### Zagadnienia wstępne

Aby skonfigurować zasięgi, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą zasięgi.

### Informacje o tym zadaniu

Jeśli aplikacje wyszukujące umożliwiają obsługę zasięgów, podczas wyszukiwania w kolekcji użytkownicy mogą wyszukiwać tylko te dokumenty, które odpowiadają identyfikatorom URI zdefiniowanym w ramach zasięgu.

Podczas konfigurowania zasięgów konieczne może być dwukrotne zreorganizowanie indeksu, aby wprowadzone zmiany zostały uwzględnione. Jeśli zasięgi zostaną skonfigurowane przed wykonaniem pierwszej reorganizacji indeksu dla kolekcji, użytkownicy będą mogli wyszukiwać dokumenty w kolekcji, lecz w wynikach wyszukiwania nie będą mogli wyświetlić danych o zasięgu. Aby mieć pewność, że wyniki wyszukiwania odpowiadają zakresowi identyfikatorów URI w zasięgu, należy ponownie zreorganizować indeks.

Jeśli zasięgi zostaną skonfigurowane po przynajmniej jednej reorganizacji indeksu, zamiany zaczną obowiązywać po kolejnej reorganizacji.

## Procedura

Aby skonfigurować zasięg, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Konfiguracja zasięgów**.
2. Na stronie Zasięgi kliknij przycisk **Utwórz zasięg**.
3. Podaj nazwę zasięgu oraz identyfikatory URI i wzorce URI, które definiują granice zasięgu. Można określić również identyfikatory URI i wzorce URI, które mają być wykluczone z zasięgu.
4. Kliknij przycisk **OK**.

Nowy zasięg jest umieszczony na stronie Zasięgi wraz z innymi należącymi do tej kolekcji zasięgami.

### Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79 Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

## Zwinięte identyfikatory URI

Wyniki wyszukiwania korporacyjnego mogą być zorganizowane tak, aby dokumenty ze źródeł WWW i NNTP mające ten sam przedrostek URI były zgrupowane i zwijane podczas wyświetlania w wynikach wyszukiwania.

Gdy wyniki są zwinięte, pierwszy z nich zazwyczaj jest wyrównany do lewego marginesu. Jeden lub kilka dokumentów o niższych ocenach są zgrupowane i wyświetlane poniżej (wyróżnione wcięciem).

Aby zwinąć w jednej grupie dokumenty wynikowe mające różne przedrostki URI, można skojarzyć przedrostki URI z nazwą tworzonej grupy. Na przykład, jeśli używane są trzy serwery do zarządzania danymi finansowymi, w wynikach wyszukiwania można pogrupować dokumenty ze wszystkich trzech serwerów i zwinąć dokumenty o niższym stopniu zgodności z kryteriami wyszukiwania pod dokumentami o wyższym stopniu zgodności.

Aplikacje wyszukujące mogą zwijać dokumenty w wynikach wyszukiwania na podstawie przedrostka URI lub nazwy grupy. W przykładowej aplikacji wyszukującej dla wyszukiwania korporacyjnego wyświetlane są dwa najwyżej ocenione dokumenty wynikowe wyszukiwania. Jeśli w wynikach znajdują się więcej niż dwa dokumenty o tym samym przedrostku URI (lub należące do tej samej grupy URI), można włączyć opcję powodującą zwijanie wyników.

Użytkownicy mogą wykorzystać składnię zapytania wyszukiwania korporacyjnego (`samegroupas:przedrostek URI`) do wyszukiwania wszystkich dokumentów, które są w tej samej grupie co określony przedrostek URI.

## Jak zorganizować przedrostki URI i nazwy grup

Podczas korzystania z konsoli administracyjnej do konfigurowania reguł sterujących zwijaniem wyników wyszukiwania należy określić przedrostki URI dokumentów, które mają być zwijane, oraz skojarzyć przedrostki URI z nazwą grupy (opcjonalnie).

Istotne znaczenie ma kolejność konfigurowanych przedrostków URI. Serwer indeksowania uwzględnia kolejność przedrostków URI podczas obliczania wartości poszczególnych identyfikatorów URI w kolekcji. Dla każdego identyfikatora URI:

1. Serwer indeksowania przegląda kolejno przedrostki URI w regułach dotyczących zwijania wyników wyszukiwania.

2. Kiedy serwer indeksowania znajdzie pierwszy przedrostek URI pasujący do przedrostka dokumentu w indeksie, przypisuje nazwę grupy (lub przedrostek URI, jeśli w regule nie określono nazwy grupy) jako dodatkowy termin wyszukiwania dla tego dokumentu.
3. Jeśli dokumentu nie uda się dopasować do przedrostka URI, to:
  - Dla identyfikatorów URI typu WWW serwer indeksowania stosuje nazwę hosta adresu URL jako przedrostek URI.
  - Dla identyfikatorów URI typu NNTP serwer indeksowania stosuje pierwszy identyfikator wiadomości z nagłówka odwołania jako przedrostek URI.

Po dodaniu przedrostka URI do listy tych, które mają być zwijane w wynikach wyszukiwania, należy jeszcze ustawić dany przedrostek URI w takiej kolejności, w jakiej serwer indeksowania ma go przeglądać i ewentualnie kojarzyć z dokumentami w indeksie jako dodatkowy termin wyszukiwania:

- Po dodaniu przedrostka URI bez kojarzenia go z nazwą grupy można wybrać indywidualny przedrostek URI i przenieść go w górę lub w dół listy.
- Po dodaniu przedrostka URI i skojarzeniu go z nazwą grupy przesunięcie jednego przedrostka z tej grupy w górę lub w dół listy powoduje przesunięcie całej grupy. Kolejność przedrostków URI w obrębie grupy nie ma znaczenia; wybranie jednego przedrostka powoduje automatyczne wybranie całej grupy.

## Zwijanie identyfikatorów URI w wynikach wyszukiwania

Można określić opcje grupowania i zwijania dokumentów wynikowych ze źródeł WWW lub NNTP mających ten sam przedrostek URI. Można także określić nazwę grupy, która umożliwi zwijanie razem dokumentów wynikowych o różnych przedrostkach identyfikatora URI.

### Zagadnienia wstępne

Aby podać opcje zwijania wyników wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienie administratora kolekcji dla tej kolekcji.



### Informacje o tym zadaniu

Zmiany wprowadzone do zwijania wyników wyszukiwania zostaną uwzględnione dopiero po reorganizacji indeksu.

### Procedura

Aby określić opcje zwijania wyników wyszukiwania, należy wykonać następujące czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Zwijanie wyników wyszukiwania**.
2. Na stronie Zwijanie wyników wyszukiwania kliknij opcję **Dodaj przedrostek URI**.
3. Na stronie Dodawanie przedrostka URI dla zwijania wyników wpisz przedrostek URI dla dokumentów, które mają być zwijane w wynikach wyszukiwania. Na przykład:  
`http://finanse/ROI/`  
`http://serwer1.com/finanse/`
4. Można wpisać opisową nazwę grupy, która ma być powiązana z tym przedrostkiem URI. Aby zwinąć w jedną grupę dokumenty wynikowe z kilku źródeł, należy wpisać tę samą nazwę grupy podczas dodawania każdego przedrostka URI.
5. Kliknij przycisk **OK**.
6. Na stronie Zwijanie wyników wyszukiwania umieść nową regułę w takiej kolejności, w jakiej ma być skanowana przez serwer indeksowania:

- Jeśli został dodany nowy przedrostek URI, który nie jest powiązany z nazwą grupy, pojawia się on na dole listy. Przenieś go na właściwą pozycję za pomocą klawiszy strzałek.
  - Jeśli z nowym przedrostkiem URI została powiązana nazwa grupy, pojawia się on na końcu zestawu przedrostków URI należących do tej samej grupy. Do przeniesienia całej grupy przedrostków URI na właściwą pozycję należy użyć klawiszy strzałek.
7. Aby zmienić przedrostek URI lub nazwę grupy, wybierz przedrostek URI i kliknij opcję  **Edytuj**.
  8. Aby usunąć przedrostek URI z listy, wybierz przedrostek URI i kliknij opcję  **Usuń**.

## Usuwanie z indeksu identyfikatorów URI

Aby zapobiec wyszukiwaniu dokumentów w kolekcji przez użytkowników, należy usunąć z indeksu identyfikatory URI tych dokumentów.

### Zagadnienia wstępne

Aby usunąć identyfikatory URI z indeksu, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

### Informacje o tym zadaniu

Jeśli został określony pełny identyfikator URI, użytkownicy nie zobaczą go w wynikach wyszukiwania. Jeśli jednak użytkownik uruchomi to samo zapytanie, a dokumenty wynikowe dla tego zapytania znajdują się w pamięci podręcznej wyszukiwania, to buforowana strona wynikowa dla usuniętego identyfikatora URI nadal będzie zwracana w wynikach wyszukiwania. Pamięć podręczna wyszukiwania nie jest odświeżana, a identyfikator URI nie jest usuwany z indeksu aż do momentu, w którym indeks zostanie odświeżony lub zreorganizowany.

Jeśli został określony wzorzec do usunięcia wielu identyfikatorów URI, pasujące do wzorca identyfikatory URI będą nadal widoczne w wynikach wyszukiwania aż do momentu, w którym indeks zostanie odświeżony lub zreorganizowany.

Usunięcie identyfikatora URI z indeksu nie usuwa go z przestrzeni przeszukiwania. Przy następnym przeszukiwaniu dokumentu przez przeszukiwacz identyfikator URI zostanie wbudowany w indeks i będzie dostępny do ponownego wyszukiwania. Aby usunąć identyfikator URI z przestrzeni przeszukiwania, należy zaktualizować reguły przeszukiwania, w celu wykluczenia dokumentu, po czym zatrzymać i zrestartować przeszukiwacz.

### Procedura

Aby usunąć z indeksu identyfikatory URI określonych dokumentów, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Usuwanie identyfikatorów URI z indeksu**.
2. Na stronie Usuwanie identyfikatorów URI z indeksu wpisz identyfikatory URI (lub wzorce URI), które mają zostać usunięte z indeksu.

Na przykład:

```
http://domena.org/kadry/*
db2://zarzadzanieInformacja/ROI*
cm://przedsiębiorstwo/finanse*
```

3. Kliknij przycisk **OK**.

### Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79  
Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego  
wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.



---

# Administrowanie serwerem wyszukiwania dla wyszukiwania korporacyjnego

Opcje określone dla serwerów wyszukiwania to między innymi: używanie obszaru pamięci podręcznej do zwracania wyników wyszukiwania, definiowanie maksymalnej długości wyświetlanych podsumowań dokumentów w wynikach wyszukiwania, powiązanie niestandardowych słowników w celu zwiększenia jakości wyszukiwania, a także zwracanie predefiniowanych identyfikatorów URI w wynikach wyszukiwania, gdy w zapytaniu zostały użyte określone terminy.

Gdy użytkownik wprowadza zapytanie, serwery wyszukiwania znajdują szybko odpowiednie dokumenty, korzystając z indeksu. Serwery wyszukiwania korzystają ze składnicy danych wyszukiwania korporacyjnego (zawierającej przeanalizowane dane podzielone na leksemy) do pobierania metadanych dla odpowiednich dokumentów. Metadane mogą zawierać identyfikatory URI dokumentów, tytuły, opisy, daty, informacje o typach danych itp.

Podczas konfigurowania serwerów wyszukiwania dla kolekcji określa się opcje determinujące sposób przetwarzania zapytań, w tym także opcje, które mogą mieć wpływ na wydajność zapytań:

## **Konfiguracja pamięci podręcznej wyszukiwania**

Aby zoptymalizować wydajność zapytań, można określić, że dokumenty wynikowe mają być przechowywane w pamięci podręcznej. Można także ustalić wielkość pamięci przeznaczonej na buforowanie dokumentów.

## **Konfiguracja maksymalnej długości wyświetlanych podsumowań dokumentów**

Dla większości dokumentów zwracanych w wynikach wyszukiwania wyświetlane jest podsumowanie treści dokumentu pomagające użytkownikom zdecydować, czy jest to dokument, który chcą pobrać. Istnieje możliwość określenia, jak duży obszar wyników wyszukiwania ma być przeznaczony na wyświetlanie podsumowania.

## **Określanie innego języka domyślnego**

Domyślny język wyszukiwania dokumentów w kolekcji jest określany podczas tworzenia kolekcji, jednak w razie potrzeby można wybrać inny język.

## **Powiązanie niestandardowych słowników**

Jeśli twórcy aplikacji utworzyli niestandardowe słowniki synonimów, słów ignorowanych i słów-kryteriów, można określić, które z nich będą używane podczas przeszukiwania kolekcji przez użytkowników.

## **Konfiguracja szybkich odsyłaczy**

Można predefiniować identyfikatory URI, które mają być zwracane po podaniu określonych słów kluczowych i fraz. Gdy użytkownik wprowadzi określone słowo kluczowe lub frazę w zapytaniu, w wynikach wyszukiwania zostanie zwrócony predefiniowany identyfikator URI. Identyfikatory URI związane z szybkimi odsyłaczami są zwracane dodatkowo, oprócz identyfikatorów URI zwracanych przez serwery wyszukiwania w wyniku przeszukania indeksu.

## **Pojęcia pokrewne**

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

“Niestandardowe słowniki kryteriów” na stronie 144

Aby podnieść jakość wyników wyszukiwania, można wpłynąć na oceny dokumentów w wynikach wyszukiwania przez utworzenie niestandardowego słownika kryteriów.

---

## Pamięci podręczne wyszukiwania

Gdy obciążenie serwerów wyszukiwania jest stosunkowo duże, wydajność procesów wyszukiwania można zwiększyć, używając funkcji buforowania wyników wyszukiwania.

Gdy serwery wyszukiwania przetwarzają zapytania wyszukiwania, najpierw sprawdzana jest zawartość pamięci podręcznej w poszukiwaniu gotowych wyników takiego zapytania. Jeśli serwery wyszukiwania znajdą odpowiednie dokumenty stanowiące wyniki wyszukiwania, mogą je szybko zwrócić użytkownikowi. Jeśli serwery wyszukiwania nie znajdą odpowiednich dokumentów, zaczną wyszukiwanie w indeksie.

W miarę zapełniania się pamięci podręcznej najstarsze dokumenty wynikowe i dokumenty wynikowe rzadko wpisywanych zapytań są usuwane w celu zwolnienia pamięci dla wyników nowych wyszukiwań.

Konsola administracyjna wyszukiwania korporacyjnego umożliwia włączenie buforowania wyników wyszukiwania i określenie pojemności pamięci podręcznej (liczby zapytań, których wyniki mogą być zapisane jednocześnie w pamięci podręcznej).

Aby zmiany wprowadzone do ustawień pamięci podręcznej wyszukiwania zaczęły obowiązywać, należy zrestartować serwery wyszukiwania.

---

## Konfiguracja pamięci podręcznej wyszukiwania

Dla każdej kolekcji można włączyć lub wyłączyć pamięć podręczną wyszukiwania. Istnieje również możliwość określenia opcji sterujących wielkością pamięci podręcznej wyszukiwania.

### Zagadnienia wstępne

Aby skonfigurować pamięć podręczną wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

### Procedura

Aby skonfigurować pamięć podręczną wyszukiwania, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. Na stronie Opcje serwera wyszukiwania zaznacz pole wyboru **Użyj pamięci podręcznej wyszukiwania**.
3. W polu **Maksymalna liczba pozycji w pamięci podręcznej** wpisz maksymalną liczbę zestawów wynikowych wyszukiwania, która może być przechowywana w pamięci podręcznej.
4. Kliknij przycisk **OK**.
5. Aby zmiany zostały uwzględnione, przejdź do monitorowania serwerów wyszukiwania i zrestartuj procesy serwera.

---

## Niestandardowe słowniki synonimów

Aby podnieść jakość wyników wyszukiwania, podczas przeszukiwania kolekcji można włączyć funkcję wyszukiwania synonimów terminów wprowadzonych w zapytaniu.

Po utworzeniu słownika synonimów, dodaniu go do systemu wyszukiwania korporacyjnego i skojarzeniu go z kolekcją użytkownicy mogą wyszukiwać w kolekcjach dokumenty



zawierające synonimy określonych terminów z zapytania. Dzięki takiemu poszerzeniu zakresu zapytań zwiększa się prawdopodobieństwo znalezienia wszystkich potrzebnych dokumentów, a nie tylko tych, które zawierają podane sformułowanie w dosłownej postaci. Ponieważ w słowniku synonimów definiuje się, które słowa są dokładnymi odpowiednikami innych, użytkownicy mogą odszukać wszystkie interesujące ich dokumenty bez konieczności samodzielnego wprowadzania każdej odmiany wyrazu oddzielnie.

Na przykład w danej organizacji mogą funkcjonować akronimy i skrócone nazwy działów, sprzętu itp. Dokumenty w niektórych kolekcjach mogą też zawierać słownictwo specyficzne dla danej branży. Tworząc słownik synonimów, można spowodować, że po uruchomieniu zapytania zawierającego akronim (na przykład ACL) zwrócone zostaną dokumenty zawierające rozwinięcie tego akronimu (lista kontroli dostępu, listy ACL, kontrola dostępu itd.).

W składni zapytań wyszukiwania korporacyjnego synonimy są uwzględniane przez umieszczenie tyldy przed szukanym terminem. Na przykład zapytanie ~WAS może zwrócić dokumenty zawierające pełną nazwę produktu WebSphere Application Server. Twórcy aplikacji mogą także wbudować obsługę synonimów na poziomie właściwości zapytania, co eliminuje konieczność stosowania specjalnej składni.

Słowniki synonimów zawierają odmiany wyrazów. Mają one następujące właściwości:

- Wyrazy nie są związane z jednym językiem, mogą być używane w różnych językach. Dla każdej kolekcji istnieje tylko jeden słownik synonimów.
- Wyrazy nie są odmieniane automatycznie. Wszystkie formy gramatyczne wyrazu należy dodać do listy synonimów. Na przykład liczba pojedyncza i mnoga wyrazu będą stanowiły osobne wpisy w słowniku.

Większość terminów umieszczanych w słowniku stanowi swoje dokładne odpowiedniki semantyczne, co oznacza, że jeśli termin A jest synonimem terminu B, to B jest też synonimem A. Terminy A i B mogą być więc w zapytaniu stosowane zamiennie.

Jednak można także dodawać terminy różne znaczeniowo, na przykład wyrazy o bardziej ogólnym lub bardziej konkretnym znaczeniu. Na przykład w jednej grupie synonimów znajdują się wyrazy budowanie i dom, a w drugiej bank, fundusz i towarzystwo kredytowe.

Im dalsze są związki między terminami, tym większy będzie zestaw wyników wyszukiwania, ale niektóre wyniki mogą nie mieć związku z terminem z zapytania. Interfejs API wyszukiwania i indeksowania udostępnia metody pozwalające wybierać użytkownikom właściwe synonimy przy wprowadzaniu zapytania, a także opcje pokazujące, które kryteria zapytania zostały poszerzone o jakie synonimy.

Słownik synonimów powinien zostać przygotowany przez znawcę tematyki danej kolekcji. Słownik ma postać listy synonimów w formacie XML, w przygotowaniu której może być potrzebna pomoc twórcy aplikacji. Następnie przy użyciu narzędzia dostarczonego z produktem WebSphere II OmniFind Edition plik XML należy przekształcić do postaci pliku binarnego (.dic).

Administrator wyszukiwania korporacyjnego przesyła plik binarny do systemu i przypisuje mu nazwę wyświetlaną. Korzystając z opcji serwera wyszukiwania dla kolekcji, administratorzy kolekcji mogą wybrać słownik synonimów stosowany do wyszukiwania dokumentów w danej kolekcji.

**Ograniczenie:** Niestandardowy słownik synonimów nie może być już modyfikowany po dodaniu go do systemu. Aby zmienić synonimy w słowniku kolekcji, należy:

1. Zmodyfikować odpowiednio źródłowy plik XML.

2. Przekształcić źródło XML na nowy plik słownika.
3. Usunąć poprzedni słownik synonimów z kolekcji, które go używają.
4. Usunąć poprzedni słownik synonimów z systemu.
5. Dodać nowy słownik synonimów do systemu.
6. Powiązać nowy słownik synonimów z kolekcjami, które mają go używać.

#### Pojęcia pokrewne

"Synonym support in search applications" (Obsługa synonimów w aplikacjach wyszukiwujących) w podręczniku "Text Analysis Integration"

#### Zadania pokrewne

"Creating an XML file for synonyms" (Tworzenie pliku XML synonimów) w podręczniku "Text Analysis Integration"

"Creating a synonym dictionary" (Tworzenie słownika synonimów) w podręczniku "Text Analysis Integration"

## Dodawanie słowników synonimów do systemu


Po utworzeniu niestandardowych słowników synonimów do wyszukiwania w dokumentach kolekcji należy je powiązać z systemem wyszukiwania korporacyjnego. Później można wybrać słownik synonimów, który ma być wykorzystywany podczas wyszukiwania w kolekcji.

#### Zagadnienia wstępne

Aby udostępnić niestandardowe słowniki synonimów do wykorzystywania w zapytaniach wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego.

#### Procedura

Aby powiązać synonimy z systemem wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Wyszukiwanie kliknij opcję **Konfigurowanie słowników synonimów**.
4. Na stronie Konfigurowanie słowników synonimów kliknij opcję **Dodaj słownik synonimów**.
5. Na stronie Dodawanie słownika synonimów wpisz unikalną nazwę wyświetlaną dla słownika synonimów i opcjonalnie podaj opis.
6. Określ położenie pliku .dic. Jeśli plik ten jest mniejszy niż 8 MB, może znajdować się w systemie użytkownika. Jeśli plik jest większy niż 8 MB, musi znajdować się na serwerze indeksowania.
7. Kliknij przycisk **OK**. Niestandardowy słownik synonimów zostanie dodany do systemu wyszukiwania korporacyjnego i stanie się dostępny podczas wyszukiwania w kolekcjach.

## Powiązanie słownika synonimów z kolekcją

Jeśli z systemem wyszukiwania korporacyjnego są powiązane słowniki synonimów, można wybrać jeden z nich do wykorzystania podczas wyszukiwania w kolekcji. Jeśli termin z zapytania odpowiada terminowi ze słownika, wynikowe dokumenty zawierające synonimy również zostaną zwrócone w wynikach wyszukiwania.

#### Zagadnienia wstępne

Aby wybrać słownik synonimów dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienie administratora kolekcji dla tej kolekcji.

### Procedura

Aby powiązać słownik synonimów z kolekcją, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. W polu **Nazwa słownika synonimów** na stronie Opcje serwera wyszukiwania wybierz słownik synonimów, który ma być używany podczas wyszukiwania w kolekcji.  
Lista dostępnych słowników synonimów zawiera wszystkie słowniki synonimów dodane do systemu wyszukiwania korporacyjnego.
3. Kliknij przycisk **OK**.

---

## Niestandardowe słowniki słów ignorowanych

Aby podnieść jakość wyników wyszukiwania, można określić, że pewne słowa mają być automatycznie usuwane z terminów użytych w zapytaniu podczas przetwarzania zapytania.

Słownik słów ignorowanych zawiera często używane terminy specyficzne dla danego przedsiębiorstwa, które nie nadają się z tego względu do zastosowania jako terminy w zapytaniu. Dzięki wyłączeniu tych słów z zapytań można zapewnić, że użytkownicy nie będą zasypywani nieodpowiednimi dokumentami wynikowymi (zostaną zwrócone tylko dokumenty zgodne z innymi terminami użytymi w zapytaniu). Podczas przetwarzania zapytania serwery wyszukiwania usuną słowa ignorowane z zapytań. Do usuwanych słów należą słowa ignorowane ze słownika niestandardowego i słowa ignorowane predefiniowane dla wyszukiwania korporacyjnego (takie jak powszechnie używane przyimki i przedimki).

Program WebSphere II OmniFind Edition domyślnie rozpoznaje słowa ignorowane specyficzne dla danego języka. Ten proces usuwa z zapytania często używane słowa typu **to** i **ten**. Niestandardowy słownik słów ignorowanych należy zdefiniować tylko dla terminów charakterystycznych dla danego przedsiębiorstwa lub domeny.

Słownik słów ignorowanych powinien zostać przygotowany przez znawcę tematyki danej kolekcji. Słownik ma postać listy słów ignorowanych w formacie XML, w przygotowaniu której może być potrzebna pomoc twórcy aplikacji. Do przekształcenia tego pliku XML w plik binarny (.dic) należy użyć narzędzia udostępnianego wraz z produktem WebSphere II OmniFind Edition.

Administrator wyszukiwania korporacyjnego przesyła plik binarny do systemu i przypisuje mu nazwę wyświetlaną. Korzystając z opcji serwera wyszukiwania dla kolekcji, administratorzy kolekcji mogą wybrać słownik słów ignorowanych stosowany do wyszukiwania dokumentów w danej kolekcji.

**Ograniczenie:** Niestandardowy słownik słów ignorowanych nie może być już modyfikowany po dodaniu go do systemu. Aby zmodyfikować słowa ignorowane dostępne podczas przetwarzania zapytania, należy:

1. Zmodyfikować odpowiednio źródłowy plik XML.
2. Przekształcić źródło XML na nowy plik słownika.
3. Usunąć poprzedni słownik słów ignorowanych z kolekcji, które go używają.
4. Usunąć poprzedni słownik słów ignorowanych z systemu.
5. Dodać nowy słownik słów ignorowanych do systemu.

6. Powiązać nowy słownik słów ignorowanych z kolekcjami, które mają go używać.

#### **Pojęcia pokrewne**

"Custom stop word dictionaries" (Niestandardowe słowniki słów ignorowanych) w podręczniku "Text Analysis Integration"

#### **Zadania pokrewne**

"Creating an XML file for stop words" (Tworzenie pliku XML słów ignorowanych) w podręczniku "Text Analysis Integration"

"Creating a stop word dictionary" (Tworzenie słownika słów ignorowanych) w podręczniku "Text Analysis Integration"

## **Dodawanie do systemu słowników słów ignorowanych**


Po utworzeniu niestandardowych słowników słów ignorowanych do usuwania słów z zapytań należy je powiązać z systemem wyszukiwania korporacyjnego. Później można wybrać słownik słów ignorowanych, który ma być wykorzystywany podczas wyszukiwania w kolekcji.

#### **Zagadnienia wstępne**

Aby dodać słowniki słów ignorowanych do systemu, wymagana jest rola administratora wyszukiwania korporacyjnego.

#### **Procedura**

Aby powiązać słowa ignorowane z systemem wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Wyszukiwanie kliknij opcję **Konfiguruj słowniki słów ignorowanych**.
4. Na stronie Konfigurowanie słowników słów ignorowanych kliknij opcję **Dodaj słownik słów ignorowanych**.
5. Na stronie Dodawanie słownika słów ignorowanych wpisz unikalną nazwę wyświetlaną dla słownika.
6. Określ położenie pliku .dic. Jeśli plik ten jest mniejszy niż 8 MB, może znajdować się w systemie użytkownika. Jeśli plik jest większy niż 8 MB, musi znajdować się na serwerze indeksowania.
7. Kliknij przycisk **OK**. Niestandardowy słownik słów ignorowanych zostanie dodany do systemu wyszukiwania korporacyjnego i stanie się dostępny podczas wyszukiwania w kolekcjach.

## **Powiązanie słownika słów ignorowanych z kolekcją**

Jeśli z systemem wyszukiwania korporacyjnego są powiązane słowniki słów ignorowanych, można wybrać jeden z nich do wykorzystania podczas wyszukiwania w kolekcji. Jeśli termin użyty w zapytaniu odpowiada terminowi ze słownika, termin ten zostanie usunięty z zapytania przed jego przetworzeniem.

#### **Zagadnienia wstępne**

Aby wybrać słownik słów ignorowanych dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

## Procedura

Aby powiązać słownik słów ignorowanych z kolekcją, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. W polu **Nazwa słownika słów ignorowanych** na stronie Opcje serwera wyszukiwania wybierz słownik słów ignorowanych, który ma być używany podczas wyszukiwania w kolekcji.  
Lista dostępnych słowników zawiera wszystkie słowniki słów ignorowanych dodane do systemu wyszukiwania korporacyjnego.
3. Kliknij przycisk **OK**.

---

## Podsumowywanie dynamiczne

Dynamiczne podsumowywanie to technika umożliwiająca określenie w dokumencie wynikowym fraz, które najlepiej reprezentują pojęcia wyszukiwane przez użytkownika.

W wyszukiwaniu korporacyjnym technika dynamicznego podsumowywania polega na próbach przechwycenia w dokumentach zdań, w których występuje wiele poszukiwanych wyrazów. Wybieranych jest kilka zdań lub części zdań, które są wyświetlane w wynikach wyszukiwania. Poszukiwane wyrazy są podświetlone w dokumencie HTML z wynikami wyszukiwania.

Podczas konfigurowania opcji serwera wyszukiwania dla kolekcji można określić maksymalną długość podsumowań dokumentów wyświetlanych w wynikach wyszukiwania. Ponieważ podsumowanie obejmuje także znaki podświetlone, bufor zwrócony do aplikacji wyszukującej będzie większy niż określona wartość maksymalna. Jednak wyświetlana długość nie przekroczy określonej wartości maksymalnej. Podsumowanie może też być krótsze, co zależy od danych podsumowania wyodrębnionych z dokumentu źródłowego.

## Dostosowywanie podsumowań dokumentów w konsoli administracyjnej

Można dostosować zakres informacji wyświetlanych w podsumowaniach dokumentów, określając opcje dla serwera wyszukiwania w konsoli administracyjnej wyszukiwania korporacyjnego.

### Zagadnienia wstępne

Aby sterować wyświetlaną długością podsumowań dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora danej kolekcji.

### Informacje o tym zadaniu

Wartość określona dla maksymalnej wyświetlanej długości podsumowań dokumentów jest powiązana z wartością określoną dla liczby zdań, które może zawierać każde podsumowanie. Pierwszeństwo ma wartość, która daje najkrótsze podsumowanie dokumentu.

Jeśli na przykład zostanie określony limit czterech zdań, podsumowanie dokumentu będzie zawierać tylko cztery zdania, nawet jeśli wyświetlana długość dopuszcza większą liczbę znaków niż łączna liczba znaków w tych zdaniach. Kolejny przykład: limit 10 zdań w połączeniu z limitem wyświetlanej długości równym 500 znaków może w efekcie dać podsumowanie dokumentu zawierające mniej niż 10 zdań.

## Procedura

Aby skonfigurować wyświetlaną długość podsumowań dokumentów, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. Na stronie Opcje serwera wyszukiwania określ maksymalną wyświetlaną długość podsumowań dokumentów. Podczas przeglądania wyników wyszukiwania przez użytkowników podsumowania dokumentów nie będą przekraczały określonej wielkości.
3. Określ liczbę zdań, które może zawierać każde podsumowanie dokumentu (podsumowania mogą zawierać do 10 zdań).
4. Kliknij przycisk **OK**.
5. Aby zmiany zostały uwzględnione, przejdź do monitorowania serwerów wyszukiwania i zrestartuj procesy serwera.

## Dostosowywanie podsumowań dokumentów za pomocą edycji właściwości

Każdy dokument wynikowy zwracany w wyniku zapytania wyszukiwania korporacyjnego zawiera podsumowanie. Można zmienić zakres informacji, które zawiera każde podsumowanie, edytując plik właściwości.

### Informacje o tym zadaniu

Opisy wyników wyszukiwania można dostosować, zmieniając wartości następujących właściwości w pliku

`ES_NODE_ROOT/master_config/identyfikator_kolekcji.runtime.node1/runtime-generic.properties`:

#### **MinWordsPerSentence**

Minimalna liczba słów, które może zawierać zdanie opisu. Wartością domyślną jest 4.

#### **MaxWordsPerSentence**

Maksymalna liczba słów, które może zawierać zdanie opisu. Wartością domyślną jest 20.

#### **NumberOfReturnedSentences**

Liczba zdań, które tworzą opis dokumentu. Wartością domyślną jest 5.

#### **MaxSentencesPerDocument**

Maksymalna liczba zdań w dokumencie, które będą brane pod uwagę w procesie tworzenia opisu. Wartością domyślną jest 1000.

### Procedura

Aby dostosować podsumowania dokumentów w wynikach wyszukiwania, należy wykonać następujące czynności:

1. Zaloguj się jako administrator wyszukiwania korporacyjnego na serwerach wyszukiwania. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
2. Korzystając z edytora tekstu zmień następujący plik, gdzie zmienna *identyfikator\_kolekcji* jest identyfikatorem, który został określony (lub przypisany przez system) podczas tworzenia kolekcji:

```
ES_NODE_ROOT/master_config/identyfikator_kolekcji.runtime.node1/
runtime-generic.properties
```

**Wskazówka:** Informację o tym, jaki identyfikator kolekcji jest przypisany do danej nazwy kolekcji, znajdziesz w pliku ES\_NODE\_ROOT/master\_config/collections.ini.

3. Zmień właściwości, które należy dostosować, a następnie zapisz i zamknij plik.
4. Zatrzymaj i zrestartuj serwery wyszukiwania, aby zastosować wprowadzone zmiany.

---

## Praca z szybkimi odsyłaczami

Szybkie odsyłacze są to dokumenty zwracane w wynikach wyszukiwania, gdy użytkownik wprowadzi zapytanie, w którym występują konkretne słowa i frazy.

Do konfigurowania szybkich odsyłaczy dla kolekcji służy konsola administracyjna wyszukiwania korporacyjnego.

### Szybkie odsyłacze

Użycie szybkich odsyłaczy umożliwia udostępnienie użytkownikom odsyłaczy do dokumentów, które wstępnie zaklasyfikowano jako związane z tematem wyszukiwania.

Szybki odsyłacz jest to identyfikator URI umieszczany automatycznie przez mechanizm wyszukiwania korporacyjnego w wynikach wyszukiwania, gdy w zapytaniu występują określone słowa lub frazy. Zazwyczaj identyfikatory URI szybkich odsyłaczy są wyświetlane na początku listy wynikowej w celu zapewnienia użytkownikom łatwego dostępu do dokumentów, które zostały wstępnie określone jako odpowiednie dla danego zapytania.

Szybkie odsyłacze są zwracane dodatkowo oprócz pozostałych wyników wyszukiwania. Procesy wyszukiwania przeszukują indeksy w celu znalezienia dokumentów zgodnych z terminami określonymi w zapytaniu i oprócz identyfikatorów URI szybkich odsyłaczy zwracają identyfikatory URI tych dokumentów.

Podczas konfigurowania szybkiego odsyłacza można dla każdego identyfikatora URI określić opisowy tytuł i podsumowanie, aby ułatwić użytkownikom rozpoznanie dokumentów i podjęcie szybkiej decyzji, czy określony dokument ma zostać pobrany.

Na przykład dla identyfikatora URI <http://www.ibm.com/education/us/> może zostać użyty tytuł taki, jak Szkolenia firmy IBM w Stanach Zjednoczonych, a także następujące podsumowanie: Rozwiązania, produkty i zasoby dla profesjonalistów, instruktorów i studentów w Stanach Zjednoczonych.

Aby można było używać szybkich odsyłaczy w kolekcjach wyszukiwania korporacyjnego, w aplikacji wyszukującej musi być dostępna opcja wyświetlania szybkich odsyłaczy. W niektórych aplikacjach wyszukujących użytkownicy mogą mieć możliwość włączenia i wyłączenia funkcji zwracania szybkich odsyłaczy podczas wyszukiwania kolekcji.

### Konfiguracja szybkich odsyłaczy

Aby utworzyć szybki odsyłacz do kolekcji wyszukiwania korporacyjnego, należy powiązać identyfikator URI dokumentu ze słowami kluczowymi, które spowodują jej włączenie do wyników wyszukiwania.

#### Zagadnienia wstępne

Aby skonfigurować szybkie odsyłacze, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji, do której należą te odsyłacze.

#### Informacje o tym zadaniu

Przykłady określania słów kluczowych i identyfikatorów URI dla szybkich odsyłaczy dostępne są po kliknięciu przycisku **Pomoc** w trakcie tworzenia lub edytowania szybkiego odsyłacza.

Nie ma potrzeby restartowania serwerów wyszukiwania, aby zastosować wprowadzone zmiany.

### **Procedura**

Aby skonfigurować szybki odsyłacz, należy wykonać poniższe czynności:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj szybkie odsyłacze**.
2. Na stronie Szybkie odsyłacze kliknij opcję **Utwórz szybki odsyłacz**.
3. Określ słowa kluczowe i frazy, dzięki którym ten szybki odsyłacz będzie zwracany w wynikach wyszukiwania, identyfikator URI dokumentu określonego uprzednio jako odpowiedni dla tego zapytania oraz opcje szybkiego odsyłacza.

W jednej linii można określić jedno lub kilka słów kluczowych bądź jedną frazę (dwa lub więcej słów ujętych w cudzysłów). Słowa kluczowe należy rozdzielać spacją (nie należy oddzielać słów kluczowych za pomocą przecinka). Naciśnij klawisz Enter, aby rozpocząć nowy wiersz.

4. Kliknij przycisk **OK**.

Nowy szybki odsyłacz jest wyświetlany - wraz z innymi należącymi do kolekcji - na stronie Szybkie odsyłacze.

#### **Odsyłacze pokrewne**

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79  
Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.



---

## Ocenianie dokumentów w wyszukiwaniu korporacyjnym

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

Serwery wyszukiwania obsługują wzbogaconą składnię zapytań i korzystają z kilku technik zapewniających zwrócenie najlepiej dopasowanych wyników wyszukiwania, takich jak określanie wyników na podstawie tekstu i ocena statyczna. Można rozszerzyć domyślny sposób oceny, konfigurując opcje wpływające na ważność dokumentów w wynikach wyszukiwania:

- Można utworzyć niestandardowe słowniki kryteriów, aby wpłynąć na sposób oceny dokumentów zawierających określone słowa-kryteria w wynikach wyszukiwania.
- Można wpływać na wyniki dokumentów zgodnych z określonym wzorcem URI.
- Można wpływać na wyniki dokumentów zawierających pola odwzorowane na klasy kryteriów.

### Pojęcia pokrewne

“Ocenianie dokumentów na podstawie klas kryteriów” na stronie 148

Przez odwzorowanie pól na klasy kryteriów można wpłynąć na ocenę dokumentów w wynikach wyszukiwania.

“Ocena dokumentów na podstawie wzorców URI” na stronie 146

Ważność dokumentów można zwiększyć lub zmniejszyć, przypisując współczynniki kryteriów do wzorców URI.

“Niestandardowe słowniki kryteriów” na stronie 144

Aby podnieść jakość wyników wyszukiwania, można wpłynąć na oceny dokumentów w wynikach wyszukiwania przez utworzenie niestandardowego słownika kryteriów.

---

## Wartościowanie danych tekstowych

Wyszukiwanie korporacyjne w sposób dynamiczny oblicza wynik dla każdego dokumentu, który spełnia warunki zapytania.

Przy obliczaniu wyniku tekstowego poszczególnych dokumentów spełniających warunki zapytania wyszukiwanie korporacyjne uwzględnia wiele czynników, takich jak:

- Częstotliwość występowania każdego z warunków w całej kolekcji. Na ogół warunki zapytania występujące w większości dokumentów w mniejszym stopniu wpływają na wynik dokumentu niż te, które spotykane są w wybranym zbiorze dokumentów.
- Liczba wystąpień każdego z warunków zapytania w pasującym dokumencie. Na ogół wynik jest tym większy, im więcej razy warunek występuje w danym dokumencie.
- Odległość występowania warunków zapytania w każdym z pasujących dokumentów. Na ogół warunki zapytania znajdujące się blisko siebie w ramach dokumentu bardziej wpływają na wynik dokumentu niż te same warunki występujące dalej od siebie.
- Kontekst, w jakim poszczególne warunki zapytania występują w każdym z dokumentów odpowiadających tym warunkom. Na przykład warunek zapytania występujący w tytule dokumentu w większym stopniu wpływa na wynik dokumentu niż ten sam warunek występujący w zwykłym tekście.

Na wynik wpływają także długości poszczególnych dokumentów i bogactwo zawartego w nich słownictwa.

---

## Ocena statyczna

W przypadku niektórych typów dokumentów możliwe jest przypisanie statycznego współczynnika oceniania, który zwiększa jego ważność w wynikach wyszukiwania.

Decyzję o przypisaniu statycznego współczynnika oceniania dokumentom w kolekcji podejmuje się przy tworzeniu kolekcji. W przypadku dokumentów internetowych czynnikami zwiększającymi ważność dokumentu w wynikach wyszukiwania są liczba odsyłaczy do tego dokumentu z innych dokumentów oraz źródła tych odsyłaczy.

W przypadku dokumentów zawierających pola daty lub metadane z datą ważność dokumentu może być powiększana na podstawie daty dokumentu. Na przykład nowsze artykuły w grupach dyskusyjnych NNTP mogą być bardziej adekwatne niż artykuły starsze. Jeśli źródło danych zawiera wiele wartości daty, można wybrać, która z tych dat jest najważniejsza, aby określić ważność dokumentów w tym źródle danych.

Korzystając z oceny statycznej w odniesieniu do kolekcji, należy pamiętać, aby nie mieszać ze sobą w ramach jednej kolekcji źródeł danych stosujących różne typy ocen. Na przykład, jeśli statycznym współczynnikiem oceniania mają być odsyłacze do dokumentu, kolekcja powinna zawierać wyłącznie dokumenty WWW. Połączenie w ramach jednej kolekcji źródeł o zróżnicowanych modelach oceniania może ujemnie wpłynąć na jakość wyników wyszukiwania.

Należy także zadbać o to, aby dokumenty zawierały pola i wartości umożliwiające stosowanie ocen statycznych. Na przykład, jeśli współczynnikiem oceniania dokumentów w kolekcji ma być data dokumentu, a dokumenty nie zawierają pól ani atrybutów daty, jakość wyników wyszukiwania na tym ucierpi.

### Zadania pokrewne

“Migrowanie kolekcji z portalu WebSphere Portal” na stronie 205

Aby przeprowadzić migrację kolekcji z portalu WebSphere Portal do wyszukiwania korporacyjnego, należy przygotować kolekcje w portalu WebSphere Portal, a następnie użyć Kreatora migracji do wykonania migracji.

---

## Niestandardowe słowniki kryteriów

Aby podnieść jakość wyników wyszukiwania, można wpłynąć na oceny dokumentów w wynikach wyszukiwania przez utworzenie niestandardowego słownika kryteriów.

Jeśli w zapytaniu określono słowo występujące w słowniku kryteriów, ważność dokumentów zawierających to słowo zostanie zwiększona lub zmniejszona zgodnie ze współczynnikiem kryterium skonfigurowanym dla tego słowa w słowniku.

Współczynniki kryteriów mają wartości z zakresu od -10 do 10. Podczas przetwarzania zapytania serwery wyszukiwania zwiększają ważność dokumentów zawierających słowa z dodatnimi współczynnikami kryteriów i zmniejszają ważność dokumentów zawierających słowa z ujemnymi współczynnikami kryteriów.

Na przykład dokument zgodny z użytymi w zapytaniu terminami o wysokich współczynnikach kryteriów zostanie oceniony wyżej niż w przypadku, gdy współczynnik kryterium nie zostałby zastosowany (na ocenę dokumentu w wynikach wyszukiwania wpływają tylko terminy z zapytania mające określone wartości kryteriów, a współczynnik kryterium jest tylko jednym z czynników wpływających na wynik dokumentu).

Tworząc słownik można przypisać ten sam współczynnik kryterium do dowolnej liczby słów. Słownik taki może zawierać terminy złożone z pojedynczych słów, jak również z wielu słów (terminy złożone z wielu słów są dopasowywane jako frazy).

Jeśli słowo podlegające wartości kryterium jest określone w zapytaniu z użyciem operatora OR (na przykład: *jeden | dwa*), dla terminów użytych w zapytaniu obliczana jest średnia ważona. Zagregowany wynik jest używany dla wszystkich wystąpień operandów zapytania OR (nie są obliczane różne wyniki dla różnych operandów).

Wpływanie na wyniki przy użyciu słowników kryteriów nie jest obsługiwane w przypadku terminów użytych w zapytaniu w odniesieniu pól. Podczas analizowania terminów użytych w zapytaniu do obliczenia wyniku dokumentu używany jest wyłącznie tekst zapytania, a nie nazwa pola (aby zastosować współczynniki kryteriów do terminów użytych w zapytaniu i występujących w polach, można odwzorować nazwy pól na klasy kryteriów).

Słownik kryteriów powinien zostać przygotowany przez znawcę tematyki danej kolekcji. Słownik ma postać listy słów-kryteriów w formacie XML, w przygotowaniu której może być potrzebna pomoc twórcy aplikacji. Do przekształcenia tego pliku XML w plik binarny (.dic) należy użyć narzędzia udostępnianego wraz z produktem WebSphere II OmniFind Edition.

Administrator wyszukiwania korporacyjnego przesyła plik binarny do systemu i przypisuje mu nazwę wyświetlaną. Administratorzy kolekcji mogą wybrać słownik kryteriów używany przy wyszukiwaniu dokumentów w kolekcji podczas konfigurowania opcji serwera wyszukiwania dla kolekcji.

**Ograniczenie:** Niestandardowy słownik kryteriów nie może być już modyfikowany po dodaniu go do systemu. Aby zweryfikować listę słów-kryteriów dostępnych do przetwarzania zapytań, należy:

1. Zmodyfikować odpowiednio źródłowy plik XML.
2. Przekształcić źródło XML na nowy plik słownika.
3. Usunąć poprzedni słownik kryteriów z kolekcji, które go używają.
4. Usunąć ten słownik kryteriów z systemu.
5. Dodać do systemu nowy słownik kryteriów.
6. Powiązać nowy słownik kryteriów z kolekcjami, które mają z niego korzystać.

#### **Pojęcia pokrewne**

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

“Custom boost word dictionaries” (Niestandardowe słowniki kryteriów) w podręczniku “Text Analysis Integration”

#### **Zadania pokrewne**

“Creating an XML file for boost words” (Tworzenie pliku XML słów-kryteriów) w podręczniku “Text Analysis Integration”

“Creating a boost word dictionary” (Tworzenie słownika kryteriów) w podręczniku “Text Analysis Integration”

## **Dodawanie do systemu słowników kryteriów**


Po utworzeniu niestandardowych słowników kryteriów należy je powiązać z systemem wyszukiwania korporacyjnego. Później można wybrać słownik kryteriów, który ma być używany podczas wyszukiwania w kolekcji.

## Zagadnienia wstępne

Do dodawania niestandardowych słowników kryteriów do systemu wymagana jest rola administratora wyszukiwania korporacyjnego.

## Procedura

Aby powiązać niestandardowe słowa-kryteria z systemem wyszukiwania korporacyjnego:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Wyszukiwanie kliknij opcję **Konfiguruj słowniki kryteriów**.
4. Na stronie Konfigurowanie słowników kryteriów kliknij opcję **Dodaj słownik kryteriów**.
5. Na stronie Dodawanie słownika kryteriów wpisz unikalną nazwę wyświetlaną słownika i opcjonalnie podaj jego opis.
6. Określ położenie pliku .dic. Jeśli plik ten jest mniejszy niż 8 MB, może znajdować się w systemie użytkownika. Jeśli plik jest większy niż 8 MB, musi znajdować się na serwerze indeksowania.
7. Kliknij przycisk **OK**. Niestandardowy słownik kryteriów zostanie dodany do systemu wyszukiwania korporacyjnego i będzie dostępny na potrzeby wyszukiwania w kolekcjach.

## Powiązanie słownika kryteriów z kolekcją

Jeśli z systemem wyszukiwania korporacyjnego są powiązane słowniki kryteriów, można wybrać jeden z nich do użycia podczas wyszukiwania w kolekcji. Jeśli termin użyty w zapytaniu odpowiada terminowi ze słownika, wówczas ważność dokumentów zawierających ten termin zostanie zwiększona lub zmniejszona zgodnie ze współczynnikiem kryterium przypisanym do danego terminu w słowniku.

## Zagadnienia wstępne

Do wybrania słownika kryteriów dla kolekcji wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora do danej kolekcji.

## Procedura

Aby powiązać słownik kryteriów z kolekcją:

1. Rozpocznij edycję kolekcji, wybierz stronę Wyszukiwanie i kliknij opcję **Konfiguruj opcje serwera wyszukiwania**.
2. W polu **Nazwa słownika kryteriów** na stronie Opcje serwera wyszukiwania wybierz słownik kryteriów, który ma być używany podczas wprowadzania przez użytkowników zapytań dla tej kolekcji.  
Lista dostępnych słowników zawiera wszystkie słowniki kryteriów dodane do systemu wyszukiwania korporacyjnego.
3. Kliknij przycisk **OK**.

---

## Ocena dokumentów na podstawie wzorców URI

Ważność dokumentów można zwiększyć lub zmniejszyć, przypisując współczynniki kryteriów do wzorców URI.

Do wszystkich dokumentów dodawanych do indeksu przypisywane są domyślne wyniki oceny statycznej. Domyślny wynik różni się w zależności od tego, czy ocena statyczna została

włączona dla kolekcji, a jeśli tak, od typu oceny statycznej (według daty dokumentu lub, dla dokumentów sieci WWW, liczby innych dokumentów zawierających odsyłacze do danego dokumentu).

Użytkownik może wpływać na względną ważność dokumentu, przypisując współczynniki kryteriów do wzorców URI. Współczynnik kryterium jest używany wraz z domyślnym wynikiem oceny statycznej i innymi czynnikami do określenia ostatecznego wyniku statycznego dokumentu.

Istotne znaczenie ma kolejność konfigurowanych wzorców URI. Podczas obliczania wartości każdego dokumentu w kolekcji serwer indeksowania ocenia wzorce URI w kolejności, w jakiej zostały podane. Dla każdego identyfikatora URI:

1. Serwer indeksowania skanuje kolejno poszczególne wzorce URI.
2. Gdy serwer indeksowania znajdzie pierwszy wzorec URI zgodny z dokumentem w indeksie, zastosuje do dokumentu współczynnik kryterium skonfigurowany dla tego wzorca URI.
3. Jeśli dokumentu nie uda się dopasować do wzorca URI, zostanie użyty domyślny wynik oceny statycznej.

Po skonfigurowaniu współczynnika kryterium dla wzorca URI należy określić kolejność, w jakiej będzie go przeglądał serwer indeksowania.

#### **Pojęcia pokrewne**

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

## **Wpływanie na wyniki dokumentów zgodnych z wzorcami URI**

Ważność dokumentów zgodnych z wzorcem URI można zwiększyć lub zmniejszyć, stosując współczynnik kryterium do domyślnego wyniku oceny statycznej.

#### **Zagadnienia wstępne**

Aby zmienić ważność dokumentów zgodnych ze wzorcem URI, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora do danej kolekcji.

#### **Informacje o tym zadaniu**

Skonfigurowany współczynnik kryterium jest używany wraz z domyślnym wynikiem oceny statycznej do obliczenia nowego wyniku statycznego dla wszystkich dokumentów zgodnych z określonym wzorcem URI.

Współczynniki kryteriów poprawiają tylko wyniki statyczne i stanowią tylko jeden z elementów wpływających na obliczenia określające ostateczną ocenę dokumentu. Jeśli na przykład istnieje wiele odsyłaczy do danego dokumentu (które dają wysoki wynik początkowy), inny dokument, do którego nie ma żadnych odsyłaczy, zostanie oceniony niżej.

#### **Procedura**



Aby wpłynąć na wyniki dokumentów zgodnych z wzorcem URI:

1. Rozpocznij edycję kolekcji, wybierz stronę Indeksowanie i kliknij opcję **Wpływaj na wyniki poprzez dopasowywanie wzorców identyfikatorów URI**.
2. Na stronie Oddziaływanie na wyniki poprzez dopasowywanie wzorców identyfikatorów URI kliknij opcję **Dodaj wzorec URI**.

3. Wpisz wzorzec URI dla dokumentów, których ważność w wynikach wyszukiwania ma zostać zwiększona lub zmniejszona. Na przykład:  
`http://domena.org/kadry/*`  
`db2://*ROI*`  
`*/afs/*`
4. Wpisz wartość współczynnika kryterium z zakresu od -10 do 10. Ostateczny wynik statyczny dla wszystkich dokumentów zgodnych ze wzorcem URI zostanie obliczony na podstawie tego współczynnika kryterium.
5. Kliknij przycisk **OK**.
6. Na stronie Oddziaływanie na wyniki poprzez dopasowywanie wzorców identyfikatorów URI ustaw nowy wzorzec URI w takiej kolejności, w jakiej ma go przeglądać serwer indeksowania.

Serwer indeksowania oblicza wyniki oceny statycznej w kolejności, w jakiej wymienione są identyfikatory URI. Aby uzyskać najlepsze rezultaty, bardziej szczegółowe identyfikatory URI należy wymienić jako pierwsze. W poniższym przykładzie podkatalog `/forms` jest zgodny ze wzorcem identyfikatora URI `http://www.ibm.com/hr/*`. Aby mieć pewność, że wyniki dla dokumentów z podkatalogu `/forms` są obliczane poprawnie, wzorzec identyfikatora URI dla podkatalogu `/forms` należy umieścić jako pierwszy:

```
http://www.ibm.com/hr/forms/* 8
http://www.ibm.com/hr/* -2
```

7. Aby zmienić wzorzec URI lub współczynnik kryterium, należy wybrać wzorzec URI i kliknąć opcję  **Edytuj**.
8. Aby usunąć wzorzec URI z listy, należy wybrać wzorzec URI i kliknąć opcję  **Usuń**.
9. Aby zastosować współczynniki kryteriów do dokumentów poindeksowanych wcześniej, należy zreorganizować indeks.

#### Pojęcia pokrewne

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

#### Odsyłacze pokrewne

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79

Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

---

## Ocenianie dokumentów na podstawie klas kryteriów

Przez odwzorowanie pól na klasy kryteriów można wpłynąć na ocenę dokumentów w wynikach wyszukiwania.

Podczas analizowania dokumentów analizator składni przypisuje *klasy kryteriów* do leksemów dokumentu w zależności od pól, do których należą te leksemy. Te klasy kryteriów są włączane do indeksu i używane podczas wartościowania zapytania do obliczania wyników mających wpływ na ocenę zwracanych dokumentów.

Aby wpłynąć na sposób obliczania wyników, dla klas kryteriów można skonfigurować liczbowe współczynniki kryteriów. Jeśli termin użyty w zapytaniu jest zgodny z leksemem w polu odwzorowanym na klasę kryteriów, wkład tego wystąpienia leksemu wpływa na łączny wynik dokumentu. Wynik jest obliczany przez zastosowanie współczynnika kryterium skonfigurowanego w tej klasie kryteriów.

Można na przykład uwydatnić wyniki pól tytułu. Gdy termin użyty w zapytaniu wystąpi w tytule, takie wystąpienie będzie miało duży wkład w wynik tego dokumentu i pomoże mu w uzyskaniu lepszej oceny w wynikach wyszukiwania.

W celu oddziaływania na oceny dokumentów należy określić współczynniki kryteriów dla klas kryteriów oraz odwzorować pola na klasy kryteriów za pomocą konsoli administracyjnej wyszukiwania korporacyjnego. Na potrzeby wyszukiwania korporacyjnego udostępniono szesnaście predefiniowanych klas kryteriów. Osiem z nich zostało zaprojektowanych do użycia z polami treści, a drugie osiem dla pól metadanych. Wyniki powiązane z domyślnymi klasami kryteriów można edytować. Można również powiązać z klasami kryteriów inne lub dodatkowe pola.

W przypadku zmiany odwzorowań pól konieczne jest ponowne przeszukanie i przeanalizowanie dokumentów, aby zmiany zostały zastosowane do dokumentów, które zostały już wcześniej poindeksowane. Jeśli zmienione zostaną współczynniki określone dla klasy kryteriów, należy przejść do monitorowania serwerów wyszukiwania i zatrzymać, a następnie zrestartować procesy serwera wyszukiwania, aby zmiany te odniosły skutek.

## Wykrywanie duplikatów dokumentów

Odwzorowując pole na klasę kryteriów, należy określić, czy pole to jest używane do wykrywania zduplikowanych dokumentów:

- Jeśli pole jest używane do wykrywania zduplikowanych dokumentów, jest to pole treści i dostępne do wybrania są tylko klasy kryteriów zaprojektowane dla pól treści.
- Jeśli pole nie jest używane do wykrywania zduplikowanych dokumentów, jest to pole metadanych i dostępne do wybrania są tylko klasy kryteriów zaprojektowane dla pól metadanych. W tym przypadku za duplikaty uważa się dwa dokumenty, które są dokładnie takie same za wyjątkiem określonego pola.

## Wysoka i niska wartość znajdawalności

Przy wartościowaniu zapytania procesy wyszukiwania szacują liczbę dokumentów wynikowych, które zostaną zwrócone. Wartości progowe określają wówczas, czy zapytanie jest uważane za mające niską wartość znajdawalności czy też wysoką wartość znajdawalności:

### Niska znajdawalność

Jeśli szacowana liczba dokumentów wynikowych jest mniejsza niż wartość dolnego progu, zapytanie jest uznawane za zapytanie o niskiej znajdawalności.

### Wysoka znajdawalność

Jeśli szacowana liczba dokumentów wynikowych jest większa niż wartość górnego progu, zapytanie jest uznawane za zapytanie o wysokiej znajdawalności.

### Mieszana

Jeśli szacowana liczba dokumentów znajduje się między wartościami obu progów, wartość znajdawalności tego zapytania jest połączeniem tych dwóch wartości progowych.

Każda klasa kryteriów określa współczynniki kryteriów powiązane podczas przetwarzania zapytań z zapytaniami o niskiej znajdawalności i zapytaniami o wysokiej znajdawalności. Niski współczynnik kryterium wpływa na względną ważność zapytań o niskiej znajdawalności, a wysoki współczynnik kryterium wpływa na względną ważność zapytań o wysokiej znajdawalności. Połączenie tych dwóch współczynników wpływa na względną ważność zapytań o mieszanej wartości znajdawalności.

Wartości współczynników kryteriów kontrolują względną ważność każdego wystąpienia terminu z zapytania w dokumencie. Każde wystąpienie w dokumencie terminu użytego w zapytaniu jest zliczane zgodnie z odpowiednim współczynnikiem kryterium.

Domyślne współczynniki kryteriów można edytować podczas konfigurowania klas kryteriów dla kolekcji. Na przykład można określić współczynniki kryteriów w taki sposób, aby terminy użyte w zapytaniu i występujące w polu tytułu liczyły się pięć razy bardziej niż terminy występujące w zwykłym tekście.

#### Pojęcia pokrewne

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

## Odwzorowywanie pól na klasy kryteriów

Istnieje możliwość wpływania na względną ważność pól przez odwzorowywanie nazw pól na klasy kryteriów.

### Zagadnienia wstępne

Do odwzorowywania pól na klasy kryteriów wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora do danej kolekcji.

### Informacje o tym zadaniu

System używa współczynnika kryterium do wpływania na ocenę dokumentów zawierających terminy z zapytania w obrębie pól odwzorowanych na klasy kryteriów.

Pewne nazwy pól są zarezerwowane na potrzeby wyszukiwania korporacyjnego, aby możliwe było obliczanie wyników dla pól i tekstu pozbawionych dodatkowych definiujących je charakterystyk (na przykład dla pola niebędącego polem tytułu lub dla zwykłego, niewyróżnionego tekstu). Na klasy kryteriów używane przez pola zarezerwowane można odwzorowywać też inne pola, ale nie można edytować ani usuwać pól zarezerwowanych.

### Procedura

Aby odwzorować pola na klasy kryteriów:


1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij przycisk **Odwzoruj pola na klasy kryteriów**.
2. Na stronie Odwzorowywanie pól na klasy kryteriów kliknij przycisk **Dodaj pole**.
3. Na stronie Dodawanie pola do klasy kryteriów wpisz nazwę pola, które ma zostać odwzorowane na klasę kryteriów.

Określić można nazwę pola istniejącego w przeszukiwanym źródle lub źródle zewnętrznym, nazwę pola odwzorowanego z elementu XML, nazwę pola odwzorowanego z elementu metadanych HTML lub jedną z predefiniowanych nazw pól.


4. Określ, czy to pole jest używane do wykrywania zduplikowanych dokumentów. Po zaznaczeniu tego pola wyboru lista dostępnych klas kryteriów zawiera klasy odnoszące się do pól treści. Po anulowaniu zaznaczenia tego pola wyboru lista dostępnych klas kryteriów zawiera klasy odnoszące się do pól metadanych.

5. Wybierz klasę kryteriów i kliknij przycisk **OK**.

Dodane pole zostanie wyświetlone na stronie Odwzorowywanie pól na klasy kryteriów. Istnieje możliwość wyboru opcji służącej do edycji klasy kryteriów i konfigurowania innych współczynników kryteriów w celu określenia wyników dokumentów zawierających to pole.

6. Aby zmienić opcję dotyczącą stosowania pola do wykrywania zduplikowanych dokumentów lub aby odwzorować to pole na inną klasę kryteriów, kliknij opcję  **Edytuj** (nie można edytować pól zarezerwowanych do użycia przez wyszukiwanie korporacyjne).



7. Aby usunąć pole z klasy kryteriów, kliknij opcję  **Usuń** (nie można usuwać pól zarezerwowanych do użycia przez wyszukiwanie korporacyjne).
8. Aby zastosować zmiany do poindeksowanych już wcześniej dokumentów, należy je ponownie przeszukać, przeanalizować i poindeksować.

#### **Pojęcia pokrewne**

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

## **Konfigurowanie współczynników kryteriów dla klas kryteriów**

Współczynniki kryteriów konfigurowane dla klas kryteriów reprezentują oszacowanie odzwierciedlające, jak ważna jest dla danego zapytania obecność wybranych pól w dokumentach wynikowych. Klasy kryteriów z wysokimi współczynnikami kryteriów mogą zwiększać ważność dokumentów wynikowych zawierających pola odwzorowane na daną klasę kryteriów.

#### **Zagadnienia wstępne**


Do konfigurowania współczynników kryteriów dla klas kryteriów wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora do danej kolekcji.

#### **Informacje o tym zadaniu**

System używa współczynników kryteriów skonfigurowanych dla klasy kryteriów, domyślnego wyniku oceny statycznej oraz innych czynników do obliczenia nowego wyniku dla zwracanych dokumentów zawierających pola odwzorowane na tę klasę kryteriów.

#### **Procedura**

Aby skonfigurować współczynniki kryteriów dla klas kryteriów:

1. Rozpocznij edycję kolekcji, wybierz stronę Analiza składni i kliknij przycisk **Odwzoruj pola na klasy kryteriów**.
2. Na stronie Odwzorowywanie pól na klasy kryteriów kliknij przycisk **Edytuj klasy kryteriów**.
3. Na stronie Klasy kryteriów znajdź klasę kryteriów do zmodyfikowania i kliknij opcję  **Edytuj**.
4. Na stronie Edycja klasy kryteriów określ nowe wartości dla wysokich i niskich współczynników kryteriów. Dla obu współczynników można podać te same wartości.
5. Kliknij przycisk **OK**.
6. Aby zmiany zostały zastosowane, przejdź do monitorowania serwerów wyszukiwania i wybierz ikony służące do zatrzymywania, a następnie restartowania procesów wyszukiwania. Przy wprowadzaniu zapytań przez użytkowników względna ważność dokumentów wynikowych zawierających pola odwzorowane na tę klasę kryteriów będzie określana przez nowe współczynniki kryteriów.

#### **Pojęcia pokrewne**

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.

## Domyślne wartości klas kryteriów

Wyszukiwanie korporacyjne udostępnia 16 klas kryteriów, za pomocą których można wpływać na oceny dokumentów w wynikach wyszukiwania.

Do obliczania wyników dla pól i tekstu pozbawionych dodatkowych definiujących je charakterystyk zarezerwowane zostały do użycia w wyszukiwaniu korporacyjnym następujące pola:

```
es_special_field.regular_text
es_special_field.default_field
```

Na klasy kryteriów używane przez pola zarezerwowane można odwzorowywać też inne pola, ale nie można edytować ani usuwać pól zarezerwowanych.

W przypadku wszystkich innych pól istnieje możliwość edytowania współczynników kryteriów używanych przez system do obliczania oceny dokumentu. Można również odwzorować dowolną liczbę pól na dowolną z klas kryteriów, w tym na klasy kryteriów używane przez pola zarezerwowane.

W poniższej tabeli przedstawiono listę nazw klas kryteriów, domyślnych współczynników kryteriów dla zapytań o niskiej wartości znajdowalności, domyślnych współczynników kryteriów dla zapytań o wysokiej wartości znajdowalności, a także nazw pól predefiniowanych odwzorowanych na klasy kryteriów w konfiguracji domyślnej.

Domyślne współczynniki kryteriów różnią się w zależności od metody oceny statycznej wybranej dla danej kolekcji podczas jej tworzenia. Dostępne opcje to brak oceny statycznej, ocena określana na podstawie liczby odsyłaczy do dokumentu (w przypadku źródeł WWW) oraz ocena określana na podstawie daty dokumentu.

Tabela 3. Domyślne wartości klas kryteriów

Domyślne niskie i wysokie współczynniki kryteriów				
Nazwa klasy kryteriów	Brak oceny statycznej	Odsyłacze do dokumentu	Data dokumentu	Odwzorowania pól predefiniowanych
Klasa treści A	Niski: 4 Wysoki: 2	Niski: 6 Wysoki: 1	Niski: 4 Wysoki: 2	es_special_field.regular_text
Klasa treści B	Niski: 5 Wysoki: 4	Niski: 7 Wysoki: 3	Niski: 5 Wysoki: 4	es_special_field.html_emphasized_text  Uwzględnia następujące elementy języka HTML: b, big, caption, dfn, em, h4, h5, h6, strong
Klasa treści C	Niski: 7 Wysoki: 4	Niski: 9 Wysoki: 3	Niski: 7 Wysoki: 4	es_special_field.html_headers  Uwzględnia następujące elementy języka HTML: h1, h2, h3
Klasa treści D	Niski: 2 Wysoki: 5	Niski: 1 Wysoki: 5	Niski: 2 Wysoki: 5	title
Klasa treści E	Niski: 1 Wysoki: 1	Niski: 5 Wysoki: 10	Niski: 1 Wysoki: 1	es_special_field.anchor
Klasa treści F	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.anchor_same_dir
Klasa treści G	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.anchor_same_host
Klasa treści H	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	

Tabela 3. Domyślne wartości klas kryteriów (kontynuacja)

Domyślne niskie i wysokie współczynniki kryteriów				
Nazwa klasy kryteriów	Brak oceny statycznej	Odsyłacze do dokumentu	Data dokumentu	Odwzorowania pól predefiniowanych
Klasa metadanych A	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.default_field
Klasa metadanych B	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	
Klasa metadanych C	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	
Klasa metadanych D	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	
Klasa metadanych E	Niski: 1 Wysoki: 1	Niski: 5 Wysoki: 1	Niski: 1 Wysoki: 1	keywords
Klasa metadanych F	Niski: 1 Wysoki: 1	Niski: 3 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.urlhost
Klasa metadanych G	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	es_special_field.urlpath
Klasa metadanych H	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	Niski: 1 Wysoki: 1	opis

#### Pojęcia pokrewne

“Ocenianie dokumentów w wyszukiwaniu korporacyjnym” na stronie 143

Gdy użytkownik przeszukuje kolekcję, procesy wyszukiwania zwracają wyniki najlepiej odpowiadające terminom i warunkom użytym w zapytaniu.



---

## Aplikacje wyszukujące dla wyszukiwania korporacyjnego

Aplikacja wyszukująca umożliwia wyszukiwanie w kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Liczba możliwych do utworzenia aplikacji wyszukujących nie jest ograniczona, a pojedyncza aplikacja wyszukująca pozwala przeszukać dowolną liczbę kolekcji i źródeł zewnętrznych.

### Przykładowa aplikacja wyszukująca

Przykładowa aplikacja wyszukująca zawiera wiele funkcji wyszukiwania i pobierania dostępnych dla wyszukiwania korporacyjnego. Aplikacja przykładowa jest także działającym przykładem wykorzystania funkcji interfejsu API wyszukiwania i indeksowania (Search and Index API, SI-API) firmy IBM do konstruowania interaktywnych, niestandardowych aplikacji wyszukujących, które odpowiadają celom przedsiębiorstwa.

Jeśli nie zmieniono właściwości w domyślnym pliku konfiguracyjnym, przykładowa aplikacja wyszukująca pozwala wyszukiwać we wszystkich aktywnych kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Dzięki temu przykładowa aplikacja wyszukująca może służyć do testowania nowych kolekcji i źródeł zewnętrznych przed udostępnieniem ich użytkownikom.

Przykładowa aplikacja wyszukująca jest automatycznie powiązana ze wszystkimi kolekcjami i źródłami zewnętrznymi. W środowisku produkcyjnym o tym, które aplikacje obsługują wyszukiwanie w poszczególnych kolekcjach, decydują administratorzy wyszukiwania korporacyjnego.

### Niestandardowe aplikacje wyszukujące

Aplikacje wyszukujące mogą być uruchamiane jako samodzielne aplikacje WWW w środowisku serwera IBM WebSphere Application Server albo jako portlety w środowisku IBM WebSphere Portal. Korzystając z funkcji API wyszukiwania i indeksowania, można projektować aplikacje wyszukujące, które będą bezproblemowo działały w obu środowiskach, podobnie jak przykładowa aplikacja wyszukująca.

#### Pojęcia pokrewne

“Indeksowane opcje wyszukiwania dokumentów” na stronie 121

Podczas konfigurowania opcji wyszukiwania przeszukanych danych lub źródeł zewnętrznych (albo w przypadku odwzorowania elementów metadanych XML i HTML na pola wyszukiwania) należy określić sposób wyszukiwania dokumentów i ich prezentacji w wynikach wyszukiwania.

“Zabezpieczenia przy użyciu identyfikatorów aplikacji wyszukującej” na stronie 183  
Zabezpieczenia na poziomie kolekcji są realizowane przez określenie, które aplikacje wyszukujące mogą wyszukiwać w poszczególnych kolekcjach lub źródłach zewnętrznych.

“Zabezpieczenia na poziomie dokumentu” na stronie 184

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentu. Mechanizm zabezpieczeń na poziomie dokumentu jest odpowiedzialny za to, aby użytkownicy wyszukujący w kolekcji mieli dostęp tylko do tych dokumentów, względem których przydzielono im stosowne uprawnienia.

“Zabezpieczenia na poziomie dokumentu przy użyciu mechanizmu wyszukiwania środowiska Portal” na stronie 197

Gdy użytkownicy przeszukują kolekcje wyszukiwania korporacyjnego, mechanizm

wyszukiwania środowiska IBM WebSphere Portal może zostać wykorzystany do wymuszenia zabezpieczeń na poziomie dokumentu.

"Search and index API overview" (Przegląd interfejsu API wyszukiwania i indeksowania) w podręczniku "Programming Guide and API Reference for Enterprise Search"

"Składnia zapytania" w podręczniku "Programming Guide and API Reference for Enterprise Search"

---

## Przypisywanie aplikacji wyszukiwujących do kolekcji

Przed rozpoczęciem korzystania z nowej aplikacji wyszukiwującej należy powiązać ją z kolekcjami, w których może ona wyszukiwać dokumenty.

### Zagadnienia wstępne

Aby powiązać aplikacje wyszukiwujące z kolekcjami, w których mogą prowadzić wyszukiwanie, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Procedura

Aby powiązać aplikację wyszukiwującą z jedną lub z większą liczbą kolekcji, należy wykonać poniższe czynności:

1. Kliknij opcję **Zabezpieczenia** na pasku narzędzi konsoli administracyjnej.
2. Na stronie Aplikacje wyszukiwujące kliknij opcję **Konfiguruj aplikacje wyszukiwujące**.
3. Na stronie Konfigurowanie aplikacji wyszukiwujących kliknij opcję **Dodaj aplikację wyszukiwującą**.
4. Wpisz nazwę aplikacji wyszukiwującej.
5. Wybierz kolekcje, w których aplikacja będzie mogła wyszukiwać:
  - Kliknij opcję **Wszystkie kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukiwująca ma mieć dostęp do wszystkich kolekcji dodanych do systemu.
  - Kliknij opcję **Konkretne kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukiwująca ma mieć dostęp tylko do określonych kolekcji.

Po wybraniu tej opcji zostanie wyświetlona lista nazw kolekcji i źródeł zewnętrznych. Zaznacz pole wyboru **Wybierz** dla każdej kolekcji, w której aplikacja będzie mogła wyszukiwać.
6. Kliknij przycisk **OK**.

---

## Funkcje przykładowej aplikacji wyszukiwującej

Przykładowa aplikacja wyszukiwująca przeznaczona do wyszukiwania korporacyjnego zawiera większość funkcji wyszukiwujących, jakie można wbudować w tworzone aplikacje.

Za pomocą tej aplikacji można wyszukiwać w jednej, kilku lub wszystkich kolekcjach i źródłach zewnętrznych jednocześnie. O ile nie zostaną zmodyfikowane domyślne właściwości aplikacji, można jej użyć do wyszukiwania we wszystkich kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego.

### Funkcje operujące na zapytaniach

Te funkcje umożliwiają:

- Definiowanie prostych zapytań pełnotekstowych.

- Definiowanie bardziej złożonych zapytań w celu poprawienia dokładności wyników wyszukiwania. Przykładem może być wyszukiwanie konkretnych pól w dokumencie lub użycie składni zapytania do wyszukania dokumentów, które zawierają konkretne słowa lub frazy albo które ich nie zawierają.
- Określenie, które kolekcje i źródła zewnętrzne należy przeszukać.
- Przeszukanie konkretnych typów źródeł lub wszystkich typów źródeł.
- Wyszukiwanie dokumentów określonych typów. Na przykład można szukać tylko dokumenty programu Microsoft Word lub tylko dokumenty w formacie PDF.
- Określanie języka, w jakim definiowane są warunki zapytania. Można również określić języki dokumentów, które należy szukać.
- Wyszukiwanie określonych podzbiorów kolekcji. Przykładem może być aplikacja wyszukiwująca, która ogranicza zawartość widoku do predefiniowanej grupy dokumentów (zasięg) albo umożliwia wprowadzanie zapytań wyszukujących tylko dokumenty należące do pewnej kategorii.
- Rozszerzanie zapytania o synonimy terminów z zapytania. Jeśli z kolekcją skojarzony jest słownik synonimów, w wynikach zapytania zwrócone zostaną dokumenty zawierające synonimy terminów użytych w zapytaniu.

## Funkcje operujące na wynikach wyszukiwania

Te funkcje umożliwiają:

- Wyświetlenie wyników wyszukiwania zgodnych z terminami użytymi w zapytaniu.
- Określanie liczby dokumentów wynikowych wyświetlanych na każdej stronie, a także przeglądanie tabeli wynikowej w obie strony.
- Ukrywanie i wyświetlanie szczegółowych informacji o dokumentach wynikowych. Na przykład można przejrzeć krótkie opisy dokumentów lub szczegółowe informacje, takie jak nazwy pól w każdym dokumencie wynikowym.
- Zwijanie dokumentów z jednego źródła. Na przykład, jeśli z jednego źródła zwracanych jest 100 dokumentów, dwa najbardziej trafne dokumenty będą w tabeli wynikowej zgrupowane. Pozostałe 98 dokumentów można wyświetlić, wybierając opcję przeglądania większej liczby dokumentów z tego samego źródła.
- Sortowanie dokumentów według trafności lub według daty w kolejności malejącej lub rosnącej.
- Uzyskiwanie propozycji poprawienia pisowni słów, które mogły zostać błędnie zapisane w zapytaniu.
- Wyświetlanie informacji o kategoriach, do których należy dokument wynikowy (jeśli w kolekcji są stosowane kategorie), oraz przeglądanie wyłącznie dokumentów należących do wybranej kategorii.
- Określanie dodatkowych terminów zapytania przeznaczonych do wyszukiwania w ramach wyników wyszukiwania.

## Funkcje pobierania dokumentów

Te funkcje umożliwiają:

- Pobieranie dokumentów przez kliknięcie identyfikatora URI dokumentu. Jeśli są dostępne aplikacje klienckie, można również przejrzeć dokument wynikowy za pomocą rodzimej przeglądarki. Jeśli na przykład została zainstalowana aplikacja kliencka programu Notes, można jej użyć do przeglądania dokumentów z bazy danych Lotus Notes.  
Jeśli w przeszukiwaczu zostały skonfigurowane zabezpieczenia na poziomie dokumentu, pobierać dokumenty będą mogli tylko użytkownicy mający uprawnienia dostępu do zabezpieczonych treści.

- Pobieranie dokumentów przez kliknięcie szybkich odsyłaczy. Szybki odsyłacz wiąże słowa kluczowe z identyfikatorami URI. Jeśli w zapytaniu użyto określonych słów kluczowych, powiązane z nimi identyfikatory URI (które predefiniowano jako w wysokim stopniu adekwatne dla tych słów kluczowych) są wyświetlane w górnej części listy wyników wyszukiwania.

---

## Właściwości przykładowej aplikacji wyszukującej

Plik `config.properties` przykładowej aplikacji wyszukującej można edytować w celu określenia opcji dotyczących używanego środowiska, zmiany wyglądu aplikacji, a także określenia opcji dostępnych dla użytkowników po uruchomieniu przez nich aplikacji wyszukującej.

### Parametry środowiska

Istnieje możliwość określania opcji sterujących działaniem aplikacji wyszukującej.

#### **applicationName**

Określa nazwę poprawnej aplikacji wyszukującej. Wartością domyślną jest nazwa `Default`.

Wartość domyślną można zmienić, aby jako domyślna aplikacja używana była inna aplikacja wyszukująca.

**Wskazówka:** Jeśli nazwa aplikacji to `Default`, przykładowej aplikacji wyszukującej można używać do wyszukiwania danych we wszystkich kolekcjach i źródłach zewnętrznych w pojedynczym zapytaniu.

#### **hostname**

Określa pełną nazwę hosta odpowiadającą serwerowi WWW skonfigurowanemu do obsługi używanej instancji serwera WebSphere Application Server. Wartością domyślną jest `localhost`.

Aby mieć pewność, że aplikacja wyszukująca działa poprawnie, wartość domyślną można zmienić na pełną nazwę hosta, do używania którego skonfigurowany jest serwer WebSphere Application Server. Na przykład, jeśli nazwą hosta komputera jest `MojKomputer`, a serwer WWW ma korzystać z nazwy hosta `www.ibm.com`, należy podać nazwę `www.ibm.com`.

**port** Określa numer portu serwera WWW skonfigurowanego do obsługi używanej instancji serwera WebSphere Application Server. Wartością domyślną jest `80`. Ustawienie to należy zmienić tylko w przypadku, gdy zmieniono port używany przez instancję serwera WWW.

#### **timeout**

Określa liczbę sekund oczekiwania na odpowiedź z serwera wyszukiwania, zanim żądanie wyszukiwania przekroczy limit czasu. Musi to być liczba całkowita (np. `60`, a nie `60,5`). Jeśli limit czasu nie zostanie określony, wartością domyślną jest `30` sekund.

#### **username**

Określa nazwę użytkownika umożliwiającą systemowi wyszukiwania korporacyjnego wymuszanie uwierzytelniania użytkowników, a także uwierzytelnianie użytkowników przy użyciu serwera WebSphere Application Server podczas wprowadzania przez użytkowników żądań wyszukiwania. To pole jest używane tylko wtedy, gdy włączono zabezpieczenia globalne na serwerze WebSphere Application Server.

#### **password**

Określa hasło dla podanej nazwy użytkownika.



**filter** Określa klasę, która ma być używana do pobierania dokumentów wymienionych w wynikach wyszukiwania. Domyślna klasa to `com.ibm.es.api.filters.SetDocumentURIFilterFetch`. Wartość tę należy zmienić tylko wówczas, gdy dysponuje się niestandardową klasą, która ma zostać użyta do pobierania dokumentów.

#### **logging.level**

Określa ilość rejestrowanych szczegółów:

**OFF** Nie są rejestrowane żadne komunikaty.

**INFO** Rejestrowane są komunikaty informacyjne.

**FINE** Rejestrowane są komunikaty o błędach.

**ALL** Rejestrowane są wszystkie komunikaty.

## **Ikony typu źródła**

Istnieje możliwość dostosowania obrazów reprezentujących typ źródła danych, do którego należy dokument wynikowy. Poniższe ikony typu źródła identyfikujące przeszukiwacze i źródła zewnętrzne obsługiwane po zainstalowaniu produktu WebSphere Information Integrator OmniFind Edition są predefiniowane w pliku `config.properties`.



#### **documentSource.vbr.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Content Edition. Domyślna ikona to `/images/sourceVBR.gif`.



#### **documentSource.db2.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz DB2. Domyślna ikona to `/images/sourceDB2.gif`.



#### **documentSource.cm.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz DB2 Content Manager. Domyślna ikona to `/images/sourceCM.gif`.



#### **documentSource.dominodoc.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Domino Document Manager. Domyślna ikona to `/images/sourceDominoDoc.gif`.



#### **documentSource.exchange.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz Exchange Server. Domyślna ikona to `/images/sourceExchange.gif`.



#### **documentSource.nntp.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez przeszukiwacz NNTP. Domyślna ikona to `/images/sourceNNTP.gif`.

**documentSource.notes.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez wyszukiwacz Notes. Domyślna ikona to /images/sourceNotes.gif.

**documentSource.quickplace.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez wyszukiwacz QuickPlace. Domyślna ikona to /images/sourceWorkplace.gif.

**documentSource.unixfs.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez wyszukiwacz system plików UNIX. Domyślna ikona to /images/sourceUnixFS.gif.

**documentSource.web.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez wyszukiwacz sieci WWW. Domyślna ikona to /images/sourceWeb.gif.

**documentSource.wps.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez wyszukiwacz WebSphere Portal. Domyślna ikona to /images/sourceWPS.gif.

**documentSource.winfs.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument był przeszukiwany przez wyszukiwacz system plików Windows. Domyślna ikona to /images/sourceWindowsFS.gif.

**documentSource.ldap.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument należy do zewnętrznego źródła utworzonego dla serwera LDAP. Domyślna ikona to /images/sourceLDAP.gif.

**documentSource.jdbc.icon**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dokument należy do zewnętrznego źródła utworzonego dla tabeli bazy danych JDBC (Java Database Connectivity). Domyślna ikona to /images/sourceJDBC.gif.

## Ikony przeglądarki klienckiej

Dokumenty wynikowe mogą być wyświetlane w przeglądarce WWW. Jeśli dostępna jest rodzima aplikacja kliencka dla danego typu dokumentu wynikowego, dokument taki może zostać również wyświetlony w przeglądarce klienckiej. Obrazy reprezentujące typ przeglądarki klienckiej używanej do wyświetlania dokumentu można dostosowywać. W poniższym przykładzie ikona Lotus Notes służy do wskazania, że dokument może być wyświetlany w aplikacji klienckiej Lotus Notes:

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

W wynikach wyszukiwania ikona oraz odsyłacz do aplikacji klienckiej są wyświetlane następująco:



## Przeglądarka kliencka

### Pola dokumentu

W przypadku typów źródeł danych zawierających pola można określić, które pola będą wyświetlane w dokumentach wynikowych.

#### **fields.przedrostek URI=lista\_nazw\_pól\_oddzielonych\_znakami\_spacji**

Znaczenie znaku dwukropka (:) w przedrostku URI należy zmienić, poprzedzając go znakiem odwrotnego ukośnika (\). Aby kontynuować listę nazw pól w kolejnym wierszu, poprzedni wiersz należy zakończyć znakiem odwrotnego ukośnika (\). Na przykład:

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

### Ikony pól

W przypadku typów źródeł danych i dokumentów zawierających pola można dostosować obrazy reprezentujące pola. W pliku config.properties predefiniowane są następujące ikony pól.



#### **field.icon.databasetitle**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera tytuł dokumentu. Domyślna ikona to /images/notesdb.gif.



#### **field.icon.databasename**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera nazwę bazy danych, do której należy dokument. Domyślna ikona to /images/db2.gif.



#### **field.icon.tablename**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera nazwę tabeli, do której należy dokument. Domyślna ikona to /images/table.gif.



#### **field.icon.directory**

Określa nazwę i ścieżkę do pliku graficznego wskazującego, że dane pole zawiera nazwę katalogu, do którego należy dokument. Domyślna ikona to /images/closedFolder.gif.

### Domyślna ikona pola

Dla pól wyświetlanych w wynikach wyszukiwania można określić obraz używany, gdy nie skonfigurowano żadnych ikon pól. W pliku config.properties predefiniowana jest następująca domyślna ikona pola.



#### **field.defaultIcon**

Określa nazwę i ścieżkę do pliku graficznego będącego domyślną ikoną dla pól w wynikach wyszukiwania. Domyślna ikona to /images/database.gif.

## Pola daty

Istnieje możliwość określenia, które pola są polami daty. Pola, których nazwy zostaną określone w tym miejscu, są formatowane w wynikach wyszukiwania tak jak daty. Format daty jest zgodny z ustawieniami narodowymi przeglądarki WWW.

**date.fields=***lista\_nazw\_pól\_oddzielonych\_znakami\_spacji*

Aby kontynuować listę nazw pól w kolejnym wierszu, poprzedni wiersz należy zakończyć znakiem odwrotnego ukośnika (\).

Przykład:

```
date.fields=modifieddate createddate
```

## Tytuły dokumentów

Istnieje możliwość określenia alternatywnych tytułów dokumentów przez zastąpienie tekstu tytułu bardziej zrozumiałymi danymi (pozwala to *wyklarować* tytuły). Na przykład zamiast wyświetlania tytułów dokumentów w postaci niewiele mówiącej etykiety Slajd 1, można określić, że ma ona zostać zastąpiona w wynikach wyszukiwania (zamiast niej do zidentyfikowania dokumentu wynikowego można użyć bardziej zrozumiałego pola, takiego jak nazwa pliku).

Istnieje również możliwość określenia alternatywnych tytułów dokumentów przez usunięcie z tytułów dokumentów słów zbędnych (co pozwala *skracać* tytuły). Jeśli na przykład wiele dokumentów wynikowych rozpoczyna się tekstem Microsoft Word -, można poprawić czytelność wyników wyszukiwania, usuwając powtarzający się tekst.

**titles.clean=***lista\_tytułów\_oddzielonych\_przecinkami*

**titles.truncatePrefix=***lista\_przedrostków\_oddzielonych\_przecinkami*

Pozycje na liście rozdzielanej przecinkami mogą zawierać znaki spacji oraz inne znaki za wyjątkiem przecinka. Aby kontynuować listę w kolejnym wierszu, poprzedni wiersz należy zakończyć znakiem odwrotnego ukośnika (\).

Na przykład:

```
titles.clean=Slajd 1, Układ 1, Szablon prezentacji IBM Software Group, \
bez tytułu, Dokument bez tytułu, Prezentacja programu PowerPoint, \
(brak tytułu tej strony)
```

```
titles.truncatePrefix=Microsoft Word -, Microsoft PowerPoint -
```

## Wartości domyślne preferencji użytkownika

Istnieje możliwość określenia wartości domyślnych dla strony Preferencje w aplikacji wyszukiwającej. Jeśli użytkownik zmieni te preferencje, nowe wartości będą obowiązywać tylko na czas bieżącej sesji użytkownika. W pliku config.properties predefiniowane są następujące preferencje.

**preferences.resultsRange=10**

Określa, że na każdej stronie wyników wyszukiwania może znaleźć się lista 10 dokumentów wynikowych.

**preferences.siteCollapsing=Yes**

Określa, że jeśli w konsoli administracyjnej wyszukiwania korporacyjnego skonfigurowano reguły zwijania serwisu, wówczas identyfikatory URI zgodne z regułą przedrostka URI mają być zwijane w wynikach wyszukiwania.

**preferences.spellCorrections=Yes**

Określa, że gdy użytkownik wprowadzi zapytanie zawierające potencjalnie niepoprawnie wpisane słowo, wyświetlane mają być proponowane poprawki dotyczące pisowni.

### **preferences.extendedHighlighting=No**

Określa, że oprócz pola podsumowania dokumentu terminy użyte w zapytaniu nie będą wyróżniane w polach dodatkowych (takich jak tytuł dokumentu).

## **Domyślne kolekcje i źródła zewnętrzne**

Można określić kolekcje i źródła zewnętrzne wybrane wstępnie na stronach Preferencje oraz Wyszukiwanie zaawansowane. Użytkownicy mogą edytować zestaw domyślny, aby wyszukiwać w mniejszej liczbie kolekcji i źródeł zewnętrznych niż udostępniona domyślnie. Jeśli zestaw kolekcji i źródeł zewnętrznych zostanie tu ograniczony, użytkownicy mogą wybrać dowolną kolekcję lub źródło zewnętrzne dostępne dla aplikacji wyszukującej podczas modyfikowania swoich preferencji lub opcji wyszukiwania zaawansowanego.

### **preferences.defaultCollections=\***

#### **preferences.defaultCollections=lista\_identyfikatorów\_kolekcji\_oddzielonych\_znakami\_spacji**

Określenie znaku gwiazdki (\*) umożliwia wyszukiwanie we wszystkich kolekcjach i źródłach zewnętrznych (kolekcje i źródła zewnętrzne muszą być powiązane z aplikacją wyszukującą w konsoli administracyjnej wyszukiwania korporacyjnego). Jest to domyślne ustawienie w pliku config.properties.

Aby ograniczyć elementy, w których użytkownicy mogą wyszukiwać bez modyfikowania swoich preferencji lub opcji wyszukiwania zaawansowanego, należy określić identyfikatory kolekcji dla kolekcji i źródeł zewnętrznych, które mają być domyślnie udostępnione użytkownikom do wyszukiwania.

Na przykład:

```
preferences.defaultCollections=*\npreferences.defaultCollections=id_kolekcji1 id_kolekcji2
```

## **Informacje dodatkowe w wynikach wyszukiwania**

Ilość informacji podawanych w wynikach wyszukiwania można dostosować. Można również kontrolować możliwość filtrowania wyników wyszukiwania przez użytkowników. Poniższe ustawienia to domyślne ustawienia w pliku config.properties.

### **refreshButton.show=false**

Steruje wyświetlaniem przycisku **Odśwież** oraz możliwością odświeżania aplikacji wyszukującej. Po ustawieniu tej opcji na wartość true użytkownicy mogą odświeżać listę kolekcji i źródeł zewnętrznych dostępnych do wyszukiwania (na przykład gdy aplikacja wyszukująca została powiązana z dodatkowymi kolekcjami lub źródłami zewnętrznymi w konsoli administracyjnej wyszukiwania korporacyjnego).

Przycisk **Odśwież** można wyświetlić na przykład podczas testowania zmian wprowadzanych w pliku config.properties. Po zapisaniu wprowadzonych zmian można wówczas kliknąć przycisk **Odśwież** w celu sprawdzenia, jak wpłynęły one na aplikację wyszukującą. Bez widocznego przycisku **Odśwież** konieczne jest zrestartowanie aplikacji korporacyjnej ESSearchApplication na serwerze WebSphere Application Server, aby zmiany odniosły skutek.

Jeśli żadne kolekcje i źródła zewnętrzne nie są dostępne do wyszukiwania (na przykład jeśli określono niewłaściwą nazwę hosta, nie uruchomiono serwerów wyszukiwania lub nie uruchomiono aplikacji korporacyjnej ESSearchServer na serwerze WebSphere Application Server), przycisk **Odśwież** jest wyświetlany automatycznie, aby ułatwić rozwiązanie zaistniałego problemu.

### **extraMessages.show=false**

Steruje wyświetlaniem obszaru u dołu wyników wyszukiwania, w którym wyświetlane są ostrzeżenia i komunikaty informacyjne. Komunikaty o błędach są

zawsze wyświetlane w górnej części strony. Aby wyświetlane były dodatkowe komunikaty, należy ustawić tę opcję na wartość true.

#### **builtQueryString.show=false**

Steruje wyświetlaniem w pełni rozwiniętej składni zapytania w obszarze poprzedzającym listę dokumentów wynikowych. Aby przetworzone właśnie zapytanie wyświetlane było w pełnej postaci, należy ustawić tę opcję na wartość true.

#### **refineResults.show=true**

Steruje możliwością zawężania przez użytkowników wyników wyszukiwania przez określanie dodatkowych terminów dla zapytania. Po ustawieniu tej opcji na wartość true w dolnej części strony wyników wyszukiwania wyświetlane jest pole zapytania z etykietą **Szukaj w wynikach**.

#### **filter.showOnTwoLines=true**

Określa, czy opcje filtrowania wyników według typu źródła i typu pliku mają być wyświetlane w wynikach wyszukiwania w jednym czy też w dwóch wierszach (podczas przeglądania wyników wyszukiwania użytkownicy mogą wybrać typ źródła oraz typ pliku, aby wyświetlić tylko dokumenty wynikowe zgodne z wybranymi filtrami).

Aby maksymalnie zwiększyć ilość miejsca dostępnego do wyświetlania wyników wyszukiwania, należy ustawić tę właściwość na wartość false. Aby zwiększyć czytelność filtrów, zwłaszcza jeśli dostępne filtry są dłuższe niż jeden wiersz, można ustawić tę właściwość na wartość true, dzięki czemu każdy filtr będzie wyświetlany w osobnym wierszu.

## **Niestandardowy baner i logo**

Istnieje możliwość dostosowania obrazów wyświetlanych w obszarze banera w górnej części aplikacji wyszukującej. Można na przykład zastąpić domyślne obrazy produktu WebSphere II OmniFind Edition obrazami związanymi z marką firmy. Jeśli baner nie ma być wyświetlany, jeden lub oba poniższe wiersze należy przekształcić w komentarz. Właściwość `banner.icon` identyfikuje grafikę wyświetlaną po lewej stronie obszaru banera. Właściwość `banner2.icon` identyfikuje grafikę wyświetlaną po prawej stronie obszaru banera.

```
banner.icon=/images/WS_II_OFEdition.gif
banner2.icon=/images/WS_II_mosaic.gif
```

## **Niestandardowy obraz tła**

Istnieje możliwość dostosowania obrazów wyświetlanych w tle stron aplikacji wyszukującej. Można na przykład zastąpić domyślne obrazy produktu WebSphere II OmniFind Edition obrazami związanymi z marką firmy. Jeśli obraz tła ma nie być wyświetlany na stronach, jeden lub większą liczbę poniższych wierszy należy przekształcić w komentarz.

```
search.backgroundImage=/images/IIOF_search.gif
preferences.backgroundImage=/images/IIOF_options.gif
advanced.backgroundImage=/images/IIOF_advanced.gif
browse.backgroundImage=/images/IIOF_tree.gif
myProfile.backgroundImage=/images/IIOF_profile.gif
logoff.backgroundImage=/images/IIOF_logout.gif
```

## **Odsyłacze**

Właściwości w sekcji pliku `config.properties` dotyczącej odsyłaczy umożliwiają wyświetlanie nazw stron aplikacji wyszukującej na każdej stronie jako odsyłaczy, a nie na pasku narzędzi ani na stronach z kartami. Wyświetlanie odsyłaczy jest przydatne, gdy aplikacja wyszukująca działa jako portlet i pożądane jest zminimalizowanie obszaru używanego do wyświetlania aplikacji wyszukującej na stronie portalu.

Jeśli użytkownik preferuje nawigowanie po aplikacji wyszukującej przez wybór opcji na pasku narzędzi i na stronach z kartami, należy przekształcić te wiersze w komentarz.

## Karty wyszukiwania

Właściwości w sekcji pliku `config.properties` dotyczącej kart wyszukiwania określają nazwy stron JSP (Java Server Pages) używanych jako strony z kartami w widoku Wyszukiwania aplikacji wyszukującej (Wyszukiwanie, Wyszukiwanie zaawansowane i Drzewo kategorii). Nie należy edytować tych stron, jeśli nie ma się doświadczenia w programowaniu w języku Java i z obsługą stron JSP.

Przykładowe sposoby dostosowania tego obszaru to:

- Kierowanie aplikacji wyszukującej do niestandardowych stron JSP udostępniających inny wygląd stron z kartami.
- Przekształcanie w komentarz pozycji drzewa kategorii. Na przykład jeśli nie zostaną skonfigurowane kategorie dla kolekcji, nie ma potrzeby wyświetlania strony Drzewo kategorii w aplikacji wyszukującej.
- Kopiowanie pozycji dla stron z kartami do sekcji pliku `config.properties` dotyczącej pasków narzędzi i przekształcanie tych wierszy w komentarz. Na przykład może być przydatne wyświetlenie samego paska narzędzi bez wyświetlania stron z kartami.

## Paski narzędzi

Właściwości w sekcji pliku `config.properties` dotyczącej pasków narzędzi określają nazwy stron JSP (Java Server Pages) używanych jako paski narzędzi w aplikacji wyszukującej. Nie należy edytować tych stron, jeśli nie ma się doświadczenia w programowaniu w języku Java i z obsługą stron JSP.

Przykładowe sposoby dostosowania tego obszaru to:

- Kierowanie aplikacji wyszukującej do niestandardowych stron JSP udostępniających inny wygląd paska narzędzi.
- Przekształcanie w komentarz pozycji paska narzędzi dla elementów, które nie mają być wyświetlane. Na przykład na pasku narzędzi może nie być potrzebny odsyłacz do strony Informacje.
- Przenoszenie funkcji wyświetlania strony Wyszukiwanie zaawansowane z sekcji pliku `config.properties` dotyczącej kart wyszukiwania, tak aby opcja ta była dostępna jedynie z poziomu paska narzędzi.

## Rozszerzone etykiety typu dokumentu

Istnieje możliwość zwiększenia czytelności filtra typu dokumentu przez odwzorowanie rzeczywistych nazw typów dokumentów na terminy, które są zwięźlejsze i bardziej adekwatne. Typy dokumentów dostępne dla aplikacji wyszukujących zdefiniowane są w klasie `AvailableDocumentTypes` interfejsu API wyszukiwania i indeksowania (SI-API). Dla wygody dostępne typy dokumentów zostały również wymienione na końcu pliku `config.properties`.

**`documentType.etykieta=lista_typów_dokumentów_oddzielonych_znakami_spacji`**

Określa nazwę wyświetlaną w wynikach wyszukiwania w wierszu filtra typu dokumentu oraz listę rzeczywistych typów dokumentów, które mają być wyświetlane po wybraniu tego filtra przez użytkownika.

Na przykład można określić etykietę `html` i odwzorować na tę nazwę rozszerzenia plików i typu MIME różnych dokumentów WWW. Po kliknięciu przez użytkownika

pozycji **html** w celu odfiltrowania wyników wyszukiwania wyświetlone zostaną tylko dokumenty z określonymi rozszerzeniami i typami MIME.

W pliku `config.properties` predefiniowane są następujące odwzorowania typów dokumentów:

```
documentType.html=shtml text/html html xhtml htm
documentType.doc=doc application/msword
documentType.ppt=application/mspowerpoint ppt
documentType.xls=xls application/x-excel application/msexcel \
application/x-msexcel application/excel application/vnd.ms-excel
documentType.xml=xml text/xml
documentType.txt=txt text/plain
documentType.pdf=pdf application/pdf
```

## Filtry niestandardowe

Do filtrowania wyświetlanych wyników wyszukiwania można określić zapytania niestandardowe.

### **filterCustom.etykieta=terminy\_dla\_zapytania**

Określa nazwę wyświetlaną w wynikach wyszukiwania w wierszu filtru niestandardowego oraz zapytanie zawężające wyniki wyszukiwania po wybraniu tego filtru przez użytkownika (podczas przeglądania wyników wyszukiwania użytkownicy mogą wybrać filtr niestandardowy, aby wyświetlić tylko dokumenty wynikowe zgodne z predefiniowanym zapytaniem).

W poniższym przykładzie wyniki wyszukiwania są filtrowane w celu wyświetlenia tylko dokumentów należących do bazy danych kadr (ka):

```
filterCustom.HR_database_only=databasename::ka
```

Po kliknięciu przez użytkownika pozycji **HR\_database\_only** w celu odfiltrowania wyników wyszukiwania przetwarzane jest zapytanie `databasename::ka`. Podczas wyświetlenia wyników wyszukiwania wymienione zostaną tylko dokumenty z bazy danych ka.

W pliku `config.properties` udostępnionych jest kilka przykładowych filtrów niestandardowych przekształconych w komentarz.

---

## Edytowanie właściwości przykładowej aplikacji wyszukiującej

Przykładowa aplikacja wyszukiująca dla wyszukiwania korporacyjnego umożliwia wyszukiwanie we wszystkich aktywnych kolekcjach i źródłach zewnętrznych w systemie. Edytując plik właściwości, można określić opcje dotyczące środowiska serwera WWW, użyć innej aplikacji wyszukiującej jako aplikacji domyślnej lub sterować wyświetlaniem opcji podczas uruchamiania aplikacji wyszukiującej.

### Informacje o tym zadaniu

Program instalacyjny instaluje przykładową aplikację wyszukiującą dla systemu wyszukiwania korporacyjnego na serwerze IBM WebSphere Application Server dla serwerów wyszukiwania korporacyjnego. Aby skonfigurować aplikację wyszukiującą, należy wyedytować plik właściwości `config.properties` instalowany wraz z aplikacją.

Aby zmiany zostały wprowadzone, zatrzymaj i zrestartuj aplikację korporacyjną `ESSearchApplication` na serwerze WebSphere Application Server.

### Procedura



Aby zmienić właściwości przykładowej aplikacji wyszukiwującej, należy wykonać następujące czynności:

1. Zmodyfikuj plik `config.properties` przy użyciu standardowego edytora tekstu.  
Plik `config.properties` jest instalowany w następującym miejscu (`ES_INSTALL_ROOT` to katalog instalacyjny programu WebSphere II OmniFind Edition na serwerze wyszukiwania):  

```
ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/
ESSearchApplication.war/WEB-INF/config.properties
```
2. Zmień właściwości w celu określenia informacji o środowisku serwera WWW i preferencjach wyszukiwania, a następnie zapisz i zamknij plik. (Znak krzyżyka (#) w pliku oznacza wiersz komentarza).
3. Zatrzymaj i zrestartuj aplikację `ESSearchApplication`:
  - a. Na serwerze wyszukiwania uruchom konsolę administracyjną serwera WebSphere Application Server.  
Konsolę administracyjną można otworzyć na kilka sposobów:
    - W systemie Windows należy użyć menu **Start**, aby wybrać odpowiedni program.
    - W przypadku serwera WebSphere Application Server w wersji 5 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/admin`, gdzie `nazwa_hosta` jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a `port` to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9090.
    - W przypadku serwera WebSphere Application Server w wersji 6 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/ibm/console`, gdzie `nazwa_hosta` jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a `port` to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9060.
  - b. Gdy pojawi się prośba o wprowadzenie identyfikatora i hasła użytkownika, wprowadź nazwę i hasło zapisane w rejestrze użytkowników serwera WebSphere Application Server. Jeśli rejestr użytkowników jest obsługiwany przez lokalny system operacyjny, można określić identyfikator i hasło administratora wyszukiwania korporacyjnego.
  - c. Po zalogowaniu się do konsoli administracyjnej kliknij opcję **Aplikacje**, a następnie opcję **Aplikacje korporacyjne**.
  - d. Zaznacz pole wyboru `ESSearchApplication` i kliknij opcję **Stop** (Zatrzymaj).
  - e. Gdy aplikacja zostanie zatrzymana, ponownie zaznacz pole wyboru i kliknij opcję **Start** (Uruchom).

---

## Uzyskiwanie dostępu do przykładowej aplikacji wyszukiwującej

Dostęp do przykładowej aplikacji wyszukiwującej odbywa się przez wprowadzenie adresu URL w przeglądarce WWW.

### Zagadnienia wstępne

Przykładową aplikację wyszukiwującą należy skonfigurować dla używanego środowiska serwera WWW.

### Informacje o tym zadaniu

Przykładowa aplikacja wyszukiwująca jest instalowana na serwerach wyszukiwania korporacyjnego. Z aplikacji tej można korzystać w udostępnionej postaci do testowania kolekcji i źródeł zewnętrznych przed ich udostępnieniem użytkownikom. Ponadto aplikacja przykładowa może służyć jako model przy tworzeniu własnych aplikacji wyszukiwujących.

## Procedura

Aby uruchomić przykładową aplikację wyszukującą, należy wykonać poniższe czynności:

1. Wpisz adres URL aplikacji wyszukującej w przeglądarce WWW. Na przykład:  
`http://SerwerWyszukiwania.com/ESSearchApplication/`  
`SerwerWyszukiwania.com` to nazwa hosta serwera wyszukiwania.  
Jeśli konfiguracja serwera WWW nie przewiduje korzystania z portu 80, wówczas trzeba podać poprawny numer portu. Na przykład:  
`http://SerwerWyszukiwania.com:9080/ESSearchApplication/`
2. Jeśli włączono zabezpieczenia serwera WebSphere Application Server, zaloguj się w aplikacji przy użyciu poprawnego identyfikatora i hasła użytkownika.  
Jeśli dowolna z kolekcji dostępnych dla aplikacji wyszukującej ma włączone zabezpieczenia i jeśli zabezpieczone kolekcje zawierają przeszukiwacze skonfigurowane do sprawdzania poprawności referencji użytkownika podczas przetwarzania zapytania, można skonfigurować profil użytkownika. Na stronie *Mój profil* określ referencje umożliwiające dostęp do zabezpieczonych domen. Następnie można wyszukiwać w tych domenach bez konieczności logowania się do nich.
3. Na stronie Wyszukiwanie wybierz kolekcje i źródła zewnętrzne, w których chcesz wyszukiwać, i wprowadź zapytanie.

---

## Włączanie zabezpieczeń dla przykładowej aplikacji wyszukującej

Jeśli włączono zabezpieczenia globalne na serwerze WebSphere Application Server, a przykładowa aplikacja wyszukująca ma być używana do wyszukiwania w zabezpieczonych kolekcjach, należy zmienić ustawienia konfiguracyjne w aplikacji przykładowej oraz na serwerze WebSphere Application Server.

### Zagadnienia wstępne

- Wymagana jest przynależność do roli administratorów wyszukiwania korporacyjnego.
- Należy włączyć zabezpieczenia globalne serwera WebSphere Application Server. Jeśli produkt WebSphere II OmniFind Edition zainstalowano w konfiguracji wieloserwerowej, należy włączyć zabezpieczenia globalne na serwerach wyszukiwania. Instrukcje dotyczące włączania zabezpieczeń globalnych można znaleźć w dokumentacji produktu WebSphere Application Server.
- Jeśli w konfiguracji serwera WebSphere Application Server jako rejestr użytkowników wybrano katalog LDAP, należy dodać identyfikator i hasło administratora wyszukiwania korporacyjnego (są to identyfikator i hasło zdefiniowane podczas instalacji produktu WebSphere II OmniFind Edition).

Na przykład, jeśli identyfikator administratora wyszukiwania korporacyjnego ma postać `adminUser`, to wpis w rejestrze LDAP może wyglądać następująco:

`uid=adminUser,ou=Employees,o=IBM,c=US`. Szczegółowe instrukcje zawiera dokumentacja serwera LDAP.

## Procedura

Aby włączyć zabezpieczenia w przykładowej aplikacji w systemie wyszukiwania korporacyjnego:

1. Zmodyfikuj właściwości aplikacji wyszukującej w konsoli administracyjnej:
  - a. Zaloguj się w konsoli administracyjnej wyszukiwania korporacyjnego jako użytkownik z uprawnieniami administratora wyszukiwania korporacyjnego.
  - b. Kliknij przycisk **Zabezpieczenia** na pasku narzędzi.
  - c. Na stronie Aplikacje wyszukujące kliknij opcję **Konfiguruj aplikacje wyszukujące**.

- d. Kliknij opcję **Dodaj aplikację wyszukującą** i w polu **Nazwa aplikacji wyszukującej** wpisz identyfikator administratora wyszukiwania korporacyjnego podany podczas instalowania produktu WebSphere II OmniFind Edition.
  - e. Upewnij się, że włączona jest opcja **Wszystkie kolekcje i zewnętrzne źródła** i kliknij przycisk **OK**.
2. Zmodyfikuj plik config.properties:
    - a. W przypadku systemu UNIX otwórz okno konsoli. W przypadku systemu Microsoft Windows otwórz okno wiersza komend.
    - b. Przejdź do katalogu WEB-INF dla przykładowej aplikacji wyszukującej. W poniższych przykładach dla przejrzystości komendy zostały przedstawione w dwóch wierszach. W rzeczywistości należy je wpisać w jednym wierszu:
 

**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. Zmodyfikuj plik config.properties w edytorze tekstu.
    - d. Jako wartość właściwości username wpisz nazwę użytkownika serwera WebSphere Application Server.
    - e. Jako wartość właściwości password wpisz hasło wpisanego wcześniej użytkownika.
    - f. Zapisz i zamknij plik.
  3. Uruchom ponownie aplikację ESSearchApplication na serwerze WebSphere Application Server:
    - a. Na serwerze wyszukiwania uruchom konsolę administracyjną serwera WebSphere Application Server.  
Konsolę administracyjną można otworzyć na kilka sposobów:
      - W systemie Windows należy użyć menu **Start**, aby wybrać odpowiedni program.
      - W przypadku serwera WebSphere Application Server w wersji 5 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/admin`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9090.
      - W przypadku serwera WebSphere Application Server w wersji 6 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/ibm/console`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9060.
    - b. Po wyświetleniu monitu o podanie identyfikatora użytkownika i hasła należy podać identyfikator administratora i hasło, które zostały określone podczas włączania zabezpieczenia globalnego w programie WebSphere Application Server.
    - c. Po zalogowaniu się do konsoli administracyjnej kliknij opcję **Aplikacje**, a następnie opcję **Aplikacje korporacyjne**.
    - d. Zaznacz pole wyboru ESSearchApplication i kliknij opcję **Stop** (Zatrzymaj).
    - e. Gdy aplikacja zostanie zatrzymana, ponownie zaznacz pole wyboru ESSearchApplication i kliknij opcję **Start** (Uruchom).



---

## Zewnętrzne źródła wyszukiwania korporacyjnego

*Źródło zewnętrzne* jest źródłem danych udostępnianym do wyszukiwania przy użyciu aplikacji wyszukiwania korporacyjnego bez konieczności przeszukiwania, analizowania i indeksowania dokumentów w tym źródle danych.

Jako źródła zewnętrzne na potrzeby wyszukiwania mogą służyć następujące typy źródeł danych:

- Tabele baz danych JDBC (Java database connectivity): tylko bazy danych IBM DB2 Universal Database (DB2 UDB) oraz Oracle. Dla każdej tabeli w bazie danych JDBC tworzone jest oddzielne źródło zewnętrzne.
- Serwery LDAP (Lightweight Directory Access Protocol). Dla każdego serwera LDAP tworzone jest jedno źródło zewnętrzne.

Po skonfigurowaniu informacji na temat źródła zewnętrznego konieczne jest powiązanie go z co najmniej jedną aplikacją wyszukującą. Użytkownicy mogą następnie wyszukiwać dane w tym źródle zewnętrznym jednocześnie z wprowadzaniem zapytań dla kolekcji utworzonych przez przeszukiwanie, przeanalizowanie i poindeksowanie danych na potrzeby wyszukiwania korporacyjnego.

### Pojęcia pokrewne

"Search and index API federators" (Programy stowarzyszające interfejsu API wyszukiwania i indeksowania) w podręczniku "Programming Guide and API Reference for Enterprise Search"

---

## Dodawanie do systemu źródeł zewnętrznych

Podczas dodawania zewnętrznego źródła do systemu wyszukiwania korporacyjnego należy określić typ dodawanego źródła. Kreator pomaga w określeniu informacji na temat źródła danych i sposobu wyszukiwania w nim danych.

### Zagadnienia wstępne

Do dodawania źródeł zewnętrznych do systemu wymagana jest rola administratora wyszukiwania korporacyjnego.

### Ograniczenia

Aby możliwe było przeszukiwanie bazy danych Oracle jako źródła zewnętrznego, na serwerach wyszukiwania korporacyjnego musi być zainstalowany program kliencki Oracle.

### Informacje o tym zadaniu

Dodanie do systemu informacji o źródle zewnętrznym pozwala użytkownikom wprowadzać zapytania dla tego źródła za pomocą aplikacji wyszukiwania korporacyjnego. Włączyć można wyszukiwanie na serwerach LDAP (Lightweight Directory Access Protocol) i w tabelach baz danych JDBC (Java database connectivity).

Podczas konfigurowania serwera LDAP kreator pomaga w określeniu informacji umożliwiających nawiązanie przez system połączenia z tym serwerem oraz w określeniu opcji dotyczących sposobów wyszukiwania na tym serwerze.

Podczas konfigurowania bazy danych JDBC kreator pomaga w określeniu informacji umożliwiających nawiązanie przez system połączenia z tą bazą danych, w wybraniu tabel, które mają zostać udostępnione do wyszukiwania, a także w określeniu opcji dotyczących sposobów wyszukiwania danych w tych tabelach. Dla każdej tabeli dodawanej do systemu tworzone jest źródło danych podlegające odrębnym operacjom wyszukiwania.

### Procedura

Aby dodać do systemu źródło zewnętrzne:

1. Kliknij opcję **Źródła zewnętrzne**, aby otworzyć widok Źródła zewnętrzne.
2. Kliknij opcję **Dodaj źródło zewnętrzne**.
3. Wybierz typ źródła zewnętrznego, które ma zostać dodane (serwer LDAP lub baza danych JDBC).
4. Kliknij przycisk **Dalej**, aby rozpocząć konfigurowanie źródła zewnętrznego.

Otwarty zostanie kreator odpowiedni dla typu tworzonego źródła. Postępuj zgodnie z instrukcjami wyświetlanymi w kreatorze, aby skonfigurować źródło zewnętrzne. Kliknij przycisk **Pomoc** na dowolnej stronie kreatora, aby dowiedzieć się więcej na temat opcji dostępnych do określenia.

Poniższe domyślne nazwy i położenia sterownika JDBC mogą okazać się pomocne podczas konfigurowania informacji na temat połączeń dla baz danych DB2 Universal Database (DB2 UDB) i Oracle:

#### DB2: Legacy JDBC Driver

Nazwa sterownika: COM.ibm.db2.jdbc.app.DB2Driver

Położenie: *główny\_katalog\_instalacyjny\_db2/java/db2java.zip*

#### DB2: Universal JDBC Driver

Nazwa sterownika: com.ibm.db2.jcc.DB2Driver

Położenie:

*główny\_katalog\_instalacyjny\_db2/java/db2jcc.jar*

*główny\_katalog\_instalacyjny\_db2/java/db2jcc\_license\_cu.jar*

#### Oracle

Nazwa sterownika: oracle.jdbc.driver.OracleDriver

Położenie:

*główny\_katalog\_oracle/jdbc/lib/classes12.zip*

*główny\_katalog\_oracle/jdbc/lib/nls\_charset12.zip*

5. Po określeniu opcji dotyczących wyszukiwania w źródle zewnętrznym kliknij przycisk

**Zakończ**.

Nowe źródło zewnętrzne zostanie wymienione w widoku Źródła zewnętrzne wraz z innymi źródłami zewnętrznymi dodanymi do systemu.

#### Pojęcia pokrewne

"Search and index API federators" (Programy stowarzyszające interfejsu API wyszukiwania i indeksowania) w podręczniku "Programming Guide and API Reference for Enterprise Search"

---

## Powiązanie aplikacji wyszukujących ze źródłami zewnętrznymi

Zanim możliwe będzie wyszukiwanie w źródle zewnętrznym należy powiązać z nim co najmniej jedną aplikację wyszukującą.

### Zagadnienia wstępne

Do powiązania aplikacji wyszukujących ze źródłami zewnętrznymi, w których mogą one wykonywać wyszukiwanie, wymagana jest rola administratora wyszukiwania korporacyjnego.

## Procedura

Aby powiązać aplikację wyszukującą z jednym lub większą liczbą źródeł zewnętrznych:

1. Kliknij opcję **Zabezpieczenia** na pasku narzędzi konsoli administracyjnej.
2. Na stronie Aplikacje wyszukujące kliknij opcję **Konfiguruj aplikacje wyszukujące**.
3. Na stronie Konfigurowanie aplikacji wyszukujących kliknij opcję **Dodaj aplikację wyszukującą**.
4. Wpisz nazwę aplikacji wyszukującej.
5. Wybierz źródła zewnętrzne, w których aplikacja będzie mogła wyszukiwać:
  - Kliknij opcję **Wszystkie kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukująca ma mieć dostęp do wszystkich źródeł zewnętrznych dodanych do systemu.
  - Kliknij opcję **Konkretne kolekcje i zewnętrzne źródła**, jeśli aplikacja wyszukująca ma mieć dostęp tylko do określonych źródeł zewnętrznych.

Po wybraniu tej opcji zostanie wyświetlona lista nazw kolekcji i źródeł zewnętrznych. Należy zaznaczyć pole wyboru **Wybierz** dla każdego źródła zewnętrznego, w którym aplikacja będzie mogła wyszukiwać.
6. Kliknij przycisk **OK**.

### Pojęcia pokrewne

"Search and index API federators" (Programy stwarzające interfejsu API wyszukiwania i indeksowania) w podręczniku "Programming Guide and API Reference for Enterprise Search"





---

## Zabezpieczenia wyszukiwania korporacyjnego

Mechanizmy zabezpieczeń w wyszukiwaniu korporacyjnym pozwalają chronić źródła przed nieautoryzowanym wyszukiwaniem oraz ograniczać grupę użytkowników, którzy będą mogli korzystać z funkcji administracyjnych.

Wyszukiwanie korporacyjne umożliwia wyszukiwanie w szerokim zakresie źródeł danych. Wyszukiwanie korporacyjne oferuje kilkupoziomowy mechanizm koordynowania i egzekwowania zabezpieczeń pozwalający zagwarantować, że dostęp do danych mają wyłącznie użytkownicy do tego uprawnieni oraz że tylko uprawnieni użytkownicy mają dostęp do konsoli administracyjnej.

### Serwer WWW

Pierwszym poziomem zabezpieczeń jest serwer WWW. Jeśli włączony jest globalny mechanizm zabezpieczeń serwera WebSphere Application Server, można przypisywać użytkownikom role administracyjne oraz uwierzytelniać użytkowników próbujących administrować systemem. Po zalogowaniu się do konsoli administracyjnej użytkownikowi udostępniane są tylko te funkcje i kolekcje, którymi może on administrować.

Aplikacje wyszukujące także mogą korzystać z mechanizmów dostępnych na serwerze WebSphere Application Server do uwierzytelniania użytkowników korzystających z aplikacji do wyszukiwania w kolekcjach.

### Zabezpieczenia na poziomie kolekcji

Podczas tworzenia kolekcji można włączyć zabezpieczenia na poziomie kolekcji. Ustawienia tego nie można zmienić po utworzeniu kolekcji. Jeśli nie włączono zabezpieczeń na poziomie kolekcji, nie można później ustawiać opcji zabezpieczeń na poziomie dokumentów.

Kiedy zabezpieczenia na poziomie kolekcji są włączone, procesy analizy globalnej przestrzegają następujących specjalnych reguł:

- Aby upewnić się, że mechanizmy zabezpieczeń dla każdego dokumentu są wartościowane, dokumenty mające identyczną (lub prawie identyczną) treść są indeksowane niezależnie od siebie, zamiast indeksowania ich wspólnej treści w postaci reprezentacji kanonicznej.
- W fazie analizy globalnej odpowiedzialnej za przetwarzanie tekstu zakotwiczenia standardowo tekst znajdujący się w jednym dokumencie (dokument źródłowy) kojarzony jest z innym dokumentem (dokumentem docelowym), w którym wcale nie musi on występować. W efekcie dokument docelowy może być pobrany w wyniku zapytań dotyczących tekstu występującego w dokumencie źródłowym. Taki model przetwarzania tekstu zakotwiczenia stwarza niebezpieczeństwo w sytuacji, gdy użytkownicy mają prawo do przeglądania dokumentu docelowego, lecz nie źródłowego. Kiedy zabezpieczenia kolekcji są włączone, tekst zakotwiczenia z odsyłaczy w zabronionych dokumentach jest wyłączony z indeksu. Dokument jest zwracany w wynikach wyszukiwania tylko wtedy, gdy jego własna treść lub metadane pasują do warunków zapytania.

Jest to kompromis pomiędzy włączeniem zabezpieczeń kolekcji a jakością wyszukiwania. Włączenie zabezpieczeń kolekcji zmniejsza ilość informacji indeksowanej dla każdego dokumentu. Efektem ubocznym jest to, że w wyniku niektórych zapytań zwracana będzie mniejsza liczba trafień.

Zabezpieczenia na poziomie kolekcji są także dostępne dla aplikacji wyszukujących za pośrednictwem identyfikatora aplikacji. Aby można było wyszukiwać w kolekcji, administrator wyszukiwania korporacyjnego musi powiązać aplikację wyszukującą z kolekcjami dostępnymi do wyszukiwania. Następnie, korzystając ze standardowych mechanizmów kontroli dostępu, można zezwolić lub zabronić użytkownikom dostępu do poszczególnych aplikacji wyszukujących.

### **Zabezpieczenia na poziomie dokumentu**

Podczas konfigurowania przeszukiwaczy dla kolekcji można uaktywnić zabezpieczenia na poziomie dokumentów. W przypadku wybrania tej opcji przeszukiwacz kojarzy tokeny zabezpieczeń z każdym przeszukiwanym przez niego dokumentem. Tokeny zabezpieczeń są przechowywane wraz z dokumentami w indeksie. W przypadku niektórych przeszukiwaczy można także włączyć opcję sprawdzania poprawności referencji użytkownika podczas przetwarzania zapytania, które jest realizowane przez porównanie ich z aktualnymi danymi kontroli dostępu (zgodnie z konfiguracją w rodzimym źródle danych).

Do kontroli dostępu aplikacje wyszukujące mogą korzystać z tokenów zabezpieczeń i referencji użytkowników. Aby zagwarantować, że użytkownicy wyszukują i pobierają tylko te dokumenty, do których mają uprawnienia dostępu, aplikacja wyszukująca może dołączać referencje zalogowanego użytkownika do zapytań przekazywanych do serwerów wyszukiwania.

Zabezpieczenie kolekcji wychodzi poza mechanizmy uwierzytelniania i kontroli dostępu używane przez wyszukiwanie korporacyjne do zabezpieczania indeksowanych treści. Istnieją także zabezpieczenia przed przechwytywaniem danych przez nieupoważnionych użytkowników w trakcie ich przesyłania. Na przykład do komunikowania się z serwerem indeksowania i z aplikacją wyszukującą serwery wyszukiwania wykorzystują takie protokoły, jak SSL, SSH i HTTPS.

Dodatkowym zabezpieczeniem jest szyfrowanie. Na przykład hasło administratora wyszukiwania korporacyjnego określone podczas instalowania produktu jest przechowywane w formie zaszyfrowanej. Hasła zapisane przez użytkowników w ich profilach są także zaszyfrowane.

Dla dodatkowego bezpieczeństwa należy się upewnić, że komponenty sprzętowe serwera są należycie izolowane i chronione przed ingerencją nieupoważnionych osób. Instalując firewall, można zabezpieczyć serwery wyszukiwania korporacyjnego przed ingerencją za pośrednictwem innego segmentu sieci. Należy również się upewnić, że serwery wyszukiwania korporacyjnego nie mają żadnych otwartych portów zapasowych. System należy skonfigurować w taki sposób, aby nasłuchiwanie nadchodzących żądań odbywało się wyłącznie na portach w sposób jawny przypisanych aplikacjom i operacjom wyszukiwania korporacyjnego.

---

## **Role administracyjne**

Wyszukiwanie korporacyjne wykorzystuje role jako metodę kontroli dostępu do różnych funkcji konsoli administracyjnej.

Podczas instalacji produktu WebSphere Information Integrator OmniFind Edition (WebSphere II OmniFind Edition) instalator konfiguruje identyfikator i hasło administratora wyszukiwania korporacyjnego. Korzystając z konsoli administracyjnej po raz pierwszy, należy się zalogować jako użytkownik o skonfigurowanym identyfikatorze i hasle. Jeśli globalny mechanizm zabezpieczeń serwera IBM WebSphere Application Server nie został uruchomiony, identyfikator ten będzie jedynym identyfikatorem użytkownika umożliwiającym dostęp do konsoli administracyjnej wyszukiwania korporacyjnego.

Jeśli globalny mechanizm zabezpieczeń serwera WebSphere Application Server zostanie włączony, można zarejestrować dodatkowych użytkowników jako użytkowników z administracyjnymi uprawnieniami wyszukiwania korporacyjnego. Przypisując użytkowników do ról, można ograniczyć dostęp do wybranych kolekcji i kontrolować funkcje, z których każdy z użytkowników z uprawnieniami administracyjnymi będzie mógł korzystać. Identyfikatory użytkowników przypisywane do ról administracyjnych w wyszukiwaniu korporacyjnym muszą istnieć w rejestrze użytkowników serwera WebSphere Application Server.

Podczas logowania użytkownika z uprawnieniami administracyjnymi wyszukiwanie korporacyjne uwierzytelnia jego identyfikator. W konsoli będą dostępne wyłącznie te kolekcje i funkcje, którymi dany użytkownik ma prawo administrować.

Użytkowników można przypisywać do następujących ról administracyjnych:

#### **Administrator wyszukiwania korporacyjnego**

Ci użytkownicy tworzą kolekcje i mają prawo do administrowania wszystkimi aspektami funkcjonowania systemu wyszukiwania korporacyjnego. W trakcie instalacji produktu WebSphere II OmniFind Edition, określone są identyfikator i hasło użytkownika dla pierwszego użytkownika wyszukiwania korporacyjnego z uprawnieniami administracyjnymi. Po pierwszym zalogowaniu się użytkownik ten będzie mógł przypisać rolę administratora wyszukiwania korporacyjnego innym użytkownikom.

#### **Administrator kolekcji**

Ci użytkownicy mogą modyfikować, monitorować i sterować działaniem określonych kolekcji albo wszystkich kolekcji. Nie mogą oni tworzyć kolekcji. Administratorzy kolekcji mogą monitorować i obsługiwać działania systemowe, tylko jeśli takie uprawnienie zostanie im nadane przez administratora wyszukiwania korporacyjnego.

#### **Operator**

Tacy użytkownicy mogą monitorować działanie konkretnych lub wszystkich kolekcji i sterować nim. Mogą oni na przykład uruchamiać i zatrzymywać działania kolekcji, ale nie mogą tworzyć ani edytować kolekcji. Operator może monitorować i obsługiwać działania systemowe, tylko jeśli takie uprawnienie zostanie mu nadane przez administratora wyszukiwania korporacyjnego.

#### **Monitor**

Tacy użytkownicy mogą monitorować konkretne lub wszystkie kolekcje. Nie mogą oni sterować operacjami (takimi jak uruchamianie i zatrzymywanie serwerów), tworzyć kolekcji ani ich edytować. Użytkownik monitorujący może obserwować, ale nie obsługiwać, działania systemowe, tylko jeśli takie uprawnienie zostanie mu nadane przez administratora wyszukiwania korporacyjnego.

#### **Zadania pokrewne**

“Logowanie się do konsoli administracyjnej” na stronie 5

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

“Uruchamianie serwerów wyszukiwania korporacyjnego” na stronie 209

Aby umożliwić użytkownikom wyszukiwanie w kolekcji, należy uruchomić procesy systemowe, a następnie uruchomić serwery przeszukiwania, analizy, indeksowania i wyszukiwania w kolekcji.

---

## Konfiguracja użytkowników z uprawnieniami administracyjnymi

Konfigurując role administracyjne można ograniczyć dostęp do kolekcji i sterować dostępem użytkowników z uprawnieniami administracyjnymi do poszczególnych funkcji.

### Zagadnienia wstępne

Przed przypisaniem użytkownika do roli administracyjnej należy się upewnić, że na serwerze WebSphere Application Server jest włączone zabezpieczenie. Należy również upewnić się, że w rejestrze użytkowników serwera WebSphere Application Server istnieje identyfikator użytkownika.

Aby konfigurować użytkowników administracyjnych, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Procedura

Aby przypisać użytkowników do ról administracyjnych, należy wykonać poniższe czynności:

1. Kliknij przycisk **Zabezpieczenia**, aby otworzyć widok Zabezpieczenia.
2. Na stronie Role administracyjne kliknij przycisk **Dodaj użytkownika**.
3. Wpisz identyfikator użytkownika, który ma zostać zarejestrowany, i wybierz odpowiednią rolę administracyjną.
4. Jeśli użytkownik nie jest rejestrowany jako administrator wyszukiwania korporacyjnego, określ, czy może on uzyskiwać dostęp do stron z paska narzędzi **System**.  
Na przykład można umożliwić niektórym operatorom lub administratorom kolekcji monitorowanie systemowych plików dziennika.
5. Jeśli użytkownik nie jest rejestrowany jako administrator wyszukiwania korporacyjnego, wybierz kolekcje i źródła zewnętrzne, którymi będzie mógł administrować.  
Można zaznaczyć pola wyboru dla pojedynczych kolekcji i źródeł zewnętrznych lub umożliwić użytkownikowi administrowanie wszystkimi kolekcjami i źródłami zewnętrznymi.

---

## Uwierzytelnianie a kontrola dostępu

Ochrona danych i funkcji administracyjnych przed dostępem nieupoważnionych osób odbywa się przy wyszukiwaniu korporacyjnym za pośrednictwem dwóch mechanizmów: uwierzytelniania użytkowników i kontroli dostępu.

### Uwierzytelnianie

Uwierzytelnianie jest to proces, w drodze którego system sprawdza, czy użytkownicy są faktycznie kimś lub czymś, za kogo się podają. Ponieważ dostęp do zasobów jest realizowany na podstawie identyfikacji tożsamości użytkownika, który zgłasza żądanie dostępu, uwierzytelnianie stanowi fundament skutecznego systemu zabezpieczeń.

W celu uwierzytelnienia użytkowników próbujących uzyskać dostęp do konsoli administracyjnej wyszukiwanie korporacyjne korzysta z obsługi uwierzytelniania będącej częścią serwera WebSphere Application Server.

W celu uwierzytelnienia użytkowników wykonujących wyszukiwanie w kolekcji wyszukiwania korporacyjnego aplikacje wyszukujące mogą korzystać z mechanizmów zabezpieczeń serwera WebSphere Application Server i posługiwać się wybraną metodą

uwierzytelniania referencji użytkownika. Zazwyczaj referencje użytkownika obejmują identyfikator i hasło przekazywane do aplikacji wyszukującej, gdy użytkownik się loguje lub próbuje uzyskać dostęp do tej aplikacji.

Uwierzytelnianie użytkownika można zaimplementować na inne sposoby zależnie od zasobów i protokołów dostępnych w przedsiębiorstwie. Na przykład można wymusić na użytkownikach potwierdzanie tożsamości przy użyciu kart Smart Card, certyfikatów cyfrowych i infrastruktury klucza publicznego albo przez przypisywanie biletów w momencie logowania.

## Kontrola dostępu

Kontrola dostępu to technika ograniczania zestawu czynności, jakie użytkownikom wolno wykonywać po potwierdzeniu tożsamości i uwierzytelnieniu. Lista kontroli dostępu (ACL) to najbardziej rozpowszechniona metoda ograniczania dostępu do zasobów. ACL jest to lista identyfikatorów użytkowników (nazwy użytkowników, nazwy grup, role użytkowników itd.). Każdy identyfikator użytkownika jest powiązany z zestawem uprawnień definiujących prawa i uprawnienia danego użytkownika.

Na przykład mechanizm kontroli dostępu może umożliwiać albo zabraniać dostępu do plików na serwerze plików oraz sterować, czy użytkownik dysponujący prawem dostępu może odczytywać, modyfikować lub usuwać pliki na tym serwerze.

Wszelka kontrola dostępu w wyszukiwaniu korporacyjnym jest uzależniona od tego, czy użytkownik ma uprawnienie do odczytu danych w indeksie. W zależności od sposobu obsługi kontroli dostępu w aplikacjach wyszukiwanych oraz od reguł określonych dla kolekcji i przeszukiwaczy przy administrowaniu wyszukiwania korporacyjnego można:

- Umożliwić wszystkim użytkownikom wyszukiwanie wszystkich dokumentów w kolekcji.
- Umożliwić wszystkim użytkownikom wyszukiwanie wszystkich dokumentów, które zostały przeszukane przez określone przeszukiwacze oraz ograniczyć dostęp do dokumentów przeszukanych przez inne przeszukiwacze.
- Umożliwić wybranym użytkownikom wyszukiwanie wybranych dokumentach. Na przykład podczas określania baz danych, które mają być przeszukane przez przeszukiwacz Notes, można określić opcje zezwalające niektórym użytkownikom na wyszukiwanie dokumentów w określonych widokach i folderach i zabraniające innym użytkownikom wyszukiwania tych dokumentów.

---

## Wyłączanie zabezpieczenia dla aplikacji korporacyjnej na serwerze WebSphere Application Server

Aby określić, które działania programu WebSphere II OmniFind Edition wymagają uwierzytelniania użytkowników, na serwerze WebSphere Application Server można wyłączyć zabezpieczenie globalne dla pojedynczych aplikacji korporacyjnych.

### Informacje o tym zadaniu

Program instalacyjny WebSphere II OmniFind Edition instaluje trzy aplikacje korporacyjne serwera WebSphere Application Server:

- Aplikacja ESAdmin zawiera interfejs konsoli administracyjnej wyszukiwania korporacyjnego.
- Aplikacja ESSearchApplication zawiera interfejs przykładowej aplikacji wyszukiwującej.
- Aplikacja ESSearchServer udostępnia zdalną komunikację dla implementacji SIAPI programu WebSphere II OmniFind Edition i pozwala interfejsom SIAPI komunikować się z serwerami wyszukiwania.

Domyślnie wszystkie trzy aplikacje korporacyjne obsługują globalne zabezpieczenie programu WebSphere Application Server. Gdy te aplikacje wykryją, że zostało włączone zabezpieczenie globalne, rozpoczną uwierzytelnianie wszystkich otrzymywanych żądań.

W niektórych organizacjach może zaistnieć potrzeba włączenia lub wyłączenia zabezpieczenia dla konkretnych aplikacji korporacyjnych programu WebSphere II OmniFind Edition. Na przykład można uwierzytelniać wszystkich użytkowników, którzy uzyskują dostęp do konsoli administracyjnej wyszukiwania korporacyjnego, i nie uwierzytelniać użytkowników, którzy korzystają z interfejsów SI-API programu WebSphere II OmniFind Edition lub z przykładowej aplikacji wyszukiwującej.

## Procedura

Aby wyłączyć zabezpieczenie dla konkretnej aplikacji korporacyjnej, należy wykonać następujące czynności:

1. Na serwerze wyszukiwania uruchom konsolę administracyjną serwera WebSphere Application Server.  
Konsolę administracyjną można otworzyć na kilka sposobów:
  - W systemie Windows należy użyć menu **Start**, aby wybrać odpowiedni program.
  - W przypadku serwera WebSphere Application Server w wersji 5 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/admin`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9090.
  - W przypadku serwera WebSphere Application Server w wersji 6 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/ibm/console`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9060.
2. Po wyświetleniu monitu o podanie identyfikatora użytkownika i hasła należy podać identyfikator administratora i hasło, które zostały określone podczas włączania zabezpieczenia globalnego w programie WebSphere Application Server.
3. Po zalogowaniu się do konsoli administracyjnej kliknij opcję **Aplikacje**, a następnie opcję **Aplikacje korporacyjne**.
4. Zaznacz pole wyboru obok nazwy aplikacji korporacyjnej, dla której ma zostać wyłączone zabezpieczenie.
5. Przejdź niżej i kliknij odsyłacz **Przypisz użytkownikom/grupom role dotyczące zabezpieczenia**.
6. Znajdź rolę **WszyscyUwierzytelniani** i zaznacz pole wyboru w kolumnie **Wszyscy?**.
7. Kliknij przycisk **OK**.
8. Kliknij odsyłacz **Zapisz**, aby zapisać zmiany.
9. W przypadku korzystania z programu WebSphere Network Deployment zaznacz pole wyboru **Synchronizuj zmiany z węzłami**.
10. Kliknij przycisk **Zapisz**.
11. Ponownie kliknij opcję **Aplikacje**, a następnie opcję **Aplikacje korporacyjne**.
12. Zaznacz pole wyboru obok właśnie zmodyfikowanej aplikacji i kliknij przycisk **Zatrzymaj**.
13. Po zatrzymaniu aplikacji ponownie zaznacz pole wyboru dla tej aplikacji i kliknij przycisk **Uruchom**.

---

## Zabezpieczenia na poziomie kolekcji

W celu zapewnienia zabezpieczeń na poziomie kolekcji należy skonfigurować opcje dotyczące indeksowania treści oraz opcje umożliwiające aplikacjom wyszukującym realizację wyszukiwania w określonych kolekcjach.

Podczas tworzenia kolekcji można wybrać opcję powodującą włączenie zabezpieczeń kolekcji. Wybranie tej opcji umożliwia późniejsze skonfigurowanie mechanizmów zabezpieczeń na poziomie dokumentów. Kiedy włączone są zabezpieczenia kolekcji, globalne procesy analizy w wyszukiwaniu korporacyjnym stosują także odmienne reguły przy indeksowaniu duplikatów dokumentów i tekstu zakotwiczenia.

Po utworzeniu aplikacji wyszukującej przypisany jej identyfikator pozwala określić, w których kolekcjach aplikacja wyszukująca będzie mogła wyszukiwać oraz którzy użytkownicy będą mieli dostęp do tej aplikacji.

### Analiza duplikatów dokumentów

W przypadku włączonego mechanizmu zabezpieczeń kolekcji globalne procesy analizowania nie identyfikują duplikatów dokumentów w kolekcji.

W trakcie analizy globalnej procesy indeksowania identyfikują dokumenty będące duplikatami lub prawie duplikatami innych dokumentów. Następnie wszystkie te dokumenty zostają powiązane z pojedynczą kanoniczną reprezentacją treści. Zezwalając na identyfikowanie duplikatów dokumentów, można się upewnić, że wyniki wyszukiwania nie zawierają wielu dokumentów o identycznej (lub niemal identycznej) treści.

Jeśli przy tworzeniu kolekcji zostanie włączony mechanizm zabezpieczeń, duplikaty dokumentów nie będą identyfikowane i nie będą powiązane z ogólną reprezentacją kanoniczną. Każdy dokument będzie indeksowany niezależnie od innych. Ma to służyć zapewnieniu, że użytkownicy będą wyszukiwać wyłącznie dokumenty z tokenami zabezpieczeń dopasowanymi do ich referencji. Na przykład dwa dokumenty mogą mieć prawie identyczną treść, ale bezpieczeństwo może być zapewniane przy zastosowaniu różnych list kontroli dostępu.

Wyłączenie analizy duplikatów dokumentów może zwiększyć bezpieczeństwo dokumentów w kolekcji, jednak może się to odbyć kosztem pogorszenia jakości wyszukiwania, gdy użytkownicy w wynikach wyszukiwania otrzymywać będą wiele kopii tego samego dokumentu.

### Analiza tekstu zakotwiczenia

Jeśli włączony jest mechanizm zabezpieczeń kolekcji, w procesach analizy globalnej stosowane są specjalne reguły dotyczące indeksowania tekstu zakotwiczenia w dokumentach przeszukiwanych przez przeszukiwacze sieci WWW. Jeśli nie włączono mechanizmu zabezpieczeń kolekcji, podczas konfigurowania indywidualnych przeszukiwaczy sieci WWW można określić, czy ma być indeksowany tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów.

Tekst zakotwiczenia jest informacją w obrębie odsyłacza hipertekstowego opisującą stronę, do której ten odsyłacz prowadzi. Na przykład w poniższym odsyłaczu tekst **Składnia zapytania** jest tekstem zakotwiczenia w odsyłaczu prowadzącym do strony `syntax.htm` w serwisie WWW:

```
Składnia zapytania
```

Zazwyczaj przeszukiwacz sieci WWW korzysta z odsyłaczy w dokumentach w celu przeszukania dodatkowych dokumentów i dopisuje te dowiązane strony do indeksu. W trakcie analizy globalnej procesy indeksowania kojarzą tekst zakotwiczenia nie tylko z dokumentem, w którym jest on osadzony (dokumentem źródłowym), lecz także z dokumentem docelowym. W przedstawionym powyżej przykładzie tekst zakotwiczenia Składnia zapytania jest kojarzony ze stroną docelową syntax.htm oraz ze stroną zawierającą konstrukcję zakotwiczenia.

Jeśli podczas tworzenia kolekcji zostaną włączone zabezpieczenia kolekcji, przetwarzanie tekstu zakotwiczenia jest wyłączone. Tekst zakotwiczenia nie jest już indeksowany z dokumentem, chyba że występuje w jawnej postaci w treści lub w metadanych dokumentu. Taki mechanizm uniemożliwia użytkownikom oglądanie zawartości dokumentów, do których nie mają prawa dostępu, ponieważ tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów nie jest kojarzony z dokumentami, do których mają dostęp.

Włączenie zabezpieczeń kolekcji umożliwia zwiększenie bezpieczeństwa dokumentów WWW, umożliwiając użytkownikom wyszukiwania wyłącznie dokumentów z tokenami zabezpieczeń pasującymi do ich referencji. Jednak w przypadku rezygnacji z przetwarzania tekstu zakotwiczenia może dojść do sytuacji, w której wyniki wyszukiwania mogą nie zawierać wszystkich dokumentów potencjalnie odpowiadających warunkom zapytania.

Jeśli nie włączono mechanizmu zabezpieczeń kolekcji, podczas konfigurowania indywidualnych przeszukiwaczy sieci WWW można określić, czy tekst zakotwiczenia w odsyłaczach wiodących do zabronionych dokumentów ma być przetwarzany przez procesy indeksowania. Jeśli tekst zakotwiczenia ma być objęty indeksowaniem, procesy analizy i indeksowania indeksują tekst zakotwiczenia na wszystkich stronach, które pobiera przeszukiwacz sieci WWW. Jeśli tekst zakotwiczenia ma nie być przetwarzany, tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów jest wyłączony z indeksu.

## **Indeksowanie tekstu zakotwiczenia w odsyłaczach do dokumentów zabronionych**

Jeśli dokument zawiera odsyłacze do dokumentów, do których przeszukiwacz sieci WWW nie ma dostępu, podczas konfigurowania przeszukiwacza sieci WWW można określić, czy tekst zakotwiczenia odsyłaczy ma być zachowywany w indeksie.

### **Zagadnienia wstępne**

Opcje indeksowania tekstu zakotwiczeń może skonfigurować uczestnik roli administratorów wyszukiwania korporacyjnego oraz administrator kolekcji dla konfigurowanego przeszukiwacza WWW.

### **Informacje o tym zadaniu**


Dyrektwy zawarte w pliku robots.txt lub w metadanych dokumentów WWW mogą uniemożliwić przeszukiwaczowi sieci WWW dostęp do dokumentów serwisu. Jeśli dokument, do którego przeszukiwacz sieci WWW ma dostęp, zawiera odsyłacze do zabronionych dokumentów, można zdecydować o sposobie obsługi tekstu zakotwiczeń dla tych odsyłaczy.

Podczas konfigurowania przeszukiwaczy sieci WWW można określić, czy ma być indeksowany tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów. Ze względów bezpieczeństwa zalecane jest włączenie opcji blokującej indeksowanie tekstu zakotwiczenia w odsyłaczach do zabronionych dokumentów. Jednak w wyniku pominięcia tekstu zakotwiczeń przy indeksowaniu wyniki wyszukiwania mogą nie zawierać niektórych dokumentów potencjalnie spełniających warunki zapytania.



## Procedura

Aby włączyć lub wyłączyć indeksowanie tekstu zakotwiczenia w odsyłaczach do zabronionych dokumentów:

1. Przejdź do edycji kolekcji, wybierz stronę Przeszukiwanie, znajdź przeszukiwacz sieci WWW, który chcesz skonfigurować, i kliknij ikonę  **Właściwości przeszukiwacza**.
2. Kliknij opcję **Edytuj zaawansowane właściwości przeszukiwacza sieci WWW**.
3. Aby indeksować tekst zakotwiczenia we wszystkich dokumentach objętych działaniem danego przeszukiwacza, zaznacz pole wyboru **Indeksuj tekst zakotwiczenia w odsyłaczach do dokumentów zabronionych**. Użytkownicy mogą uzyskać informacje o stronach niedostępnych dla przeszukiwacza sieci WWW dzięki wyszukiwaniu tekstu zakotwiczenia w odsyłaczach prowadzących do tych stron.  
Aby wykluczyć z indeksowania tekst zakotwiczenia w odsyłaczach do zabronionych dokumentów, usuń zaznaczenie tego pola wyboru. Użytkownicy nie będą mieli możliwości uzyskania informacji o stronach niedostępnych dla przeszukiwacza sieci WWW. Tekst zakotwiczenia zostanie wykluczony z indeksu razem z zabronionymi dokumentami.
4. Kliknij przycisk **OK**, a następnie - na stronie Właściwości przeszukiwacza sieci WWW - kliknij ponownie przycisk **OK**.
5. Aby zmiany zostały uwzględnione, zatrzymaj i ponownie uruchom przeszukiwacz.

Aby zastosować zmiany do dokumentów poindeksowanych wcześniej, należy ponownie przeszukać i poindeksować te dokumenty. Jeśli w poprzednim przeszukiwaniu do indeksu trafiły informacje o zabronionych dokumentach, zostaną one usunięte.

## Zabezpieczenia przy użyciu identyfikatorów aplikacji wyszukującej

Zabezpieczenia na poziomie kolekcji są realizowane przez określenie, które aplikacje wyszukujące mogą wyszukiwać w poszczególnych kolekcjach lub źródłach zewnętrznych.

Wszystkie aplikacje wyszukujące muszą przekazywać swoje identyfikatory do funkcji API wyszukiwania korporacyjnego. Administrator wyszukiwania korporacyjnego i aplikacje wyszukujące mogą korzystać z tego identyfikatora w celu wymuszenia zabezpieczeń kolekcji i źródeł zewnętrznych.

Aby aplikacja wyszukująca mogła uzyskać dostęp do kolekcji lub źródła zewnętrznego, administrator wyszukiwania korporacyjnego musi powiązać tę aplikację z określonymi kolekcjami i źródłami, w których może ona wyszukiwać. Aplikacja wyszukująca może przeszukiwać wszystkie kolekcje i źródła zewnętrzne w systemie wyszukiwania korporacyjnego lub tylko określone przez użytkownika.

W celu wymuszenia kontroli dostępu można powiązać tokeny zabezpieczeń (np. identyfikatory użytkowników i grup lub role użytkowników) z aplikacją wyszukującą i zezwolić na dostęp do tej aplikacji tylko wyznaczonym użytkownikom. Na przykład można ograniczyć dostęp do adresu URL uruchamiającego aplikację wyszukującą.

Więcej informacji na temat identyfikatorów aplikacji wyszukującej oraz metod wbudowywania mechanizmów zabezpieczeń w niestandardowe aplikacje wyszukujące zawiera opis funkcji API wyszukiwania i indeksowania dla wyszukiwania korporacyjnego.

### Pojęcia pokrewne

“Aplikacje wyszukujące dla wyszukiwania korporacyjnego” na stronie 155  
Aplikacja wyszukująca umożliwia wyszukiwanie w kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Liczba możliwych do utworzenia aplikacji

wyszukujących nie jest ograniczona, a pojedyncza aplikacja wyszukująca pozwala przeszukać dowolną liczbę kolekcji i źródeł zewnętrznych.

"Search and index API overview" (Przegląd interfejsu API wyszukiwania i indeksowania) w podręczniku "Programming Guide and API Reference for Enterprise Search"

"Search and index API security" (Zabezpieczenia interfejsu API wyszukiwania i indeksowania) w podręczniku "Programming Guide and API Reference for Enterprise Search"

---

## Zabezpieczenia na poziomie dokumentu

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentu. Mechanizm zabezpieczeń na poziomie dokumentu jest odpowiedzialny za to, aby użytkownicy wyszukujący w kolekcji mieli dostęp tylko do tych dokumentów, względem których przydzielono im stosowne uprawnienia.

Aby sterować dostępem do dokumentów w kolekcji, przeszukiwacze mogą gromadzić dane dotyczące bezpieczeństwa dla indeksu. W przypadku niektórych typów źródeł danych można też podczas wprowadzania zapytania sprawdzać bieżące referencje użytkownika. Aby weryfikować aktualne referencje użytkownika, w tworzone aplikacje wyszukujące można wbudować obsługę profili użytkowników. Zapisanie profili użytkowników pozwala użytkownikom na dostęp do dokumentów bez konieczności wielokrotnego podawania referencji.

### Pojęcia pokrewne

"Aplikacje wyszukujące dla wyszukiwania korporacyjnego" na stronie 155

Aplikacja wyszukująca umożliwia wyszukiwanie w kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Liczba możliwych do utworzenia aplikacji wyszukujących nie jest ograniczona, a pojedyncza aplikacja wyszukująca pozwala przeszukać dowolną liczbę kolekcji i źródeł zewnętrznych.

"Zabezpieczenia na poziomie dokumentu przy użyciu mechanizmu wyszukiwania środowiska Portal" na stronie 197

Gdy użytkownicy przeszukują kolekcje wyszukiwania korporacyjnego, mechanizm wyszukiwania środowiska IBM WebSphere Portal może zostać wykorzystany do wymuszenia zabezpieczeń na poziomie dokumentu.

"Search and index API security" (Zabezpieczenia interfejsu API wyszukiwania i indeksowania) w podręczniku "Programming Guide and API Reference for Enterprise Search"

## Sprawdzanie poprawności na podstawie zapisanych tokenów zabezpieczeń

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentów przez zapisywanie danych zabezpieczeń w indeksie.

Domyślnie każdy dokument jest uważany za dokument publiczny, co oznacza, że może on być wyszukiwany przez wszystkich użytkowników. W przypadku większości typów dokumentów bezpieczeństwo na poziomie dokumentu można uzyskać, kojarząc jeden lub większą liczbę tokenów zabezpieczeń z dokumentami i zapisując te elementy razem z dokumentami w indeksie. Podczas konfigurowania przeszukiwacza można określić, że ma on korzystać z tokenów zabezpieczeń, aby umożliwić dostęp do przeszukiwanych dokumentów tylko wybranym użytkownikom.

Jeśli typ źródła danych obejmuje pola, można zdecydować, że kontrola dostępu ma się odbywać w oparciu o wartość jednego z tych pól. Jeśli źródło danych nie zawiera pól, jeśli wartość pola nie powinna być używana do zabezpieczeń lub jeśli wskazane pole nie zawiera

danych umożliwiającą kontrolę dostępu, pozostaje możliwość zdefiniowania tokenów zabezpieczeń, które przeszukiwacz powiąże z dokumentami.

Administrator dla każdej kolekcji definiuje tokeny zabezpieczeń, które przeszukiwacz ma kojarzyć z dokumentami. Na przykład token zabezpieczeń może odpowiadać identyfikatorowi użytkownika lub grupy, roli użytkownika lub dowolnej innej wartości poprawnej w odniesieniu do odpowiedniego źródła danych. Jeśli administrator źródła danych zaktualizuje rodzimą listę kontroli dostępu, zaktualizowane mechanizmy zabezpieczeń staną się dostępne po następnym odświeżeniu lub reorganizacji indeksu.

Tokeny zabezpieczeń towarzyszą dokumentom przy kolejnych etapach analizy składni, analizy powiązań i indeksowania. Jeśli używane aplikacje wyszukujące umożliwiają korzystanie z zabezpieczeń, można posłużyć się tokenami zabezpieczeń w celu kontrolowania dostępu do dokumentów. Użytkownicy wyszukujący w kolekcji mogą wyszukiwać tylko te dokumenty, na przeglądanie których zezwalają przyznane im uprawnienia. Jeśli referencje użytkowników nie spełniają warunków zabezpieczeń, użytkownik nie może wyszukiwać dokumentów zabezpieczonych danymi tokenami.

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

## Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań

Jeśli podczas tworzenia kolekcji włączono opcje zabezpieczeń, w niektórych typach domen możliwe jest sprawdzanie poprawności bieżących referencji użytkownika podczas wprowadzania zapytania.

Podczas konfigurowania poniższych typów przeszukiwaczy można włączyć opcję sprawdzania referencji użytkownika przez porównanie ich z bieżącymi ustawieniami kontroli dostępu obsługiwanyymi przez rodzime repozytorium:

- przeszukiwacz Content Edition (tylko repozytoria typu: Documentum, FileNet Panagon Content Services oraz Portal Document Manager)
- przeszukiwacz Domino Document Manager
- przeszukiwacz Notes
- przeszukiwacz QuickPlace
- przeszukiwacz WebSphere Portal
- przeszukiwacz system plików Windows

Przed wysłaniem odpowiedzi na zapytanie serwery wyszukiwania nawiązują kontakt z rodzimymi repozytoriami w celu zweryfikowania bieżących uprawnień użytkownika, a następnie usuwają z wyników wyszukiwania wszystkie dokumenty, których użytkownik nie może przeglądać.

Taka metoda egzekwowania zabezpieczeń na poziomie dokumentów zapewnia wysoki poziom bezpieczeństwa, ponieważ referencje użytkownika są porównywane z bieżącymi ustawieniami, a nie z ustawieniami zabezpieczeń, które zapisano wcześniej w indeksie. Ponadto zapewnia, że kontrolą dostępu sterują mechanizmy zabezpieczeń rodzimego repozytorium bez względu na stopień ich złożoności. Ponieważ filtrowanie dokumentów odbywa się w czasie rzeczywistym, wyniki wyszukiwania odzwierciedlają najnowsze ustawienia kontroli dostępu dla każdego dokumentu, który spełnia kryteria wyszukiwania.

Inna zaleta tej metody polega na tym, że nie wpływa ona na wielkość indeksu (nie jest potrzebne dodatkowe miejsce na indeksowanie tokenów zabezpieczeń). Jednak ponieważ sprawdzanie poprawności wymaga połączeń z rodzimymi repozytoriami, może się to niekorzystnie odbić na wydajności przetwarzania zapytań.

Aby uzyskać maksymalny poziom bezpieczeństwa i zminimalizować spowolnienie przetwarzania zapytań, można połączyć opcję przechowywania tokenów zabezpieczeń w indeksie z opcją sprawdzania poprawności bieżących praw dostępu. Kiedy użytkownik wysyła zapytanie, sprawdzenie poprawności wykonywane jest w dwóch etapach:

- Najpierw na podstawie danych dotyczących zabezpieczeń z indeksu serwery wyszukiwania szybko ustalają, czy użytkownik ma uprawnienia dostępu do serwera i bazy danych zawierającej przeszukany dokument. Indeks jest optymalizowany pod kątem szybkości i zapewnia czasy odpowiedzi poniżej sekundy.
- Następnie serwery wyszukiwania tworzą tymczasową listę zawierającą tylko dokumenty znajdujące się w domenach na serwerach, do których użytkownik ma prawo dostępu. Serwery używają tej listy przy nawiązywaniu połączeń z rodzimymi repozytoriami w celu ustalenia, czy użytkownik ma uprawnienia do przeglądania żadanego dokumentu.

Jeśli użytkownik ma dostęp do serwera i domeny, istnieje wysokie prawdopodobieństwo, że ma też dostęp do dokumentów. Jednak ten ostatni etap filtrowania pozwala zagwarantować, że w wynikach wyszukiwania znajdują się wyłącznie dokumenty zgodne z bieżącymi uprawnieniami użytkownika.

#### **Pojęcia pokrewne**

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows” na stronie 188

Aby włączyć opcję sprawdzania bieżących referencji podczas wyszukiwania przez użytkownika dokumentów przeszukanych przez przeszukiwacz system plików Windows, należy skonfigurować konto domeny zarówno na serwerze przeszukiwacza, jak i na serwerze Microsoft Windows.

“Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino” na stronie 190

Jeśli na serwerze Lotus Notes, który ma być przeszukiwany, używany jest protokół NRPC (Notes remote procedure call), należy skonfigurować serwer przeszukiwacza w sposób powodujący wymuszenie zabezpieczeń na poziomie dokumentu.

#### **Zadania pokrewne**

“Konfigurowanie serwerów Lotus Domino Trusted Servers do sprawdzania poprawności referencji użytkownika” na stronie 191

Aby wymusić zabezpieczenia dokumentów, które zostały przeszukane przez przeszukiwacz Notes używający protokołu NRPC (Notes remote procedure call), serwery Domino, które mają być przeszukiwane, należy skonfigurować jako zaufane serwery Lotus Domino.

## **Profile użytkowników i zarządzanie tożsamościami**

Tworząc profile dla wyszukiwania korporacyjnego, użytkownicy mogą zapisywać referencje umożliwiające im wyszukiwanie w zabezpieczonych domenach.

Aby móc przeszukiwać domenę wymagającą sprawdzenia referencji podczas wprowadzania zapytania, użytkownicy muszą udostępnić aplikacji wyszukującej referencje, za pomocą których logują się do tej domeny. Dzięki funkcji *zarządzania tożsamościami* w programie WebSphere Information Integrator OmniFind Edition użytkownicy mogą zapisywać w swoich profilach referencje dla dowolnej liczby domen. Referencje są zaszyfrowane i bezpiecznie zapisane w systemie wyszukiwania korporacyjnego.

Użytkownicy mogą tworzyć profile i rejestrować w nich swoje referencje, kiedy korzystają z aplikacji wyszukującej. W przykładowej aplikacji wyszukującej dla systemu wyszukiwania

korporacyjnego możliwość ta jest dostępna przy użyciu opcji **Mój profil** (niestandardowe aplikacje wyszukujące mogą udostępniać tę opcję w inny sposób).

W profilu użytkownika zapisane są różne referencje, które musi on podawać podczas logowania się do domen, które chce przeszukiwać. Użytkownicy mogą tworzyć profile użytkowników, jeśli spełnione są następujące warunki:

- Na serwerze WebSphere Application Server włączone są zabezpieczenia globalne.
- W konsoli administracyjnej wyszukiwania korporacyjnego włączone jest zarządzanie tożsamościami programu WebSphere II OmniFind Edition.
- Zabezpieczenia są włączone w co najmniej jednej kolekcji, którą może przeszukiwać aplikacja wyszukująca.
- Przynajmniej jedna zabezpieczona kolekcja zawiera dokumenty, które zostały przeszukane przez przeszukiwacz wymuszający kontrolę dostępu polegającą na sprawdzaniu referencji podczas wprowadzania zapytania.
- Podczas konfigurowania zabezpieczeń na poziomie dokumentu dla przynajmniej jednego z poniższych typów przeszukiwacz wybrano opcję sprawdzania bieżących referencji podczas przetwarzania zapytania:
  - Content Edition (w przypadku niektórych typów repozytoriów)
  - Domino Document Manager
  - Notes
  - QuickPlace
  - WebSphere Portal
  - system plików Windows

W profilu wymieniono wszystkie domeny dostępne dla aplikacji wyszukującej, które wymagają sprawdzania referencji użytkowników podczas przetwarzania zapytań. Użytkownicy mogą określić, dla których domen mają być przechowywane referencje. Jeśli użytkownik nie określi referencji dla domeny, dokumenty ze źródeł danych w tej domenie będą wykluczone z wyników wyszukiwania.

Gdy nie jest używane zarządzanie tożsamościami programu WebSphere II OmniFind Edition, aplikacja wyszukująca musi udostępnić łańcuch kontekstu zabezpieczeń użytkownika (USC) podczas realizowania zapytań kierowanych do domen wymagających sprawdzania bieżących referencji.

## **Konfigurowanie zarządzania tożsamościami**

Profile użytkowników można zapisywać, korzystając z zarządzania tożsamościami programu WebSphere II OmniFind Edition. Profile umożliwiają użytkownikom wyszukiwanie w domenach, w których podczas przetwarzania zapytań wymagane jest sprawdzanie referencji użytkowników.

### **Zagadnienia wstępne**

Aby konfigurować opcje zarządzania tożsamościami, wymagana jest rola administratora wyszukiwania korporacyjnego.

### **Informacje o tym zadaniu**

Jeśli podczas konfigurowania opcji zabezpieczeń na poziomie dokumentu dla przeszukiwacza zostanie określony wymóg sprawdzania referencji użytkowników podczas przetwarzania zapytań, użytkownicy muszą udostępniać swoje referencje, kiedy kierują zapytanie do domeny wymagającej ich sprawdzania. Za pomocą zarządzania tożsamościami programu WebSphere II OmniFind Edition użytkownicy mogą tworzyć swoje profile i rejestrować

referencje używane podczas logowania się do zabezpieczonych domen. Referencje są zaszyfrowane i znajdują się w zabezpieczonej bazie danych, którą zarządza program WebSphere II OmniFind Edition.

Serwery wyszukiwania używają zapisanych referencji do uwierzytelniania użytkowników, którzy przeszukują zabezpieczone domeny. Gdy referencje są niepoprawne, dokumenty znajdujące się w zabezpieczonej domenie są wykluczane z wyników wyszukiwania.

### Procedura

Aby skonfigurować zarządzanie tożsamościami:

1. Kliknij przycisk **Zabezpieczenia**, aby otworzyć widok Zabezpieczenia.
2. Na stronie Aplikacje wyszukujące kliknij opcję **Konfiguruj zarządzanie tożsamościami**.
3. Na stronie Konfigurowanie zarządzania tożsamościami zaznacz pole wyboru umożliwiające programowi WebSphere II OmniFind Edition zarządzanie referencjami użytkowników znajdującymi się w ich profilach.
4. Kliknij przycisk **OK**.

## Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów systemu plików Windows

Aby włączyć opcję sprawdzania bieżących referencji podczas wyszukiwania przez użytkownika dokumentów przeszukanych przez przeszukiwacz system plików Windows, należy skonfigurować konto domeny zarówno na serwerze przeszukiwacza, jak i na serwerze Microsoft Windows.

Podczas konfigurowania przeszukiwacza system plików Windows należy określić, czy przeszukiwane mają być podkatalogi na komputerze lokalnym, czy podkatalogi na komputerze zdalnym. Jeśli zabezpieczenia kolekcji są włączone, można też określić opcje kontroli dostępu do dokumentów w przeszukiwanych podkatalogach.

W przypadku wymuszenia kontroli dostępu polegającej na sprawdzaniu bieżących referencji użytkownika podczas wprowadzania przez niego zapytania należy także zadbać o poprawną konfigurację kont domen. Wymagania dotyczące kont domen dla plików, które zostały przeszukane na komputerze lokalnym, różnią się od wymagań dla plików, które zostały przeszukane na zdalnym serwerze Windows.

### Sprawdzanie poprawności za pomocą lokalnych danych dotyczących kontroli dostępu

Sprawdzając bieżące referencje użytkownika, system korzysta zarówno z informacji o lokalnym koncie użytkownika, jak i z informacji o koncie domeny (jeśli komputer należy do domeny Windows). Aby możliwe było sprawdzenie referencji podczas przetwarzania zapytania, wśród informacji na temat zabezpieczeń dla dokumentów, które mają być wyszukiwane, muszą się znajdować obie nazwy użytkownika.

#### Konta lokalne

W przypadku konta lokalnego nazwa użytkownika ma następujący format:

*NAZWA KOMPUTERA\NAZWA UŻYTKOWNIKA*

Aby się zalogować, użytkownicy podają tylko swoje nazwy, ale do właściwego określenia uprawnień użytkownika systemu Windows wykorzystywana jest pełna nazwa. Na przykład jeśli nazwą lokalnego konta użytkownika jest abcuser, pełną nazwą konta może być WINSERVER1\abcuser.

Kiedy użytkownicy korzystają z aplikacji wyszukującej i konfiguruje profile używane do wyszukiwania zabezpieczonych dokumentów w systemie lokalnym, muszą podawać nazwę użytkownika, której używają do logowania się w systemie Windows (np. abcuser).

### **Konta domen**

W przypadku konta domeny nazwa użytkownika ma następujący format:

*NAZWA\_DOMENY\NAZWA\_UŻYTKOWNIKA*

Aby się zalogować, użytkownicy podają te informacje w następującym formacie:

*NAZWA\_UŻYTKOWNIKA@NAZWA\_DOMENY*

Na przykład jeśli podczas konfigurowania praw dostępu użytkownika do pliku wybrana zostanie domena WIN1\abcuser, konto jest wyświetlane jako abcuser@win1.firma.com.

Kiedy użytkownicy korzystają z aplikacji wyszukującej i konfiguruje profile umożliwiające im wyszukiwanie dokumentów w zabezpieczonej domenie, muszą podawać nazwę użytkownika, której używają do logowania się w systemie Windows (np. abcuser@win1.firma.com).

Aby możliwe było wymuszenie sprawdzania bieżących referencji na komputerach lokalnych, konta użytkowników używane przez serwer przeszukiwacza muszą mieć następujące uprawnienia użytkowników Windows. (Do określania uprawnień użytkowników służą narzędzia administracyjne systemu Windows: **Narzędzia administracyjne** → **Zasady zabezpieczeń lokalnych** → **Zasady lokalne** → **Przypisywanie praw użytkownika lokalnego**).

- Identyfikator użytkownika, którym posługuje się serwer przeszukiwacza, musi mieć ustawione prawo **Działanie jako element systemu operacyjnego**. (Prawo to konfigurowane jest dla użytkownika administracyjnego wyszukiwania korporacyjnego na serwerze przeszukiwacza podczas instalowania programu WebSphere Information Integrator OmniFind Edition).
- Użytkownicy muszą mieć uprawnienie **Logowanie lokalne**.

### **Sprawdzanie poprawności za pomocą danych dotyczących kontroli dostępu do domeny zdalnej**

W systemie operacyjnym Windows każdy katalog o nazwie rozpoczynającej się od łańcucha *\\nazwa\_serwera* uważany jest za katalog zdalny. Na przykład:

*\\software\utilities\IBM*

Aby uzyskać dostęp do katalogu zdalnego, użytkownicy muszą podawać swoje nazwy w następującym formacie:

*NAZWA\_UŻYTKOWNIKA@NAZWA\_DOMENY*

Kiedy użytkownicy korzystają z aplikacji wyszukującej i konfiguruje profile umożliwiające im wyszukiwanie zabezpieczonych dokumentów w systemie zdalnym, muszą podawać nazwę użytkownika, której używają w celu uzyskania dostępu do zdalnego systemu Windows (np. abcuser@win1.firma.com).

Aby możliwe było wymuszenie sprawdzania bieżących referencji na komputerach zdalnych, konta użytkowników muszą mieć następujące uprawnienia użytkowników systemu Windows. (Do określania uprawnień użytkowników służą narzędzia administracyjne systemu Windows: **Narzędzia administracyjne** → **Zasady zabezpieczeń domeny**).

- Serwer przeszukiwacza i serwer Windows, który ma być przeszukiwany, muszą należeć do tej samej domeny.
- Identyfikator użytkownika, którym posługuje się serwer przeszukiwacza, musi mieć ustawione prawo **Działanie jako element systemu operacyjnego**. (Prawo to skonfigurowane jest dla użytkownika administracyjnego wyszukiwania korporacyjnego na serwerze przeszukiwacza podczas instalowania programu WebSphere Information Integrator OmniFind Edition).
- Użytkownicy muszą mieć uprawnienie **Logowanie w trybie wsadowym**.

#### Pojęcia pokrewne

“Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań” na stronie 185

Jeśli podczas tworzenia kolekcji włączono opcje zabezpieczeń, w niektórych typach domen możliwe jest sprawdzanie poprawności bieżących referencji użytkownika podczas wprowadzania zapytania.

“przeszukiwacze systemu plików Windows” na stronie 76

Aby do kolekcji wyszukiwania korporacyjnego włączyć dokumenty zapisane w systemie plików Microsoft Windows, należy skonfigurować przeszukiwacz system plików Windows.

## Wymuszanie zabezpieczeń na poziomie dokumentu dla dokumentów produktu Lotus Domino

Jeśli na serwerze Lotus Notes, który ma być przeszukiwany, używany jest protokół NRPC (Notes remote procedure call), należy skonfigurować serwer przeszukiwacza w sposób powodujący wymuszenie zabezpieczeń na poziomie dokumentu.

Aby wymusić zabezpieczenia na poziomie dokumentu w dokumentach, które zostały przeszukane na serwerze Lotus Notes wykorzystującym protokół NRPC, na serwerze przeszukiwacza należy zainstalować serwer Domino. Ten serwer Domino musi należeć do domeny Domino użytkownika. Aby zainstalować i skonfigurować serwer Domino, należy postępować zgodnie z instrukcjami znajdującymi się w dokumentacji produktu Lotus Domino.

Aby umożliwić serwerom wyszukiwania sprawdzanie, czy użytkownik przeszukujący zabezpieczoną kolekcję ma uprawnienia do przeglądania dokumentów Lotus Notes spełniających kryteria wyszukiwania, należy także wykonać poniższe zadania. Dokumenty, których użytkownikowi nie wolno przeglądać, zostają usunięte z wyników wyszukiwania przed ich zwróceniem.

- “Konfigurowanie serwerów Lotus Domino Trusted Servers do sprawdzania poprawności referencji użytkownika” na stronie 191.
- Włączenie zabezpieczeń globalnych na serwerze WebSphere Application Server i skonfigurowanie aplikacji wyszukującej do korzystania z zabezpieczeń. Ten krok zapewnia, że użytkownicy będą proszeni o podanie referencji podczas próby użycia aplikacji wyszukującej. Serwery wyszukiwania mogą następnie używać tych referencji podczas sprawdzania uprawnień dostępu każdego użytkownika do dokumentów Lotus Notes.

#### Pojęcia pokrewne

“Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań” na stronie 185

Jeśli podczas tworzenia kolekcji włączono opcje zabezpieczeń, w niektórych typach domen możliwe jest sprawdzanie poprawności bieżących referencji użytkownika podczas wprowadzania zapytania.



“przeszukiwacze Notes” na stronie 43

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

## **Konfigurowanie serwerów Lotus Domino Trusted Servers do sprawdzania poprawności referencji użytkownika**

Aby wymusić zabezpieczenia dokumentów, które zostały przeszukane przez przeszukiwacz Notes używający protokołu NRPC (Notes remote procedure call), serwery Domino, które mają być przeszukiwane, należy skonfigurować jako zaufane serwery Lotus Domino.

### **Zagadnienia wstępne**

Poniższa procedura jest wymagana, jeśli planuje się wymuszenie zabezpieczeń na poziomie dokumentu podczas przeszukiwania zdalnych baz danych. W przypadku lokalnych baz danych na serwerze przeszukiwacza procedura ta nie jest konieczna.

Aby można było skonfigurować serwery zaufane, na przeszukiwaczu musi być zainstalowany serwer Domino. Ten serwer Domino musi należeć do domeny Domino użytkownika.

### **Informacje o tym zadaniu**

Podczas konfigurowania opcji zabezpieczeń na poziomie dokumentu dla przeszukiwacza Notes określa się, czy ma być stosowana kontrola dostępu polegająca na sprawdzaniu bieżących referencji użytkownika podczas wprowadzania przez niego zapytania. Aby możliwe było wymuszenie zabezpieczeń tego typu, serwery Domino przeznaczone do przeszukiwania muszą być zaufanymi serwerami Lotus Domino.

Gdy użytkownicy przeszukują domeny wymagające sprawdzenia ich bieżących referencji, serwer zaufany umożliwi identyfikatorowi serwera Domino przełączenie kontekstu na bieżący identyfikator użytkownika. Baza danych Domino zostaje otwarta w taki sposób, jakby otworzył ją bieżący użytkownik, i stosowane są wszystkie informacje z listy kontroli dostępu do bazy danych, które odnoszą się do tego użytkownika.

Możliwość przełączenia kontekstu w ten sposób jest zwykle dostępna wyłącznie dla baz danych zapisanych w katalogu data lokalnego serwera Domino. Począwszy od oprogramowania Lotus Domino w wersji 6.5.1 możliwość taką zapewnia serwer zaufany. W celu skonfigurowania serwera zaufanego administrator Domino określa serwery Domino, które mają mieć możliwość wykonywania operacji wymagających szczególnej uwagi z punktu widzenia bezpieczeństwa, takich jak działanie w imieniu innego użytkownika podczas uzyskiwania dostępu do baz danych ze zdalnego komputera.

### **Procedura**

Aby skonfigurować serwer zaufany, należy wykonać następujące czynności na wszystkich serwerach Domino przeszukiwanych za pomocą przeszukiwacza Notes:

1. Na serwerze Domino otwórz klienta Lotus Domino Administrator za pomocą pliku identyfikacyjnego administratora domeny Domino.
2. Wybierz opcję **Plik** → **Otwórz serwer**.
3. Wpisz nazwę serwera Domino, dla którego chcesz aktywować możliwości serwera zaufanego.
4. Wybierz kartę **Konfiguracja**.
5. Rozwiń obiekt **Serwer**, wybierz dokument **Bieżący serwer** i kliknij przycisk **Edytuj serwer**.
6. Wybierz kartę **Zabezpieczenia**, przewiń dokument do końca, aby znaleźć pozycję **Serwery zaufane**, i kliknij przycisk strzałki w dół.

7. Podaj jedną z następujących opcji:

**LokalneSerweryDomeny**

Wybierz tę opcję, jeśli wszystkie serwery w domenie Domino mają być traktowane jako serwery zaufane.

*nazwa\_serwera*

Podaj nazwę serwera Domino, który ma mieć możliwość przeszukiwania i wyszukiwania jako serwer zaufany.

Jeśli serwer Domino, który ma być przeszukiwany, znajduje się w innej domenie Domino, należy podać jego nazwę lub wybrać grupę **InneSerweryDomeny**. Należy też wykonać procedury Domino dotyczące certyfikacji pliku identyfikacyjnego serwera Domino używanego przez program WebSphere II OmniFind Edition z drugą domeną Domino. Informacje na temat tych procedur można znaleźć w dokumentacji serwera Domino.

8. Kliknij przycisk **Zapisz i zamknij**, aby zapisać zmiany.

9. Zatrzymaj i zrestartuj zdalne serwery Domino, które skonfigurowano jako serwery zaufane.

**Pojęcia pokrewne**

“Sprawdzanie poprawności bieżących referencji podczas przetwarzania zapytań” na stronie 185

Jeśli podczas tworzenia kolekcji włączono opcje zabezpieczeń, w niektórych typach domen możliwe jest sprawdzanie poprawności bieżących referencji użytkownika podczas wprowadzania zapytania.

“przeszukiwacze Notes” na stronie 43

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

---

## Wyłączanie zabezpieczeń na poziomie dokumentu

Użytkownikom można umożliwić wyszukiwanie w kolekcji niezależnie od praw dostępu powiązanych z dokumentami w indeksie. W przypadku przeszukiwaczy obsługujących sprawdzanie bieżących referencji można pozwolić użytkownikom na wyszukiwanie w kolekcji także bez sprawdzania bieżących praw dostępu podczas przetwarzania zapytania.

### Zagadnienia wstępne

Aby skonfigurować opcje zabezpieczeń na poziomie dokumentu, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

### Ograniczenia

Opcje zabezpieczeń na poziomie dokumentu można określać tylko wtedy, gdy podczas tworzenia kolekcji zostały włączone zabezpieczenia.

### Informacje o tym zadaniu

Podczas konfigurowania przeszukiwacza istnieje możliwość powiązania tokenów zabezpieczeń z przeszukiwanymi dokumentami. Aplikacje wyszukujące mogą używać tych - przechowywanych w indeksie - tokenów do wymuszania kontroli dostępu podczas wyszukiwania w kolekcji przez użytkowników. W przypadku niektórych przeszukiwaczy można też określić, że podczas wprowadzania zapytania należy sprawdzać bieżące prawa dostępu związane z dokumentami znajdującymi się w ich własnych repozytoriach.

Aby usunąć ograniczenia związane z bezpieczeństwem, można określić, że serwery wyszukiwania mają ignorować wszystkie tokeny zabezpieczeń przekazywane z zapytaniem. Można także umożliwić użytkownikom wyszukiwanie dokumentów bez porównywania ich referencji z bieżącymi prawami dostępu.

Do testowania nowej kolekcji lub rozwiązywania problemów z aplikacją wyszukującą przydatne może być czasowe wyłączenie zabezpieczeń na poziomie dokumentu.

### Procedura

Aby wyłączyć kontrolę dostępu na poziomie dokumentu:

1. Przejdź do edycji kolekcji, wybierz stronę Ogólne, a następnie kliknij opcję **Konfiguruj zabezpieczenia na poziomie dokumentu**.
2. Jeśli tokeny zabezpieczeń powiązane przez przeszukiwacz z dokumentami nie mają być wykorzystywane podczas wyszukiwania przez użytkowników w kolekcji, na stronie Zabezpieczenia na poziomie dokumentu zaznacz pole wyboru **Ignoruj w indeksie prawa dostępu na poziomie dokumentu**.

Przeszukiwacze będą nadal dodawały tokeny zabezpieczeń do dokumentów, ale serwery wyszukiwania zignorują te tokeny i umożliwią użytkownikom wyszukiwanie we wcześniej zabezpieczonych dokumentach.

3. Jeśli podczas wprowadzania przez użytkowników zapytań bieżące prawa dostępu powiązane z dokumentami w ich rodzimych repozytoriach nie mają być sprawdzane, zaznacz pole wyboru **Nie sprawdzaj poprawności bieżących referencji podczas przetwarzania zapytania**. To pole wyboru jest dostępne tylko dla dokumentów przeszukanych przez przeszukiwacze, które obsługują tę możliwość.

Jeśli to pole wyboru zostanie zaznaczone, inne opcje zabezpieczeń na poziomie dokumentu pozostają aktywne. Na przykład jeśli podczas konfigurowania przeszukiwacza została określona opcja zapisywania praw dostępu w indeksie, to prawa te są stosowane, o ile nie zostanie zaznaczone także pole wyboru **Ignoruj w indeksie prawa dostępu na poziomie dokumentu**.

4. Kliknij przycisk **OK**.



---

# Integracja wyszukiwania korporacyjnego z portalem WebSphere Portal

Można rozszerzyć możliwości wyszukiwania przy użyciu portalu IBM WebSphere Portal przez wdrożenie portletów wyszukiwania korporacyjnego w portalu WebSphere Portal i w Centrum wyszukiwania portalu WebSphere Portal.

## Punkty integracji

Istnieje kilka sposobów integrowania portletów wyszukiwania korporacyjnego z portalem WebSphere Portal:

### Portal WebSphere Portal

W portalu WebSphere Portal istnieje pojedynczy punkt dostępu umożliwiający wzajemne oddziaływanie między aplikacjami, treścią, procesami i użytkownikami. Struktura portalu WebSphere Portal umożliwia zintegrowanie i wdrożenie nowych aplikacji nazywanych portletami bez oddziaływania na pozostałe aplikacje używane w portalu.

Jeśli w portalu WebSphere Portal zostanie wdrożony portlet wyszukiwania korporacyjnego, można będzie przeszukiwać kolekcje wyszukiwania korporacyjnego przy użyciu interfejsu portalu WebSphere Portal. Definiując odpowiednio ustawienia konfiguracyjne portalu WebSphere Portal, można zapewnić, że wygląd i działanie portletu wyszukiwania korporacyjnego będzie takie samo, jak pozostałych portletów w używanym środowisku portalu WebSphere Portal.

### Mechanizm wyszukiwania środowiska Portal

Mechanizm wyszukiwania produktu WebSphere Portal zajmuje się przeszukiwaniem serwisów WWW. Portlety administracyjne umożliwiają administratorom tworzenie poindeksowanych kolekcji, a portlety wyszukiwania są używane przez użytkowników do ich przeszukiwania.

Jeśli używany jest portal WebSphere Portal, wersja 5.0.2 lub nowsza, można przeprowadzić migrację informacji konfiguracyjnych kolekcji i systematyk mechanizmu wyszukiwania środowiska Portal do wyszukiwania korporacyjnego.

Jeśli używany jest portal WebSphere Portal, wersja 5.1 lub nowsza, kolekcje wyszukiwania korporacyjnego mogą być przeszukiwane przy użyciu portletu wyszukiwania dokumentów mechanizmu wyszukiwania środowiska Portal. Użycie właściwości konfiguracyjnych umożliwia administratorom wybranie w prosty sposób jednego z dwóch mechanizmów wyszukiwania.

### Centrum wyszukiwania portalu WebSphere Portal

Centrum wyszukiwania portalu WebSphere Portal stanowi centralny punkt początkowy dla wyszukiwania we wszystkich źródłach, które zostały udostępnione dla wyszukiwania przy użyciu portalu WebSphere Portal. Centrum wyszukiwania i uniwersalny portlet wyszukiwania umożliwiają użytkownikom przeszukiwanie treści portalu WebSphere Portal i wszystkich pozostałych kolekcji zarejestrowanych przez administratora przy użyciu Centrum wyszukiwania.

Interfejs Centrum wyszukiwania składa się z wielu stron. Istnieje możliwość przeszukiwania wszystkich dostępnych kolekcji przy użyciu jednej wspólnej strony, lub możliwość wybrania strony służącej do wyszukiwania w pojedynczej kolekcji. Na przykład dostępna jest strona związana z indeksami mechanizmu wyszukiwania środowiska Portal i strona związana z bibliotekami zarządzania dokumentami w portalu.

Aby umożliwić wyszukiwanie w kolekcjach wyszukiwania korporacyjnego z poziomu Centrum wyszukiwania, produkt WebSphere Information Integrator OmniFind Edition (WebSphere II OmniFind Edition) udostępnia adapter oraz portlet rejestracji. Adapter umożliwia dodanie karty wyszukiwania korporacyjnego do interfejsu Centrum wyszukiwania, a portlet rejestracji - zarejestrowanie w Centrum wyszukiwania portletu wyszukiwania korporacyjnego.

Adapter oraz portlet rejestracji muszą zostać zainstalowane, aby możliwe było korzystanie z Centrum wyszukiwania do wyszukiwania danych w kolekcjach wyszukiwania korporacyjnego.

### **przeszukiwacz WebSphere Portal**

Aplikacja korporacyjna ESPACServer.ear jest instalowana na serwerach wyszukiwania podczas instalacji produktu WebSphere II OmniFind Edition. Po wdrożeniu tej aplikacji korporacyjnej na serwerze WebSphere Portal można użyć konsoli administracyjnej wyszukiwania korporacyjnego do skonfigurowania przeszukiwacza WebSphere Portal i dodania serwisów WebSphere Portal do kolekcji wyszukiwania korporacyjnego.

## **Korzyści wynikające z integracji**

Wyszukiwanie korporacyjne rozszerza środowisko wyszukiwania portalu WebSphere Portal, ponieważ umożliwia wyszukiwanie w szerszym zakresie typów źródeł danych. Portlet wyszukiwania korporacyjnego umożliwia wyszukiwanie danych w serwisach WWW oraz dodatkowo we wszystkich innych typach źródeł danych obsługiwanych przez produkt WebSphere II OmniFind Edition.

Użycie wyszukiwania korporacyjnego wiąże się także z korzyściami dotyczącymi skalowalności. Mechanizm wyszukiwania środowiska Portal jest przydatny w przypadku małych i średnich przedsiębiorstw, w których do obsłużenia obciążenia związanego z wyszukiwaniem i pobieraniem wystarczy pojedynczy serwer. Aby możliwe było obsłużenie obciążeń spotykanych w dużym przedsiębiorstwie, w wyszukiwaniu korporacyjnym obciążenie jest rozdzielane między cztery serwery (jeden służy do przeszukiwania danych, drugi - do analizowania i indeksowania danych, a pozostałe dwa obsługują procesy wyszukiwania i pobierania).

## **Przegląd informacji o wdrażaniu portletów**

To, które portlety mogą zostać użyte razem z wyszukiwaniem korporacyjnym, zależy od wersji używanego portalu WebSphere Portal:

### **WebSphere Portal, wersja 5.0.2**

Możliwe jest wdrożenie portletu wyszukiwania korporacyjnego i używanie go do przeszukiwania kolekcji wyszukiwania korporacyjnego. Portlet ten może współistnieć z portletami mechanizmu wyszukiwania środowiska Portal.

### **WebSphere Portal, wersja 5.1 lub nowsza**

- Możliwe jest wdrożenie portletu wyszukiwania korporacyjnego i używanie go do przeszukiwania kolekcji wyszukiwania korporacyjnego. Portlet ten może współistnieć z portletami mechanizmu wyszukiwania środowiska Portal.
- Po wdrożeniu portletu wyszukiwania korporacyjnego można skonfigurować portlet wyszukiwania dokumentów w mechanizmie wyszukiwania środowiska Portal pod kątem przeszukiwania kolekcji wyszukiwania korporacyjnego.
- Po zarejestrowaniu portletu wyszukiwania korporacyjnego w Centrum wyszukiwania portalu WebSphere Portal można go używać do przeszukiwania kolekcji wyszukiwania korporacyjnego. Użytkownik może wybrać stronę służącą do wyszukiwania tylko w kolekcji wyszukiwania korporacyjnego lub wprowadzić

zapytanie służące do wyszukiwania zarówno w kolekcji wyszukiwania korporacyjnego, jak i we wszystkich pozostałych kolekcjach dostępnych w Centrum wyszukiwania.

#### **Pojęcia pokrewne**

“Migracja z portalu WebSphere Portal do wyszukiwania korporacyjnego” na stronie 203  
W wyszukiwaniu korporacyjnym dostępny jest kreator migracji, za pomocą którego można dokonać migracji systematyk i kolekcji z portalu IBM WebSphere Portal do wyszukiwania korporacyjnego.

#### **Zadania pokrewne**

“Wdrażanie aplikacji korporacyjnej dla przeszukiwacza WebSphere Portal” na stronie 74  
Przed utworzeniem przeszukiwacza WebSphere Portal konieczne jest wdrożenie aplikacji korporacyjnej ESPACServer.ear na serwerze WebSphere Portal.

---

## **Zabezpieczenia na poziomie dokumentu przy użyciu mechanizmu wyszukiwania środowiska Portal**

Gdy użytkownicy przeszukują kolekcje wyszukiwania korporacyjnego, mechanizm wyszukiwania środowiska IBM WebSphere Portal może zostać wykorzystany do wymuszenia zabezpieczeń na poziomie dokumentu.

Jeśli przeszukiwacz wyszukiwania korporacyjnego przypisze związany z identyfikatorem grupy token zabezpieczeń do przeszukiwanych dokumentów, a użytkownik skonfiguruje portlet wyszukiwania dokumentów dla mechanizmu wyszukiwania środowiska Portal do przeszukiwania kolekcji wyszukiwania korporacyjnego, wówczas mechanizm wyszukiwania środowiska Portal będzie mógł utworzyć identyfikator grupy dla zalogowanego użytkownika i przekazać związany z tym identyfikatorem grupy token zabezpieczeń w zapytaniu dla wyszukiwania korporacyjnego. Token zabezpieczeń zapewnia, że w wynikach wyszukiwania zwrócone zostaną tylko dokumenty, do których określony użytkownik ma odpowiednie uprawnienia.

Jeśli przeszukiwacz przypisze do dokumentów inny typ tokenu zabezpieczeń (taki jak identyfikator użytkownika lub rola użytkownika), a użytkownik chce wymusić użycie zabezpieczeń na poziomie dokumentu podczas przeszukiwania kolekcji wyszukiwania korporacyjnego, wówczas konieczne jest utworzenie niestandardowego portletu wyszukiwania. Mechanizm wyszukiwania środowiska Portal tworzy tokeny zabezpieczeń tylko dla identyfikatorów grup.

#### **Pojęcia pokrewne**

“Aplikacje wyszukujące dla wyszukiwania korporacyjnego” na stronie 155  
Aplikacja wyszukująca umożliwia wyszukiwanie w kolekcjach i źródłach zewnętrznych w systemie wyszukiwania korporacyjnego. Liczba możliwych do utworzenia aplikacji wyszukujących nie jest ograniczona, a pojedyncza aplikacja wyszukująca pozwala przeszukać dowolną liczbę kolekcji i źródeł zewnętrznych.

“Zabezpieczenia na poziomie dokumentu” na stronie 184

Jeśli podczas tworzenia kolekcji zostały włączone zabezpieczenia, można skonfigurować opcje zabezpieczeń na poziomie dokumentu. Mechanizm zabezpieczeń na poziomie dokumentu jest odpowiedzialny za to, aby użytkownicy wyszukujący w kolekcji mieli dostęp tylko do tych dokumentów, względem których przydzielono im stosowne uprawnienia.

---

## Wdrażanie portletu Wyszukiwanie

Portlet Wyszukiwanie (ESSearchPortlet.war) umożliwia korzystanie z mechanizmu wyszukiwania produktu WebSphere Portal do wyszukiwania danych w kolekcjach wyszukiwania korporacyjnego.

### Zagadnienia wstępne

Portlet Wyszukiwanie musi zostać wdrożony dla wyszukiwania korporacyjnego na serwerze WebSphere Portal, aby można było skonfigurować mechanizm wyszukiwania portalu do korzystania z tego portletu.

### Informacje o tym zadaniu

Po wdrożeniu portletu Wyszukiwanie można w dalszym ciągu korzystać z mechanizmu wyszukiwania portalu do wyszukiwania poindeksowanych danych w produkcie WebSphere Portal. Aby zapewnić użytkownikom możliwość przeszukiwania kolekcji wyszukiwania korporacyjnego, administrator portalu WebSphere Portal powinien dokonać edycji odpowiednich właściwości w ustawieniach konfiguracyjnych mechanizmu wyszukiwania środowiska Portal. Możliwość przełączania się między tymi funkcjami wyszukiwania pozwala na korzystanie z rozwiązania najbardziej odpowiedniego w konkretnej sytuacji.

Plik ESSearchPortlet.war jest instalowany na serwerach wyszukiwania w katalogu ES\_INSTALL\_ROOT/bin podczas instalacji produktu WebSphere II OmniFind Edition. Domyślne ścieżki instalacyjne to:

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

### Procedura

Aby wdrożyć portlet Wyszukiwanie w produkcie WebSphere Portal:

1. Zainstaluj portlet:
  - a. Zaloguj się na serwerze WebSphere Portal przy użyciu identyfikatora administratora serwera WebSphere Portal.
  - b. Wybierz opcję **Administracja** z paska narzędzi.
  - c. Wybierz opcję **Portlety** w obszarze nawigacyjnym po lewej stronie, a następnie wybierz opcję **Instaluj** z menu **Portlety**.
  - d. Kliknij przycisk **Przełóżaj**, wybierz z systemu plik ESSearchPortlet.war i kliknij przycisk **Dalej**.
  - e. Na następnej stronie kliknij przycisk **Instaluj**. Po zainstalowaniu portletu wyświetlany jest następujący komunikat: APIN00051: Portlety zostały pomyślnie zainstalowane.
2. Zmodyfikuj parametry portletu:
  - a. Wybierz opcję **Zarządzaj portletami** w obszarze nawigacyjnym po lewej stronie, a następnie wybierz opcję **Wyszukiwanie korporacyjne IBM**.
  - b. Po wyświetleniu nowych ikon po prawej stronie okna wyboru wybierz opcję **Modyfikowanie parametrów**.
  - c. Na liście parametrów portletu zmodyfikuj następujące parametry:

<b>port</b>	Ustaw tę wartość na numer portu serwera WWW dla serwera wyszukiwania produktu WebSphere II OmniFind Edition. Wartością domyślną jest 80.
<b>applicationName</b>	Ustaw tę wartość na nazwę poprawnej aplikacji wyszukującej wyszukiwania



korporacyjnego (nazwy dostępnych aplikacji wyszukiwujących są wymienione na stronie Aplikacje wyszukiwujące w konsoli administracyjnej wyszukiwania korporacyjnego). Wartość domyślna to Default i oznacza ona aplikację wyszukiwującą, która jest zawsze skonfigurowana na serwerach wyszukiwania.

#### **hostname**

Ustaw tę wartość na pełną nazwę hosta odpowiadającą serwerowi wyszukiwania produktu WebSphere II OmniFind Edition.

- d. Kliknij przycisk **Zapisz**, aby zapisać wprowadzone zmiany, a następnie kliknij przycisk **Anuluj**, aby zamknąć stronę Modyfikowanie parametrów.
3. Utwórz stronę portletu dla portletu:
  - a. Wybierz opcję **Interfejs użytkownika portalu** w obszarze nawigacyjnym po lewej stronie, a następnie wybierz opcję **Zarządzaj stronami**.
  - b. Wybierz odsyłacz **Mój portal**.
  - c. Kliknij opcję **Nowa strona**, określ tytuł strony i kliknij przycisk **OK**. Po utworzeniu strony wyświetlany jest następujący komunikat: APPR00101: *tytuł\_strony* została pomyślnie utworzona.
  - d. Kliknij przycisk **OK**. Utworzona strona zostanie wyświetlona na liście dostępnych stron.
4. Dodaj do strony portlet Wyszukiwanie:
  - a. Wybierz ikonę edycji (ołówka) odpowiadającą utworzonej stronie.
  - b. Kliknij opcję **Dodaj portlety**.
  - c. W polu tekstowym **Szukaj** wpisz Enterprise i kliknij przycisk **Szukaj**.
  - d. Zaznacz pole wyboru obok pozycji **Wyszukiwanie korporacyjne IBM** i kliknij przycisk **OK**.
  - e. Kliknij przycisk **Gotowe**. Ten krok oznacza, że portlet Wyszukiwanie został pomyślnie zainstalowany i dodany do strony produktu WebSphere Portal.
5. Przejdź do portletu:
  - a. Wybierz opcję **Mój portal** na pasku nawigacyjnym w górnej części strony.
  - b. Wybierz stronę o tytule określonym podczas dodawania portletu do systemu.  
Przy pierwszej próbie otwarcia strony portletu jej wyświetlanie może być powolne, ponieważ system musi skompilować pliki JSP (Java Server Pages) portletu.

---

## **Konfigurowanie portletu Wyszukiwanie i przeglądanie produktu WebSphere Portal dla wyszukiwania korporacyjnego**

Produkt WebSphere Portal udostępnia portlet, przy użyciu którego można wyszukiwać i przeglądać rodzime kolekcje produktu WebSphere Portal. Portlet ten można skonfigurować do wyszukiwania danych w kolekcjach wyszukiwania korporacyjnego.

### **Procedura**

Aby skonfigurować portlet Wyszukiwanie i przeglądanie do wyszukiwania danych w kolekcjach wyszukiwania korporacyjnego:

1. Zatrzymaj instancję serwera WebSphere\_Portal.
2. Skopiuj następujące pliki z katalogu ES\_INSTALL\_ROOT/lib produktu WebSphere II OmniFind Edition do katalogu /WebSphere/PortalServer/shared/app na serwerze WebSphere Portal:  
esapi.jar  
siapi.jar

3. Uruchom instancję serwera WebSphere\_Portal i zaloguj się na serwerze WebSphere Portal przy użyciu identyfikatora administratora serwera WebSphere Portal.
4. Skopiuj portlet:
  - a. Kliknij opcję **Administracja** w prawym górnym rogu, a następnie rozwiń obiekt **Zarządzanie portletami** i wybierz opcję **Portlety**.
  - b. Wyszukaj słowo **search**.
  - c. Kliknij opcję **Kopiuj portlet** dla portletu Wyszukiwanie i przeglądanie, podaj unikalną nazwę portletu, a następnie kliknij przycisk **OK**.
5. Skonfiguruj portlet:
  - a. Kliknij ikonę **Konfiguruj portlet** obok określonej nazwy portletu.
  - b. Kliknij strzałkę w prawo, aby przejść do drugiej strony z parametrami.
  - c. W obszarze **Nowy parametr** wpisz **ApplicationInfoId**, określ nazwę poprawnej aplikacji wyszukującej produktu WebSphere II OmniFind Edition w polu **Nowa wartość**, a następnie kliknij przycisk **Dodaj**, aby dodać nowy parametr.  
 Nazwy dostępnych aplikacji wyszukujących są wymienione na stronie Aplikacje wyszukujące w konsoli administracyjnej wyszukiwania korporacyjnego. Wartość domyślna to Default i oznacza ona aplikację wyszukującą, która jest zawsze skonfigurowana na serwerach wyszukiwania.
  - d. Usuń następujące parametry:
 

```
ApplicationInfoId_EXAMPLE
QueryFactoryImp
BrowseFactoryImp
SearchFactoryImp
IndexName
EJB_Example
IIOP_URL_Example
SOAP_URL_Example
EJB
SOAP_URL
IIOP_URL
```
  - e. Dodaj następujące parametry i określ wartości przedstawione poniżej. Kliknij przycisk **Dodaj**, aby dodać nazwę i wartość każdego parametru. W nazwach parametrów rozróżniana jest wielkość liter.
 

```
IMPLEMENT = ES
ApplicationInfoId = nazwa_aplikacji_wyszukujacej (np. Default)
SearchFactoryImp = com.ibm.es.api.search.RemoteSearchFactory
BrowseFactoryImp = com.ibm.es.api.browse.RemoteBrowseFactory
hostname = nazwa_hosta_serwera_wyszukiwania (np. omnifind.ibm.com)
port = numer_portu_serwera_wyszukiwania (np. 80)
```
6. Przypisz nową kopię portletu Wyszukiwanie i przeglądanie do strony portalu. Szczegółowe instrukcje można znaleźć w dokumentacji dotyczącej administrowania produktem WebSphere Portal.
7. Przejdź do portletu:
  - a. Wybierz opcję **Mój portal** na pasku nawigacyjnym w górnej części strony portalu.
  - b. Wybierz stronę o nazwie określonej podczas kopiowania portletu w kroku 4.

---

## Instalowanie adaptera wyszukiwania korporacyjnego dla Centrum wyszukiwania

Zanim możliwe będzie korzystanie z Centrum wyszukiwania produktu WebSphere Portal do przeprowadzania wyszukiwania w kolekcjach wyszukiwania korporacyjnego konieczne jest zainstalowanie adaptera dla wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Adapter ESSearchAdapter.ear dodaje stronę wyszukiwania korporacyjnego do interfejsu Centrum wyszukiwania. Po zainstalowaniu tej aplikacji oraz portletu rejestracji możliwe jest dodawanie kolekcji wyszukiwania korporacyjnego do Centrum wyszukiwania i wyszukiwanie w tych kolekcjach przy użyciu uniwersalnego portletu wyszukiwania.

Plik ESSearchAdapter.ear jest instalowany na serwerach wyszukiwania w katalogu ES\_INSTALL\_ROOT/bin podczas instalacji produktu WebSphere II OmniFind Edition.

Domyślne ścieżki instalacyjne to:

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

## Procedura

Aby zainstalować adapter dla wyszukiwania korporacyjnego w Centrum wyszukiwania:

1. Zatrzymaj instancję serwera WebSphere\_Portal.
2. Jeśli nie jest jeszcze uruchomiona, uruchom instancję serwera WebSphere Application Server o nazwie server1.
3. Skopiuj następujące pliki z katalogu ES\_INSTALL\_ROOT/lib produktu WebSphere II OmniFind Edition do katalogu /WebSphere/PortalServer/shared/app na serwerze WebSphere Portal:

```
esapi.jar
siapi.jar
```

Jeśli pojawi się pytanie o zastąpienie pliku siapi.jar, należy wybrać opcję **Tak**.

4. Na serwerze WebSphere Portal uruchom konsolę administracyjną serwera WebSphere Application Server. Jeśli pojawi się odpowiednie wezwanie, zaloguj się.  
Konsolę administracyjną można otworzyć na kilka sposobów:
  - W systemie Windows należy użyć menu **Start**, aby wybrać odpowiedni program.
  - W przypadku serwera WebSphere Application Server w wersji 5 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/admin`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9090.
  - W przypadku serwera WebSphere Application Server w wersji 6 należy otworzyć przeglądarkę WWW i przejść pod adres `http://nazwa_hosta:port/ibm/console`, gdzie *nazwa\_hosta* jest nazwą hosta odpowiadającą serwerowi wyszukiwania, a *port* to numer portu konsoli administracyjnej serwera WebSphere Application Server. Zazwyczaj port konsoli administracyjnej to port 9060.
5. Kliknij przycisk **Aplikacje**, a następnie kliknij przycisk **Zainstaluj nowe aplikacje**.
6. Kliknij przycisk **Przełóżaj**, wybierz z systemu plik ESSearchAdapter.ear i kliknij dwukrotnie przycisk **Dalej**. Jeśli pojawi się ostrzeżenie na temat plików strategii, kliknij przycisk **Kontynuuj**.
7. Klikaj przycisk **Dalej** do czasu wyświetlenia strony Odwzorowanie modułów na serwery aplikacji, dokonaj poniższych wyborów i kliknij przycisk **Zastosuj**:
  - W polu **Klasy i serwery** wybierz opcję **WebSphere:cell=nazwa\_komórki, node=nazwa\_węzła, server=server1**.
  - Zaznacz pole wyboru obok modułu **ESSearchAdapterEJB**.
8. Klikaj przycisk **Dalej** do czasu wyświetlenia strony Podsumowanie, a następnie kliknij przycisk **Zakończ**.
9. Kliknij odsyłacz **Zapisz do głównej konfiguracji**, a następnie kliknij przycisk **Zapisz**, aby zapisać wprowadzone zmiany w konfiguracji serwera WebSphere Application Server.

10. Zrestartuj serwer WebSphere Portal.

---

## Instalowanie portletu rejestracji wyszukiwania korporacyjnego dla Centrum wyszukiwania

Zanim możliwe będzie korzystanie z Centrum wyszukiwania produktu WebSphere Portal do wyszukiwania danych w kolekcjach wyszukiwania korporacyjnego konieczne jest zainstalowanie portletu rejestracji dla wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Portlet rejestracji o nazwie ESSearchAdapterPortlet.war rejestruje portlet wyszukiwania korporacyjnego w Centrum wyszukiwania. Po zainstalowaniu tego portletu (oraz aplikacji adaptera dla wyszukiwania korporacyjnego) możliwe jest dodawanie kolekcji wyszukiwania korporacyjnego do Centrum wyszukiwania i wyszukiwanie danych w tych kolekcjach przy użyciu uniwersalnego portletu wyszukiwania.

Plik ESSearchAdapterPortlet.war jest instalowany na serwerach wyszukiwania w katalogu ES\_INSTALL\_ROOT/bin podczas instalacji produktu WebSphere II OmniFind Edition. Domyślne ścieżki instalacyjne to:

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

### Procedura

Aby zainstalować portlet rejestracji dla wyszukiwania korporacyjnego:

1. Zaloguj się na serwerze WebSphere Portal przy użyciu identyfikatora administratora serwera WebSphere Portal.
2. Wybierz opcję **Administracja** z paska narzędzi.
3. Wybierz opcję **Portlety** w obszarze nawigacyjnym po lewej stronie, a następnie wybierz opcję **Instaluj** z menu **Portlety**.
4. Kliknij przycisk **Przeglądaj**, wybierz z systemu plik ESSearchAdapterPortlet.war i kliknij przycisk **Dalej**.
5. Na następnej stronie kliknij przycisk **Instaluj**. Po zainstalowaniu portletu wyświetlany jest następujący komunikat: APIN0005I: Portlety zostały pomyślnie zainstalowane.
6. Przypisz portlet (o nazwie IBM WebSphere II OmniFind Edition registration portlet for enterprise search) do strony portalu. Szczegółowe instrukcje można znaleźć w dokumentacji dotyczącej administrowania produktem WebSphere Portal.

---

## Migracja z portalu WebSphere Portal do wyszukiwania korporacyjnego

W wyszukiwaniu korporacyjnym dostępny jest kreator migracji, za pomocą którego można dokonać migracji systematyk i kolekcji z portalu IBM WebSphere Portal do wyszukiwania korporacyjnego.

Aby przeprowadzić migrację systematyk i kolekcji, należy uruchomić kreatora migracji na serwerze indeksowania w wyszukiwaniu korporacyjnym. Po wykonaniu migracji systematyki można jej używać względem kolekcji wyszukiwania korporacyjnego. Wyszukiwanie korporacyjne umożliwia także administrowanie i wyszukiwanie kolekcji, które zostały poddane migracji z portalu WebSphere Portal.

W wyszukiwaniu korporacyjnym systematyka jest określana jako drzewo kategorii. Po wykonaniu migracji systematyki opartej na regułach drzewem kategorii można administrować przy użyciu konsoli administracyjnej wyszukiwania korporacyjnego. Aby w wyszukiwaniu korporacyjnym można było używać systematyk opartych na modelach, na serwerze indeksowania musi być zainstalowane oprogramowanie WebSphere Portal.

Jeśli mają zostać poddane migracji zarówno systematyki, jak i kolekcje, należy zawsze migrować najpierw pliki systematyk opartych na modelach, a dopiero później kolekcje. W przeciwnym razie nie będzie można użyć klasyfikacji opartych na modelach względem kolekcji, które zostały zmigrowane z portalu WebSphere Portal.

### Pojęcia pokrewne

“Integracja wyszukiwania korporacyjnego z portalem WebSphere Portal” na stronie 195  
Można rozszerzyć możliwości wyszukiwania przy użyciu portalu IBM WebSphere Portal przez wdrożenie portletów wyszukiwania korporacyjnego w portalu WebSphere Portal i w Centrum wyszukiwania portalu WebSphere Portal.

### Zadania pokrewne

“Konfiguracja kategorii” na stronie 93  
Dla kolekcji można utworzyć dowolną liczbę kategorii, a każda kategoria może zawierać dowolną liczbę reguł. Reguły określają, które dokumenty będą automatycznie powiązane z daną kategorią.

---

## Migrowanie systematyki opartej na modelach z portalu WebSphere Portal

Za pomocą portletu zarządzania systematykami WebSphere Portal można wybrać systematykę opartą na modelach, która ma być używana względem kolekcji wyszukiwania korporacyjnego. Migracja nowej systematyki nie wpływa na kolekcje, które migrowano już wcześniej do wyszukiwania korporacyjnego.

### Zagadnienia wstępne

Przed pierwszym uruchomieniem kreatora migracji należy zatrzymać system wyszukiwania korporacyjnego, aby dokonać zmian w plikach konfiguracyjnych. Podczas pierwszego uruchomienia kreatora migracji trzeba określić ścieżkę lokalizacji, w której zainstalowano serwer WebSphere Application Server oraz portal WebSphere Portal. W przypadku każdego następnego uruchomienia kreatora migracji system wyszukiwania korporacyjnego może być aktywny.

## Informacje o tym zadaniu

Aby przeprowadzić migrację systematyki opartej na modelach, należy wybrać i wyeksportować tę systematykę w portalu WebSphere Portal. Następnie należy użyć Kreatora migracji w wyszukiwaniu korporacyjnym, aby dokonać migracji tej systematyki.

## Procedura

Aby przeprowadzić migrację systematyki opartej na modelach z portalu WebSphere Portal do wyszukiwania korporacyjnego:

1. Wyeksportuj bieżącą systematykę opartą na modelach z portletu zarządzania systematykami w portalu WebSphere Portal. Systematyka składa się z następujących plików XML:

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

2. Skopiuj te pliki na serwer indeksowania w wyszukiwaniu korporacyjnym.
3. Na serwerze indeksowania wyszukiwania korporacyjnego zaloguj się jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
4. Przejdź do katalogu instalacyjnego wyszukiwania korporacyjnego:

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

5. Jeśli kreator migracji uruchamiany jest po raz pierwszy, wprowadź następującą komendę, aby zatrzymać system wyszukiwania korporacyjnego:

```
esadmin stop
```

6. Wprowadź następującą komendę, aby uruchomić Kreatora migracji, a następnie kliknij przycisk **Dalej**.

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

Jeśli nie jest to pierwsze uruchomienie kreatora migracji i system wyszukiwania korporacyjnego jest aktywny, należy zadbać o to, aby podczas działania kreatora nie tworzyć kolekcji wyszukiwania korporacyjnego.

7. Wybierz opcję **Importowanie plików systematyki opartej na modelach z produktu WebSphere Portal**, a następnie kliknij przycisk **Dalej**.
8. Jeśli Kreator migracji został uruchomiony po raz pierwszy, określ ścieżki, w których jest zainstalowany serwer WebSphere Application Server i portal WebSphere Portal. Kliknij przycisk **Dalej**.
9. Wyszukaj katalog, w którym znajdują się pliki systematyki opartej na modelach, wybierz pliki XML, które mają zostać poddane migracji w celu zastosowania systematyki opartej na modelach (synonyms.xml, titles.xml i treenodes.xml), a następnie kliknij przycisk **Dalej**.
10. W przypadku pierwszego uruchomienia kreatora migracji wprowadź po jego zakończeniu następującą komendę, aby zrestartować system wyszukiwania korporacyjnego:

```
esadmin start
```

Jeśli wystąpią błędy, wyszukaj odpowiednie informacje w pliku MigrationWizard.log znajdującym się w katalogu, w którym zainstalowany jest Kreator migracji.

### Pojęcia pokrewne

“Kategorie oparte na modelach” na stronie 91

Jeśli używane są kategorie oparte na modelach zdefiniowanych w systemie IBM WebSphere Portal, kategorie te mogą być używane również w wyszukiwaniu korporacyjnym.

### Zadania pokrewne

“Konfiguracja kategorii” na stronie 93

Dla kolekcji można utworzyć dowolną liczbę kategorii, a każda kategoria może zawierać dowolną liczbę reguł. Reguły określają, które dokumenty będą automatycznie powiązane z daną kategorią.

---

## Migrowanie kolekcji z portalu WebSphere Portal

Aby przeprowadzić migrację kolekcji z portalu WebSphere Portal do wyszukiwania korporacyjnego, należy przygotować kolekcje w portalu WebSphere Portal, a następnie użyć Kreatora migracji do wykonania migracji.

### Zagadnienia wstępne

Planując migrację kolekcji oraz systematyk opartych na modelach, należy przed skorzystaniem z procedury migracji kolekcji dokonać migracji plików systematyk opartych na modelach. W przeciwnym razie nie będzie można użyć klasyfikacji opartych na modelach względem kolekcji, które zostały zmigrowane z portalu WebSphere Portal.

### Procedura

Aby przeprowadzić migrację kolekcji z portalu WebSphere Portal do wyszukiwania korporacyjnego:

1. W mechanizmie wyszukiwania środowiska WebSphere Portal zatrzymaj wszystkie procesy przeszukiwaczy w kolekcjach, które mają zostać poddane migracji, i zatwierdź lub odrzuć wszystkie oczekujące dokumenty. (W wyszukiwaniu korporacyjnym nie są obsługiwane oczekujące dokumenty).
2. Dla każdej kolekcji, która ma zostać poddana migracji, użyj portletów mechanizmu wyszukiwania środowiska Portal, aby wyeksportować ustawienia do plików XML.
3. Jeśli serwer indeksowania wyszukiwania korporacyjnego jest zainstalowany na oddzielnym serwerze, skopiuj wyeksportowane pliki XML na ten serwer indeksowania.
4. Na serwerze indeksowania wyszukiwania korporacyjnego zaloguj się jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
5. Przejdź do katalogu instalacyjnego wyszukiwania korporacyjnego:

UNIX: `cd $ES_INSTALL_ROOT/bin`

Windows: `cd %ES_INSTALL_ROOT%\bin`

6. Wprowadź następującą komendę, aby uruchomić Kreatora migracji, a następnie kliknij przycisk **Dalej**.

UNIX: `./eswpsmigrate.sh`

Windows: `eswpsmigrate.bat`

7. Wybierz opcję **Migrowanie ustawień wyszukiwania z mechanizmu wyszukiwania Portalu produktu WebSphere Portal**, a następnie kliknij przycisk **Dalej**.
8. Wyszukaj katalog, w którym znajdują się wyeksportowane pliki konfiguracyjne mechanizmu wyszukiwania środowiska Portal, wybierz pliki, które mają zostać poddane

migracji, a następnie kliknij przycisk **Dalej**. Wybrane pliki konfiguracyjne zostaną przeanalizowane i sprawdzone pod kątem poprawności.

9. Dla każdej kolekcji wprowadź poniższe informacje, a następnie kliknij przycisk **Dalej**, aby rozpocząć migrowanie kolekcji do wyszukiwania korporacyjnego:

- Nazwa kolekcji, która ma być używana w wyszukiwaniu korporacyjnym.
- Kryteria określania ważności dokumentu w kolekcji. Można wyłączyć statyczny współczynnik oceniania lub zastosować czynnik oceniania określany na podstawie daty dokumentu albo liczby odsyłaczy do określonego dokumentu WWW w innych dokumentach WWW.
- Typ kategorii, które mają zostać użyte dla określonej kolekcji. Można nie wybierać żadnej kategorii lub wybrać kategorie oparte na regułach albo kategorie oparte na modelach.

Po wybraniu kategorii opartych na regułach systematyka i reguły kolekcji portalu WebSphere Portal zostaną poddane migracji do wyszukiwania korporacyjnego.

Kategorie oparte na modelach mogą zostać wybrane tylko wówczas, gdy wcześniej wykonano migrację systematyki opartej na modelach z portalu WebSphere Portal do wyszukiwania korporacyjnego.

Jeśli podczas wykonywania migracji wystąpią błędy, należy przejrzeć plik MigrationWizard.log znajdujący się w katalogu, w którym zainstalowany jest Kreator migracji.

Teraz można użyć konsoli administracyjnej wyszukiwania korporacyjnego i skonfigurować dodatkowe ustawienia dla zmigrowanych kolekcji.

**Wymaganie:** Podczas konfigurowania właściwości przeszukiwacza sieci WWW dla zmigrowanej kolekcji należy określić adres e-mail do otrzymywania komentarzy o przeszukiwaczu oraz nazwę agenta użytkownika (w celu uzyskania pomocy w trakcie konfigurowania właściwości przeszukiwacza sieci WWW należy kliknąć przycisk **Pomoc**).

10. Używając konsoli administracyjnej wyszukiwania korporacyjnego, uruchom procesy przeszukiwania, analizowania i indeksowania dla zmigrowanej kolekcji.
11. Po sprawdzeniu, że zmigrowana kolekcja jest dostępna dla operacji przeszukiwania w wyszukiwaniu korporacyjnym, usuń oryginalną kolekcję w mechanizmie wyszukiwania środowiska Portal.
12. Opcjonalnie: Jako administrator portalu WebSphere Portal wykonaj następujące kroki, o ile chcesz, aby użytkownicy mogli przeszukiwać zmigrowaną kolekcję przy użyciu portalu WebSphere Portal.

- a. Wykonaj wdrożenie portletu wyszukiwania korporacyjnego w lokalnej instalacji portalu WebSphere Portal.

W klastrze serwerów WebSphere Portal należy dokonać wdrożenia na serwerze, na którym zainstalowany jest program WebSphere Application Server Deployment Manager. Program Deployment Manager rozprowadza portlet wyszukiwania korporacyjnego na pozostałe serwery w klastrze serwerów WebSphere Portal.

- b. Dodaj portlet wyszukiwania korporacyjnego do odpowiednich stron portalu.

W portalu WebSphere Portal dostęp portletu wyszukiwania jest określany na podstawie uprawnień dostępu do konkretnych stron i portletów. Choć ustawienia konfiguracyjne kolekcji są migrowane, administrator serwera WebSphere Portal musi ręcznie określić pozycję portletu.

#### Pojęcia pokrewne

“Kolekcje wyszukiwania korporacyjnego” na stronie 13

W kolekcji wyszukiwania korporacyjnego dostępny jest cały zestaw źródeł, w których



użytkownicy mogą przeprowadzać wyszukiwanie za pomocą pojedynczego zapytania. Dzięki stowarzyszeniu użytkownicy mogą wyszukiwać w wielu kolekcjach za pomocą jednego zapytania.

“Kategorie oparte na regułach” na stronie 89

Użytkownik konfigurując reguły określa dokumenty przypisywane do kategorii w kolekcji wyszukiwania korporacyjnego.

“Drzewa kategorii” na stronie 92

Drzewo kategorii umożliwia wyświetlenie wszystkich - używanych w kolekcji - kategorii opartych na regułach. Drzewo kategorii służy do tworzenia i usuwania kategorii oraz do edytowania reguł przypisywania dokumentów do kategorii.

“Ocena statyczna” na stronie 144

W przypadku niektórych typów dokumentów możliwe jest przypisanie statycznego współczynnika oceniania, który zwiększa jego ważność w wynikach wyszukiwania.

---

## Ustawienia zmigrowanej kolekcji

Podczas migrowania kolekcji z portalu IBM WebSphere Portal Kreator migracji tworzy domyślne ustawienia dla kolekcji i przeszukiwaczy.

Jeśli to samo ustawienie istnieje w kolekcjach mechanizmu wyszukiwania środowiska Portal i kolekcjach wyszukiwania korporacyjnego, wówczas podczas migrowania kolekcji do wyszukiwania korporacyjnego w kreatorze zostanie użyte ustawienie mechanizmu wyszukiwania środowiska Portal. W przypadku ustawień, które istnieją tylko w wyszukiwaniu korporacyjnym, w kreatorze użyte zostaną ustawienia określone przez użytkownika podczas migrowania kolekcji lub domyślne ustawienia dla kolekcji w wyszukiwaniu korporacyjnym.

### Ustawienia istniejące w mechanizmie wyszukiwania środowiska Portal i wyszukiwaniu korporacyjnym

Dla każdej migrowanej przez użytkownika kolekcji Kreator migracji migruje następujące ustawienia:

- Występujące w kolekcji mechanizmu wyszukiwania środowiska Portal serwisy mechanizmu wyszukiwania.
- Język kolekcji.
- Systematyka (lub drzewo kategorii) i reguły dla kategorii opartych na regułach, o ile w kolekcji wyszukiwania korporacyjnego jest używana klasyfikacja oparta na regułach.

Każdy występujący w kolekcji serwis mechanizmu wyszukiwania środowiska Portal jest konsolidowany z przeszukiwaczem sieci WWW wyszukiwania korporacyjnego. Kreator migracji migruje następujące ustawienia przeszukiwacza:

- początkowy adres URL
- liczba równoległych procesów przeszukiwacza
- głębokość przeszukiwania
- limit czasu (w sekundach) dla operacji pobierania dokumentu
- domyślny zestaw znaków
- reguły włączania i wykluczania dla operacji przeszukiwania

### Ustawienia istniejące wyłącznie w wyszukiwaniu korporacyjnym

Podczas migrowania kolekcji należy określić informacje dotyczące tej kolekcji. Kreator migracji migruje te ustawienia i wykorzystuje domyślne ustawienia kolekcji w wyszukiwaniu korporacyjnym do skonfigurowania każdej migrowanej kolekcji.

Użytkownik może zmodyfikować ustawienia konfiguracyjne kolekcji i przeszukiwacza sieci WWW za pomocą konsoli administracyjnej wyszukiwania korporacyjnego. Wartości ujęte w nawiasy ( ) to domyślne ustawienia dla migrowanych danych.

- nazwa kolekcji
- strategia statycznego oceniania dokumentów
- typ używanej klasyfikacji - na przykład klasyfikacja oparta na regułach albo brak klasyfikacji
- ustawienie określające, czy ma być używana pamięć podręczna wyszukiwania i ustawienie określające liczbę zapytań z wynikami wyszukiwania, które mogą zostać zapisane w pamięci podręcznej wyszukiwania (tak, 5000)
- ustawienie określające, czy ma być monitorowany czas uzyskania odpowiedzi z wynikami wyszukiwania i generowany alert, gdy przekroczony zostanie określony limit (tak, 5 sekund)
- ustawienie określające, czy mają być używane mechanizmy kontroli dostępu (nie)
- harmonogram odświeżania indeksu
- harmonogram reorganizacji indeksu
- poziom szczegółowości informacji w dzienniku (wszystkie komunikaty)

Kreator migracji tworzy także dla każdego przeszukiwacza sieci WWW następujące ustawienia:

- nazwa przeszukiwacza
- opis przeszukiwacza
- maksymalna długość strony
- ustawienia zabezpieczeń dokumentu
- typy MIME (Multipurpose Internet Mail Extensions) dokumentów, które mają zostać przeszukane, o ile mają one zastosowanie dla używanego typu źródła danych

Przed uruchomieniem nowo zmigrowanego przeszukiwacza sieci WWW należy przejrzeć wszystkie właściwości przeszukiwacza, ustawienia przestrzeni przeszukiwania oraz sprawdzić, czy wszystkie wymagane wartości zostały określone (wymagane pola są oznaczone czerwoną gwiazdką). W szczególności należy sprawdzić, czy został określony adres e-mail do otrzymywania komentarzy o przeszukiwaczu oraz nazwa agenta użytkownika dla przeszukiwacza. W celu uzyskania pomocy podczas konfigurowania właściwości przeszukiwacza sieci WWW należy kliknąć przycisk **Pomoc**.

---

## Plik dziennika Kreatora migracji

Kreator migracji zapisuje wszystkie komunikaty w pliku WpsMigratorLog.log znajdującym się w katalogu, w którym zainstalowany jest Kreator migracji.

Dla każdej migrowanej kolekcji w pliku dziennika WpsMigratorLog.log zapisywane są wartości wszystkich ustawień odczytanych z mechanizmu wyszukiwania środowiska WebSphere Portal, a także informacje o miejscach w kolekcjach wyszukiwania korporacyjnego, do których ustawienia te zostały zaimportowane.

---

## Uruchamianie i zatrzymywanie serwerów wyszukiwania korporacyjnego

Po utworzeniu kolekcji należy uruchomić serwery odpowiedzialne za przeszukiwanie, analizę składni, indeksowanie i wyszukiwanie danych. Po wprowadzeniu zmian w kolekcji należy zatrzymać i zrestartować serwery.

Większość serwerów wyszukiwania korporacyjnego może działać w sposób nieprzerwany lub zgodnie z określonym harmonogramem. Na przykład istnieje możliwość określenia harmonogramu reorganizacji lub odświeżania indeksu. Po uruchomieniu serwerów odpowiedzialnych za analizę składni danych i wyszukiwanie w indeksie konieczność ich zatrzymania i restartu zachodzi tylko w przypadku zmiany ustawień konfiguracyjnych (np. w przypadku aktualizacji kategorii lub zwiększenia pamięci podręcznej wyszukiwania).

W przypadku wprowadzenia zmian w treści kolekcji lub po zmianie reguł dotyczących sposobu gromadzenia danych ze źródeł w zasobach przedsiębiorstwa przez przeszukiwacze uwzględnienie tych zmian zazwyczaj wymaga zatrzymania i zrestartowania przeszukiwaczy. Jeśli reguły przeszukiwania nie są zmieniane, przeszukiwacze działają albo nieprzerwanie (w przypadku przeszukiwaczy sieci WWW i NNTP), albo zgodnie z określonym harmonogramem.

---

## Uruchamianie serwerów wyszukiwania korporacyjnego

Aby umożliwić użytkownikom wyszukiwanie w kolekcji, należy uruchomić procesy systemowe, a następnie uruchomić serwery przeszukiwania, analizy, indeksowania i wyszukiwania w kolekcji.

### Zagadnienia wstępne

Należy skonfigurować źródła danych, które mają być przeszukiwane, i podać opcje określające sposób analizowania, indeksowania i wyszukiwania danych. Na przykład, jeśli użytkownicy mają mieć możliwość wyświetlania w wynikach wyszukiwania szczegółów kategorii, przed uruchomieniem analizatora składni należy skonfigurować kategorie.

Aby uruchomić serwery wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji. Uruchomić serwery może również operator z uprawnieniami do administrowania daną kolekcją.

Serwery wyszukiwania korporacyjnego dla kolekcji należy uruchomić w poprawnej kolejności. Na przykład należy uruchomić przeszukiwacz i przeszukać dane przed analizowaniem i indeksowaniem dokumentów.

### Procedura

Aby uruchomić serwery wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Aby skorzystać z konsoli administracyjnej wyszukiwania korporacyjnego lub z aplikacji wyszukiwującej:
  - a. Uruchom serwer IBM HTTP Server, jeśli nie jest uruchomiony.
  - b. Za pomocą programu WebSphere Application Server Administrative Console uruchom aplikacje korporacyjne ESAdmin i ESSearchApplication, jeśli nie zostały jeszcze uruchomione.

2. Jeśli warstwa CCL wyszukiwania korporacyjnego nie działa, uruchom ją:
  - a. Na serwerze indeksowania zaloguj się z identyfikatorem administratora wyszukiwania korporacyjnego, który został określony podczas instalowania programu WebSphere II OmniFind Edition.

#### Systemy UNIX






Wprowadź następującą komendę:

```
startccl.sh -bg
```

#### Systemy Windows

Wybierz kolejno opcje: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi**, a następnie zrestartuj usługę programu WebSphere Information Integrator OmniFind Edition.

3. Jeśli system wyszukiwania korporacyjnego nie działa, uruchom go:
  - a. Na serwerze indeksowania zaloguj się z identyfikatorem administratora wyszukiwania korporacyjnego, który został określony podczas instalowania programu WebSphere II OmniFind Edition.
  - b. Wprowadź następującą komendę:  

```
esadmin start
```
4. Uruchom konsolę administracyjną wyszukiwania korporacyjnego i zaloguj się jako administrator wyszukiwania korporacyjnego. Jeśli korzystasz z roli administracyjnej, możesz zalogować się jako administrator kolekcji lub jako operator mający uprawnienia do uruchamiania kolekcji.
5. W widoku Kolekcje znajdź kolekcję, którą chcesz administrować, i kliknij ikonę  **Monitor**.
6. Na stronie Przeszukiwanie dla każdego przeszukiwacza, który ma zostać uruchomiony, kliknij opcję  **Uruchom**.
  - Jeśli uruchamiany jest przeszukiwacz sieci WWW lub NNTP, rozpocznie on przeszukiwanie danych natychmiast. Te typy przeszukiwaczy działają nieprzerwanie, przeszukując i ponownie przeszukując dokumenty w serwisach WWW i grupach dyskusyjnych NNTP.
  - Jeśli uruchamiany jest jeden z pozostałych typów przeszukiwaczy, zostanie uruchomiona sesja przeszukiwacza. Przeszukiwacz rozpocznie przeszukiwanie w zaplanowanym terminie. Jeśli przeszukiwacz nie figuruje w harmonogramie lub jeśli trzeba wcześniej uruchomić przeszukiwacz, przejdź do monitorowania przeszukiwacza i kliknij ikonę uruchomienia dla każdego źródła danych, które ma zostać przeszukane.  
Po uruchomieniu przeszukiwacza można zezwolić mu na nieprzerwane działanie. Jeśli utworzono harmonogram dla przeszukiwacza, zostanie on ponownie uruchomiony w zaplanowanym dniu i o zaplanowanej godzinie.
7. Po zakończeniu przeszukiwania danych otwórz stronę Analiza składni i kliknij ikonę  **Uruchom**, aby uruchomić analizator składni.  
Można zezwolić analizatorowi składni na nieprzerwane działanie. Zwykle nie ma potrzeby zatrzymywania analizatora, chyba że wprowadzone zostały zmiany w sposobie analizowania danych (takie jak aktualizacja kategorii lub odwzorowań pól XML).
8. Opcjonalnie: Aby wymusić rozpoczęcie procesów indeksowania, zamiast czekać ich na rozpoczęcie w zaplanowanym terminie, otwórz stronę Indeksowanie i w obszarze **Reorganizacja** kliknij opcję  **Uruchom**.  
Można zezwolić procesom indeksowania na nieprzerwane działanie. Indeks zostanie odświeżony i zreorganizowany w zaplanowanym dniu i o zaplanowanej godzinie.
9. Na stronie Szukaj kliknij ikonę  **Uruchom**.

Można zezwolić serwerom wyszukiwania na nieprzerwane działanie. Zwykle nie ma potrzeby zatrzymywania serwerów wyszukiwania, chyba że zostały wprowadzone zmiany w ustawieniach pamięci podręcznej wyszukiwania lub ustawieniach podsumowania dokumentów.

#### **Pojęcia pokrewne**

“Administrowanie wyszukiwaniem korporacyjnym - przegląd” na stronie 1  
Konsola administracyjna wyszukiwania korporacyjnego służy do tworzenia kolekcji oraz źródeł zewnętrznych i administrowania nimi, uruchamiania i zatrzymywania komponentów, monitorowania działań w systemie, monitorowania plików dziennika, konfigurowania ustawień dotyczących użytkowników z uprawnieniami administracyjnymi, tworzenia powiązań aplikacji wyszukiwujących z kolekcjami i źródłami zewnętrznymi, a także do określania informacji umożliwiających wymuszanie stosowania zabezpieczeń.

“Role administracyjne” na stronie 176

Wyszukiwanie korporacyjne wykorzystuje role jako metodę kontroli dostępu do różnych funkcji konsoli administracyjnej.

#### **Zadania pokrewne**

“Logowanie się do konsoli administracyjnej” na stronie 5

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

#### **Odsyłacze pokrewne**

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## **Zatrzymywanie serwerów wyszukiwania korporacyjnego**

Może wystąpić konieczność zatrzymania i zrestartowania serwera wyszukiwania korporacyjnego, jeśli dokonano zmian w jego konfiguracji lub w celu rozwiązania problemów.

#### **Zagadnienia wstępne**

Aby zatrzymać serwery wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji. Zatrzymać serwery może również operator z uprawnieniami do administrowania daną kolekcją.

#### **Informacje o tym zadaniu**


Serwery wyszukiwania korporacyjnego można zatrzymywać niezależnie od siebie. Na przykład zatrzymanie i zrestartowanie przeszukiwacza w celu uaktywnienia zmian dokonanych w jego konfiguracji nie wymaga zatrzymania i zrestartowania analizatora składni.

Aby zatrzymać cały system wyszukiwania korporacyjnego, a nie jego poszczególne serwery, można się zalogować na serwerze indeksowania przy użyciu identyfikatora administratora wyszukiwania korporacyjnego (ten identyfikator został określony podczas instalowania programu WebSphere II OmniFind Edition). Następnie należy wpisać następującą komendę:

```
esadmin stop
```

#### **Procedura**

Aby zatrzymać serwery wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję, którą chcesz administrować, i kliknij ikonę  **Monitoruj**.

2. Na stronie Przeszukiwanie znajdź przeszukiwacz, którym chcesz administrować, i zatrzymaj lub wstrzymaj go.


Jeśli zmieniona została przestrzeń przeszukiwania lub właściwości przeszukiwacza, aby zmiany zostały uwzględnione, należy zatrzymać i zrestartować przeszukiwacz. Jeśli zmieniona została przestrzeń przeszukiwania i zmiany mają być zastosowane do dokumentów już znajdujących się w kolekcji, należy ponownie przeszukać dokumenty.


**Wskazówka:** Może zostać wyświetlony komunikat o przekroczeniu limitu czasu przez żadaną operację, mimo że proces nadal działa w tle. Aby dowiedzieć się, czy zadanie zostało ukończone, kliknij przycisk **Odśwież** w konsoli administracyjnej (nie klikaj przycisku **Odśwież** w przeglądarce WWW). Proces jest zakończony, gdy ikona statusu przeszukiwacza wskazuje, że został on zatrzymany.

3. Na stronie Analiza składni kliknij ikonę  **Zatrzymaj**, aby zatrzymać analizator składni.

Jeśli zmienione zostały reguły analizowania danych, aby zmiany zostały uwzględnione, należy zatrzymać i zrestartować analizator składni. Zmiany będą stosowane tylko w odniesieniu do nowo przeszukanych dokumentów. Jeśli należy zastosować zmiany do dokumentów, które już znajdują się w indeksie, uruchom pełne przeszukiwanie w celu ponownego przeszukania wszystkich dokumentów, co umożliwi ich ponowne przeanalizowanie i umieszczenie w indeksie.

4. Na stronie Indeks kliknij opcję  **Zatrzymaj**, aby zatrzymać indeks, który jest odświeżany lub reorganizowany.

Można również zatrzymać budowanie indeksu podczas monitorowania kolejki indeksów. Aby to zrobić, wybierz opcję **System** na pasku narzędzi, otwórz stronę Indeksowanie i kliknij opcję  **Zatrzymaj** dla indeksu, którego tworzenie ma być zatrzymane.

5. Na stronie Wyszukiwanie kliknij opcję  **Zatrzymaj**, aby zatrzymać serwery wyszukiwania. Zwykle nie ma potrzeby zatrzymywania i restartowania serwerów wyszukiwania, chyba że zmienione zostały ustawienia pamięci podręcznej wyszukiwania lub ustawienia podsumowania dokumentu.

### Pojęcia pokrewne

“Administrowanie wyszukiwaniem korporacyjnym - przegląd” na stronie 1  
Konsola administracyjna wyszukiwania korporacyjnego służy do tworzenia kolekcji oraz źródeł zewnętrznych i administrowania nimi, uruchamiania i zatrzymywania komponentów, monitorowania działań w systemie, monitorowania plików dziennika, konfigurowania ustawień dotyczących użytkowników z uprawnieniami administracyjnymi, tworzenia powiązań aplikacji wyszukujących z kolekcjami i źródłami zewnętrznymi, a także do określania informacji umożliwiających wymuszanie stosowania zabezpieczeń.

### Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## Monitorowanie działań związanych z wyszukiwaniem korporacyjnym

Monitorując działania systemu i działania związane z kolekcjami, można wyświetlić informacje o statusie różnych procesów, obserwować zmienne w celu wykrycia potencjalnych problemów lub dopasować ustawienia konfiguracyjne w celu poprawienia wydajności.

Konsola administracyjna wyszukiwania korporacyjnego umożliwia monitorowanie systemu i dostosowanie operacji do własnych potrzeb. Można także wyświetlić szczegółowe informacje statystyczne dotyczące każdego podstawowego działania (przeszukiwanie, analizowanie, indeksowanie i wyszukiwanie). Dane statystyczne zawierają między innymi informacje o średnim czasie odpowiedzi i o postępie, takie jak liczba dokumentów przeszukanych lub poindeksowanych podczas sesji przeszukiwania lub tworzenia indeksu.

Klikając odpowiednie ikony, użytkownik może zatrzymać lub uruchomić większość działań. Istnieje możliwość wstrzymania operacji, zmodyfikowania związanych z nią ustawień konfiguracyjnych lub rozwiązania zaistniałych problemów, a następnie zrestartowania procesu przetwarzania w odpowiednim momencie, gdy możliwe będzie kontynuowanie operacji.

### Pojęcia pokrewne

“Administrowanie wyszukiwaniem korporacyjnym - przegląd” na stronie 1  
Konsola administracyjna wyszukiwania korporacyjnego służy do tworzenia kolekcji oraz źródeł zewnętrznych i administrowania nimi, uruchamiania i zatrzymywania komponentów, monitorowania działań w systemie, monitorowania plików dziennika, konfigurowania ustawień dotyczących użytkowników z uprawnieniami administracyjnymi, tworzenia powiązań aplikacji wyszukujących z kolekcjami i źródłami zewnętrznymi, a także do określania informacji umożliwiających wymuszanie stosowania zabezpieczeń.

### Zadania pokrewne

“Uruchamianie serwerów wyszukiwania korporacyjnego” na stronie 209  
Aby umożliwić użytkownikom wyszukiwanie w kolekcji, należy uruchomić procesy systemowe, a następnie uruchomić serwery przeszukiwania, analizy, indeksowania i wyszukiwania w kolekcji.

“Zatrzymywanie serwerów wyszukiwania korporacyjnego” na stronie 211  
Może wystąpić konieczność zatrzymania i zrestartowania serwera wyszukiwania korporacyjnego, jeśli dokonano zmian w jego konfiguracji lub w celu rozwiązania problemów.

### Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## Szacowanie liczby dokumentów w kolekcji

Podczas tworzenia lub edycji kolekcji wyszukiwania korporacyjnego należy podać szacunkową liczbę dokumentów, które mają być w niej przechowywane. Menedżer zasobów używa tej liczby do oszacowania wielkości zasobów pamięci i dysku, które są wymagane przez kolekcję, ale nie wprowadza ograniczenia na wielkość kolekcji.

### Zagadnienia wstępne

Aby zmienić szacowaną wielkość kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienie administratora danej kolekcji.

### Informacje o tym zadaniu

Gdy kolekcja osiągnie szacowaną wielkość, system nie przestanie dodawać dokumentów do indeksu. Menedżer zasobów będzie generował ostrzeżenia, gdy oszacowane zapotrzebowanie na zasoby pamięci lub dysku przekroczy zasoby dostępne w danej chwili w systemie. Te ostrzeżenia umożliwiają uniknięcie przyszłych problemów związanych z brakiem zasobów.

Jeśli dla kolekcji zostaną skonfigurowane alerty i będzie włączona opcja powiadamiania w przypadku przekroczenia limitu liczby dokumentów w indeksie, domyślny limit będzie równy podanej szacowanej liczbie dokumentów w kolekcji. Monitor używa tej liczby łącznie z określoną przez użytkownika wartością progu alertu - wyrażoną w procentach - i wysyła wiadomość e-mail w przypadku zbliżania się do maksymalnej liczby dokumentów skonfigurowanych dla danej kolekcji.

### Procedura

Aby podać oszacowanie potencjalnej wielkości kolekcji, należy wykonać następujące czynności:

1. Przejdź do edycji kolekcji, wybierz stronę Ogólne, a następnie kliknij opcję **Konfiguruj opcje ogólne**.
2. W polu **Szacowana liczba dokumentów** wpisz przewidywaną maksymalną wielkość kolekcji. Wartością domyślną jest 1 000 000 dokumentów.

---

## Sprawdzanie dostępności zasobów systemu

Jeśli sprawdzenie dostępności zasobów systemu zostanie wykonane po utworzeniu kolekcji lub przeszukiwacza, ale przed uruchomieniem przeszukiwania dokumentów, można określić, czy dostępne zasoby systemu są odpowiednie do uruchomienia produktu WebSphere II OmniFind Edition przy pełnej pojemności wynikającej z bieżących ustawień.

### Zagadnienia wstępne

Aby sprawdzić zasoby systemowe, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Podczas tworzenia kolekcji lub przeszukiwacza system automatycznie sprawdza dostępność zasobów. Opcję sprawdzania dostępności zasobów systemu można wybrać w dowolnym momencie.

System porównuje szacowaną ilość wykorzystywanego miejsca podaną podczas konfigurowania właściwości kolekcji i przeszukiwacza z miejscem dostępnym w systemie, a następnie wyświetla komunikat informujący o dostępności zasobów. Komunikat wskazuje, który serwer może mieć za mało zasobów, ile miejsca wymaga szacowana wielkość kolekcji i ile jest dostępnego wolnego miejsca.

Jeśli dostępne zasoby są niewystarczające, należy wykonać jedno z następujących działań naprawczych:

- Zwiększyć wielkość systemu plików podanego w komunikacie.
- Zmodyfikować opcje ogólne dla kolekcji i podać mniejszą szacowaną liczbę dokumentów.



- Zmodyfikować właściwości przeszukiwacza i podać mniejszą maksymalną liczbę dokumentów do przeszukania, mniejszy maksymalny rozmiar strony i mniejszą maksymalną liczbę wątków.
- Zmodyfikować właściwości przeszukiwacza sieci WWW i podać mniejszą liczbę aktywnych hostów, nowych dokumentów i mniejszą maksymalną liczbę dokumentów w pamięci tymczasowej.
- Zmodyfikować właściwości indeksu dla systemu i zmniejszyć liczbę procesów tworzenia indeksu działających równocześnie.
- Unikać uruchamiania jednocześnie zbyt wielu procesów w wielu kolekcjach. Na przykład uruchomienie jednocześnie przeszukiwaczy i analizatorów składni dla wielu kolekcji może wpłynąć na ograniczenie ilości dostępnych zasobów.
- Usunąć z systemu przeszukiwacze.
- Usunąć z systemu kolekcje.

Jeśli ostrzeżenia o możliwym braku zasobów nie pojawiają się, oznacza to, że dostępne zasoby są wystarczające na potrzeby przeszukiwania i indeksowania danych.

### Procedura

Aby określić, czy dla bieżącej konfiguracji systemu wyszukiwania korporacyjnego dostępne są wystarczające zasoby:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Na stronie Ogólne kliknij opcję **Sprawdź zasoby systemowe**.

---

## Monitorowanie kolekcji


Użytkownik może wyświetlić ogólne informacje na temat statusu poszczególnych komponentów kolekcji albo wybrać opcje dotyczące wyświetlania szczegółowych informacji o wybranych komponentach i identyfikatorach URI.


### Zagadnienia wstępne

Do monitorowania kolekcji uprawnieni są wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego. Aby edytować właściwości przeszukiwacza, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora kolekcji dla tej kolekcji.

### Procedura

Aby monitorować kolekcję, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję, którą chcesz monitorować, i kliknij opcję  **Monitoruj**. Wyświetlone zostaną informacje na temat bieżącego statusu każdego z komponentów kolekcji.

**Wskazówka:** W przypadku, gdy wykonywana jest edycja kolekcji i strona Ogólne jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby wyświetlić widok umożliwiający monitorowanie kolekcji.

2. Aby wyświetlić szczegółowe informacje o identyfikatorze URI, kliknij ikonę  **Szczegóły dotyczące identyfikatorów URI**.

Na przykład można wyświetlić informacje o tym, czy dany identyfikator URI znajduje się w indeksie lub czy indeks, w którym identyfikator URI się znajduje, został skopiowany na serwery wyszukiwania.

3. Aby monitorować wybrany komponent i wyświetlić szczegółową statystykę na temat jego działania, kliknij ikonę **Status**.

#### Pojęcia pokrewne

“Kolekcje wyszukiwania korporacyjnego” na stronie 13

W kolekcji wyszukiwania korporacyjnego dostępny jest cały zestaw źródeł, w których użytkownicy mogą przeprowadzać wyszukiwanie za pomocą pojedynczego zapytania.

Dzięki stowarzyszeniu użytkownicy mogą wyszukiwać w wielu kolekcjach za pomocą jednego zapytania.

---

## Wyświetlanie szczegółów dotyczących identyfikatora URI

Użytkownik może wyświetlić szczegółowe informacje na temat identyfikatora URI.

Wyświetlić można aktualne i historyczne informacje odnośnie przeszukiwania, indeksowania i wyszukiwania dokumentu reprezentowanego przez dany identyfikator URI.

#### Zagadnienia wstępne

Przed wprowadzeniem żądania wyświetlenia raportu o identyfikatorze URI lub wysłania raportu na wskazany adres e-mail należy się upewnić, że komponent, który ma być źródłem pochodzenia informacji, jest aktywny. Na przykład, aby wyświetlić informacje szczegółowe na temat przeszukiwania, indeksowania i wyszukiwania dokumentu, należy się upewnić, że uruchomiony jest przeszukiwacz sieci WWW, serwer indeksowania i serwer wyszukiwania. Aby możliwe było śledzenie usuniętego dokumentu, muszą być skonfigurowane opcje rejestrowania dotyczące funkcji śledzenia dokumentów.

#### Informacje o tym zadaniu


Gromadzenie informacji na temat identyfikatora URI jest procesem czasochłonnym. Można wybrać opcję umożliwiającą wyświetlenie informacji objętych żądaniem, a następnie poczekać na ich wyświetlenie. Metodą bardziej efektywną jest wysłanie raportu na podany przez użytkownika adres e-mail.


Przed wysłaniem raportu należy się upewnić, że dla wyszukiwania korporacyjnego zostały skonfigurowane informacje o serwerze poczty. Informacje te należy podać podczas konfigurowania opcji poczty elektronicznej na stronie Rejestrowanie w widoku System.


Serwer indeksowania i serwer wyszukiwania mogą udostępnić informacje na temat wszystkich identyfikatorów URI (na przykład, czy identyfikator URI znajduje się w indeksie oraz czy został skopiowany na serwer wyszukiwania). Aby wyświetlić informacje na temat przeszukiwania dokumentu, należy podać identyfikator URI dokumentu, który został przeszukany przez przeszukiwacz sieci WWW.

#### Procedura

Aby wyświetlić szczegóły na temat identyfikatora URI, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję, którą chcesz monitorować, i kliknij opcję  **Monitoruj**.

**Wskazówka:** W przypadku, gdy wykonywana jest edycja kolekcji i strona Ogólne jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby wyświetlić widok umożliwiający monitorowanie kolekcji.

2. Kliknij opcję  **Szczegóły dotyczące identyfikatorów URI**.
3. Na stronie Szczegóły dotyczące identyfikatorów URI wpisz identyfikator URI, którego mają dotyczyć wyświetlane informacje.

4. Zaznacz pola wyboru odpowiadające rodzajowi informacji, które mają być wyświetlone:

**Szczegóły przeszukiwacza (dostępne tylko dla przeszukiwaczy sieci WWW)**

Zaznacz to pole wyboru, aby wyświetlić informacje o operacji przeszukiwania dokumentu przez przeszukiwacz sieci WWW, a także o jego aktualnym statusie w przestrzeni przeszukiwania.

**Szczegóły indeksu**

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został poindeksowany i skopiowany na serwer wyszukiwania.

**Szczegóły wyszukiwania**

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, w jaki sposób dokument może być wyszukiwany i czy jest on w ogóle dostępny do wyszukiwania.

**Dokumenty usunięte przez analizator składni**

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty z systemu wyszukiwania korporacyjnego podczas analizowania, a jeśli tak, to jaka była tego przyczyna.

**Dokumenty usunięte z indeksu**

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty z systemu wyszukiwania korporacyjnego podczas indeksowania lub analizowania, a jeśli tak, to jaka była tego przyczyna.

5. Aby zaczekać na wyświetlenie raportu, kliknij opcję **Wyświetl raport**.
6. Aby wysłać raport na podany adres e-mail w celu przejrzania go w późniejszym terminie, kliknij opcję **Wyślij raport**.
- a. Na stronie Wysłanie szczegółowego raportu o identyfikatorach URI, w polu **Adres poczty elektronicznej do wysłania powiadomienia** wpisz adres poczty elektronicznej, na który ma zostać wysłany raport.
- b. Kliknij przycisk **Wyślij raport**.

**Pojęcia pokrewne**

“Śledzenie dokumentu” na stronie 231

Dokumenty mogą być usuwane z systemu na różnych etapach przetwarzania. Istnieje możliwość określenia opcji powiadamiania o usuwaniu dokumentów oraz o problemach, które spowodowały usunięcie.

**Zadania pokrewne**

“Wyświetlanie raportów o usuniętych dokumentach” na stronie 233

Użytkownik może wyświetlić szczegółowe informacje na temat dokumentów, które zostały usunięte z systemu wyszukiwania korporacyjnego. Informacje te dostępne są tylko wtedy, gdy włączona jest opcja śledzenia dokumentów danej kolekcji.

**Odsyłacze pokrewne**

“Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79

Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

---

## Monitorowanie przeszukiwaczy


Użytkownik może wyświetlić ogólne informacje na temat statusu poszczególnych przeszukiwaczy w kolekcji albo wybrać opcje dotyczące wyświetlania szczegółowych informacji o działaniu przeszukiwacza.


### Zagadnienia wstępne


Jeśli uprawnienia przyznane użytkownikowi ograniczone są do monitorowania kolekcji, ma on możliwość przeglądania statystyki przeszukiwacza, ale nie może wpływać na jego działanie (uruchamiać lub zatrzymywać przeszukiwacz).

## Procedura


Aby monitorować przeszukiwacz, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję, którą chcesz monitorować, i kliknij opcję  **Monitoruj**.
2. Otwórz stronę Przeszukiwanie.

**Wskazówka:** W przypadku, gdy jest wykonywana edycja kolekcji i strona Przeszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby wyświetlić widok umożliwiający monitorowanie przeszukiwaczy.

3. Jeśli przeszukiwacz jest uruchomiony lub wstrzymany, a chcesz uzyskać szczegółowe informacje o statusie przeszukiwacza, kliknij ikonę  **Szczegóły**. Rodzaj wyświetlanej statystyki zależy od typu przeszukiwacza.

Jeśli uprawnienia przypisane użytkownikowi umożliwiają administrowanie procesami dotyczącymi kolekcji, może on uruchamiać, zatrzymywać i wstrzymywać przeszukiwacz podczas wyświetlania szczegółów na temat jego działania. Jeśli dla przeszukiwacza został utworzony harmonogram, można także włączyć lub wyłączyć ten harmonogram przeszukiwania.

4. Jeśli w sytuacji, gdy przeszukiwacz został zatrzymany lub wstrzymany, zajdzie potrzeba uruchomienia sesji przeszukiwacza, kliknij ikonę  **Uruchom** lub **Wznów**.

### W przypadku przeszukiwaczy sieci WWW:

Jeśli przeszukiwacz był zatrzymany, rozpoczyna na nowo pracę i przeszukuje całą przestrzeń przeszukiwania. Jeśli przeszukiwacz był wstrzymany, wznawia przeszukiwanie od początku elementu, na którym został wstrzymany.

Aby wymusić natychmiastowe rozpoczęcie przez przeszukiwacz pełnego przeszukiwania, kliknij ikonę **Szczegóły**, a następnie ikonę **Uruchom ponownie pełne przeszukiwanie**. Przeszukiwacz rozpocznie przeszukiwanie całej przestrzeni przeszukiwania, w tym stron, które nie uległy zmianie od czasu ich ostatniego przeszukiwania. Ponowne przeszukiwanie wszystkich dokumentów może być potrzebne na przykład wtedy, gdy zmianie uległy reguły analizowania dokumentów i trzeba zastosować te reguły do dokumentów, które zostały poindeksowane wcześniej.

### W przypadku przeszukiwaczy NNTP:



Jeśli przeszukiwacz był zatrzymany, rozpoczyna na nowo pracę i przeszukuje całą przestrzeń przeszukiwania. Jeśli przeszukiwacz był wstrzymany, wznawia przeszukiwanie od początku elementu, na którym został wstrzymany.

### W przypadku przeszukiwaczy wszystkich innych typów:

Jeśli przeszukiwacz był zatrzymany, rozpocznie przeszukiwanie w zaplanowanym dniu i o zaplanowanej godzinie. Podczas pierwszego przeszukiwania źródła danych przez przeszukiwacz zostaje wykonane pełne przeszukiwanie. Gdy zaplanowane przeszukiwanie jest powtarzane, przeszukiwacz przeszukuje albo wszystkie aktualizacje źródła danych (dodane, usunięte i zmodyfikowane dokumenty), albo tylko dodane i zmodyfikowane dokumenty. Typ przeszukiwania można skonfigurować w harmonogramie przeszukiwacza.

Jeśli przeszukiwacz nie figuruje w harmonogramie lub jeśli zachodzi potrzeba wcześniejszego uruchomienia przeszukiwacza, kliknij ikonę **Szczegóły**. Następnie w obszarze szczegółów przestrzeni przeszukiwania kliknij ikonę typu

przeszukiwania, które chcesz uruchomić: pełne przeszukiwanie, wszystkie aktualizacje lub tylko nowe i zmodyfikowane dokumenty. Należy kliknąć odpowiednią ikonę uruchamiającą dla każdego źródła danych, które ma zostać przeszukane (takiego jak serwer, baza danych lub podfolder).

5. Jeśli uruchomiony przeszukiwacz ma zostać zatrzymany, kliknij ikonę  **Zatrzymaj** lub  **Wstrzymaj**. Przeszukiwacz zatrzymuje przeszukiwanie do momentu ponownego uruchomienia lub wznowienia jego działania.

Jeśli przeszukiwacz był wstrzymany, wznowia przeszukiwanie od początku elementu, na którym został wstrzymany. Na przykład przeszukiwacz DB2 wznowi przeszukiwanie od pierwszego wiersza w tabeli, która była przeszukiwana w chwili jego wstrzymania.

#### Pojęcia pokrewne

“Administrowanie przeszukiwaczem wyszukiwania korporacyjnego” na stronie 19  
Konfiguracja przeszukiwaczy dla różnych typów danych, które mają zostać włączone do kolekcji. Pojedyncza kolekcja może zawierać dowolną liczbę przeszukiwaczy.

#### Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpocząć i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## Wyświetlanie szczegółów na temat działalności przeszukiwacza sieci WWW


Zapoznając się ze szczegółami na temat działań przeszukiwacza sieci WWW, można ocenić jego ogólną wydajność i odpowiednio dopasować właściwości przeszukiwacza oraz definicje przestrzeni przeszukiwania.


### Zagadnienia wstępne


Do monitorowania działań przeszukiwacza uprawnieni są wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego. Aby uruchamiać lub zatrzymywać przeszukiwacz, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora danej kolekcji.

### Procedura

Aby wyświetlić szczegółowe informacje na temat przeszukiwacza sieci WWW, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję będącą właścicielem przeszukiwacza, który ma być monitorowany, i kliknij ikonę  **Monitoruj**.
2. Otwórz stronę Przeszukiwanie.

**Wskazówka:** W przypadku, gdy jest wykonywana edycja kolekcji i strona Przeszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby wyświetlić widok umożliwiający monitorowanie przeszukiwaczy.

3. Jeśli przeszukiwacz sieci WWW, który ma być monitorowany, jest uruchomiony lub wstrzymany, kliknij ikonę  **Szczegóły**.
4. Na stronie szczegółowych informacji o przeszukiwaczu sieci WWW przejrzyj lub wybierz poniższe opcje, aby wyświetlić szczegółową statystykę o aktualnej i dotychczasowej aktywności przeszukiwacza.
  - a. Kliknij opcję **Szczegóły wątku**, aby zobaczyć, ile obecnie wątków aktywnie przeszukuje serwisy WWW, a ile znajduje się w stanie nieaktywnym.

- b. Kliknij opcję **Aktywne serwisy**, aby wyświetlić informacje o serwisach WWW, które przeszukiwacz aktualnie przeszukuje.
- c. Kliknij opcję **Ostatnio przeszukane adresy URL**. Powoduje to wyświetlenie informacji o tym, co przeszukiwacz ostatnio przeszukał. Jeśli elementy listy nie ulegają zmianom przy odświeżaniu widoku, oznacza to, że przeszukiwanie nie jest aktualnie prowadzone.
- d. Kliknij opcję **Historia przeszukiwacza**, aby wyświetlić raporty o dotychczasowej aktywności przeszukiwacza.
- e. W obszarze **Status adresu URL** wpisz adres URL serwisu WWW, którego mają dotyczyć wyświetlane informacje, i kliknij przycisk **Wyświetl**.

Na przykład opcja ta umożliwi sprawdzenie, czy wybrany adres URL znajduje się w przestrzeni przeszukiwania, czy został przeszukany czy tylko wykryty, kiedy powinien zostać przeszukany ponownie oraz kiedy przeprowadzono ostatnią próbę przeszukania danego serwisu WWW.

Po wyświetleniu szczegółowych informacji o wybranym adresie URL można kliknąć opcję **Historia serwisu**, aby zobaczyć dodatkowe informacje statystyczne o aktywności przeszukiwacza odnośnie tego adresu.

## Szczegółowe informacje o wątkach przeszukiwacza sieci WWW

Przeszukiwacz sieci WWW można monitorować w celu sprawdzenia liczby wątków aktywnie przeszukujących serwisy WWW oraz liczby wątków, które pozostają nieaktywne.

Przy przeglądaniu szczegółów dotyczących przeszukiwacza sieci WWW podczas monitorowania kolekcji możliwe jest wyświetlenie statusu wątków przeszukiwacza. Najbardziej prawdopodobne statusy to:

### Oczekiwanie

Wskazuje, że wątek nie ma adresu URL do przeszukania. Taka sytuacja może wystąpić, gdy wątek skończył przeszukiwanie, a przeszukiwacz nie może wystarczająco szybko znaleźć więcej adresów URL do przeszukania. Na przykład, jeśli właściwość przeszukiwacza decydująca o tym, jak długo przeszukiwacz musi odczekać, zanim będzie mógł pobrać inną stronę z tego samego serwisu, ma zbyt dużą wartość, może to uniemożliwić wystarczająco szybkie dostarczanie adresów URL.

### Pobieranie

Wskazuje, że wątek pobiera stronę z serwisu WWW.

### Zakończone

Wskazuje, że wątek przesyła do przeszukiwacza przeszukane strony, ale nie jest jeszcze gotowy do przeszukania następnego adresu URL.

### Zawieszono

Wskazuje, że przeszukiwacz został wstrzymany.

W idealnej sytuacji wszystkie wątki pobierają strony przez cały czas. Jeśli jednak wątki często są w stanie zakończenia, mogą wystąpić problemy z przepustowością bazy danych.

Jeśli wątki często znajdują się w stanie oczekiwania, należy sprawdzić, jaka wartość została określona we właściwościach przeszukiwacza w polu **Maksymalna liczba aktywnych hostów**. Jeśli wartość ta jest mała, w przestrzeni przeszukiwania może nie być wystarczającej liczby serwisów do utrzymywania wszystkich wątków w stanie zajętości lub może brakować adresów URL dostępnych do przeszukania. Do sytuacji mogących powodować obniżenie aktywności należą błędy wyszukiwania na serwerze DNS oraz błędy wyszukiwania przez roboty.

## Aktywne serwisy przeszukiwacza sieci WWW

Istnieje możliwość monitorowania przeszukiwacza sieci WWW w celu uzyskania informacji o serwisach WWW, aktywnie przeszukiwanych przez przeszukiwacz.

Przy przeglądaniu szczegółów dotyczących przeszukiwacza sieci WWW podczas monitorowania kolekcji możliwe jest wyświetlenie danych statystycznych dotyczących aktywnych serwisów. Dane statystyczne przedstawiają:

- liczbę adresów URL, które przeszukiwacz pobiera ze swojej wewnętrznej bazy danych do pamięci do przeszukiwania w danej chwili;
- liczbę adresów URL, które przeszukiwacz próbował przeszukać do tej pory;
- ile czasu pozostało do dezaktywowania serwisu i usunięcia go z pamięci dla tej iteracji przeszukiwacza;
- jak długo serwis znajduje się już w pamięci.

Te informacje zmieniają się co chwila, w miarę jak przeszukiwacz kolejno realizuje skonfigurowane dla niego reguły przeszukiwania. W idealnej sytuacji liczba aktywowanych adresów URL jest bliska wartości skonfigurowanej w polu **Maksymalna liczba aktywnych hostów** we właściwościach przeszukiwacza.

Jeśli liczba aktywowanych adresów URL jest bliska zeru, przeszukiwacz nie znajduje właściwych adresów URL. Warunki, które mogą powodować tak niską aktywność, obejmują niepowodzenia wyszukiwania na serwerze DNS, zagadnienia związane z połączeniem z siecią, błędy bazy danych oraz problemy związane z definicją przestrzeni przeszukiwania. Na przykład:

- Jeśli od dłuższego czasu w pamięci znajduje się wiele serwisów, a przeszukano tylko kilka adresów URL, należy sprawdzić, czy nie występują problemy z nawiązaniem połączenia z siecią.
- Jeśli na liście jest niewystarczająca liczba serwisów, należy sprawdzić, czy nie występują problemy związane z definicją przestrzeni przeszukiwania lub z wyszukiwaniem DNS.
- Jeśli serwisy są przeszukiwane z odpowiednią szybkością, ale pozostawiają w pamięci wiele nie przeszukanych adresów URL, należy zmodyfikować właściwości przeszukiwacza i dopasować czas oczekiwania w polu **Maksymalny czas przechowywania adresu URL w pamięci**, aby serwisy dłużej były przechowywane w pamięci.

## Szybkość przeszukiwania przeszukiwacza sieci WWW

Działanie przeszukiwacza sieci WWW można monitorować, aby uzyskać informacje dotyczące szybkości, z jaką przeszukiwacz pobiera strony z serwisów WWW.

Przy przeglądaniu szczegółów dotyczących przeszukiwacza sieci WWW podczas monitorowania kolekcji możliwe jest wyświetlenie statystyki dotyczącej szybkości, z jaką przeszukiwacz przeszukuje dane (szybkość przeszukiwania). Można również wyświetlić statystykę dotyczącą liczby adresów URL przeszukanych przez przeszukiwacz WWW od momentu rozpoczęcia bieżącej sesji.

Szybkość przeszukiwania to liczba stron przeszukiwanych w ciągu sekundy. Ta liczba jest skorelowana z różnymi właściwościami przeszukiwacza sieci WWW, które użytkownik może konfigurować:

- liczba wątków przeszukiwacza
- liczba aktywnych serwisów
- czas, który przeszukiwacz musi odczekać przed pobraniem innej strony z tego samego serwera WWW

Jeśli przeszukiwacz ma jeden aktywny serwis na wątek przeszukiwacza i musi odczekać dwie sekundy, zanim będzie mógł pobrać inną stronę z tego samego serwera WWW, to przeszukiwacz nie może przeszukiwać szybciej niż jedna strona na wątek na dwie sekundy. Na przykład, jeśli przeszukiwacz używa domyślnej liczby wątków (200), to może przeszukiwać z szybkością 100 stron na sekundę dla 200 wątków.

Jeśli jest dwa razy więcej aktywnych serwisów niż wątków przeszukiwacza, a przeszukiwacz musi odczekać dwie sekundy, zanim będzie mógł pobrać inną stronę z tego samego serwera WWW, to przeszukiwacz może uzyskać jedną stronę na wątek w ciągu sekundy. Jednak czynnikami ograniczającymi są wówczas szybkość pobierania danych w sieci oraz przepustowość bazy danych. Wysoka wydajność przeszukiwacza wskazywana jest wtedy, gdy szybkość przeszukiwania jest równa liczbie wątków przeszukiwacza, aktywnych serwisów oraz czasowi oczekiwania przeszukiwacza.

Innym czynnikiem, który można przejrzeć podczas monitorowania wydajności przeszukiwacza sieci WWW, jest liczba adresów URL, które przeszukiwacz przeszukał od momentu uruchomienia bieżącej sesji przeszukiwacza. Podzielenie tej liczby przez całkowity czas działania przeszukiwacza daje wynik w postaci średniej długookresowej przepustowości. Jeśli ta liczba się nie zwiększa, przeszukiwacz kończy działanie lub nie może go kontynuować. Na przykład błędy związane z nawiązaniem połączenia z siecią, błędy bazy danych i niepowodzenia wyszukiwania na serwerze DNS mogą uniemożliwić kontynuowanie pracy przeszukiwacza.

## Tworzenie raportów przeszukiwacza sieci WWW

Zapoznając się z raportami dotyczącymi dotychczasowych działań przeszukiwacza sieci WWW, można ocenić jego ogólną wydajność i odpowiednio dopasować właściwości przeszukiwacza oraz definicje przestrzeni przeszukiwania.

### Zagadnienia wstępne

Jeśli uprawnienia przyznane użytkownikowi ograniczone są do monitorowania kolekcji, ma on możliwość przeglądania statystyki przeszukiwacza lub tworzenia raportów o działaniu przeszukiwacza, ale nie może wpływać na jego działanie (uruchamiać lub zatrzymywać przeszukiwacz).


### Informacje o tym zadaniu

Różne typy raportów zapewniają dostęp do informacji o różnych aspektach działania przeszukiwacza sieci WWW. W przypadku niektórych typów raportów informacje są zwracane, jak tylko zostaną one pobrane z wewnętrznej bazy danych przeszukiwacza. Z kolei utworzenie raportu o serwisie i raportów o kodach powrotu HTML wymaga czasu. Podczas tworzenia raportów tego typu można podać adres e-mail, na który ma zostać przesłany raport; wówczas nie trzeba czekać, aż wyniki zostaną zwrócone na konsolę administracyjną wyszukiwania korporacyjnego.


Aby uzyskać informacje o właściwej interpretacji danych statystycznych zawartych w raportach, można użyć przycisku **Pomoc** w trakcie monitorowania przeszukiwacza sieci WWW i tworzenia raportu.


### Procedura

Aby utworzyć raporty o przeszukiwaczu sieci WWW, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję będącą właścicielem przeszukiwacza, który ma być monitorowany, i kliknij ikonę  **Monitoruj**.
2. Otwórz stronę Przeszukiwanie.



**Wskazówka:** W przypadku, gdy wykonywana jest edycja kolekcji i strona Przeszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby zmienić widok na umożliwiający monitorowanie przeszukiwaczy.

3. Jeśli przeszukiwacz sieci WWW, dla którego mają być tworzone raporty, jest uruchomiony lub wstrzymany, kliknij ikonę  **Szczegóły**.
4. Na stronie szczegółowych informacji o przeszukiwaczu sieci WWW wybierz opcję określającą typ raportu, który ma być utworzony.
  - W obszarze **Podsumowanie statusu przeszukiwacza** kliknij opcję **Historia przeszukiwacza**, aby wygenerować raporty na temat przeszukiwacza i wszystkich wykrytych lub przeszukanych przez niego serwisów.
  - W obszarze **Status adresu URL** określ adres URL wybranego serwisu, którego ma dotyczyć tworzony raport, kliknij opcję **Widok**, a następnie opcję **Historia serwisu**.
5. W przypadku obu typów raportu - historii przeszukiwacza i historii serwisu - zaznacz pole wyboru obok każdej informacji statystycznej, jaka ma trafić do raportu, i kliknij opcję **Wyświetl raport**.

Dla tych typów danych statystycznych przeszukiwacz zwraca raport na konsolę administracyjną natychmiast po pobraniu informacji z wewnętrznej bazy danych.

6. Przy tworzeniu raportów z historii przeszukiwacza określ opcje dotyczące tworzenia raportu o serwisie i kliknij opcję **Uruchom raport**.

Raport zawierający wybrane dane statystyczne zostaje wygenerowany i zapisany w pliku o wskazanej nazwie (nazwa pliku musi być bezwzględna). Zakończenie generowania raportu może być sygnalizowane wysłaniem wiadomości e-mail.

7. Przy tworzeniu raportów z historii przeszukiwacza określ opcje dotyczące tworzenia raportu o kodach powrotu HTTP i kliknij opcję **Uruchom raport**.

Ten raport zawiera informacje o liczbie kodów powrotu HTTP przypadających na poszczególne serwisy. Raport zostaje zapisany w pliku o podanej nazwie (nazwa pliku musi być bezwzględna). Zakończenie generowania raportu może być sygnalizowane wysłaniem wiadomości e-mail.

Ten raport pozwala ustalić, które serwisy zwracają dużą liczbę kodów powrotu typu 4xx (które oznaczają niemożliwość znalezienia stron), typu 5xx (które wskazują na problem w działaniu serwera), typu 6xx (wskazujących na problem z połączeniem) itd.

Raport ten jest najbardziej użyteczny w przypadku przeszukiwaczy, które przez pewien czas są aktywne (na przykład przeszukiwacz działający całymi tygodniami). W ten sposób można zidentyfikować zaginione serwisy, serwisy nowo powstałe, serwisy o dużej liczbie adresów URL (co może wskazywać na nadmiarowe przeszukiwanie bazy danych Lotus Notes) oraz serwisy z rekurencyjnym systemem plików obsługiwane przez serwer HTTP. Jeśli serwisy z dużą liczbą kodów powrotu HTTP nie są dostarczane do indeksu, można poprawić wydajność przeszukiwacza, usuwając te serwisy z przestrzeni przeszukiwania.

## Kody powrotu HTTP przeszukiwacza sieci WWW

Podczas monitorowania przeszukiwacza sieci WWW można wyświetlić informacje dotyczące kodów powrotu HTTP otrzymywanych przez przeszukiwacz od stron, które próbuje on przeszukiwać.

### Podsumowanie tabeli

Podczas monitorowania historii przeszukiwacza sieci WWW lub statusu konkretnego adresu URL można uzyskać informację dotyczącą zwracanych do przeszukiwacza kodów powrotu HTTP. Jest ona przydatna do zarządzania przestrzenią przeszukiwania oraz optymalizowania wydajności przeszukiwacza. Na przykład, jeśli przeszukiwacz otrzymuje dużą liczbę kodów

powrotu HTTP dla adresu URL, które wskazują, że strony w tej lokalizacji nie mogą być przeszukiwane, można zwiększyć wydajność, usuwając ten adres URL z przestrzeni przeszukiwania.

Poniższa tabela zawiera kody powrotu HTTP i sposób ich interpretacji przez przeszukiwacz sieci WWW. Wartości od 100 do 505 są standardowymi kodami powrotu HTTP (więcej informacji można znaleźć na stronie <http://www.w3.org/Protocols/rfc2616/rfc2616.html>). Pozostałe kody powrotu HTTP są zastrzeżone dla wyszukiwania korporacyjnego i przeszukiwacza sieci WWW.

Tabela 4. Kody powrotu HTTP przeszukiwacza sieci WWW

Kod	Opis	Kod	Opis	Kod	Opis	Kod	Opis
NULL	Nie przeszukany	401	Brak uprawnień	500	Błąd wewnętrzny serwera	700	Błąd analizatora składni (brak zakończenia nagłówka)
100	Kontynuacja	402	Wymagana opłata	501	Nie zaimplementowano	710	Błąd analizatora składni (nagłówki)
200	Powodzenie	403	Zabronione	502	Nieprawidłowy gateway	720	Błąd analizatora składni (brak kodu HTTP)
201	Utworzone	404	Nie znaleziono	503	Usługa niedostępna	730	Błąd analizatora składni (treść)
202	Zaakceptowane	405	Niedozwolona metoda	504	Przekroczenie limitu czasu gatewaya	740	Wykluczone przez plik robots.txt
203	Niemiarodajna informacja	406	Nie akceptowalne	505	Nieobsługiwana wersja HTTP	741	Roboty tymczasowo niedostępne
204	Brak treści	407	Wymagane uwierzytelnienie serwera proxy	611	Błąd odczytu	760	Wykluczone w definicji przestrzeni przeszukiwania
205	Zerowanie treści	408	Przekroczenie limitu czasu żądania	612	Błąd połączenia	770	Nieprawidłowy protokół lub niestandardowy port systemu
206	Częściowa treść	409	Konflikt	613	Przekroczenie limitu czasu odczytu	780	Wykluczone przez wykluczenia typów plików
300	Wielokrotne wybory	410	Niedostępne	614	Uzgadnianie SSL zakończyło się niepowodzeniem	2004	Brak znacznika META indeksu
301	Przeniesione na stałe	411	Wymagana długość	615	Inny błąd odczytu	3020	Przekierowanie programowe
302	Znalezione	412	Warunek wstępny nie został spełniony	616	Anomalia FBA	4044	Wykluczone przez plik robots.txt
303	Patrz inny	413	Zbyt duża jednostka żądania	617	Błąd kodowania		

Tabela 4. Kody powrotu HTTP przeszukiwacza sieci WWW (kontynuacja)

Kod	Opis	Kod	Opis	Kod	Opis	Kod	Opis
304	Nie zmodyfikowane	414	Identyfikator URI w żądaniu jest za długi	680	Niepowodzenie wyszukiwania na serwerze DNS		
305	Użyj serwera proxy	415	Nieobsługiwany typ nośnika				
306	(Nieużywane)	417	Oczekiwania nie zostały spełnione				
307	Tymczasowe przekierowanie						

## Uwagi do tabeli

### Kody powrotu 4xx

Kod 400 (błędne żądanie) występuje rzadko. Zgodnie ze standardem kodów powrotu HTTP kody 4xx wskazują błąd klienta (przeszukiwacza). Jednak zazwyczaj problem dotyczy serwera lub adresu URL, który przeszukiwacz otrzymał jako odsyłacz. Na przykład niektóre serwery WWW nie tolerują adresów URL przechodzących z katalogu głównego serwisu (takich jak <http://xyz.ibm.com/.../sales>). Inne serwery WWW nie mają problemu z przechodzeniem do katalogu macierzystego i ignorują operator katalogu macierzystego (.), gdy przeszukiwacz już jest w katalogu głównym.

Niektóre serwery traktują żądanie dostępu do katalogu głównego serwisu jako błąd, a niektóre przestarzałe odsyłacze mogą żądać operacji, które nie są już rozpoznawane ani implementowane. Przy próbie uzyskania dostępu do strony, która już nie działa, serwer aplikacji zgłasza wyjątek, który powoduje zwrócenie przez serwer WWW kodu powrotu HTTP 400, ponieważ żądanie to nie jest już traktowane jako poprawne.

- 615** Wskazuje, że serwer przeszukiwacza pobierający dane z serwisów WWW napotkał nieoczekiwany wyjątek. Duża liczba tego typu kodów powrotu może wskazywać na wystąpienie problemu z przeszukiwaczem.

### Kody powrotu 6xx

Poza kodem 615 kody powrotu 6xx wskazują na problemy, których wystąpienia podczas przeszukiwania można się spodziewać, takich jak przekroczenia limitu czasu. Następujące kody powrotu mogą wymagać działań naprawczych:

#### 611, 612 i 613

Oznaczają powolne serwisy lub niską wydajność sieci.

- 614** Wskazuje, że przeszukiwacz nie może przeszukiwać bezpiecznych serwisów (HTTPS). Jeśli użytkownik jest przekonany, że te serwisy powinny być dostępne, należy sprawdzić prawidłowość instalacji certyfikatów na serwerze przeszukiwacza i na docelowym serwerze WWW. Na przykład, jeśli serwis jest certyfikowany przez autoryzowane ośrodki certyfikacji (CA), można dodać nowe autoryzowane ośrodki certyfikacji do zaufanych certyfikatów używanych przez przeszukiwacz.

Można również sprawdzić, w jaki sposób samopodpisane certyfikaty są skonfigurowane w przeszukiwanych serwisach. Przeszukiwacz jest skonfigurowany do akceptowania samopodpisanych certyfikatów. Niektóre serwisy tworzą samopodpisany certyfikat dla głównego adresu URL (na przykład <http://sales.ibm.com/>), a następnie próbują go używać w poddomenach (na przykład <http://internal.sales.ibm.com/>). Przeszukiwacz

nie może akceptować używanych w ten sposób certyfikatów. Przeszukiwacz akceptuje samopodpisane certyfikaty tylko w nazwie domeny będącej podmiotem certyfikatu (sales.ibm.com), a osoba podpisująca certyfikat sprawdza zgodność nazwy domeny strony, do której zażądano dostępu.

- 616** Wskazuje, że formularz logowania nadal pojawia się podczas pobierania po ponownym uwierzytelnieniu.
- 617** Wskazuje, że nie można utworzyć łańcucha w oparciu o bajtową treść dokumentu, ponieważ łańcuch kodowania (zestaw znaków) jest nieprawidłowy lub dokument zawiera niepoprawne bajty.
- 680** Wskazuje, że przeszukiwacz nie mógł uzyskać w przestrzeni przeszukiwania adresów IP hostów, być może z powodu problemów z dostępem do sieci. Ten typ błędu oznacza, że przeszukiwacz nie może przeszukiwać całych serwisów, a nie tylko niektórych adresów URL. Duża liczba kodów powrotu tego typu znacząco ogranicza przepustowość.

### Kody powrotu 7xx

Kody 7xx w większości podlegają regułom przestrzeni przeszukiwania:

#### 710 - 730

Wskazują, że problemy uniemożliwiły przeszukiwaczowi pełne pobranie danych lub że przeszukiwacz napotkał w serwisie niepoprawne dane HTML. Jeśli występuje duża liczba kodów powrotnych tych typów, należy skontaktować się z przedstawicielem oferującym wsparcie wyszukiwania korporacyjnego.

#### 740 lub 4044

Wskazują, że treść pliku nie może zostać poindeksowana, ponieważ dokument jest wykluczony przez ograniczenia zawarte w pliku robots.txt serwisu.

**740** Wskazuje, że w indeksie mogą być uwzględnione zakotwiczone odsyłacze wskazujące na wykluczony dokument.

**4044** Wskazuje, że zakotwiczone odsyłacze w dokumentach wskazujące na wykluczony dokument są wykluczone także z indeksu.

- 741** Wskazuje, że plik robots.txt serwisu pozwala na przeszukiwanie, ale pobranie danych zakończyło się niepowodzeniem. Jeśli w sposób powtarzalny nie można przeszukać tego adresu URL, adres ten jest usuwany z przestrzeni przeszukiwania. Jeśli występuje duża liczba tego typu kodów powrotu, należy sprawdzić, czy serwis docelowy jest tymczasowo czy trwale niedostępny. Jeśli serwis docelowy nie jest już dostępny, należy go usunąć z przestrzeni przeszukiwania.

Pozostałe kody powrotu 7xx występują zwykle, gdy dokonano zmian w przestrzeni przeszukiwania po uruchomieniu przeszukiwacza na chwilę. Te kody powrotu zazwyczaj nie wskazują problemów, którymi trzeba się zająć.

- 3020** Wskazuje, że dokument z kodem powrotu 200 zawiera nagłówek położenia, który kieruje agenta użytkownika na inny adres URL.

---

## Monitorowanie analizatora składni

Monitorowanie analizatora składni służy do wyświetlania informacji o dokumentach, które są analizowane, przed dodaniem ich do indeksu wyszukiwania korporacyjnego. Istniejące opcje umożliwiają przeglądanie danych statystycznych i administrowanie działaniem analizatora składni.

## Zagadnienia wstępne

Jeśli rola administracyjna użytkownika jest ograniczona do monitorowania kolekcji, może on wyświetlać status analizatora składni, ale nie może uruchamiać ani zatrzymywać analizatora.

### Informacje o tym zadaniu


Podczas monitorowania szczegółowych informacji o analizatorze składni jest wyświetlany obraz stanu działania analizatora, który zawiera dane statystyczne dotyczące działania analizatora składni w określonym momencie. Informacje statystyczne obejmują: liczbę dokumentów przeszukanych, analizowanych lub oczekujących na analizę i liczbę przeanalizowanych dokumentów oczekujących na zapisanie w indeksie.


Kiedy analizator składni jest aktywny, komunikaty zawierają dodatkowe informacje o jego stanie. Na przykład:


- Analizator składni może być w trakcie analizowania dokumentów.
- Analizator składni może być bezczynny. Analizator oczekuje w stanie uśpienia, aż pojawi się więcej dokumentów do analizy. W przypadku błędów analizator składni oczekuje na restart. Analizator składni jest restartowany automatycznie, gdy nie są dostępne usługi analizatora (automatyczny restart następuje na przykład wtedy, gdy nie można nawiązać połączenia z usługą analizatora lub gdy wszystkie wirtualne maszyny języka Java analizatora są zajęte innymi kolekcjami).
- Analizator składni może być wstrzymany (na przykład wtedy, gdy oczekuje na zakończenie reorganizacji indeksu).

### Procedura

Aby monitorować analizator składni dla kolekcji, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję, którą chcesz monitorować, i kliknij opcję  **Monitoruj**.
2. Otwórz stronę Analiza składni.


**Wskazówka:** W przypadku, gdy wykonywana jest edycja kolekcji i strona Analiza składni jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

3. Jeśli analizator składni jest uruchomiony i chcesz wyświetlić szczegółowe informacje o działaniu analizatora, kliknij opcję  **Szczegóły**.

Jeśli rola administracyjna użytkownika zezwala na administrowanie procesami kolekcji, może on uruchamiać i zatrzymywać analizator składni podczas wyświetlania szczegółowych informacji o jego działaniu.

4. Jeśli analizator składni został zatrzymany, a trzeba go uruchomić, kliknij ikonę  **Uruchom**.

Jeśli najpierw jest tworzona kolekcja, analizator składni należy uruchamiać dopiero po tym, jak przeszukiwacz rozpocznie przeszukiwanie danych. W ten sposób można dostarczyć analizatorowi składni dane do analizy i klasyfikowania. Jeśli reguły analizy nie są zmieniane, analizator może działać bez przerwy.

5. Jeśli analizator jest uruchomiony i chcesz go zatrzymać, kliknij przycisk  **Zatrzymaj**.

Po zmianie reguł analizowania należy zatrzymać i zrestartować analizator składni. Na przykład jeśli zostanie zmieniona konfiguracja analizatora składni, zanim zmiany zaczną obowiązywać, należy go zatrzymać i zrestartować.

---

## Monitorowanie aktywności indeksu dla kolekcji


Monitorowanie operacji na indeksie związanym z kolekcją pozwala wyświetlić postęp operacji budowania indeksu, włączyć lub wyłączyć harmonogram operacji na indeksie bądź uruchomić lub zatrzymać operacje indeksowania.


### Zagadnienia wstępne






Do monitorowania operacji na indeksach uprawnieni są wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego. Aby uruchomić lub zatrzymać budowanie indeksu albo włączyć lub wyłączyć harmonogram indeksu, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora kolekcji dla tej kolekcji.

### Procedura

Aby monitorować operacje na indeksie kolekcji, należy wykonać poniższe czynności:

1. W widoku Kolekcje znajdź kolekcję, którą chcesz monitorować, i kliknij opcję  **Monitoruj**.
2. Otwórz stronę Indeks.

**Wskazówka:** Jeśli wykonywana jest edycja kolekcji i strona Indeksowanie jest już wyświetlona, można kliknąć ikonę  **Monitoruj**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

3. Jeśli indeks jest zaplanowany, ale nie ma być budowany w zaplanowanym terminie, kliknij opcję  **Wyłącz harmonogram**. Indeks nie będzie budowany aż do momentu włączenia harmonogramu lub uruchomienia procesu budowania indeksu.
4. Jeśli indeks jest zaplanowany, ale harmonogram budowania jest wyłączony, kliknij opcję  **Włącz harmonogram**. Indeks zostanie umieszczony w kolejce oczekiwania na budowanie zgodnie z terminem określonym w harmonogramie operacji na indeksie.
5. Jeśli indeks został zatrzymany, a trzeba go uruchomić, kliknij ikonę  **Uruchom**.  
Zazwyczaj indeksowanie wykonywane jest regularnie zgodnie z harmonogramem. W przypadku zatrzymania indeksu w trakcie budowania lub wyłączenia harmonogramu operacji na indeksie kliknięcie ikony **Uruchom** pozwala wymusić rozpoczęcie operacji budowania indeksu.
6. Jeśli budowanie indeksu trwa i chcesz je zatrzymać, kliknij opcję  **Zatrzymaj**.  
Zatrzymanie budowania indeksu może być konieczne na przykład w celu wymuszenia reorganizacji indeksu po zmianie typu klasyfikacji w ramach kolekcji.
7. Jeśli podczas budowania indeksu wystąpiły błędy, kliknij przycisk  **Błąd**.  
Zostanie wyświetlona strona Zawartość pliku dziennika, na której znajdują się dodatkowe informacje o błędach indeksowania. Na tej stronie można wybrać poszczególne komunikaty o błędach i wyświetlić szczegóły dotyczące danego problemu.

### Pojęcia pokrewne

“Administrowanie indeksem wyszukiwania korporacyjnego” na stronie 117

Aby użytkownicy mieli zapewniony dostęp do najnowszych informacji, w wyszukiwaniu korporacyjnym dla każdej kolekcji tworzony jest indeks, a jego treść jest regularnie odświeżana i reorganizowana.

---

## Monitorowanie kolejki indeksu wyszukiwania korporacyjnego

Użytkownik może wyświetlać status wszystkich operacji budowania indeksów w kolejce indeksów, zatrzymać operację budowania indeksu, a także usunąć indeks z kolejki.

### Zagadnienia wstępne




Aby administrować kolejką indeksów, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Informacje o tym zadaniu

Jednocześnie może być budowanych wiele indeksów, ale w kolejce w danej chwili może znajdować się tylko jeden indeks dla danej kolekcji. Podczas konfigurowania w systemie opcji dotyczących indeksów podaje się liczbę indeksów, które mogą współużytkować kolejkę i współbieżnie indeksować zasoby.

### Procedura

Aby monitorować kolejkę indeksów, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Wybierz stronę Indeksowanie.  
Wyświetlona zostanie lista kolekcji, których indeksy znajdują się w kolejce indeksów. Dla każdego indeksu wyświetlane są informacje o typie budowanego indeksu (odświeżanie lub reorganizacja), godzina wprowadzenia indeksu do kolejki oraz godzina, o której budowanie indeksu zostało rozpoczęte (jeśli proces jest w trakcie realizacji).
3. Aby podjąć administrację wybranego indeksu, kliknij ikonę **Status**.  
Na przykład można sprawdzić, ile czasu potrzeba jeszcze na zakończenie operacji budowania indeksu lub ile dokumentów znajduje się w indeksie, bądź wyłączyć harmonogram operacji na indeksie.
4. Aby zatrzymać indeks będący w trakcie budowania, kliknij ikonę  **Zatrzymaj**.  
Na przykład, jeśli zmienione zostały reguły kategorii, może zaistnieć potrzeba zatrzymania odświeżania indeksu, aby zamiast tego wymusić rozpoczęcie jego reorganizacji.  
Aby uruchomić budowanie indeksu po jego zatrzymaniu, należy poczekać aż indeks znajdzie się w kolejce po nadejściu następnego zaplanowanego terminu lub kliknąć ikonę **Status**, aby monitorować indeks, a następnie kliknąć opcję  **Uruchom**, aby odświeżyć lub zreorganizować indeks.
5. Aby usunąć indeks z kolejki indeksów, kliknij opcję  **Usuń**.

### Pojęcia pokrewne

“Administrowanie indeksem wyszukiwania korporacyjnego” na stronie 117

Aby użytkownicy mieli zapewniony dostęp do najnowszych informacji, w wyszukiwaniu korporacyjnym dla każdej kolekcji tworzony jest indeks, a jego treść jest regularnie odświeżana i reorganizowana.

---

## Monitorowanie serwerów wyszukiwania


Można wyświetlać szczegółowe informacje o statusie działań serwera wyszukiwania dla określonej kolekcji lub szczegółowe informacje o statusie serwerów wyszukiwania w całym systemie wyszukiwania korporacyjnego.


### Zagadnienia wstępne



Wszyscy administratorzy wyszukiwania korporacyjnego mogą monitorować serwery wyszukiwania kolekcji, do których mają uprawnienia administracyjne. Aby monitorować wszystkie serwery wyszukiwania w systemie wyszukiwania korporacyjnego, wymagana jest rola administratora wyszukiwania korporacyjnego.

Aby zatrzymać lub uruchomić serwer wyszukiwania, wymagana jest rola administratora wyszukiwania korporacyjnego albo uprawnienia administratora lub operatora kolekcji dla tej kolekcji.

### Procedura

1. Aby monitorować serwery wyszukiwania dla pojedynczej kolekcji, należy wykonać poniższe czynności:
  - a. W widoku Kolekcje znajdź kolekcję, którą chcesz monitorować, i kliknij opcję  **Monitoruj**.
  - b. Otwórz stronę Wyszukiwanie.

**Wskazówka:** W przypadku, gdy wykonywana jest edycja kolekcji i strona Wyszukiwanie jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

2. Aby monitorować wszystkie serwery wyszukiwania w systemie wyszukiwania korporacyjnego, należy wykonać poniższe czynności:
  - a. Kliknij opcję **System**, aby otworzyć widok System.
  - b. Wybierz stronę Wyszukiwanie.
3. Jeśli serwer wyszukiwania jest zatrzymany i chcesz go uruchomić, kliknij opcję  **Uruchom**.
4. Jeśli serwer wyszukiwania jest uruchomiony i chcesz go zatrzymać, kliknij opcję  **Zatrzymaj**.

Jeśli zostanie włączona lub wyłączona pamięć podręczna wyszukiwania, zostanie zmieniona wielkość pamięci podręcznej lub gdy zostaną wprowadzone zmiany do szybkich odsyłaczy, należy zatrzymać i zrestartować serwery wyszukiwania, aby zmiany zostały uwzględnione.
5. Aby wyświetlić podsumowanie czasu przeznaczanego przez serwer wyszukiwania na przetwarzanie żądań wyszukiwania, kliknij opcję **Historia czasu odpowiedzi**.

Raport przedstawia średni czas (w milisekundach) przeznaczony przez serwer wyszukiwania na udzielanie odpowiedzi na żądania wyszukiwania w wybranym dniu. Średni czas odpowiedzi jest wskaźnikiem ogólnej sprawności systemu i zarazem jakości usług. Wydłużenie czasu odpowiedzi może wskazywać, że system jest mocno obciążony. Na przykład duża liczba objętych wyszukiwaniem kolekcji oraz rozmiar kolekcji mogą stanowić dla systemu nadmierne obciążenie.
6. Aby wyświetlić listę najczęściej wprowadzanych zapytań, kliknij opcję **Często zadawane zapytania**.

Raport zawiera słowa kluczowe użyte w 50 najczęściej wprowadzanych zapytaniach z wyszczególnieniem częstotliwości ich wprowadzania.

Przeglądając najczęściej zadawane zapytania, można zorientować się w potrzebach w zakresie tworzenia szybkich odsyłaczy. Tworzenie szybkich odsyłaczy korzystnie wpływa na jakość wyszukiwania dla wielu użytkowników. Dzięki temu dokumenty o dużym stopniu adekwatności zawsze będą zwracane w wynikach wyszukiwania.

Można również zamieszczać na stronach portalu korporacyjnego odsyłacze do zasobów odpowiadających tym zapytaniom. Na przykład, jeśli użytkownicy bardzo często wyszukują informacji na temat rozliczania wydatków, warto zamieścić odsyłacz do strony omawiającej procedury rozliczania wydatków na intranetowej stronie głównej.
7. Aby wyświetlić listę ostatnio zadawanych zapytań, kliknij opcję **Ostatnie zapytania**.



Raport przedstawia słowa kluczowe zawarte w 50 ostatnio wprowadzonych zapytaniach. Przeglądając ostatnie zapytania, można rozpoznawać aktualnie panujące tendencje oraz nagle sytuacje występujące w organizacji. Na przykład można zaobserwować wzrost zainteresowania pewnym tematem. Taki wzrost zainteresowania może świadczyć o tym, że należy utworzyć szybki odsyłacz do stron poświęconych temu tematowi albo że informacje na ten temat należy udostępnić użytkownikom w inny sposób (na przykład zamieszczając bezpośredni odsyłacz w portalu korporacyjnym).

---

## Monitorowanie nasłuchiwania danych


Monitorowanie nasłuchiwania danych służy do wyświetlania statusu i szczegółów działania klienta aplikacji nasłuchiwania danych.

### Zagadnienia wstępne

Aby monitorować nasłuchiwanie danych, wymagana jest rola administratora wyszukiwania korporacyjnego.

### Procedura

Aby monitorować program nasłuchiwania danych, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Na stronie Nasłuchiwanie danych zwróć uwagę na ikony statusu, aby sprawdzić, czy nasłuchiwanie danych jest aktywne czy zatrzymane.
3. Jeśli program nasłuchiwania danych jest uruchomiony, a chcesz uzyskać szczegółowe informacje o statusie działania aplikacji klienckiej, kliknij ikonę  **Szczegóły**.

Ikony statusu na stronie Szczegóły nasłuchiwania danych wskazują, czy nasłuchiwanie danych jest uruchomione czy zatrzymane. Na informacje statystyczne składają się: liczba żądań oczekujących na przetwarzanie, aktualny stan każdego wątku zajętego przez żądania aplikacji klienckiej oraz liczba aktywnych wątków znajdujących się w określonym stanie.

4. W przypadku zmiany numeru portu dla nasłuchiwania danych kliknij ikonę .

#### Restartuj.

Nasłuchiwanie danych jest uruchamiane podczas startu systemu wyszukiwania korporacyjnego. Nie ma potrzeby restartowania programu nasłuchiwania danych, o ile nie został zmieniony jego numer portu.

### Zadania pokrewne

“Konfiguracja obsługi aplikacji nasłuchiwania danych” na stronie 77

Istnieje możliwość rozszerzenia wyszukiwania korporacyjnego przez wykorzystanie funkcji API nasłuchiwania danych do utworzenia przeszukiwacza zewnętrznego.

Niestandardowe aplikacje nasłuchiwania danych mogą dodawać dane do kolekcji lub wskazywać sposób, w jaki przeszukiwacz sieci WWW ma odwiedzać adresy URL.

---

## Sledzenie dokumentu

Dokumenty mogą być usuwane z systemu na różnych etapach przetwarzania. Istnieje możliwość określenia opcji powiadamiania o usuwaniu dokumentów oraz o problemach, które spowodowały usunięcie.

Gdy analizator składni napotyka błąd uniemożliwiający analizę dokumentu, informacja o usunięciu dokumentu zostaje zarejestrowana w postaci odpowiedniego komunikatu wraz z kodem przyczyny. (Ten typ błędu nie powoduje usunięcia z indeksu starszych wersji danego dokumentu).

Dokumenty mogą być usuwane podczas etapów indeksowania, a informacje o tym także są rejestrowane. Na przykład identyfikatory URI i ich wzorce mogły zostać jawnie usunięte. Dokument mógł zostać przeszukany przez przeszukiwacza, który został później usunięty. Dokument źródłowy może już nie istnieć (z dokumentem związany jest ujemny kod HTTP) lub kod HTTP związany z dokumentem może być nieznan. Usunięcie jest też możliwe, gdy w dokumencie wymagającym analizy globalnej nie ma informacji o jego ocenie.

Jeśli wiadomo, że pewien dokument został przeszukany, ale nie znajduje się w indeksie, za pomocą konsoli administracyjnej wyszukiwania korporacyjnego można prześledzić drogę tego dokumentu przez system. Szczegółowe raporty mogą wskazać czas, miejsce i przyczynę usunięcia dokumentu. Na przykład raport może zawierać informację, że dokument został nieoczekiwanie usunięty podczas analizy globalnej lub że administrator usunął określony identyfikator URI z indeksu.

#### **Zadania pokrewne**

“Wyświetlanie szczegółów dotyczących identyfikatora URI” na stronie 216  
Użytkownik może wyświetlić szczegółowe informacje na temat identyfikatora URI. Wyświetlić można aktualne i historyczne informacje odnośnie przeszukiwania, indeksowania i wyszukiwania dokumentu reprezentowanego przez dany identyfikator URI.

## **Konfigurowanie plików dziennika do śledzenia dokumentów**

Aby określić, kiedy, gdzie i dlaczego dany dokument został usunięty z systemu, można skonfigurować pliki dziennika w celu śledzenia informacji o usuwanych dokumentach.

#### **Zagadnienia wstępne**

Aby skonfigurować opcje śledzenia usuwanych dokumentów, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora w odniesieniu do danej kolekcji.

#### **Informacje o tym zadaniu**

Aby uniknąć zajmowaniu zbyt dużej części dysku przez pliki dzienników, system prowadzi rotację tych plików, zaczynając nowy plik dziennika przy każdej zmianie daty. Jeśli plik dziennika osiągnie maksymalną dopuszczalną wielkość, system utworzy nowy plik dziennika, nawet jeśli data nie uległa jeszcze zmianie. Po osiągnięciu maksymalnej liczby plików dzienników najstarszy z plików zostaje usunięty, aby umożliwić utworzenie nowego.

#### **Procedura**

Aby skonfigurować pliki dziennika do śledzenia dokumentów:

1. Przejdź do edycji kolekcji, wybierz stronę Rejestrowanie, a następnie kliknij opcję **Konfiguruj śledzenie dokumentów**.
2. Na stronie Śledzenie dokumentów zaznacz pole wyboru dotyczące śledzenia dokumentów.
3. Podaj liczbę plików dziennika, które mają być używane do rejestrowania informacji na temat dokumentów usuwanych z systemu. Te pliki dziennika są współużytkowane przez wszystkie sesje, podczas których dokumenty mogą być usuwane.
4. Kliknij przycisk **OK**.

## Wyświetlanie raportów o usuniętych dokumentach

Użytkownik może wyświetlić szczegółowe informacje na temat dokumentów, które zostały usunięte z systemu wyszukiwania korporacyjnego. Informacje te dostępne są tylko wtedy, gdy włączona jest opcja śledzenia dokumentów danej kolekcji.

### Zagadnienia wstępne

Przed wprowadzeniem żądania wyświetlenia raportu o usuniętych dokumentach lub wysłania raportu na wskazany adres e-mail należy się upewnić, że sesje, które mają być źródłami informacji, są aktywne. Na przykład, aby zapoznać się z informacjami na temat dokumentów, które zostały usunięte podczas analizowania lub indeksowania, należy się upewnić, że sesje analizy i indeksowania kolekcji zostały uruchomione.

Zanim możliwe będzie odebranie raportu, należy się upewnić, że dla wyszukiwania korporacyjnego został skonfigurowany serwer poczty. Informacje te należy podać podczas konfigurowania opcji poczty elektronicznej na stronie Rejestrowanie w widoku System.


### Informacje o tym zadaniu


Gromadzenie informacji na temat usuwanych dokumentów jest procesem czasochłonnym. Można wybrać opcję umożliwiającą wyświetlenie informacji, a następnie poczekać na ich wyświetlenie. Metodą bardziej efektywną jest wysłanie raportu na podany przez użytkownika adres e-mail.


Gdy dokument został usunięty, raport zawiera datę i godzinę usunięcia, poziom istotności błędu, komponent i sesję, w których wystąpił błąd, oraz komunikat o błędzie.

### Procedura

Aby wyświetlić szczegółowe informacje o usuniętych dokumentach:

1. W widoku Kolekcje znajdź kolekcję, którą chcesz monitorować, i kliknij opcję  **Monitoruj**.

**Wskazówka:** W przypadku, gdy wykonywana jest edycja kolekcji i strona Ogólne jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby wyświetlić widok umożliwiający monitorowanie kolekcji.

2. Kliknij opcję  **Szczegóły dotyczące identyfikatorów URI**.
3. Na stronie Szczegóły dotyczące identyfikatorów URI wpisz identyfikator URI, którego mają dotyczyć wyświetlane informacje.
4. Zaznacz pola wyboru odpowiadające rodzajowi informacji, które mają być wyświetlone:

#### **Dokumenty usunięte przez analizator składni**

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty podczas analizowania, a jeśli tak, to jaka była tego przyczyna.

#### **Dokumenty usunięte z indeksu**

Zaznacz to pole wyboru, aby wyświetlić informacje o tym, czy dokument został usunięty w czasie indeksowania lub analizowania, a jeśli tak, to jaka była tego przyczyna.

5. Określ, w jaki sposób chcesz wyświetlić raport:
  - Aby zaczekać na wyświetlenie raportu, kliknij opcję **Wyświetl raport**.
  - Aby wysłać raport na podany adres e-mail w celu przejrzenia go w późniejszym terminie, kliknij opcję **Wyślij raport**.

| Na stronie Wysłanie szczegółowego raportu o identyfikatorach URI w polu **Adres**  
| **poczty elektronicznej do wysłania powiadomienia** wpisz adres e-mail, na który ma  
| zostać wysłany raport, a następnie kliknij przycisk **Wyślij raport**.

| **Zadania pokrewne**

| “Wyświetlanie szczegółów dotyczących identyfikatora URI” na stronie 216  
| Użytkownik może wyświetlić szczegółowe informacje na temat identyfikatora URI.  
| Wyświetlić można aktualne i historyczne informacje odnośnie przeszukiwania,  
| indeksowania i wyszukiwania dokumentu reprezentowanego przez dany identyfikator  
| URI.

| **Odsyłacze pokrewne**

| “Formaty identyfikatorów URI w indeksie wyszukiwania korporacyjnego” na stronie 79  
| Identyfikator URI każdego dokumentu w indeksie wyszukiwania korporacyjnego  
| wskazuje typ przeszukiwacza, który dodał ten dokument do kolekcji.

---

## Alerty i pliki dziennika wyszukiwania korporacyjnego

Użytkownik może wybrać typy komunikatów, które mają być zapisywane w dzienniku dla kolekcji i dla systemu, określić opcje dotyczące tworzenia i wyświetlania plików dzienników oraz opcje otrzymywania alertów oraz wiadomości e-mail dotyczących komunikatów.

Podczas wykonywania zwykłych operacji komponenty wyszukiwania korporacyjnego zapisują komunikaty dziennika we wspólnym pliku dziennika. Znajduje się on w katalogu ES\_NODE\_ROOT/logs na serwerze indeksowania. Wspólny plik dziennika można wyświetlić, korzystając z konsoli administracyjnej.

Po wystąpieniu problemu, takiego jak błąd komunikacji w sieci, komponent zapisuje odpowiednie komunikaty dziennika w katalogu logs na serwerze, na którym jest zainstalowany. Aby wyświetlić te lokalne pliki dzienników, należy użyć na komputerze lokalnym przeglądarki plików, takiej jak program narzędziowy tail w systemie UNIX. Do wyświetlenia tego typu plików dzienników nie można użyć konsoli administracyjnej.

Podczas konfigurowania plików dziennika użytkownik może wybrać typy komunikatów, które mają być rejestrowane (np. komunikaty o błędach lub ostrzeżenia), określić częstotliwość usuwania starych plików dzienników w celu zwolnienia miejsca dla nowych, określić maksymalną wielkość plików dzienników, a także wybrać język, w jakim mają być zapisywane komunikaty. Istnieje również możliwość określenia opcji dotyczących otrzymywania wiadomości e-mail po wystąpieniu określonych zdarzeń lub po zapisaniu w dzienniku określonych komunikatów lub komunikatów określonego typu.

Podczas monitorowania plików dzienników użytkownik może wybrać pliki dzienników, które mają zostać otwarte. Treść pliku dziennika można filtrować, aby wyświetlone zostały tylko komunikaty o określonym poziomie istotności (na przykład tylko komunikaty o błędach) lub komunikaty, które zostały wygenerowane przez konkretną sesję wyszukiwania korporacyjnego. Podczas wyświetlania pliku dziennika możliwe jest wyświetlenie szczegółowych informacji o poszczególnych komunikatach. Na przykład użytkownik może wyświetlić nazwę funkcji, która wygenerowała określony komunikat, a także, w razie potrzeby, inne informacje ułatwiające podjęcie odpowiednich działań naprawczych.

### Pojęcia pokrewne

"Messages for enterprise search" w podręczniku "Messages Reference"

---

## Alerty

Wyszukiwanie korporacyjne można skonfigurować tak, aby za każdym razem, gdy zostanie wykryte określone zdarzenie, w pliku dziennika zapisywane były odpowiednie komunikaty.

W wyzwalanych przez zdarzenia komunikatach nazywanych alertami dostępne są informacje o wystąpieniu interesujących dla użytkownika warunków, takich jak użycie przez określony zasób zbyt dużego obszaru pamięci. Podczas konfigurowania alertów dla wyszukiwania korporacyjnego należy określić warunki, które mają być monitorowane przez system. Za każdym razem, gdy wystąpi określony warunek, system automatycznie zapisze odpowiedni komunikat w pliku dziennika.

Aby otrzymywać bezpośrednio powiadomienia o wystąpieniu określonych warunków, należy wybrać odpowiednie opcje wysyłania wiadomości e-mail podczas rejestrowania monitorowanych komunikatów.

Istnieje możliwość skonfigurowania alertów związanych ze zdarzeniami na poziomie kolekcji i zdarzeniami występującymi na poziomie systemu. Dla zdarzeń na poziomie kolekcji system może wykonywać następujące operacje:

- monitorowanie liczby dokumentów przeszukiwanych przez każdego z przeszukiwaczy i generowanie komunikatu alertu po osiągnięciu maksymalnej dozwolonej liczby dokumentów
- monitorowanie liczby dokumentów dodawanych do indeksu związanego z używanymi kolekcjami i generowanie komunikatu alertu po osiągnięciu maksymalnej dozwolonej liczby dokumentów
- wysyłanie użytkownikowi informacji o przekroczeniu zdefiniowanego limitu czasu podczas uzyskiwania odpowiedzi na zapytanie wyszukiwania

Dla zdarzeń na poziomie systemu można monitorować miejsce na dyskach na każdym serwerze wyszukiwania korporacyjnego i generować odpowiedni komunikat alertu, gdy ilość wolnego miejsca na dysku jest mała.

## Konfiguracja alertów na poziomie kolekcji

Skonfigurowanie alertów gwarantuje, że w przypadku wystąpienia pewnych zdarzeń na poziomie kolekcji komunikaty są zapisywane w pliku dziennika. Możliwe jest również otrzymywanie wiadomości e-mail informujących o tym, że zdarzenia te zostały zarejestrowane.

### Zagadnienia wstępne

Aby skonfigurować alerty dla kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.

### Procedura

Aby skonfigurować alerty na poziomie kolekcji:

1. Przejdź do edycji kolekcji, wybierz stronę Rejestrowanie, a następnie kliknij opcję **Konfiguruj alerty**.
2. Jeśli system ma monitorować liczbę dokumentów przeszukiwanych przez przeszukiwacza, należy wykonać poniższe czynności:
  - a. Zaznacz pole wyboru **Gdy liczba dokumentów przeszukanych przez dowolny przeszukiwacz osiągnie procent maksymalnej dozwolonej liczby**.
  - b. W polu **Procent** określ, w którym momencie komunikat ma być rejestrowany. Liczbę tę należy określić jako procent maksymalnej liczby dokumentów, jaką przeszukiwacz może przeszukać (parametr **Maksymalna liczba dokumentów do przeszukania** jest definiowany podczas konfigurowania właściwości przeszukiwacza). Wartość domyślna to 90 procent.

Ponieważ dla poszczególnych przeszukiwaczy można skonfigurować różne limity, dla każdego przeszukiwacza rejestrowane są osobne komunikaty. Na przykład użycie domyślnego progu alertów oraz zezwolenie przeszukiwaczowi DB2 na przeszukiwanie 2 000 000 dokumentów i przeszukiwaczowi Notes na przeszukiwanie 1 000 000 dokumentów spowoduje, że jeden komunikat zostanie zarejestrowany po przeszukaniu przez przeszukiwacza DB2 1 800 000 dokumentów, a kolejny, gdy przeszukiwacz Notes przeszuka 900 000 dokumentów.
3. Jeśli system ma monitorować liczbę dokumentów dodawanych do indeksu, należy wykonać poniższe czynności:
  - a. Zaznacz pole wyboru **Gdy liczba dokumentów w kolekcji osiągnie określony procent szacowanej wielkości**.

- b. W polu **Procent** określ, w którym momencie komunikat ma być rejestrowany. Podana tu liczba oznacza procent szacowanej liczby dokumentów, które według przewidywań będą przechowywane w kolekcji. Wartością domyślną jest 85 procent.
- Bieżąca szacowana wielkość kolekcji jest wyświetlana w polu **Limit**. Aby zmienić tę wartość, otwórz stronę Ogólne kolekcji, wybierz opcję konfigurowania opcji ogólnych i w polu **Szacowana liczba dokumentów** podaj nową wartość.
- Ważne:** Ten limit i szacowana liczba dokumentów skonfigurowana dla kolekcji służą tylko do monitorowania wzrostu kolekcji. Wartości te nie oznaczają bezwzględnego limitu wzrostu wielkości indeksu.
4. Jeśli system ma informować o tym, że czas wymagany na odpowiedź na żądania wyszukiwania przekracza dopuszczalny limit, należy wykonać poniższe czynności:
- Zaznacz pole wyboru **Gdy czas odpowiedzi wyszukiwania przekroczy limit**.
  - W polu **Limit** wpisz wyrażony w sekundach czas, który można uznać za maksymalny czas odpowiedzi na żądanie wyszukiwania.
- Po przekroczeniu tej wartości system zarejestruje komunikat informujący o tym zdarzeniu. Na przykład przy zachowaniu wartości domyślnej system utworzy komunikat dziennika, gdy średni czas odpowiedzi serwera wyszukiwania osiągnie poziom przynajmniej pięciu sekund.
- Typowe czasy odpowiedzi nie przekraczają pół sekundy. Średnie czasy większe niż sekunda mogą wskazywać, że system operacyjny wymaga lepszego dostrojenia pod względem wydajności lub że istnieje problem związany z ustawieniami konfiguracyjnymi serwera wyszukiwania. Na przykład można zwiększyć ilość miejsca przeznaczonego na pamięć podręczną wyszukiwania.
5. Kliknij przycisk **OK**.

Aby otrzymywać wiadomości e-mail informujące o tym, że zdarzenia te zostały zarejestrowane, otwórz stronę Dziennik i kliknij opcję **Konfiguracja opcji elektronicznej dla komunikatów** w celu podania adresu e-mail. Identyfikatory komunikatów dla włączonych alertów zostaną automatycznie dodane do listy identyfikatorów komunikatów, dla których będą wysyłane wiadomości e-mail:

Aby otrzymywanie wiadomości e-mail było możliwe, należy pamiętać o skonfigurowaniu informacji o serwerze poczty. Aby to zrobić, administrator wyszukiwania korporacyjnego musi wybrać opcję **System** na pasku narzędzi, otworzyć stronę Dziennik, a następnie kliknąć opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów**.

#### **Zadania pokrewne**

“Odbieranie wiadomości poczty elektronicznej dotyczących zarejestrowanych komunikatów” na stronie 240

Istnieje możliwość określenia opcji w celu otrzymywania wiadomości e-mail informujących o zarejestrowaniu pewnych komunikatów lub typów komunikatów.

## **Konfiguracja alertów na poziomie systemu**


Skonfigurowanie alertów gwarantuje, że w przypadku wystąpienia pewnych zdarzeń na poziomie systemu komunikaty są zapisywane w pliku dziennika. Możliwe jest również otrzymywanie wiadomości e-mail informujących o tym, że zdarzenia te zostały zarejestrowane.

#### **Zagadnienia wstępne**

Alerty na poziomie systemu może konfigurować tylko administrator wyszukiwania korporacyjnego.

#### **Procedura**

Aby skonfigurować alerty na poziomie systemu:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Dziennik kliknij opcję **Konfiguruj alerty**.
4. Jeśli system ma monitorować ilość wolnej pamięci dostępnej na serwerach wyszukiwania korporacyjnego, zaznacz pole wyboru **Gdy ilość dostępnego miejsca w systemie plików osiągnie określony procent łącznej przestrzeni w tym systemie plików**.
5. W polu **Procent** określ, w którym momencie system ma powiadamiać o wyczerpywaniu się wolnej pamięci na serwerze. Wartość tę należy określić jako procent całkowitej pamięci systemu plików. Wartość domyślna to 80 procent.

Jeśli system wyszukiwania korporacyjnego jest skonfigurowany na wielu serwerach, dla każdego z nich tworzony jest osobny komunikat w dzienniku. Na przykład jeden komunikat może informować o małej ilości wolnej przestrzeni na serwerze przeszukiwacza, a inne o ograniczeniach dotyczących przestrzeni na serwerach wyszukiwania i indeksowania.

6. Kliknij przycisk **OK**.

Aby otrzymywać wiadomości e-mail informujące o tym, że zdarzenie to zostało zarejestrowane, otwórz stronę Dziennik i kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów** w celu podania adresu e-mail oraz informacji o serwerze poczty.

#### **Zadania pokrewne**

“Odbieranie wiadomości poczty elektronicznej dotyczących zarejestrowanych komunikatów” na stronie 240

Istnieje możliwość określenia opcji w celu otrzymywania wiadomości e-mail informujących o zarejestrowaniu pewnych komunikatów lub typów komunikatów.

---

## **Konfiguracja plików dziennika**

Istnieje możliwość określenia typów komunikatów, które mają być rejestrowane, oraz opcji dotyczących tworzenia plików dzienników.

#### **Zagadnienia wstępne**

Aby skonfigurować pliki dzienników na poziomie kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji. Pliki dzienników na poziomie systemu może konfigurować tylko administrator wyszukiwania korporacyjnego.

#### **Informacje o tym zadaniu**



Aby uniknąć zajmowaniu zbyt dużej części dysku przez pliki dzienników, system prowadzi rotację tych plików, zaczynając nowy plik dziennika przy każdej zmianie daty. Jeśli plik dziennika osiągnie maksymalną dopuszczalną wielkość, system utworzy nowy plik dziennika, nawet jeśli data nie uległa jeszcze zmianie. Po osiągnięciu maksymalnej liczby plików dzienników najstarszy z plików zostaje usunięty, aby umożliwić utworzenie nowego.

Aby otrzymywać wiadomości e-mail na temat rejestrowanych komunikatów, należy najpierw określić informacje dotyczące sposobu dostarczania tych wiadomości. Następnie należy określić, na temat jakich komunikatów mają być wysyłane wiadomości e-mail.

#### **Procedura**



Aby skonfigurować pliki dzienników wyszukiwania korporacyjnego, należy wykonać poniższe czynności:

1. Aby skonfigurować opcje dotyczące tworzenia i rotacji plików dzienników na poziomie systemu:
  - a. Kliknij opcję **System**, aby otworzyć widok System.
  - b. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
  - c. Na stronie Rejestrowanie kliknij opcję **Konfiguracja opcji pliku dziennika**. Zostanie wyświetlona strona Opcje pliku dziennika na poziomie systemu.
2. Aby skonfigurować opcje dotyczące tworzenia i rotacji plików dzienników na poziomie kolekcji:
  - a. W widoku Kolekcje znajdź kolekcję, dla której chcesz określić opcje, i kliknij ikonę  **Edytuj**.
  - b. Na stronie Rejestrowanie kliknij opcję **Konfiguracja opcji pliku dziennika**. Zostanie wyświetlona strona Opcje pliku dziennika na poziomie kolekcji.
3. W polu **Typ informacji do rejestrowania** wybierz typy komunikatów, które mają być rejestrowane:

#### **Tylko komunikaty o błędach**

Komunikaty o błędach informują o zaistnieniu niepożądanego zdarzenia lub nieoczekiwanego zachowania, skutkiem czego kontynuacja procesu nie jest możliwa. Należy przedsięwziąć działania mające na celu rozwiązanie problemu.

#### **Komunikaty o błędach i ostrzeżenia**

Komunikat ostrzegawczy wskazuje na możliwość wystąpienia konfliktu lub niespójności, jednak nie powodują zatrzymania procesu. Jest to opcja domyślna.

#### **Wszystkie komunikaty**

Komunikaty informacyjne udostępniają ogólne informacje na temat systemu lub bieżącego zadania i nie wymagają podejmowania jakichkolwiek działań naprawczych.

4. W polu **Maksymalna wielkość każdego pliku dziennika** wpisz maksymalną wielkość pliku dziennika wyrażoną w megabajtach. Wartością domyślną jest 5 MB.  
Po osiągnięciu przez plik dziennika wskazanej wielkości tworzony jest nowy plik dziennika, aż do czasu osiągnięcia maksymalnej dozwolonej liczby plików dzienników. Przechowywanie stosunkowo niewielkich plików dzienników ułatwia ich przeglądanie.
5. W polu **Maksymalna liczba plików dzienników** wpisz maksymalną liczbę plików dzienników możliwych do utworzenia. Wartością domyślną jest 10.  
Jeśli istnieje potrzeba zachowania starszych zarejestrowanych komunikatów do wglądu przez dłuższy czas, należy zwiększyć tę wartość. Jeśli użytkownika bardziej interesują bieżące komunikaty i nie ma potrzeby utrzymywania długiej historii, wartość tę można zmniejszyć.
6. W polu **Domyślne ustawienia narodowe** wybierz język, który ma być stosowany przy rejestrowaniu komunikatów. Wartością domyślną jest English (Angielski).
7. Kliknij przycisk **OK**.

---

## **Konfigurowanie informacji o serwerze SMTP**

Aby odbierać wiadomości e-mail dotyczące działań związanych z wyszukiwaniem korporacyjnym, należy skonfigurować informacje o serwerze SMTP (Simple Mail Transfer Protocol).

### **Zagadnienia wstępne**

Aby skonfigurować informacje o serwerze SMTP, wymagana jest rola administratora wyszukiwania korporacyjnego.


### Informacje o tym zadaniu

Kilka funkcji administracyjnych wyszukiwania korporacyjnego umożliwia otrzymywanie wiadomości e-mail. Aby można było odbierać pocztę elektroniczną generowaną przez te funkcje, należy podać informacje o serwerze SMTP:

- Po skonfigurowaniu alertów na poziomie kolekcji lub na poziomie systemu można otrzymywać pocztą elektroniczną powiadomienia o każdym zarejestrowaniu tych komunikatów. Można także otrzymywać powiadomienia pocztą elektroniczną o zarejestrowaniu innych komunikatów, nie tylko tych spowodowanych wystąpieniem monitorowanych zdarzeń.
- Aby zapoznać się ze szczegółowymi informacjami o identyfikatorze URI w indeksie lub o dokumencie, który został usunięty z systemu wyszukiwania korporacyjnego, można odebrać raport pocztą elektroniczną.
- Jeśli podczas monitorowania przeszukiwacza sieci WWW jest włączona opcja tworzenia raportów jego historii, można otrzymywać pocztą elektroniczną powiadomienia o utworzeniu raportów.

### Procedura

Aby skonfigurować informacje o serwerze SMTP, należy wykonać poniższe czynności:

1. Kliknij opcję **System**, aby otworzyć widok System.
2. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
3. Na stronie Dziennik kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów**.
4. Na stronie Opcje poczty elektronicznej dla komunikatów systemowych, w polu **Serwer SMTP do dostarczania poczty elektronicznej** wpisz pełną, kwalifikowaną nazwę hosta lub adres IP wykorzystywanego serwera.  
System będzie wykorzystywał ten serwer do wysyłania poczty elektronicznej na podane adresy.
5. W polu **Częstotliwość sprawdzania poczty elektronicznej** określ, jak często system ma sprawdzać obecność odpowiednich komunikatów i wysłać dotyczące ich wiadomości e-mail.  
System łączy wszystkie wiadomości kierowane na określony adres e-mail w jedną wiadomości i wysyła ją ze zdefiniowaną tutaj częstotliwością.
6. Kliknij przycisk **OK**.

---

## Odbieranie wiadomości poczty elektronicznej dotyczących zarejestrowanych komunikatów

Istnieje możliwość określenia opcji w celu otrzymywania wiadomości e-mail informujących o zarejestrowaniu pewnych komunikatów lub typów komunikatów.

### Zagadnienia wstępne

Aby konfigurować opcje poczty elektronicznej dla komunikatów na poziomie systemu, wymagana jest rola administratora wyszukiwania korporacyjnego. Aby konfigurować opcje poczty elektronicznej dla komunikatów na poziomie kolekcji, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia administratora kolekcji dla tej kolekcji.



Zanim możliwe będzie odbieranie poczty elektronicznej, należy najpierw skonfigurować parametry serwera SMTP, na który poczta będzie dostarczana.

### Informacje o tym zadaniu

Konfigurując alerty, można włączyć opcję rejestrowania komunikatów po wystąpieniu określonych zdarzeń. W przypadku włączenia tej opcji można następnie skonfigurować opcje dotyczące automatycznego otrzymywania wiadomości e-mail po zarejestrowaniu odpowiednich komunikatów. Istnieje również możliwość określenia opcji dotyczących otrzymywania wiadomości e-mail w momencie zarejestrowania innych komunikatów, a nie tylko komunikatów wyzwalanych przez zdarzenia.

### Procedura

Aby skonfigurować opcje poczty elektronicznej dla komunikatów, należy wykonać poniższe czynności:

1. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o komunikatach systemowych:
  - a. Kliknij opcję **System**, aby otworzyć widok System.
  - b. Kliknij opcję  **Edytuj**, aby przejść do widoku edycji systemu.
  - c. Na stronie Dziennik kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów**.
  - d. Na stronie Opcje poczty elektronicznej dla komunikatów systemowych zaznacz pole wyboru **Wyślij wiadomość e-mail na temat komunikatów dotyczących systemu**.
  - e. W polu **Adres poczty elektronicznej do odbierania poczty** wpisz jeden lub kilka adresów. Zwykle informacje o komunikatach systemowych powinien otrzymywać administrator wyszukiwania korporacyjnego.  
Adresy należy oddzielać przecinkiem. Na przykład:  
steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com.
  - f. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o wszystkich zarejestrowanych komunikatach o błędach, zaznacz pole wyboru **Wyślij wiadomość e-mail na temat wszystkich komunikatów o błędach**.
  - g. Jeśli chcesz otrzymywać pocztą elektroniczną tylko powiadomienia o zarejestrowaniu określonych komunikatów systemowych, wpisz identyfikatory tych komunikatów w obszarze **Wyślij wiadomość e-mail na temat wybranych komunikatów**. Każdy identyfikator komunikatu należy wpisać w osobnym wierszu. Na przykład:  
FFQC4819E  
FFQ00005E  
Niektóre identyfikatory komunikatów są wyliczone domyślnie. Opis tych komunikatów można wyświetlić za pomocą przycisku **Pomoc**.
  - h. Kliknij przycisk **OK**.
2. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o komunikatach dla kolekcji:
  - a. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.
  - b. Na liście kolekcji znajdź kolekcję, którą chcesz skonfigurować, i kliknij opcję  **Edytuj**.
  - c. Na stronie Dziennik kliknij opcję **Konfiguracja opcji poczty elektronicznej dla komunikatów**.
  - d. Na stronie Opcje poczty elektronicznej dla komunikatów kolekcji zaznacz pole wyboru **Wyślij wiadomość e-mail na temat komunikatów dotyczących kolekcji**.

- e. W polu **Adres poczty elektronicznej do odbierania poczty** wpisz jeden lub kilka adresów. Zwykle informacje o komunikatach na poziomie kolekcji powinien otrzymywać administrator kolekcji.  
Adresy należy oddzielać przecinkiem. Na przykład:  
steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com.
- f. Jeśli chcesz otrzymywać pocztą elektroniczną powiadomienia o wszystkich zarejestrowanych komunikatach o błędach, zaznacz pole wyboru **Wyślij wiadomość e-mail na temat wszystkich komunikatów o błędach**.
- g. Jeśli chcesz otrzymywać pocztą elektroniczną tylko powiadomienia o zarejestrowaniu określonych komunikatów dotyczących kolekcji, wpisz identyfikatory tych komunikatów w obszarze **Wyślij wiadomość e-mail na temat wybranych komunikatów**. Każdy identyfikator komunikatu należy wpisać w osobnym wierszu. Na przykład:  
FFQC4819E  
FFQ00005E  
Niektóre identyfikatory komunikatów są wyliczone domyślnie. Opis tych komunikatów można wyświetlić za pomocą przycisku **Pomoc**.
- h. Kliknij przycisk **OK**.

#### **Pojęcia pokrewne**

"Messages for enterprise search" w podręczniku "Messages Reference"

#### **Zadania pokrewne**

"Konfiguracja alertów na poziomie kolekcji" na stronie 236

Skonfigurowanie alertów gwarantuje, że w przypadku wystąpienia pewnych zdarzeń na poziomie kolekcji komunikaty są zapisywane w pliku dziennika. Możliwe jest również otrzymywanie wiadomości e-mail informujących o tym, że zdarzenia te zostały zarejestrowane.

"Konfiguracja alertów na poziomie systemu" na stronie 237

Skonfigurowanie alertów gwarantuje, że w przypadku wystąpienia pewnych zdarzeń na poziomie systemu komunikaty są zapisywane w pliku dziennika. Możliwe jest również otrzymywanie wiadomości e-mail informujących o tym, że zdarzenia te zostały zarejestrowane.

---

## **Wyświetlanie plików dziennika**


Użytkownik może wyświetlać komunikaty zapisywane we wspólnym pliku dziennika przez komponenty systemu i kolekcji. Na stronie tej można także określić filtry w celu wyświetlenia komunikatów o konkretnym poziomie istotności oraz komunikatów z określonych sesji wyszukiwania korporacyjnego.


### **Zagadnienia wstępne**


Wszyscy użytkownicy z uprawnieniami administracyjnymi wyszukiwania korporacyjnego mogą przeglądać zawartość plików dzienników dotyczących kolekcji, do administrowania którymi są upoważnieni. Aby móc przeglądać systemowe pliki dziennika, wymagana jest rola administratora wyszukiwania korporacyjnego lub uprawnienia dostępu do paska narzędzi **System**.

### **Procedura**

1. Aby wyświetlić pliki dzienników dla pojedynczej kolekcji, należy wykonać poniższe czynności:
  - a. Kliknij opcję **Kolekcje**, aby otworzyć widok Kolekcje.

- b. Na liście kolekcji znajdź kolekcję, którą chcesz wyświetlić, kliknij opcję  **Monitor** i otwórz stronę Rejestrowanie.

**Wskazówka:** W przypadku, gdy wykonywana jest edycja kolekcji i strona Dziennik jest już wyświetlona, można kliknąć ikonę  **Monitor**, aby zmienić widok na umożliwiający monitorowanie kolekcji.

2. Aby wyświetlić pliki dzienników na poziomie systemu, należy wykonać poniższe czynności:
  - a. Kliknij opcję **System**, aby otworzyć widok System.
  - b. Wybierz stronę Dziennik.
3. W polu **Plik dziennika** wybierz plik dziennika, który ma zostać wyświetlony.  
W nazwie każdego pliku dziennika zakodowany jest typ dziennika (taki jak identyfikator systemu lub kolekcji), data utworzenia pliku oraz przyrostek liczbowy wskazujący kolejność utworzenia tego pliku w danym dniu. Na przykład:  
*typ\_pliku\_dziennika\_2005-05-26\_1.log*  
*typ\_pliku\_dziennika\_2005-05-26\_2.log*  
*typ\_pliku\_dziennika\_2005-05-25\_1.log*  
*typ\_pliku\_dziennika\_2005-05-25\_2.log*  
*typ\_pliku\_dziennika\_2005-05-25\_3.log*
4. Aby wyświetlać wyłącznie komunikaty o określonych poziomach istotności, zaznacz odpowiednie pola wyboru w polu **Istotność**.
5. Aby wyświetlać wyłącznie komunikaty z określonych sesji, zaznacz odpowiednie pola wyboru w polu **Sesja**.
6. Kliknij przycisk **Wyświetl dziennik**.  
Dla każdego komunikatu wyświetlanego na stronie Zawartość pliku dziennika podawana jest data i godzina wygenerowania komunikatu, jego poziom istotności, nazwa sesji zgłaszającej komunikat, identyfikator komunikatu i treść błędu.  
Za pomocą odpowiednich przycisków można przechodzić do pierwszej, ostatniej, poprzedniej i następnej strony pliku dziennika. Można też podać numer strony i przejść bezpośrednio do niej.
7. Aby wyświetlić bardziej szczegółowe informacje o komunikacie, kliknij ikonę  **Szczegóły**.  
Na stronie Szczegóły komunikatu dziennika wyświetlana jest nazwa hosta serwera wyszukiwania korporacyjnego, na którym pojawił się komunikat, nazwa pliku, który spowodował błąd, nazwa funkcji i numer wiersza, w którym wystąpił błąd, identyfikator procesu i identyfikator wątku.



---

## Tworzenie i odtwarzanie kopii zapasowej systemu wyszukiwania korporacyjnego

Skrypty tworzenia i odtwarzania kopii zapasowych umożliwiają utworzenie i odtworzenie kopii zapasowej systemu wyszukiwania korporacyjnego.

Jeśli z powodu nienaprawialnego błędu wystąpi awaria systemu, konieczne będzie ponownie zainstalowanie produktu WebSphere Information Integrator OmniFind Edition, a następnie uruchomienie skryptu odtwarzania. Skrypty te można wykorzystać również do odtworzenia istotnych plików systemu na jednym lub wielu nowych serwerach.

Skrypty te służą do tworzenia kopii zapasowych i odtwarzania następujących plików:

- Plików konfiguracyjnych znajdujących się w katalogu ES\_NODE\_ROOT/master\_config.
- Plików baz danych używanych przez przeszukiwacze.
- Plików indeksów. Jeśli indeksy wyszukiwania korporacyjnego nie znajdują się w katalogu ES\_NODE\_ROOT/data, nie można utworzyć i odtworzyć kopii zapasowej plików indeksów przy użyciu skryptów wyszukiwania korporacyjnego.

Skrypt tworzenia kopii zapasowej tworzy następujące podkatalogi w katalogu określonym podczas jego uruchamiania (identyfikator administratora wyszukiwania korporacyjnego musi mieć uprawnienia do zapisu w określonym katalogu).

### **master\_config**

Zawiera pliki konfiguracyjne znajdujące się w katalogu ES\_NODE\_ROOT/master\_config.

### **database**

Zawiera pliki bazy danych pochodzące z serwera przeszukiwacza.

### **data**

Zawiera pliki indeksów pochodzące z serwera indeksowania.

Aby można było utworzyć kopie plików wyszukiwania korporacyjnego w innym katalogu, należy zapewnić wystarczającą ilość wolnego miejsca na dysku. Skrypty tworzenia kopii zapasowych i odtwarzania nie sprawdzają plików. Po uruchomieniu procesu tworzenia kopii zapasowych większość sesji systemowych będzie tymczasowo niedostępna. Procesy wyszukiwania pozostaną uruchomione. Proces tworzenia kopii zapasowych należy uruchamiać po zreorganizowaniu indeksu, aby przechowywany indeks był aktualny.

W przypadku instalacji na wielu serwerach kopie zapasowe należy tworzyć i odtwarzać przy użyciu serwera indeksowania wyszukiwania korporacyjnego. Ponieważ serwer indeksowania tworzy katalog bazy danych, może uzyskać dostęp do tabel bazy danych produktu na serwerze przeszukiwacza i utworzyć ich kopie zapasowe.

### **Odsyłacze pokrewne**

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## Tworzenie kopii zapasowej wyszukiwania korporacyjnego

Kopię zapasową systemu wyszukiwania korporacyjnego można utworzyć za pomocą skryptu `esbackup.sh` w systemie UNIX lub skryptu `esbackup.bat` w systemie Microsoft Windows.

### Ograniczenia

Podczas działania skryptów tworzenia i odtwarzania kopii zapasowych wszystkie sesje systemowe są zatrzymywane. W celu uniknięcia prezentacji niepoprawnych lub niespójnych informacji systemowych nie należy korzystać z konsoli administracyjnej wyszukiwania korporacyjnego podczas działania tych skryptów.

Identyfikator administratora wyszukiwania korporacyjnego musi mieć uprawnienia do zapisu w katalogu określonym podczas uruchamiania skryptu tworzenia kopii zapasowej.

### Procedura

Aby utworzyć kopię zapasową wyszukiwania korporacyjnego:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
2. Wprowadź następującą komendę, gdzie *katalog\_kopii\_zapasowej* to katalog, w którym ma zostać umieszczona kopia zapasowa plików:

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

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

Można podać następujące opcje:

- c Tworzone są kopie zapasowe plików konfiguracyjnych. Jest to opcja domyślna.
- d W składnicy danych przeszukiwacza tworzona jest kopia zapasowa przeszukanych dokumentów (dokumenty te nie zostały jeszcze poddane analizie i indeksowaniu).
- i Tworzone są kopie zapasowe plików indeksów. Jeśli pliki indeksów nie znajdują się w katalogu `ES_NODE_ROOT/data`, nie można utworzyć ich kopii zapasowej przy użyciu skryptu `esbackup`.

### Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## Odtwarzanie systemu wyszukiwania korporacyjnego

Pliki konfiguracyjne systemu można odtworzyć po ponownym zainstalowaniu programu WebSphere II OmniFind Edition przy użyciu skryptu `esrestore.sh` dla systemu UNIX lub skryptu `esrestore.bat` dla systemu Microsoft Windows.

### Ograniczenia

Podczas działania skryptów tworzenia i odtwarzania kopii zapasowych wszystkie sesje systemowe są zatrzymywane. W celu uniknięcia prezentacji niepoprawnych lub niespójnych informacji systemowych nie należy korzystać z konsoli administracyjnej wyszukiwania korporacyjnego podczas działania tych skryptów.



## Procedura

Aby odtworzyć system wyszukiwania korporacyjnego:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
2. Uruchom warstwę CCL, o ile jeszcze nie została uruchomiona:

UNIX: `startccl.sh -bg`

Windows: `startccl.bat`

Aby uruchomić warstwę CCL w tle w systemie Windows, kliknij kolejno opcje: **Start** → **Programy** → **Narzędzia administracyjne** → **Usługi** i zrestartuj usługę WebSphere Information Integrator OmniFind Edition.

3. Zatrzymaj kontroler, jeśli nie jest jeszcze zatrzymany:  
`esadmin stop`
4. Wprowadź następującą komendę, gdzie *katalog\_kopii\_zapasowej* jest to katalog, w którym została umieszczona kopia zapasowa plików:

UNIX: `esrestore.sh -c -d -i katalog_kopii_zapasowych`

Windows: `esrestore.bat -c -d -i katalog_kopii_zapasowych`

Można podać następujące opcje:

**-c** Powoduje odtworzenie plików konfiguracyjnych. Jest to opcja domyślna.

**-d** Powoduje odtworzenie przeszukanych dokumentów w składnicy danych przeszukiwacza. (Te dokumenty nie zostały jeszcze przeanalizowane ani poindeksowane).

**-i** Powoduje odtworzenie plików indeksu. Jeśli pliki indeksu nie znajdują się w katalogu `ES_NODE_ROOT/data`, do odtworzenia plików indeksu nie można użyć skryptu `esrestore`.

### Odsyłacze pokrewne

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## Odtwarzanie plików systemu wyszukiwania korporacyjnego na nowych serwerach

Możliwe jest wykonanie kopii zapasowej plików systemowych z jednego serwera wyszukiwania korporacyjnego i odtworzenie ich na dowolnej liczbie nowych serwerów, zamiast odtwarzania ich na serwerze oryginalnym.

### Zagadnienia wstępne

Przed uruchomieniem skryptu odtwarzającego należy zainstalować na nowych serwerach program WebSphere II OmniFind Edition.

### Ograniczenia

Podczas działania skryptów tworzenia i odtwarzania kopii zapasowych wszystkie sesje systemowe są zatrzymywane. W celu uniknięcia prezentacji niepoprawnych lub niespójnych informacji systemowych nie należy korzystać z konsoli administracyjnej wyszukiwania korporacyjnego podczas działania tych skryptów.

| Identyfikator administratora wyszukiwania korporacyjnego musi mieć uprawnienia do zapisu  
| w katalogu określonym podczas uruchamiania skryptu tworzenia kopii zapasowej.

### **Informacje o tym zadaniu**

Informacje o serwerze przechowywane w pliku ES\_NODE\_ROOT/master\_config/nodes.ini nie są uwzględnione w plikach kopii zapasowej.

### **Procedura**

Aby odtworzyć pliki systemu wyszukiwania korporacyjnego na dowolnej liczbie nowych serwerów:

1. Zaloguj się na serwerze indeksowania jako administrator wyszukiwania korporacyjnego. Ten identyfikator użytkownika został określony podczas instalowania programu WebSphere II OmniFind Edition.
2. Uruchom skrypt tworzący kopię zapasową na bieżącym (dotychczasowym) serwerze indeksowania, gdzie *katalog\_kopii\_zapasowej* jest katalogiem, w którym ma zostać utworzona kopia zapasowa plików:

System UNIX: `esbackup.sh katalog_kopii_zapasowej`

System Windows: `esbackup.bat katalog_kopii_zapasowej`

3. Użyj programu FTP w celu wysłania wszystkich plików zawartych w katalogu *katalog\_kopii\_zapasowej* na nowy serwer indeksowania.
4. Uruchom skrypt odtwarzający na nowym serwerze indeksowania:

System UNIX: `esrestore.sh katalog_kopii_zapasowej`

System Windows: `esrestore.bat katalog_kopii_zapasowej`

### **Odsyłacze pokrewne**

“Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego” na stronie 249

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

---

## Komendy, kody powrotu i identyfikatory sesji wyszukiwania korporacyjnego

Za pomocą komend można diagnozować problemy, określać status różnych części systemu, rozpoczynać i kończyć sesje oraz uruchamiać i zatrzymywać system.

W instalacji wieloserwerowej można wykonywać komendy z dowolnego serwera w systemie. Zaleca się jednak uruchamianie komend z serwera indeksowania. Serwer indeksowania lub serwer kontrolera może uzyskać dostęp do informacji znajdujących się na wszystkich innych serwerach w systemie.

Większość komend ma następujące formaty:

```
esadmin nazwa_komendy argumenty
esadmin id_sesji dzialanie -opcja
```

Aby uzyskać więcej informacji na temat wszystkich komend, należy wprowadzić komendę `esadmin help`. Aby uzyskać więcej informacji na temat konkretnej komendy, należy wprowadzić komendę `esadmin dzialanie help`.

### Komendy esadmin wyszukiwania korporacyjnego

Poniższe komendy należy wprowadzać w jednym wierszu.

Tabela 5. Komendy esadmin wyszukiwania korporacyjnego

Komenda	Opis
<code>esadmin id_sesji_przeszukiwacza start</code>	Rozpoczyna sesję przeszukiwacza. Komenda ta nie powoduje rozpoczęcia żadnych działań związanych z przeszukiwaniem.  <b>Przykładowa komenda:</b> <code>esadmin col1.WEB1.esadmin start</code>  <b>Przykładowe komunikaty i kody powrotu:</b> FFQC5310I WEBCrawler1 (sid: col1.WEB1.esadmin) nie działa. FFQC5314I Wynik: 0
<code>esadmin id_sesji_przeszukiwacza startCrawl</code>	Uruchamia przeszukiwanie.  <b>Przykładowa komenda:</b> <code>esadmin col3.DB21.esadmin startCrawl</code>  <b>Przykładowe komunikaty i kody powrotu:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0
<code>esadmin id_sesji_przeszukiwacza pause</code>	Wstrzymuje przeszukiwanie.  <b>Przykładowa komenda:</b> <code>esadmin col3.DB21.esadmin pause</code>  <b>Przykładowe komunikaty i kody powrotu:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0

Tabela 5. Komendy esadmin wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
esadmin <i>id_sesji_przeszukiwacza</i> resume	Wznawia przeszukiwanie.  <b>Przykładowa komenda:</b> esadmin col3.DB21.esadmin resume  <b>Przykładowe komunikaty i kody powrotu:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0
esadmin <i>id_sesji_przeszukiwacza</i> stopCrawl	Zatrzymuje przeszukiwanie.  <b>Przykładowa komenda:</b> esadmin col3.DB21.esadmin stopCrawl  <b>Przykładowe komunikaty i kody powrotu:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0
esadmin <i>id_sesji_przeszukiwacza</i> stop	Kończy sesję przeszukiwacza.  <b>Przykładowa komenda:</b> esadmin col3.DB21.esadmin stop  <b>Przykładowe komunikaty i kody powrotu:</b> FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650 FFQC5314I Wynik: 0
esadmin <i>id_sesji_przeszukiwacza</i> getCrawlerStatus	Pobiera status przeszukiwacza. Zwracane informacje zależą od tego, czy przeszukiwacz jest przeszukiwaczem sieci WWW czy przeszukiwaczem innych źródeł danych.  <b>Przykład dla przeszukiwacza sieci WWW:</b> esadmin col1.WEB1.esadmin getCrawlerStatus  <b>Możliwe kody powrotu i komunikaty dla przeszukiwacza sieci WWW:</b> FFQC5303I WebCrawler1 (sid: col1.WEB1.esadmin) już działa. PID: 23650  <b>Przykład dla przeszukiwacza innego niż przeszukiwacz sieci WWW:</b> esadmin col3.DB21.esadmin getCrawlerStatus  <b>Możliwe kody powrotu i komunikaty dla przeszukiwacza innego niż przeszukiwacz sieci WWW:</b> FFQC5303I db2crawler (sid: db2col.DB2_96945) już działa. PID: 5936  Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 254.

Tabela 5. Komendy esadmin wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
<pre>esadmin id_sesji_przeszukiwacza_zrodla_danych getCrawlSpaceStatus</pre>	<p>Pobiera ogólny status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW.</p>
<pre>esadmin id_sesji_przeszukiwacza_WWW getCrawlStatus -selections wartosc</pre>	<p><b>Przykładowa komenda:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatus</pre> <p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650</pre> <p>Pobiera ogólny status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW.</p> <p><b>Przykładowa komenda:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlStatus</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 254.</p>
<pre>esadmin id_sesji_przeszukiwacza_zrodla_danych getCrawlSpaceStatusDetail -ts id_serwera_docelowego</pre> <pre>esadmin id_sesji_przeszukiwacza_WWW getCrawlDetailsPerSite -url łańcuch selections liczba -threshold liczba</pre>	<p>Pobiera szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW. Jeśli użytkownik nie poda opcji określającej serwer docelowy, zwrócone zostaną dane na temat wszystkich serwerów docelowych. Na przykład, gdy przeszukiwacz DB2 przeszukuje bazy danych FOUNTAIN i SAMPLE, a użytkownik nie określi serwera docelowego, zwracany jest status wszystkich tabel w bazach danych FOUNTAIN i SAMPLE.</p> <p><b>Przykładowa komenda:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatusDetail -ts FOUNTAIN</pre> <p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) już działa. PID: 23650</pre> <p>Pobiera szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW.</p> <p><b>Przykładowa komenda:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlDetailsPerSite</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 254.</p>
<pre>esadmin startParser -cid id_kolekcji</pre>	<p>Powoduje uruchomienia analizatora.</p> <p><b>Przykładowa komenda:</b></p> <pre>esadmin startParser -cid col1</pre> <p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Kontroler (serwer1) (identyfikator SID: kontroler) już działa. PID: 25917 FFQC5314I Wynik: 0</pre>

Tabela 5. Komendy esadmin wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
esadmin stopParser <i>id_kolekcji</i>	Powoduje zatrzymanie analizatora składni.
	<p><b>Przykładowa komenda:</b></p> <pre>esadmin stopParser -cid coll</pre>
	<p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 15292 FFQC5312E Błąd podczas wykonywania działania: [stopParser] FFQC4823E Sterownik analizatora składni sesji - Kolekcja coll (węzeł1) [coll.parserdriver] nie działa.</pre>
esadmin monitor getCollectionParserMonitorStatus -cid <i>id_kolekcji</i>	Pobiera status analizatora składni.
	<p><b>Przykładowa komenda:</b></p> <pre>esadmin monitor getCollectionParserMonitorStatus -cid coll</pre>
	<p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Monitor (node1) (sid: monitor) już działa. PID: 12543</pre>
	Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 254.
esadmin startMain -cid <i>id_kolekcji</i>	Rozpoczyna reorganizację indeksu.
	<p><b>Przykładowa komenda:</b></p> <pre>esadmin startMain -cid coll</pre>
	<p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 25917 FFQC5314I Wynik: 1117671147056</pre>
esadmin startDelta -cid <i>id_kolekcji</i>	Rozpoczyna odświeżanie indeksu.
	<p><b>Przykładowa komenda:</b></p> <pre>esadmin startDelta -cid coll</pre>
	<p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 4548 FFQC5314I Wynik: 1117670603408</pre>

Tabela 5. Komendy esadmin wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
<pre>esadmin monitor getCollectionIndexMonitorStatus -cid <i>id_kolekcji</i> -buildType [main   delta] -numrecords <i>N_ostatnich_rekordów</i></pre>	<p>Pobiera status odświeżania lub reorganizacji indeksu. (Indeks <i>main</i> odwołuje się do indeksu reorganizowanego, a indeks <i>delta</i> - do indeksu odświeżanego). Opcja <i>numrecords</i> powoduje wyświetlenie rekordów statusu <i>N</i> ostatnich operacji tworzenia indeksu. Pominięcie opcji <i>numrecords</i> powoduje, że zwrócony zostaje status 20 ostatnich operacji tworzenia indeksu.</p> <p><b>Przykładowa komenda:</b></p> <pre>esadmin monitor getCollectionIndexMonitorStatus -cid coll -buildType main -numrecords 4</pre> <p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Monitor (node1) (sid: monitor) już działa. PID: 12649</pre> <p>Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu” na stronie 254.</p>
<pre>esadmin startSearch -cid <i>id_kolekcji</i></pre>	<p>Uruchamia procesy serwera wyszukiwania.</p> <p><b>Przykładowa komenda:</b></p> <pre>esadmin startSearch -cid coll</pre> <p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 25917 FFQC5314I Wynik: 0</pre>
<pre>esadmin stopSearch -cid <i>id_kolekcji</i></pre>	<p>Zatrzymuje procesy serwera wyszukiwania.</p> <p><b>Przykładowa komenda:</b></p> <pre>esadmin stopSearch -cid coll</pre> <p><b>Przykładowe komunikaty i kody powrotu:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) już działa. PID: 15292 FFQC5314I Wynik: 0</pre>

Tabela 5. Komendy esadmin wyszukiwania korporacyjnego (kontynuacja)

Komenda	Opis
esadmin monitor getCollectionSearchMonitorStatus -cid <i>id_kolekcji</i>	Pobiera status serwera wyszukiwania.  <b>Przykładowa komenda:</b> esadmin monitor getCollectionSearchMonitorStatus -cid coll
esadmin <i>id_sesji_menedżera_wyszukiwania</i> getStatus -cid <i>id_kolekcji</i>	<b>Przykładowe komunikaty i kody powrotu:</b> FFQC5303I Monitor (node1) (sid: monitor) już działa. PID: 12649  Zwraca szczegółowe informacje o statusie indeksu wyszukiwania dla kolekcji na danym serwerze wyszukiwania. Dla każdego serwera wyszukiwania istnieje jedna sesja menedżera wyszukiwania. Każda sesja menedżera wyszukiwania jest odpowiedzialna za monitorowanie indeksów wyszukiwania i działanie na nich na konkretnym serwerze wyszukiwania.  <b>Przykładowa komenda:</b> esadmin searchmanager.node1 getStatus -cid coll
	<b>Przykładowe komunikaty i kody powrotu:</b> FFQC5303I Search Manager (node1) (sid: searchmanager.node1) już działa. PID: 15711 FFQC5314I Wynik: PID=18390 CacheHits=3 QueryRate=1 Port=44008 SessionId=coll.runtime.node1 CacheHitRate=0.333 ResponseTime=70 Status=1 SessionName=coll.runtime.node1.1  Więcej informacji na temat zwracanych komunikatów o statusie można znaleźć w sekcji “Szczegółowe informacje o komendach dotyczących statusu”.

## Szczegółowe informacje o komendach dotyczących statusu

Niektóre komendy mogą zwracać bardzo obszerne informacje. W niniejszej sekcji opisane są informacje na temat statusu przeszukiwacza lub przestrzeni przeszukiwania, które mogą zostać zwrócone. Tabela w sekcji “Komendy esadmin wyszukiwania korporacyjnego” na stronie 249 zawiera możliwe informacje zwracane przez każdą z komend esadmin. W tej sekcji omówione są informacje zwracane przez następujące komendy:

- status przeszukiwacza sieci WWW
- status przeszukiwacza innego niż przeszukiwacz sieci WWW
- status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW
- status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacz sieci WWW
- szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW
- szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacz sieci WWW
- status analizatora składni



- status operacji tworzenia indeksu
- status serwera wyszukiwania
- szczegółowy status serwera wyszukiwania

**Status przeszukiwacza sieci WWW:** Komenda pobierająca status przeszukiwacza sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca status przeszukiwacza sieci WWW może zwrócić następujące informacje:

```
FFQC5314I Wynik: <?xml version='1.0' encoding='UTF-8'?>
<CrawlerStatus>
<CrawlerRunLevel Value="Running"/>
<CrawlerThreadStateDist Count="4" Total="200">
<CrawlerThreadState State="FETCHING" Count="100"/>
. . .
</CrawlerThreadState State="FETCHING" Count=100>
<ActiveBucketList Count="500">
<ActiveBucket URL="http://w3.ibm.com/"
NumActURLs="355"
NumProcURLs="350"
TimeRem="5" Duration="1195"/>
. . .
</ActiveBucketList>
<CrawlRate Value="75"/>
<RecentlyCrawledURLList Count="40">
<RecentlyCrawledURL URL="http://w3.ibm.com/foo.html"/>
<RecentlyCrawledURL URL="http://w3.ibm.com/foo.html"/>
<NumURLsThisSession Value="160000"/>
</CrawlerStatus>
```

W poniższej tabeli opisano każdy element XML (wraz z jego możliwymi atrybutami), który jest zwracany przez komendę pobierającą status przeszukiwacza sieci WWW:

Tabela 6. Informacje o statusie przeszukiwacza sieci WWW

Element	Atrybuty	Opis
CrawlerStatus	<ul style="list-style-type: none"> <li>• CrawlerThreadStateDist</li> <li>• ActiveBucketList</li> <li>• CrawlRate</li> <li>• RecentlyCrawledURLList</li> <li>• NumURLsThisSession</li> </ul>	Status przeszukiwacza.
CrawlerRunLevel Value	<ul style="list-style-type: none"> <li>• Łańcuch (w języku angielskim) “Not started”: Istnieje sesja przeszukiwacza, ale nie odebrała jeszcze komunikatu powodującego rozpoczęcie przetwarzania dokumentów.</li> <li>• “Started”: Przeszukiwacz rozpoczyna działanie.</li> <li>• “Running”: Przeszukiwacz został zainicjowany i uruchomiony i realizuje proces przeszukiwania.</li> <li>• “Paused”: Przeszukiwacz odebrał polecenie wstrzymania przeszukiwania, ale nie zakończenia działania.</li> <li>• “Stopping”: Przeszukiwacz odebrał sygnał nakazujący zatrzymanie i jest w trakcie zatrzymywania działania.</li> <li>• “Error”: Przeszukiwacz jest w stanie błędniemożliwego do naprawienia i aby wznowił przeszukiwanie, trzeba go zatrzymać i zrestartować.</li> </ul>	Informacje o działaniu wykonywanym przez przeszukiwacz.

Tabela 6. Informacje o statusie przeszukiwacza sieci WWW (kontynuacja)

Element	Atrybuty	Opis
CrawlerThreadState State	Łańcuch (w języku angielskim)	Działanie wątku przeszukiwacza. Pole to zawiera informacje o bieżącym działaniu wątku.
ActiveBucket	<ul style="list-style-type: none"> <li>URL: Łańcuch (specyfikacja adresu URL) Protokół, host i port przeszukiwanych w danej chwili adresów URL.</li> <li>NumActURLs: Liczba całkowita (dodatnia) Liczba adresów URL w porcji w momencie jej udostępnienia (aktywowania) do przeszukiwania.</li> <li>NumProcURLs: Liczba całkowita (nieujemna) Liczba adresów URL w porcji, które do chwili obecnej zostały przetworzone (przeszukane lub odrzucone).</li> <li>TimeRem: Liczba całkowita Liczba sekund, po których zostanie przekroczony limit czasu dla porcji.</li> <li>Duration: Liczba całkowita (nieujemna) Liczba sekund od chwili aktywowania porcji.</li> </ul>	Bieżąca aktywność określonego serwisu WWW.
CrawlRate	Wartość: Liczba całkowita (nieujemna)  Liczba stron przeszukiwanych w ciągu jednej sekundy (we wszystkich porcjach łącznie).	Przepustowość przeszukiwacza.
RecentlyCrawledURL	URL: Łańcuch (specyfikacja adresu URL)  Łańcuch określający protokół, hosta, port i plik, który został przeszukany.	Strona, która została ostatnio przeszukana.
NumURLsThisSession	Wartość: Liczba całkowita (nieujemna)	Liczba adresów URL, które zostały przeszukane od chwili rozpoczęcia przeszukiwania przez tę instancję (proces) przeszukiwacza.

**Status przeszukiwacza innego niż przeszukiwacz sieci WWW:** Komenda pobierająca status przeszukiwacza innego niż przeszukiwacz sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca status przeszukiwacza innego niż przeszukiwacz sieci WWW może zwrócić następujące informacje:

```
FFQC5314I Wynik: <?xml version='1.0' encoding='UTF-8'?>
<GeneralStatus>
<Status>0</Status>
<StatusMessage>Idle</StatusMessage>
<NumberOfServers>1</NumberOfServers>
<NumberOfCompletedServers>1</NumberOfCompletedServers>
<NumberOfTargets>3</NumberOfTargets>
<NumberOfCompletedTargets>3</NumberOfCompletedTargets>
<NumberOfCrawledRecords>115</NumberOfCrawledRecords>
<RunningThreads>0</RunningThreads>
</GeneralStatus>
```

Poniższa tabel zawiera opis elementów XML i ich atrybutów dla każdego przeszukiwacza wyszukiwania korporacyjnego z wyjątkiem przeszukiwacza sieci WWW. Informacje te zwracane są przez komendę pobierającą status przeszukiwacza.

Tabela 7. Elementy i atrybuty zwracane przez komendę pobierającą status przeszukiwaczy NNTP, DB2 i Notes

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacz DB2	Przeszukiwacz Notes
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd
NumberOfServers	Liczba serwerów NNTP w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.
NumberOfCompletedServers	Liczba przeszukanych serwerów NNTP.	Liczba przeszukanych baz danych.	Liczba przeszukanych baz danych.
NumberOfTargets	Liczba grup dyskusyjnych w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.	Liczba widoków i folderów w przestrzeni przeszukiwania.
NumberOfCompletedTargets	Liczba przeszukanych grup dyskusyjnych.	Liczba przeszukanych tabel.	Liczba przeszukanych widoków i folderów.
NumberOfCompletedRecords	Liczba przeszukanych artykułów.	Liczba przeszukanych rekordów.	Liczba przeszukanych dokumentów.
RunningThreads	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.

Tabela 8. Elementy i atrybuty zwracane przez komendę pobierającą status przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd
NumberOfServers	Liczba serwerów Exchange Server w przestrzeni przeszukiwania.	Liczba serwerów Content Manager w przestrzeni przeszukiwania.	Liczba repozytoriów w przestrzeni przeszukiwania.
NumberOfCompletedServers	Liczba przeszukanych serwerów Exchange Server.	Liczba przeszukanych serwerów Content Manager.	Liczba przeszukanych repozytoriów.
NumberOfTargets	Liczba podfolderów w przestrzeni przeszukiwania.	Liczba typów elementów w przestrzeni przeszukiwania.	Liczba klas w przestrzeni przeszukiwania.
NumberOfCompletedTargets	Liczba przeszukanych podfolderów.	Liczba przeszukanych typów elementów.	Liczba przeszukanych klas elementów.
NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.
RunningThreads	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.

Tabela 9. Elementy i atrybuty zwracane przez komendę pobierającą status przeszukiwaczy QuickPlace, Domino Document Manager, systemu plików UNIX i systemu plików Windows

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Status	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)	Status (0, 1, 2, -1)
StatusMessage	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd	Status: 0 - Bezczywność, 1 - Działanie, 2 - Wstrzymanie, -1 - Błąd

Tabela 9. Elementy i atrybuty zwracane przez komendę pobierającą status przeszukiwaczy QuickPlace, Domino Document Manager, systemu plików UNIX i systemu plików Windows (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
NumberOfServers	Liczba lokalizacji w przestrzeni przeszukiwania.	Liczba bibliotek w przestrzeni przeszukiwania.	Ustalona wartość równa 1.
NumberOfCompletedServers	Liczba przeszukanych lokalizacji.	Liczba przeszukanych bibliotek.	0 lub 1, gdy przeszukane są wszystkie podkatalogi.
NumberOfTargets	Liczba baz danych lokalizacji i baz danych pomieszczeń w przestrzeni przeszukiwania.	Liczba szafek w przestrzeni przeszukiwania.	Liczba podkatalogów w przestrzeni przeszukiwania.
NumberOfCompletedTargets	Liczba przeszukanych baz danych lokalizacji i baz danych pomieszczeń.	Liczba przeszukanych szafek.	Liczba przeszukanych podkatalogów.
NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych plików.
RunningThreads	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.	Liczba wątków przeszukiwacza.

**Status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW:** Komenda pobierająca status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca status przestrzeni przeszukiwania sieci WWW może zwrócić następujące informacje:

Tabela 10. Wartości maski wyboru zwracanej przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Bit maski	Wybiera
1	Liczba stron w składnicy surowych danych.
2	Liczba wykrytych serwisów.
4	Liczba serwisów z systemem DNS.
8	Liczba serwisów bez systemu DNS.
16	Liczba wykrytych adresów URL.
32	Liczba zapisanych unikalnych stron.
64	Liczba przeszukanych adresów URL.
128	Liczba adresów URL, których nie przeszukano.
256	Liczba zaległych adresów URL.
512	Rozkład kodów powrotu HTTP.

Wszystkie wartości reprezentują sumy łączne dla wszystkich sesji używających bieżącej wewnętrznej bazy danych:

```
<CrawlStatus>
 <NumPagesInRDS Value="5422386"/>
 <NumSitesDiscovered Value="15332"/>
 <NumSitesWithDNS Value="14832"/>
 <NumSitesWithoutDNS Value="500"/>
 <NumURLsDiscovered Value="15222999"/>
 <NumUniquePagesSaved Value="6234789"/>
 <NumURLsCrawled Value="7800422"/>
 <NumURLsUncrawled Value="7422577"/>
 <NumURLsOverdue Value="14000"/>
 <HTTPCodeDist Count="4" Total="1031000"/>
```

```

| <HTTPCode Code="200" Count ="1000000"/>
| <HTTPCode Code="301" Count ="1000"/>
| <HTTPCode Code="404" Count ="10000"/>
| <HTTPCode Code="780" Count="20000"/>
| </HTTPCode Code="780" Count="20000"/>
| <?CrawlStatus>

```

Zwracane dane zawierają wszystkie lub niektóre z następujących elementów (mogą nie zawierać żadnych):

Tabela 11. Informacje zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Element	Atrybut	Opis
CrawlerStatus	<ul style="list-style-type: none"> <li>• NumPagesInRDS</li> <li>• NumSitesDiscovered</li> <li>• NumSitesWithDNS</li> <li>• NumSitesWithoutDNS</li> <li>• NumURLsDiscovered</li> <li>• NumUniquePagesSaved</li> <li>• NumURLsCrawled</li> <li>• NumURLsUncrawled</li> <li>• NumURLsOverdue</li> <li>• HTTPCodeDist</li> </ul>	Łatwe do szybkiego uzyskania informacji o łącznym stanie przeszukiwania (wszystkie sesje).
NumPagesInRDS	<p>Wartość: Nieujemna liczba całkowita</p> <p>Liczba stron znajdujących się w danej chwili w obszarze pomostowym składnicy surowych danych (pochodzących tylko od danego przeszukiwacza).</p>	Stopień zapełnienia składnicy surowych danych (tylko w wyniku działania danego przeszukiwacza).
NumSitesDiscovered	<p>Wartość: Nieujemna liczba całkowita</p> <p>Liczba hostów wykrytych dzięki przeszukiwaniu (lub pochodzących z listy początkowej).</p>	Miara pokrycia przez przeszukiwacze przeszukiwanej domeny (liczba hostów).
NumSitesWithDNS	<p>Wartość: Nieujemna liczba całkowita</p> <p>Liczba hostów powiązanych z adresami IP (przetłumaczonymi przez przeszukiwacz w tle).</p>	Miara efektywności pobierania przez przeszukiwacz adresów IP hostów, które zostały wykryte na podstawie ich nazw DNS znajdujących się w adresach URL.
NumSitesWithoutDNS	<p>Wartość: Nieujemna liczba całkowita</p> <p>Liczba hostów niepowiązanych z adresami IP (przetłumaczonymi przez przeszukiwacz w tle).</p>	Miara efektywności pobierania przez przeszukiwacz adresów IP hostów, które zostały wykryte na podstawie ich nazw DNS znajdujących się w adresach URL.
NumURLsDiscovered	<p>Wartość: Nieujemna liczba całkowita</p> <p>Liczba unikalnych adresów URL, które odwiedził przeszukiwacz.</p>	Miara pokrycia przez przeszukiwacze przeszukiwanej domeny (liczba adresów URL).
NumUniquePagesSaved	<p>Wartość: Nieujemna liczba całkowita</p> <p>Liczba unikalnych stron zapisanych do składnicy surowych danych w celu dalszego przetworzenia przez inne komponenty wyszukiwania korporacyjnego.</p>	Wkład danego przeszukiwacza do wielkości indeksu.

Tabela 11. Informacje zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW (kontynuacja)

Element	Atrybut	Opis
NumURLsCrawled	Wartość: Nieujemna liczba całkowita  Liczba unikalnych adresów URL, które przeszukał przeszukiwacz.	Miara zdolności przeszukiwacza do przetwarzania danych na całej trasie. Liczba ta jest różna od liczby stron zapisanych do składnicy surowych danych, ponieważ nie wszystkie przeszukane strony zostają do niej zapisane.
NumURLsOverdue	Wartość: Nieujemna liczba całkowita  Liczba unikalnych adresów URL kwalifikujących się do przeszukania.	Miara zdolności przeszukiwacza do przemieszczania się w przestrzeni sieci WWW.

**Status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacze sieci WWW:** Komenda pobierająca status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca status przestrzeni przeszukiwania innej niż sieć WWW może zwrócić następujące informacje:

```
FFQC5314I Wynik: <?xml version='1.0' encoding='UTF-8'?>
<ServerStatus>
 <Server Name ="FOUNTAIN">
 <Status>5</Status>
 <StatusMessage>Scheduled</StatusMessage>
 <NumberOfTargets>1</NumberOfTargets>
 <NumberOfCompletedTargets>1</NumberOfCompletedTargets>
 <NumberOfErrors>0</NumberOfErrors>
 <StartTime>1118354510512</StartTime>
 <EndTime>1118354514386</EndTime>
 <ScheduleConfigured>2</ScheduleConfigured>
 <ScheduleTime>1118393377000</ScheduleTime>
 <TotalTime>3874</TotalTime>
 </Server>
</ServerStatus>
```

Poniższa tabel zawiera opis elementów XML i ich atrybutów dla każdego przeszukiwacza wyszukiwania korporacyjnego z wyjątkiem przeszukiwacza sieci WWW. Informacje te zwracane są przez komendę pobierającą status przestrzeni przeszukiwania. W przypadku przeszukiwaczy Notes, gdy poziom agregacji jest równy 0, parametr Server@Name ma wartość nazwa serwera + nazwa bazy danych. Gdy poziom agregacji jest równy 1, parametr Server@Name ma wartość nazwa serwera + nazwa katalogu.

Tabela 12. Elementy i atrybuty zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2 i Notes

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacz DB2	Przeszukiwacz Notes
Server@Name	Nazwa serwera grup dyskusyjnych	Nazwa bazy danych	Nazwa bazy danych lub nazwa katalogu
Server/Status	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>

Tabela 12. Elementy i atrybuty zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2 i Notes (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacz DB2	Przeszukiwacz Notes
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>
Server/NumberOfTargets	Liczba grup dyskusyjnych w przestrzeni przeszukiwania.	Liczba baz danych w przestrzeni przeszukiwania.	Liczba widoków i folderów w przestrzeni przeszukiwania.
Server/NumberOfCompletedTargets	Liczba przeszukanych grup dyskusyjnych.	Liczba przeszukanych tabel.	Liczba przeszukanych widoków i folderów.
Server/NumberOfErrors	Nie dotyczy.	Liczba błędów.	Liczba błędów.
Server/StartTime	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.
Server/EndTime	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.
Server/ScheduleConfigured	Nie dotyczy.	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>
Server/ScheduleTime	Nie dotyczy.	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.
Server/TotalTime	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.
Server/AggregationLevel	Nie dotyczy.	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0, 1: <ul style="list-style-type: none"> <li>0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).</li> <li>Przeszukiwacz Notes przeszukuje dokumenty w trybie katalogu.</li> </ul>

Tabela 13. Elementy i atrybuty zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Server@Name	Nazwa serwera Exchange Server.	Serwery DB2 Content Manager.	Nazwa repozytorium.
Server/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>5: Zaplanowane</li> <li>-1: Błąd</li> </ul>
Server/NumberOfTargets	Liczba podfolderów w przestrzeni przeszukiwania.	Liczba typów elementów w przestrzeni przeszukiwania.	Liczba klas elementów w przestrzeni przeszukiwania.
Server/NumberOfCompletedTargets	Liczba przeszukanych podfolderów.	Liczba przeszukanych typów elementów.	Liczba przeszukanych klas elementów.
Server/NumberOfErrors	Liczba błędów.	Liczba błędów.	Liczba błędów.
Server/StartTime	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.
Server/EndTime	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>
Server/ScheduleTime	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.



Tabela 13. Elementy i atrybuty zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Server/TotalTime	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.
Server/AggregationLevel	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).

Tabela 14. Elementy i atrybuty zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, systemu plików UNIX i systemu plików Windows

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Server@Name	Katalog lokalizacji	Baza danych biblioteki	Wartość ustalona: localhost.
Server/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• -1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• 5: Zaplanowane</li> <li>• -1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, 5, -1) <ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• 5: Zaplanowane</li> <li>• -1: Błąd</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• 5: Zaplanowane</li> <li>• -1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• 5: Zaplanowane</li> <li>• -1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• 5: Zaplanowane</li> <li>• -1: Błąd</li> </ul>
Server/NumberOfTargets	Liczba baz danych lokalizacji i baz danych pomieszczeń w przestrzeni przeszukiwania.	Liczba szafek w przestrzeni przeszukiwania.	Liczba podkatalogów w przestrzeni przeszukiwania.
Server/NumberOfCompletedTargets	Liczba przeszukanych baz danych lokalizacji i baz danych pomieszczeń.	Liczba przeszukanych szafek.	Liczba podkatalogów w przestrzeni przeszukiwania.
Server/NumberOfErrors	Liczba błędów.	Liczba błędów.	Liczba błędów.
Server/StartTime	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.	Czas rozpoczęcia, jeśli ma zastosowanie.
Server/EndTime	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.	Czas zakończenia, jeśli ma zastosowanie.

Tabela 14. Elementy i atrybuty zwracane przez komendę pobierającą status przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, systemu plików UNIX i systemu plików Windows (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Server/ScheduleConfigured	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>	0, 1, 2 <ul style="list-style-type: none"> <li>0: W plikach konfiguracyjnych nie skonfigurowano opcji planowania działania przeszukiwacza.</li> <li>1: Skonfigurowano opcję planowania działania przeszukiwacza, ale w bieżącej sesji została ona wyłączona.</li> <li>2: Skonfigurowano opcję planowania działania przeszukiwacza i w bieżącej sesji została ona włączona.</li> </ul>
Server/ScheduleTime	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.	Czas zaplanowany, jeśli ma zastosowanie.
Server/TotalTime	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.	Czas łączny, jeśli ma zastosowanie.
Server/AggregationLevel	0, 1 <ul style="list-style-type: none"> <li>0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).</li> <li>Przeszukiwacz Notes przeszukuje dokumenty w trybie katalogu.</li> </ul>	0, 1 <ul style="list-style-type: none"> <li>0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).</li> <li>Przeszukiwacz Notes przeszukuje dokumenty w trybie katalogu.</li> </ul>	0, 1 <ul style="list-style-type: none"> <li>0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).</li> <li>Przeszukiwacz Notes przeszukuje dokumenty w trybie katalogu.</li> </ul>

#### Szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW:

Komenda pobierająca szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca szczegółowy status przestrzeni przeszukiwania może zwrócić następujące informacje:

Tabela 15. Wartości maski wyboru zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Bit maski	Wybiera
1	Liczba stron w składnicy surowych danych.
2	Liczba wykrytych serwisów.
4	Liczba serwisów z systemem DNS.
8	Liczba serwisów bez systemu DNS.
16	Liczba wykrytych adresów URL.
32	Liczba zapisanych unikalnych stron.
64	Liczba przeszukanych adresów URL.

Tabela 15. Wartości maski wyboru zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW (kontynuacja)

Bit maski	Wybiera
128	Liczba adresów URL, których nie przeszukano.
256	Liczba zaległych adresów URL.
512	Rozkład kodów powrotu HTTP.

Przykładowe zwrócone informacje:

```
<CrawlDetailsPerSite>
 <Site URL=http://w3.ibm.com/">
 <NumURLsDiscovered Value="5422386"/>
 <NumURLsOverdue Value="15332"/>
 <NumURLsCrawled Value="15332"/>
 <NumURLsUncrawled Value="15332"/>
 <NumURLsOverdueBy Threshold="604800" Value="14832"/>
 <NumURLsActivated Value="2200"/>
 <LastActivationTime Value="1076227340"/>
 <LastActivationDuration Value="4300"/>
 <IPAddressList Count="1"/>
 <IPAddress Value="9.205.41.33"/>
 </IPAddressList>
 <RobotsContent>
 robots content. . .
 </RobotsContent>
 <HTTPCodeDist Count="4" Total="1031000"/>
 <HTTPCode Code="200" Count ="1000000"/>
 <HTTPCode Code="301" Count ="1000"/>
 <HTTPCode Code="404" Count ="10000"/>
 <HTTPCode Code="780" Count="20000"/>
 </HTTPCodeDist>
</CrawlDetailsPerSite>
```

Poniższa tabela zawiera opis każdego pola, które jest zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW:

Tabela 16. Informacje zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW

Element	Atrybuty	Opis
CrawlDetailsPerSite	<ul style="list-style-type: none"> <li>LastActivationTime:</li> <li>LastActivationDuration:</li> <li>IPAddressList:</li> <li>RobotsContent:</li> <li>HTTPCodeDist:</li> </ul>	Łatwe do szybkiego uzyskania informacji o szczegółowym stanie jednego serwisu.
Site	URL	Adres URL strony głównej serwisu.
NumURLsDiscovered	Wartość	Liczba adresów URL wykrytych w danym serwisie.
NumURLsOverdue	Wartość	Liczba adresów URL kwalifikujących się do ponownego przeszukania w danym serwisie.
NumURLsCrawled	Wartość	Liczba adresów URL przeszukanych w danym serwisie.
NumURLsUncrawled	Wartość	Liczba adresów URL, które nie zostały jeszcze przeszukane w danym serwisie.

Tabela 16. Informacje zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza sieci WWW (kontynuacja)

Element	Atrybuty	Opis
NumURLsOverdueBy	Próg, wartość: Liczba całkowita (dodatnia lub ujemna)  Próg dla odnajdywania ponownych przeszukiwań opóźnionych o określony czas (wartość ujemna oznacza liczbę sekund przed chwilą bieżącą) lub mających nastąpić w przeciągu określonego przedziału czasu (wartość dodatnia oznacza liczbę sekund po chwili bieżącej).	Liczba adresów URL, które kwalifikują się do przeszukania przynajmniej od określonej liczby sekund (próg) lub które będą się kwalifikować w ciągu najbliższych sekund (próg).
NumURLsActivated	Wartość	Liczba adresów URL wczytanych do pamięci podczas ostatniego skanowania danego serwisu i udostępnionych wątkom przeszukiwacza.
LastActivationTime	Wartość	Liczba sekund, które upłynęły od chwili, w której adresy URL danego serwisu zostały po raz ostatni wczytane do pamięci.
LastActivationDuration	Wartość	Liczba sekund, przez które adresy URL danego serwisu były ostatnio wczytane do pamięci i dostępne dla wątków przeszukiwacza.
IPAddressList	IPAddress	Wszystkie znane adresy IP serwera danego serwisu.
IPAddress	Wartość	Adres IPv4 serwera danego serwisu w notacji z kropkami.
RobotsContent	Tekst	Tekst znajdujący się w pliku robots, jeśli istnieje.
HTTPCodeDist	HTTPCode	Rozkład kodów HTTP z operacji pobierania, które usiłowano wykonać z tego serwisu.
HTTPCode	Kod: Liczba całkowita  Kod powrotu HTTP lub inny kod wewnętrzny.	Liczba wystąpień określonego kodu powrotu HTTP podczas przeszukiwania danego serwisu.

**Szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacze sieci WWW:** Komenda pobierająca szczegółowy status przestrzeni przeszukiwania dla przeszukiwacza innego niż przeszukiwacz sieci WWW zwraca informacje w postaci dokumentu XML. Komenda pobierająca szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy innych niż przeszukiwacze sieci WWW może zwrócić następujące informacje:

```
FFQC5314I Wynik: <?xml version='1.0' encoding='UTF-8'?>
<TargetStatus>
 <Target Name ="escmgr.crawlerinstances">
 <Status>2</Status>
 <StatusMessage>Completed</StatusMessage>
 <NumberOfRecords></NumberOfRecords>
 <NumberOfCrawledRecords>117</NumberOfCrawledRecords>
 <NumberOfInsertedRecords>21</NumberOfInsertedRecords>
 <NumberOfUpdatedRecords>45</NumberOfUpdatedRecords>
 <StartTime>1118354510727</StartTime>
```

```

| <EndTime>1118354514386</EndTime>
| <AggregationLevel>0<AggregationLevel>
| <Target>
| </TargetStatus>

```

Tabela 17. Elementy i atrybuty zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2 i Notes

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacz DB2	Przeszukiwacz Notes
Target@Name	Nazwa grupy dyskusyjnej	Nazwa tabeli	Nazwa widoku lub folderu
Target@CrawlType	Nie dotyczy.	0,1: <ul style="list-style-type: none"> <li>0: Przeszukiwanie aktywne (normalne)</li> <li>1: Przeszukiwanie pasywne (publikowanie zdarzeń DB2)</li> </ul>	0
Target/Status	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	Status: (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>
Target/NumberOfRecords	Numer ostatniego artykułu na serwerze.	Liczba przeszukanych rekordów.	Liczba przeszukanych dokumentów.
Target/NumberOfCompletedRecords	Liczba przeszukanych artykułów.	Liczba przeszukanych rekordów.	Liczba przeszukanych dokumentów.
Target/NumberOfInsertedRecords	Liczba nowo opublikowanych artykułów.	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.
Target/NumberOfUpdatedRecords	Nie dotyczy.	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.
Target/NumberOfDeletedRecords	Nie dotyczy.	Liczba usuniętych rekordów.	Liczba usuniętych rekordów.
Target/StartTime	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.
Target/EndTime	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.
Target/TotalTime	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.

Tabela 17. Elementy i atrybuty zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy NNTP, DB2 i Notes (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz NNTP	Przeszukiwacz DB2	Przeszukiwacz Notes
Target/AggregationLevel	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0, 1: <ul style="list-style-type: none"> <li>0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).</li> <li>Przeszukiwacz Notes przeszukuje dokumenty w trybie katalogu.</li> </ul>
Target/LastUpdatedTime	Nie dotyczy.	Czas ostatniej aktualizacji: <ul style="list-style-type: none"> <li>0: Przeszukiwanie aktywne (normalne)</li> <li>1: Przeszukiwanie pasywne (publikowanie zdarzeń DB2)</li> </ul>	Nie dotyczy.
Target/LastResetTime	Nie dotyczy.	Czas ostatniego resetowania statystyki: <ul style="list-style-type: none"> <li>0: Przeszukiwanie aktywne (normalne)</li> <li>1: Przeszukiwanie pasywne (publikowanie zdarzeń DB2)</li> </ul>	Nie dotyczy.

Tabela 18. Elementy i atrybuty zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Target@Name	Nazwa podfolderu	Nazwa typu elementu	Nazwa klasy elementu
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>0: Nieprzeszukane</li> <li>1: Przeszukiwanie</li> <li>2: Zakończone (niezaplanowane)</li> <li>3: Oczekiwanie</li> <li>4: Wstrzymane</li> <li>-1: Błąd</li> </ul>

Tabela 18. Elementy i atrybuty zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy Exchange Server, DB2 Content Manager i Content Edition (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz Exchange Server	Przeszukiwacz DB2 Content Manager	Przeszukiwacz Content Edition
Target/NumberOfRecords	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.
Target/NumberOfInsertedRecords	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.
Target/NumberOfUpdatedRecords	Nie dotyczy.	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.
Target/NumberOfDeletedRecords	Nie dotyczy.	Liczba usuniętych rekordów.	Liczba usuniętych rekordów.
Target/StartTime	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.
Target/EndTime	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.
Target/TotalTime	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.
Target/AggregationLevel	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).
Target/LastUpdatedTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/LastResetTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.

Tabela 19. Elementy i atrybuty zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, systemu plików UNIX i systemu plików Windows

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Target@Name	Nazwa bazy danych lokalizacji lub bazy danych pomieszczeń	Nazwa bazy danych szafek	Nazwa podkatalogu
Target@CrawlType	0	0	0
Target/Status	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• -1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• -1: Błąd</li> </ul>	Status (0, 1, 2, 3, 4, -1) <ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• -1: Błąd</li> </ul>

Tabela 19. Elementy i atrybuty zwracane przez komendę pobierającą szczegółowy status przestrzeni przeszukiwania dla przeszukiwaczy QuickPlace, Domino Document Manager, systemu plików UNIX i systemu plików Windows (kontynuacja)

Nazwa elementu i atrybutu	Przeszukiwacz QuickPlace	Przeszukiwacz Domino Document Manager	Przeszukiwacze systemów plików UNIX i Windows
Target/StatusMessage	<ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• -1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• -1: Błąd</li> </ul>	<ul style="list-style-type: none"> <li>• 0: Nieprzeszukane</li> <li>• 1: Przeszukiwanie</li> <li>• 2: Zakończone (niezaplanowane)</li> <li>• 3: Oczekiwanie</li> <li>• 4: Wstrzymane</li> <li>• -1: Błąd</li> </ul>
Target/NumberOfRecords	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/NumberOfCompletedRecords	Liczba przeszukanych dokumentów.	Liczba przeszukanych dokumentów.	Liczba przeszukanych plików.
Target/NumberOfInsertedRecords	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.	Liczba wstawionych rekordów.
Target/NumberOfUpdatedRecords	Liczba przeszukanych baz danych lokalizacji i baz danych pomieszczeń.	Liczba przeszukanych szafek.	Liczba podkatalogów w przestrzeni przeszukiwania.
Target/NumberOfDeletedRecords	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.	Liczba zaktualizowanych rekordów.
Target/StartTime	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.	Data i godzina ostatniego uruchomienia przeszukiwacza.
Target/EndTime	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.	Data i godzina zakończenia ostatniego przeszukiwania.
Target/TotalTime	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.	Czas wykorzystany przez przeszukiwacz na przeszukiwanie.
Target/AggregationLevel	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).	0: Przeszukiwacz Notes przeszukuje dokumenty w trybie normalnym. (Pozostałe przeszukiwacze, z wyjątkiem przeszukiwacza Notes, zwracają zawsze wartość 0).
Target/LastUpdatedTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.
Target/LastResetTime	Nie dotyczy.	Nie dotyczy.	Nie dotyczy.

**Status analizatora składni:** Komenda pobierająca status analizatora składni zwraca informacje w postaci dokumentu XML. Komenda pobierająca status analizatora składni może zwrócić następujące informacje:

```
FFQC5314I Wynik: <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>

```

Poniższa tabela zawiera opis elementów XML z informacjami zwracanymi przez komendę pobierającą status analizatora składni:

*Tabela 20. Elementy zwracane przez komendę pobierającą status analizatora składni*

Element	Opis
Status	1, gdy sesja analizatora składni dla danej kolekcji jest uruchomiona, i 0, gdy sesja analizatora składni jest zatrzymana.
State	<p>Możliwe są następujące stany: Idle, Restart, Parsing, Resuming, Stopped, Initializing, Pause, NoParserServiceIsAvailable.</p> <p>Status Idle wskazuje, że analizator składni oczekuje w stanie uśpienia przez N minut na kolejne dokumenty, które nadejdą od przeszukiwaczy w danej kolekcji. Wartością domyślną czasu uśpienia jest 300 sekund.</p> <p>Status Restart wskazuje, że analizator oczekuje na zrestartowanie wirtualnej maszyny języka Java, w której realizuje się analizę składni lub podział na leksemy. Maszyna JVM, w której realizuje się analizę składni lub podział na leksemy, działa w oddzielnej sesji i w niej odbywa się ostateczne przetwarzanie dokumentów.</p> <p>Status Parsing wskazuje, że analizator składni przetwarza dokumenty.</p> <p>Status Parsing wskazuje, że analizator składni przetwarza dokumenty.</p> <p>Status Pause wskazuje, że działanie analizatora składni zostało wstrzymane przez sesję tworzenia indeksu dla danej kolekcji. Status Initializing oznacza, że analizator składni zaczyna działanie i inicjuje swój stan.</p> <p>Status Resuming wskazuje, że działanie analizatora składni zostało wznowione ze stanu Pause do stanu Parsing przez sesję tworzenia indeksu dla danej kolekcji.</p> <p>Status NoParserServiceIsAvailable wskazuje, że nie można przetwarzać dokumentów danej kolekcji, ponieważ nie jest dostępna żadna maszyna JVM, w której realizuje się analizę składni lub podział na leksemy. Oznacza to, że wszystkie maszyny JVM są wykorzystywane przez inne kolekcje.</p>
SnapshotTimeStamp	Czas (wyrażony w sekundach, które upłynęły od 1970 roku) następnego odczytu przez analizator składni dokumentów z tabel przeszukiwacza.
NumberOfDocsToBeIndexed	Liczba dokumentów danej kolekcji, które znajdują się w składnicy. Liczba ta obejmuje także dokumenty, które zostały oznaczone jako dokumenty do usunięcia podczas następnego procesu tworzenia indeksu.
ParseRate	Szybkość analizy wyrażona jako liczba dokumentów na sekundę.
ParseRateMBPerHour	Szybkość analizy wyrażona w megabajtach na godzinę.
NumberOfCpmThreads	Liczba wątków CPM używanych do przetwarzania dokumentów tej kolekcji przez maszynę JVM, w której realizuje się analizę składni lub podział na leksemy.

Tabela 20. Elementy zwracane przez komendę pobierającą status analizatora składni (kontynuacja)

Element	Opis
ParserServiceSession	Nazwa maszyny JVM służącej do analizy składni lub podziału na leksemy, która przetwarza dokumenty dla danej kolekcji. Pole to jest dostępne tylko wtedy, gdy analizator składni znajduje się w stanie Parsing.
Name	Nazwa przeszukiwacza.
Type	Typ przeszukiwacza (przeszukiwacz sieci WWW, NNTP, DB2 itp.).
ParserStatus	<ul style="list-style-type: none"> <li>• 0: Dokumenty z tego przeszukiwacza nie zostały jeszcze przeanalizowane.</li> <li>• 1: Dokumenty z tego przeszukiwacza są obecnie analizowane.</li> <li>• 2: Dokumenty z tego przeszukiwacza zostały przeanalizowane.</li> </ul>
NumberOfDocsToBeParsed	Liczba dokumentów z tego przeszukiwacza oczekujących na analizę.
NubmerOfDocsAlreadyParsed	Liczba przeanalizowanych dokumentów z tego przeszukiwacza.

**Status operacji tworzenia indeksu:** Komenda pobierająca status operacji tworzenia indeksu zwraca informacje w postaci dokumentu XML. Komenda pobierająca status operacji tworzenia indeksu może zwrócić następujące informacje:

```
FFQC5314I Wynik: <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>
```

Poniższa tabela zawiera opis każdego elementu XML z informacjami zwracanymi przez komendę pobierającą status operacji tworzenia indeksu:

Tabela 21. Elementy zwracane przez komendę pobierającą status operacji tworzenia indeksu

Element	Opis
IndexStatusId	Identyfikator statusu indeksu.
StartTime	Czas (wyrażony w sekundach, które upłynęły od 1970 roku) rozpoczęcia tworzenia danego indeksu.
Progress	Procent wykonania operacji tworzenia indeksu.
CurrentPhase	<ul style="list-style-type: none"> <li>• 1: faza przepisywania składnicy danych</li> <li>• 2: faza analizy globalnej</li> <li>• 3: faza tworzenia indeksu</li> </ul>
TotalPhase	Liczba faz operacji tworzenia indeksu. Wartość ta jest aktualnie równa 3.
IndexCopyProgress	Procent wykonania kopii indeksu. Proces kopiowania indeksu kopiuje utworzony indeks z serwera tworzenia indeksu na serwery wyszukiwania.

Tabela 21. Elementy zwracane przez komendę pobierającą status operacji tworzenia indeksu (kontynuacja)

Element	Opis
CurrentServer	Serwer wyszukiwania, na który kopiowany jest indeks.
TotalServer	Liczba serwerów wyszukiwania, na które ma zostać skopiowany indeks.
IndexCopyTime	Łączny czas wszystkich faz operacji tworzenia indeksu.
Progress	0 w przypadku pomyślnego utworzenia i skopiowania indeksu, a w przypadku błędu wartość niezerowa odpowiadająca kodowi komunikatu o błędzie.
StopTime	Czas zakończenia operacji tworzenia indeksu (wszystkie fazy) oraz jego kopiowania.
TotalTime	Przedział czasu pomiędzy godziną rozpoczęcia i zakończenia.
NumberOfDocuments	Liczba dokumentów w indeksie.
CurrentIndexWildcardSupport	Ustawienie znaków zastępczych, które ma zostać użyte podczas następnej operacji tworzenia indeksu. Możliwe wartości to: None, QueryExpansion lub IndexExpansion.
ScheduleStatus	<ul style="list-style-type: none"> <li>• 0 jeśli dla danej kolekcji i typu indeksu nie aktywowano harmonogramu.</li> <li>• 1 jeśli dla danej kolekcji i typu indeksu aktywowano harmonogram.</li> </ul>
ScheduledTimeEnabled	Czas (wyrażony w sekundach, które upłynęły od 1970 roku) rozpoczęcia następnej operacji tworzenia indeksu dla danej kolekcji i danego typu indeksu.

**Status serwera wyszukiwania:** Komenda pobierająca status serwera wyszukiwania zwraca informacje w postaci dokumentu XML. Komenda pobierająca status serwera wyszukiwania może zwrócić następujące informacje:

```
FFQC5314I Wynik: <?xml version="1.0"?>
<Monitor Type="Search" Count="1">
<SearchStatus Name="Search Manager (node1)" SearchID=
"searchmanager.node1" HostName="myComputer.svl.ibm.com">
<Status>1</Status>
</SearchStatus>
</Monitor>
```

Poniższa tabela zawiera opis elementów XML z informacjami zwracanymi przez komendę pobierającą status serwera wyszukiwania:

Tabela 22. Elementy zwracane przez komendę pobierającą status serwera wyszukiwania

Element	Opis
SearchStatusName	Nazwa i identyfikator sesji menedżera wyszukiwania, która monitoruje i konserwuje indeks wyszukiwania dla danej kolekcji.
HostName	Nazwa hosta serwera, na którym działa indeks wyszukiwania.
Status	<ul style="list-style-type: none"> <li>• 0 jeśli indeks wyszukiwania dla danej kolekcji nie działa.</li> <li>• 1 jeśli indeks wyszukiwania dla danej kolekcji działa.</li> </ul>

**Szczegółowy status serwera wyszukiwania:** Komenda pobierająca status serwera wyszukiwania może zwrócić następujące informacje:

```
FFQC5303I Search Manager (node1) (sid: searchmanager.node1)
już działa. PID: 15711
FFQC5314I Wynik: PID=18390
CacheHits=3
QueryRate=1
Port=44008
SessionId=coll.runtime.node1
```

```
CacheHitRate=0.333
ResponseTime=70
Status=1
SessionName=coll.runtime.node1.1
```

Poniższa tabela zawiera opis elementów z informacjami zwracanymi przez komendę pobierającą szczegółowy status serwera wyszukiwania:

Tabela 23. Elementy zwracane przez komendę pobierającą szczegółowy status serwera wyszukiwania

Element	Opis
CacheHits	Liczba wyników pobranych z pamięci podręcznej wyszukiwania.
QueryRate	Liczba zapytań odebranych w ostatnim przedziale czasu. Domyślnie przedział czasu ma długość pięciu minut.
Port	Numer portu używanego przez indeks wyszukiwania do nasłuchiwania i odbierania zapytań.
SessionId	Identyfikator sesji indeksu wyszukiwania danej kolekcji.
CacheHitRate	Liczba wyników pobranych z pamięci podręcznej wyszukiwania wyrażona jako procent wszystkich wyników wyszukiwania.
ResponseTime	Średni czas odpowiedzi (w milisekundach) dla określonego przedziału czasu. (Wartością domyślną jest pięć minut).
Status	<ul style="list-style-type: none"> <li>• 0 jeśli indeks wyszukiwania dla danej kolekcji nie działa.</li> <li>• 1 jeśli indeks wyszukiwania dla danej kolekcji działa.</li> </ul>
SessionName	Nazwa sesji indeksu wyszukiwania danej kolekcji.

## Kody powrotu komend esadmin

Komendy esadmin mogą zwracać następujące kody:

Tabela 24. Kody powrotu komend esadmin

Kod	Name	Opis
0	CODE_ERROR_NONE	Wykonanie komendy zakończyło się pomyślnie.
102	CODE_ERROR_INSTANTIATION_EXCEPTION	Wystąpił błąd podczas tworzenia instancji programu obsługi komendy.
103	CODE_ERROR_ACCESS_EXCEPTION	Podczas tworzenia instancji programu obsługi komendy wystąpił błąd niedozwolonego dostępu.
104	CODE_ERROR_EXECUTE_EXCEPTION	
105	CODE_ERROR_THROWABLE	
106	CODE_ERROR_NO_SUCH_METHOD	
107	CODE_ERROR_INVALID_SESSION	
108	CODE_ERROR_INVALID_PARAMETER	
109	CODE_ERROR_SESSION_NOT_RUNNING	

## Uzyskiwanie identyfikatorów sesji

Do wyświetlania listy komponentów wyszukiwania korporacyjnego oraz odpowiadających im identyfikatorów sesji służy komenda `esadmin check`. Poniższa tabela zawiera listę najczęstszych sesji, ich identyfikatorów, serwera, na którym działają, oraz ich stanów.

Tabela 25. Przykłady nazw sesji, serwerów, na których je uruchomiono, identyfikatorów sesji i stanów sesji

Sesja	Serwer, na którym działa sesja	Identyfikator sesji	Stan sesji
configmanager	serwer indeksowania	10433	Started
controller	serwer indeksowania	10464	Started
customcommunication	serwer indeksowania	Nie dotyczy	Nie dotyczy
datalistener	serwer indeksowania	10582	Started
discovery	serwer indeksowania	10649	Started
monitor	serwer indeksowania	10682	Started
parserservice	serwer indeksowania	10718	Started
resource.node1	serwer indeksowania	10759	Started
samplecpp	serwer indeksowania	10827	Started
sampletest	serwer indeksowania	10857	Started
scheduler	serwer indeksowania	10889	Started
searchmanager.node1	serwer indeksowania	10927	Started
utilities.node1	serwer indeksowania	10384	Started

#### Pojęcia pokrewne

“Tworzenie i odtwarzanie kopii zapasowej systemu wyszukiwania korporacyjnego” na stronie 245

Skrypty tworzenia i odtwarzania kopii zapasowych umożliwiają utworzenie i odtworzenie kopii zapasowej systemu wyszukiwania korporacyjnego.

“Monitorowanie działań związanych z wyszukiwaniem korporacyjnym” na stronie 213  
 Monitorując działania systemu i działania związane z kolekcjami, można wyświetlić informacje o statusie różnych procesów, obserwować zmienne w celu wykrycia potencjalnych problemów lub dopasować ustawienia konfiguracyjne w celu poprawienia wydajności.

#### Zadania pokrewne

“Monitorowanie przeszukiwaczy” na stronie 217

Użytkownik może wyświetlić ogólne informacje na temat statusu poszczególnych przeszukiwaczy w kolekcji albo wybrać opcje dotyczące wyświetlania szczegółowych informacji o działaniu przeszukiwacza.

“Uruchamianie serwerów wyszukiwania korporacyjnego” na stronie 209

Aby umożliwić użytkownikom wyszukiwanie w kolekcji, należy uruchomić procesy systemowe, a następnie uruchomić serwery przeszukiwania, analizy, indeksowania i wyszukiwania w kolekcji.

“Zatrzymywanie serwerów wyszukiwania korporacyjnego” na stronie 211

Może wystąpić konieczność zatrzymania i zrestartowania serwera wyszukiwania korporacyjnego, jeśli dokonano zmian w jego konfiguracji lub w celu rozwiązania problemów.



---

## Dokumentacja wyszukiwania korporacyjnego

Dokumentacja produktu WebSphere Information Integrator OmniFind Edition jest dostępna w wersji PDF oraz HTML.

Program instalacyjny produktu WebSphere Information Integrator OmniFind Edition może automatycznie zainstalować Centrum informacyjne. Instalacja Centrum informacyjnego przez program instalacyjny odbywa się na serwerze wyszukiwania. W przypadku środowiska z wieloma serwerami Centrum informacyjne instalowane jest na obu serwerach wyszukiwania. Jeśli Centrum informacyjne nie zostanie zainstalowane, po kliknięciu przycisku pomocy otwierane jest Centrum informacyjne w serwisie WWW firmy IBM. Aby wyświetlić tematy dotyczące wyszukiwania korporacyjnego w formacie HTML, należy uruchomić Centrum informacyjne.

Aby wyświetlić dokumenty w wersji PDF, należy przejść do katalogu `docs/ustawienia_narodowe/pdf`. Na przykład aby znaleźć dokumenty w języku angielskim, należy przejść do katalogu `docs/en_US/pdf`. Dokumentację w formacie PDF, pliki do pobrania, poprawki, noty techniczne oraz Centrum informacyjne można znaleźć również w serwisie wsparcia produktu WebSphere Information Integrator OmniFind Edition.

Poniższa tabela zawiera wykaz dostępnej dokumentacji, nazwy plików oraz ich położenie.

Tabela 26. Dokumentacja w formacie PDF dotycząca wyszukiwania korporacyjnego.

Nagłówek	Nagłówek	Nagłówek
<i>Podręcznik instalowania wyszukiwania korporacyjnego</i> (tematy tego dokumentu są dostępne również w Centrum informacyjnym)	iiysi.pdf	docs/język/pdf/
<i>Administrowanie wyszukiwaniem korporacyjnym</i> (tematy tego dokumentu są dostępne również w Centrum informacyjnym)	iiysa.pdf	docs/język/pdf/
<i>Programming Guide and API Reference for Enterprise Search</i> (tematy tego dokumentu są dostępne również w Centrum informacyjnym)	iiysp.pdf	docs/język/pdf/
<i>Messages Reference</i> (tematy tego dokumentu są dostępne również w Centrum informacyjnym)	iiysm.pdf	docs/język/pdf/
<i>Installation Requirements for Enterprise Search</i> (tematy tego dokumentu są dostępne również w Centrum informacyjnym)	iiysr.txt lub iiysr.htm	docs/ustawienia_narodowe/ (ten plik można uruchomić również w programie Pierwsze kroki)
<i>Release Notes</i>	iiysn.pdf	Dostępny jedynie w serwisie WWW dokumentacji produktu IBM WebSphere Information Integrator OmniFind Edition
<i>Text Analysis Integration</i>	iiyst.pdf	docs/język/pdf/





---

## Ułatwienia dostępu w produkcie WebSphere II OmniFind Edition

Interfejs użytkownika oraz dokumentacja produktu IBM WebSphere Information Integrator OmniFind Edition zawiera szereg ułatwień dostępu.

### Program instalacyjny

Do poruszania się po programie instalacyjnym produktu WebSphere II OmniFind Edition można używać skrótów klawiszowych. Niektóre z nich zostały opisane w poniższej tabeli.

*Tabela 27. Skróty klawiszowe w programie instalacyjnym.*

Działanie	Skrót
Podświetlenie przełącznika	Klawisz strzałki
Wybór przełącznika	Klawisz Tab
Podświetlenie przycisku	Klawisz Tab
Wybór przycisku	Klawisz Enter
Przejsięcie do kolejnego lub poprzedniego okna albo anulowanie	Podświetlenie odpowiedniego przycisku klawiszem Tab i naciśnięcie klawisza Enter
Deaktywacja aktywnego okna	Ctrl + Alt + Esc

### Konsola administracyjna i Centrum informacyjne wyszukiwania korporacyjnego

Konsola administracyjna i Centrum informacyjne to interfejsy oparte na przeglądarce, które można wyświetlać w programie Microsoft Internet Explorer lub Mozilla FireFox. Listę skrótów klawiszowych i innych funkcji ułatwień dostępu przeglądarki Internet Explorer lub FireFox można znaleźć w pomocy elektronicznej tych przeglądarek.

### Dokumentacja w formacie PDF

Pełna dokumentacja wyszukiwania korporacyjnego jest dostępna w formacie PDF. Dokumenty PDF można otwierać w programie Adobe Acrobat w wersji 6.0. Dokumenty PDF charakteryzują się określoną strukturą i powinny być możliwe do odczytania przez większość lektorów ekranowych.



---

## Glosariusz terminów związanych z wyszukiwaniem korporacyjnym

W niniejszym glosariuszu zostały zdefiniowane terminy wykorzystywane w interfejsie i dokumentacji wyszukiwania korporacyjnego.

### **administrator wyszukiwania korporacyjnego**

Rola administracyjna, która umożliwia użytkownikowi administrowanie całością systemu wyszukiwania korporacyjnego.

### **adnotacja**

Informacja na temat zakresu tekstu. Na przykład adnotacja może wskazywać, że dany zakres tekstu reprezentuje nazwę firmy. W architekturze UIMA adnotacja jest specjalnym rodzajem struktury cechy.

### **adnotator**

Komponent programowy, który wykonuje określone zadania analizy lingwistycznej i generuje oraz zapisuje adnotacje. Adnotator jest komponentem logiki analizy w mechanizmie analizy.

### **adres IP**

Unikalny, 32-bitowy adres identyfikujący hosta w sieci.

### **adres URL**

Sekwencja znaków, która reprezentuje zasoby informacyjne na komputerze lub sieci, takiej jak Internet. Ta sekwencja znaków obejmuje skróconą nazwę protokołu wykorzystywanego do uzyskiwania dostępu do zasobu informacyjnego oraz informacje wykorzystywane przez dany protokół do wyszukania zasobu informacyjnego.

### **agent użytkownika**

Aplikacja, która przegląda sieć WWW i pozostawia o sobie informacje w odwiedzanych serwisach WWW. W ramach wyszukiwania korporacyjnego przeszukiwacz sieci WWW jest agentem użytkownika.

### **analiza odsyłaczy**

Metoda oparta na analizie odsyłaczy hipertekstowych między dokumentami i używana do określania, które strony w kolekcji są istotne dla użytkowników.

### **analiza tekstu**

Proces wyodrębniania z tekstu znaczenia i innych informacji mający na celu poprawę wydajności pobierania danych w kolekcji.

### **analizator składni**

Program interpretujący dokumenty, które zostały dodane do składnicy danych wyszukiwania korporacyjnego. Analizator składni wyodrębnia informacje z dokumentów i przygotowuje je do indeksowania, wyszukiwania i pobierania.

### **aplikacja wyszukująca**

Program, który przetwarza zapytania, przeszukuje indeks, zwraca wyniki wyszukiwania oraz pobiera dokumenty źródłowe do kolekcji w systemie wyszukiwania korporacyjnego.

### **Architektura zarządzania informacjami bez określonej struktury (UIMA)**

Architektura firmy IBM definiująca środowisko do implementowania systemów na potrzeby analizy danych bez określonej struktury.

**archiwum mechanizmu przetwarzania**

Plik archiwum zip o rozszerzeniu .pear zawierający mechanizm analizy UIMA oraz wszystkie zasoby wymagane do korzystania z tego mechanizmu przy niestandardowej analizie w wyszukiwaniu korporacyjnym.

**biblioteka**

Obiekt systemowy pełniący rolę katalogu innych obiektów. Patrz również biblioteka Domino Document Manager.

**biblioteka Domino Document Manager**

Baza danych produktu Domino Document Manager, która jest punktem wejścia do produktu Domino Document Manager.

**bliskie wyszukiwanie**

Typ wyszukiwania, które szuka pewnych słów w tym samym zdaniu, akapicie lub dokumencie.

**certyfiakat**

Cyfrowy dokument wiążący klucz publiczny z tożsamością posiadacza tego certyfikatu, a tym samym umożliwiający uwierzytelnienie posiadacza certyfikatu. Certyfikaty są wydawane przez ośrodki certyfikacji.

**detektor**

Funkcja przeszukiwacza, która określa, które źródła danych są dostępne dla przeszukiwacza w celu pobrania informacji.

**Domino Internet Inter-ORB Protocol (DIIOP)**

Zadanie serwera działające na serwerze i współpracujące z obiektowym brokerem zapytań Domino w celu umożliwienia komunikacji między apletami Java tworzonymi przy użyciu klas Java produktu Notes i serwera Domino. Użytkownicy przeglądarek i serwery Domino używają protokołu DIIOP do komunikacji i wymiany danych obiektowych.

**drzewo kategorii**

Hierarchia kategorii wyświetlana w konsoli administracyjnej wyszukiwania korporacyjnego.

**dyrektywa blokująca indeksowanie**

Dyrektywa na stronie WWW z instrukcją dla robotów (takich jak przeszukiwacz sieci WWW), aby nie uwzględniały treści tych stron w indeksie.

**dyrektywa blokująca wykorzystywanie odsyłaczy**

Dyrektywa na stronie WWW z instrukcją dla robotów (takich jak przeszukiwacz sieci WWW), aby nie kontynuowały działania korzystając z odsyłaczy znajdujących się na tych stronach.

**identyfikacja języka**

Funkcja wyszukiwania korporacyjnego określająca język dokumentu.

**identyfikator URI**

Minimalny łańcuch znaków identyfikujący zasób abstrakcyjny lub fizyczny.

**indeks** Patrz indeks pełnotekstowy.**indeks pełnotekstowy**

Struktura danych, która zawiera odwołania do elementów danych w celu umożliwienia szybkiego wyszukania dokumentów zawierających terminy z zapytania.

**Java Database Connectivity (JDBC)**

Niezależny od bazy danych przemysłowy standard łączności między platformą Java a szerokim zakresem baz danych. Interfejs JDBC udostępnia - oparty na języku SQL - interfejs API poziomu wywołania, umożliwiający dostęp do baz danych.

**JavaScript**

Skryptowy język WWW używany w przeglądarkach i na serwerach WWW.

**Katakana**

Zestaw znaków składający się z symboli używanych w jednym z dwóch najczęściej używanych fonetycznych alfabetów japońskich, służący przede wszystkim do fonetycznego zapisu słów w językach obcych.

**kategoria**

Grupa dokumentów o podobnych właściwościach.

**kategoria oparta na modelu**

Systematyka predefiniowanych terminów, używana do określania tematu dokumentu w celu umożliwienia poindeksowania dokumentu i wyszukania go wraz z dokumentami, które mają podobną treść.

**kategoria oparta na regule**

Kategorie tworzone w oparciu o reguły określające, które dokumenty są powiązane z danymi kategoriami. Można na przykład zdefiniować reguły wiążące dokumenty, które zawierają lub nie zawierają pewnych słów lub które są zgodne ze wzorcem identyfikatora URI, z konkretnymi kategoriami.

**klasa kryteriów**

Specyfikacja, która może mieć wpływ na względną ocenę dokumentu w wynikach wyszukiwania.

**klityka** Słowo, które składniowo funkcjonuje oddzielnie, ale fonetycznie jest połączone z innym słowem. Klityka może być zapisana łącznie lub osobno w stosunku do słowa, z którym jest powiązana. Typowe przykłady klityki to ostatnie człony kontrakcji w języku angielskim (*wouldn't* lub *you're*).

**kolejka indeksów**

Lista żądań reorganizacji lub odświeżenia indeksu, które mają być przetworzone.

**kolekcja**

Zbiór źródeł danych oraz opcji do przeszukiwania, analizowania, indeksowania i wyszukiwania w tych źródłach danych.

**lemat** Kanoniczna postać słowa. Lematy są bardzo istotne w przypadku języków o rozbudowanej fleksji, takich jak język polski.

**lematyzacja**

Proces poszukiwania w słowniku lematu danego słowa. Lematyzacja różni się od wyszukiwania rdzenia, ponieważ wyszukiwanie rdzenia jest procesem algorytmicznym i generalnie podczas działania nie używa słownika zawierającego listę słów w danym języku.

**ligatura**

Co najmniej dwa połączone znaki tworzące jeden znak (tak jak litery f oraz i składające się na ligaturę fi).

**Lightweight Directory Access Protocol (LDAP)**

Otwarty protokół używający protokołu TCP/IP w celu zapewnienia dostępu do katalogów obsługujących model X.500 i nienarzucający związanych z zasobami wymagań bardziej złożonego protokołu X.500 Directory Access Protocol.

**lista kontroli dostępu**

Lista identyfikująca użytkowników, którzy mogą uzyskać dostęp do powiązanego z nią obiektu, a także określająca prawa dostępu użytkowników do tego obiektu.

**lokalizacja**

Wirtualne miejsce widoczne w portalu, w którym osoby i grupy osób spotykają się, aby ze sobą współpracować. W portalu każdy użytkownik ma osobistą lokalizację do

pracy prywatnej, a szczególnie użytkownicy i grupy mają dostęp do różnych lokalizacji współużytkowanych, które mogą być lokalizacjami publicznymi lub zastrzeżonymi. Patrz również lokalizacja Lotus QuickPlace.

**lokalizacja Lotus QuickPlace**

Miejsce w sieci WWW udostępniane przez produkt Lotus QuickPlace i umożliwiające rozproszonym geograficznie uczestnikom współpracę nad projektami oraz komunikowanie się w trybie z połączeniem w zabezpieczonym obszarze roboczym o określonej strukturze.

**lokalny program stowarzyszający**

Program stowarzyszający klienta, stowarzyszający zestaw obiektów dostępnych do wyszukiwania.

**mechanizm analizy**

Patrz mechanizm analizy tekstu.

**mechanizm analizy tekstu**

Komponent programowy odpowiedzialny za znajdowanie i reprezentowanie kontekstu i treści semantycznej w tekście.

**mechanizm wyszukiwania**

Program, który akceptuje żądanie wyszukania i zwraca użytkownikowi listę dokumentów.

**model DOM (Document Object Model)**

System, w którym dokument o określonej strukturze (taki jak plik XML) jest wyświetlany jako drzewo obiektów, do którego można uzyskać dostęp i aktualizować je programowo.

**monitor**

Użytkownik funkcji wyszukiwania korporacyjnego z uprawnieniami do obserwowania procesów na poziomie kolekcji.

**nazwa wyróżniająca**

Nazwa jednoznacznie identyfikująca pozycję w katalogu. Nazwa wyróżniająca składa się z rozdzielonych przecinkami par atrybut:wartość. Jest to również zestaw par nazwa-wartość (takich jak CN=nazwisko osoby oraz C=kraj lub region), które jednoznacznie identyfikują podmiot w certyfikacie cyfrowym.

**normalizacja znaku**

Proces, w ramach którego różne formy znaku (takie jak wielka litera lub znaki diakrytyczne) są sprowadzane do wspólnej postaci.

**Notes remote procedure call (NRPC)**

Warstwa w architekturze produktu Lotus Notes używana zawsze podczas komunikacji między produktami Notes.

**ocena** Proces przypisywania każdemu dokumentowi w wynikach wyszukiwania zapytania wartości całkowitej. Kolejność dokumentów w wynikach wyszukiwania jest oparta na trafności dla danego zapytania. Wyższa ocena oznacza lepszą zgodność. Patrz także ocena dynamiczna i ocena statyczna.

**ocena dynamiczna**

Typ oceny, w którym warunki w zapytaniu są analizowane pod kątem wyszukiwanych dokumentów w celu określenia oceny wyników. Patrz także punktacja oparta na tekście. Patrz termin przeciwstawny ocena statyczna.

**ocena popularności**

Typ oceny, która jest dodawana do istniejącej oceny dokumentu w oparciu o jego popularność.

**ocena statyczna**

Typ oceniania, w którym na ocenę wpływają czynniki dotyczące ocenianych dokumentów, takie jak data, liczba odsyłaczy wskazujących dokument itd. Patrz termin przeciwstawny ocena dynamiczna.

**odświeżanie indeksu**

Proces dodawania nowych informacji do istniejącego indeksu w systemie wyszukiwania korporacyjnego. Patrz termin przeciwstawny reorganizacja indeksu.

**operator**

Użytkownik funkcji wyszukiwania korporacyjnego z uprawnieniami do obserwowania, uruchamiania i zatrzymywania procesów na poziomie kolekcji.

**ośrodek certyfikacji**

Organizacja wydająca certyfikaty i uwierzytelniająca podmioty (osoby lub organizacje) biorące udział w transakcjach elektronicznych. Ośrodki certyfikacji gwarantują, że obie wymieniające informacje strony są rzeczywiście tymi, za kogo się podają.

**pamięć podręczna wyszukiwania**

Bufor, który przechowuje dane i wyniki poprzednich żądań wyszukiwania.

**plik kluczy**

Plik bazy danych kluczy zawierający klucze publiczne przechowywane w postaci certyfikatów osoby podpisującej oraz klucze prywatne przechowywane w certyfikatach osobistych.

**pliki indeksu wyszukiwania**

Zestaw plików, w których przechowywany jest indeks w mechanizmie wyszukiwania.

**początkowy adres URL**

Punkt początkowy przeszukiwania.

**podsumowywanie**

Proces włączania zdań do wyników wyszukiwania w celu skróconego opisu treści dokumentu. Patrz także podsumowywanie dynamiczne i podsumowywanie statyczne.

**podsumowywanie dynamiczne**

Typ podsumowywania, w którym warunki wyszukiwania są podświetlone, a wyniki wyszukiwania zawierają frazy, które najlepiej odpowiadają pojęciom dokumentu wyszukiwanego przez użytkownika. Patrz termin przeciwstawny podsumowywanie statyczne.

**podsumowywanie statyczne**

Typ podsumowania, w którym wyniki wyszukiwania zawierają określone zapisane podsumowanie z dokumentu. Patrz termin przeciwstawny podsumowywanie dynamiczne.

**podział na ciągi n-znakowe**

Metoda analizy traktująca jako pojedyncze słowo nakładające się sekwencje podanej liczby znaków, w przeciwieństwie do opartego na kodowaniu Unicode podziału wg znaków odstępu, w którym podział na słowa następuje według znaków spacji.

**pole** Najmniejsza identyfikowalna część rekordu.

**pomieszczenie**

Program umożliwiający użytkownikom tworzenie dokumentów przeznaczonych do czytania przez inne osoby, odpowiadanie na komentarze innych osób oraz

przeglądanie statusu i terminów projektu. Użytkownicy mogą również brać udział w rozmowie sieciowej z innymi osobami z tego samego pomieszczenia. Patrz również pomieszczenie Lotus QuickPlace.

**pomieszczenie Lotus QuickPlace**

Wydzielony obszar lokalizacji Lotus QuickPlace przeznaczony tylko dla autoryzowanych członków o wspólnych zainteresowaniach i potrzebie wspólnej pracy.

**powinowactwo leksykalne**

Związek wyszukiwanych słów występujących blisko siebie w dokumencie. Powinowactwo leksykalne jest używane do obliczania trafności wyniku.

**przeźródła przeszukiwania**

Zbiór źródeł, które są zgodne z podanymi wzorcami (takimi jak nazwy baz danych, ścieżki systemu plików, nazwy domen, adresy IP i adresy URL), i które przeszukiwacz odczytuje w celu pobrania elementów do indeksowania.

**przeszukiwacz**

Program, który pobiera dane ze źródeł danych i zbiera informacje, które można wykorzystać do tworzenia indeksów wyszukiwania.

**przeszukiwacz sieci WWW**

Klasa oprogramowania robotów eksplorujących sieć WWW przez pobieranie dokumentów WWW i podążanie według odsyłaczy w obrębie kolejnych dokumentów.

**referencja**

Szczegółowe informacje uzyskiwane podczas uwierzytelniania i opisujące danego użytkownika, jego wszystkie powiązania z grupami i inne atrybuty tożsamości związane z zabezpieczeniami. Referencje mogą służyć do wykonywania całego szeregu usług, takich jak autoryzacja, kontrola i delegowanie.

**reorganizacja indeksu**

Proces budowania indeksu w systemie wyszukiwania korporacyjnego. Patrz termin przeciwstawny odświeżanie indeksu.

**Robots Exclusion Protocol**

Protokół umożliwiający administratorom serwisów WWW wskazanie robotom odwiedzającym ich serwis, które części tego serwisu nie powinny być odwiedzane.

**rola administracyjna**

Klasyfikacja użytkownika, która określa funkcje, które użytkownik może realizować w konsoli administracyjnej wyszukiwania korporacyjnego. Rola określa również, którymi kolekcjami użytkownik może administrować.

**Secure Sockets Layer (SSL)**

Protokół zabezpieczeń zapewniający poufność komunikacji.

**segmentacja**

Proces, w którym funkcja kontroli ścieżki dzieli podstawowe jednostki informacji na mniejsze jednostki (nazywane segmentami BIU) w celu dostosowania ich do mniejszych wielkości buforów w przylegających serwerach.

**segmentacja przy użyciu znaków spacji, tabulacji lub nowego wiersza oparta na kodzie Unicode**

Metoda dzielenia na leksemy, w której do odróżniania leksemów od znaków separujących stosowane są właściwości znaków w kodowaniu Unicode.



**serwer proxy**

Serwer działający jako pośrednik dla żądań WWW protokołu HTTP udostępniany przez aplikację lub serwer WWW. Serwer proxy działa jak odpowiednik serwerów treści w przedsiębiorstwie.

**serwlet** Program Java działający na serwerze WWW i rozszerzający funkcjonalność serwera przez generowanie dynamicznej treści w odpowiedzi na żądania klienta WWW. Serwlety są powszechnie stosowane do podłączania baz danych do sieci WWW.

**słownik synonimów**

Słownik umożliwiający użytkownikom wyszukiwanie synonimów dla terminów użytych w zapytaniu podczas wyszukiwania w kolekcji.

**słowo zatrzymania**

Powszechnie używane słowo, takie jak *w*, *na* lub *oraz*, które jest ignorowane przez aplikację wyszukującą.

**słowo-kryterium**

Słowo, które może mieć wpływ na ocenę istotności dokumentu w wynikach wyszukiwania. Podczas przetwarzania zapytań ważność dokumentu zawierającego słowo-kryterium może zostać zwiększona lub zmniejszona, w zależności od predefiniowanego wyniku dla danego słowa.

**stowarzyszenie**

Proces łączenia systemów nazewnictwa, tak aby zagregowany system mógł przetwarzać nazwy złożone obejmujące te systemy nazewnictwa.

**strona błędu programowego**

Specjalna strona szczegółowo wyjaśniająca problem w sytuacji, gdy serwer HTTP nie może zwrócić strony, której żąda klient, i konfigurująca serwer HTTP do zwrócenia tej strony zamiast odpowiedzi składającej się wyłącznie z nagłówka z kodem powrotu wskazującym problem.

**strony JSP (JavaServer Pages)**

Technologia skryptowa serwerów umożliwiająca dynamiczne osadzanie kodu języka Java w ramach stron WWW (plików HTML) i jego wykonywanie po udostępnieniu strony w celu zwrócenia klientowi dynamicznej treści.

**struktura cechy**

Bazowa struktura danych reprezentująca wynik analizy tekstu. Struktura cechy to struktura typu atrybut-wartość. Każda struktura cechy ma określony typ, a każdy typ ma określony zestaw poprawnych cech lub atrybutów, bardzo podobnie jak w przypadku klasy Java.

**systematyka**

Klasyfikacja obiektów według grup opartych na podobieństwie. W wyszukiwaniu korporacyjnym systematyka organizuje dane według kategorii i podkategorii. Patrz także drzewo kategorii.

**szafka Domino Document Manager**

Baza danych produktu Domino Document Manager służąca do organizowania dokumentów. W szafkach przechowywane są bazy danych Domino.

**szybki odsyłacz**

Powiązanie między identyfikatorem URI a słowami kluczowymi i frazami.

**ścieżka do cechy**

Ścieżka umożliwiająca uzyskanie dostępu do wartości cechy w strukturze cechy UIMA.

**token** Podstawowe jednostki tekstu indeksowane przez wyszukiwanie korporacyjne. Leksemami mogą być słowa języka lub inne jednostki tekstu odpowiednie do indeksowania.

**token zabezpieczeń**

Informacje o tożsamości i zabezpieczeniach, które są używane do autoryzacji dostępu do dokumentów w kolekcji. Różne typy źródeł danych obsługują różne typy tokenów zabezpieczeń. Do przykładów należą role użytkowników, identyfikatory użytkowników, identyfikatory grup oraz inne informacje, które można wykorzystać do kontroli dostępu do treści.

**tokenizer**

Program segmentacji tekstu skanujący tekst i określający, czy i kiedy seria znaków może zostać uznana za leksem.

**typ MIME**

Internetowy standard identyfikowania typu obiektu przesyłanego przez sieć Internet.

**typ źródła danych**

Zgrupowanie źródeł danych według protokołu wykorzystywanego do uzyskiwania dostępu do danych.

**Universal Resource Name (URN)**

Element protokołu internetowego składający się z krótkiego łańcucha znaków zgodnego z określoną składnią. Łańcuch ten zawiera nazwę lub adres, który umożliwia odwołanie się do zasobu.

**usuwanie słów zatrzymania**

Proces usuwania z zapytania słów ignorowanych w celu pominięcia popularnych słów i zwrócenia bardziej precyzyjnych wyników.

**usuwanie z kolejki**

Służy do usuwania elementów z kolejki.

**warstwa CCL (Common Communication Layer, wspólna warstwa komunikacji)**

Infrastruktura komunikacyjna jednocząca różne komponenty (kontroler, analizator składni, przeszukiwacz, indeksator) produktu WebSphere Information Integrator OmniFind Edition.

**wartościowanie danych tekstowych**

Proces przypisywania dokumentowi wartości całkowitej, która określa trafność dokumentu względem terminów podanych w zapytaniu. Wyższa wartość całkowita oznacza większą zgodność z zapytaniem. Patrz także ocena dynamiczna.

**wirtualna maszyna języka Java (JVM)**

Programowa implementacja procesora wykonującego skompilowany kod języka Java (aplety i aplikacje).

**wpisywanie do kolejki**

Służy do wstawiania elementów do kolejki.

**wspólna struktura analizy**

Struktura przechowująca dokument analizowany przez mechanizm analizy tekstu. Informacje we wspólnej strukturze analizy są przechowywane w formie adnotacji i innych struktur cech.

**wyniki analizy**

Informacje generowane przez adnotatory. Wyniki analizy, które odpowiadają wyszukiwanym informacjom, są zapisywane w strukturze danych nazywanej wspólną strukturą analizy.

**wyniki wyszukiwania**

Lista dokumentów, które są zgodne z żądaniem wyszukiwania.

**wyodrębnianie informacji**

Typ wyodrębniania pojęć, które automatycznie rozpoznaje w dokumentach tekstowych istotne elementy słownikowe, takie jak nazwy, terminy i wyrażenia.

**wyodrębnianie kompozycji**

Typ wyodrębniania pojęcia, które automatycznie rozpoznaje w dokumentach tekstowych istotne elementy słownikowe w celu wyodrębnienia kompozycji lub tematu dokumentu. Patrz także wyodrębnianie pojęcia.

**wyodrębnianie pojęcia**

Funkcja wyszukiwania, która identyfikuje istotne elementy słownikowe (takie jak ludzie, miejsca lub produkty) w dokumentach tekstowych i tworzy listę tych elementów. Patrz także wyodrębnianie kompozycji.

**wyszukiwanie boolowskie**

Wyszukiwanie, w ramach którego jeden lub więcej terminów wyszukiwania jest połączonych przy użyciu takich operatorów, jak AND, NOT i OR.

**wyszukiwanie dla pól**

Zapytanie, które jest ograniczone do konkretnego pola.

**wyszukiwanie hybrydowe**

Połączenie wyszukiwania boolowskiego z wyszukiwaniem pełnotekstowym.

**wyszukiwanie lingwistyczne**

Typ wyszukiwania, w którym dokumenty są przeglądane, pobierane i indeksowane z terminami zredukowanymi do form podstawowych (na przykład termin *mysz* jest indeksowany jako *mysz*) lub rozwiniętymi do form podstawowych (jak w przypadku słów złożonych).

**wyszukiwanie parametryczne**

Typ wyszukiwania, w którym wyszukiwane są obiekty zawierające liczbowe wartości lub atrybuty, takie jak data, liczba całkowita lub inne typy danych liczbowych z określonego zakresu.

**wyszukiwanie pełnotekstowe**

Wyszukiwanie, w którym wyszukiwany jest termin w postaci dowolnego tekstu.

**wyszukiwanie rdzenia**

Patrz wyszukiwanie rdzenia słowa.

**wyszukiwanie rdzenia słowa**

Proces lingwistycznej normalizacji, w ramach której warianty słowa są redukowane do wspólnej postaci. Na przykład słowa takie jak *połączenia*, *połączeniowy* i *połączony* są redukowane do słowa *połącz*.

**wyszukiwanie rozmyte**

Wyszukiwanie, które zwraca słowa wymawiane w sposób podobny do wymowy słowa będącego warunkiem wyszukiwania.

**wyszukiwanie stowarzyszone**

Funkcja wyszukiwania umożliwiająca wyszukiwanie w wielu usługach wyszukiwania i zwracająca skonsolidowaną listę wyników wyszukiwania.

**wyszukiwanie terminów ważonych**

Zapytanie, w którym pewne terminy otrzymują większe znaczenie.

**XML Path Language (XPath)**

Język jednoznacznie identyfikujący lub adresujący części źródłowego dokumentu XML. Język XPath udostępnia również podstawowe narzędzia do manipulowania łańcuchami, liczbami i operatorami boolowskimi.

**zapytanie w języku naturalnym**

Typ wyszukiwania, które analizuje zapisane wyrażenia (takie jak "Kto kieruje wydziałem finansowym?") zamiast odczytywania prostych słów kluczowych.

**zarządzanie tożsamościami**

Możliwość zaszyfrowania referencji użytkowników w bezpiecznej składnicy.

**zasięg** Grupa pokrewnych identyfikatorów URI używana do definiowania zakresu żądania wyszukiwania.

**zdalny program stowarzyszający**

Program stowarzyszający serwera, stowarzyszający zestaw obiektów dostępnych do wyszukiwania.

**zewnętrzne źródło danych**

Źródło danych do stowarzyszenia, które nie jest przeszukiwane, analizowane ani indeksowane przez produkt WebSphere Information Integrator OmniFind Edition. Operacje wyszukiwania w zewnętrznych źródłach danych są delegowane do odpowiednich interfejsów API zapytań tych źródeł danych.

**znak diakrytyczny**

Znak dodawany do litery w celu zmiany wymowy słowa lub w celu odróżnienia podobnych słów, taki jak znak akcentu lub niemiecki symbol umlaut.

**znak końcowy**

Znak zajmujący ostatnie miejsce w słowie.

**znak maskujący**

Znak, który jest wykorzystywany do reprezentowania opcjonalnych znaków na początku, w środku i na końcu wyszukiwanego terminu. Znaki maskujące są zwykle używane do znajdowania w indeksie odmian terminu. Patrz również znak zastępczy.

**znak nowego wiersza**

Znak sterujący powodujący przejście o jeden wiersz w dół na wydruku lub podczas wyświetlania. Niektóre systemy wymagają w tym celu więcej niż jednego znaku.

**znak zastępczy**

Znak, który jest wykorzystywany do reprezentowania opcjonalnych znaków na początku, w środku lub na końcu wyszukiwanego terminu.

**znak zmiany znaczenia**

Znak, który wyłącza lub wybiera specjalne znaczenie jednego lub więcej występujących po nim znaków.

**źródło danych**

Dowolne repozytorium danych, z którego można pobrać dokumenty, takie jak sieć WWW, relacyjne i nierelacyjne bazy danych oraz systemy zarządzania treścią.

---

## Uzyskiwanie dostępu do informacji o produktach WebSphere Information Integration

Informacje o produktach WebSphere Information Integration można uzyskać telefonicznie lub w sieci WWW.

Podane tutaj numery telefonów są ważne w Stanach Zjednoczonych.

- Aby zamówić produkty lub uzyskać ogólne informacje, należy zadzwonić pod numer: 1-800-IBM-CALL (1-800-426-2255)
- Aby zamówić publikacje, należy zadzwonić pod numer: 1-800-879-2755

Informacje o produktach WebSphere Information Integration można znaleźć również w sieci WWW pod adresem [www.ibm.com/software/data/integration/db2ii/](http://www.ibm.com/software/data/integration/db2ii/). Ten serwis WWW zawiera najnowsze informacje na temat:

- dokumentacji produktu
- produktów do pobrania
- pakietów poprawek
- uwag do wydania i innej dokumentacji technicznej
- wiadomości o produktach WebSphere Information Integration
- odsyłaczy do zasobów sieci WWW, takich jak raporty i dokumentacja techniczna IBM (Redbooks)
- odsyłaczy do grup dyskusyjnych i grup użytkowników
- odsyłaczy do elektronicznych Centrów informacyjnych dla produktów WebSphere Information Integration
- zamawiania podręczników

Aby uzyskać dostęp do dokumentacji produktu:

1. Przejdź do serwisu WWW pod adresem [www.ibm.com/software/data/integration/db2ii/](http://www.ibm.com/software/data/integration/db2ii/).
2. Wybierz produkt z listy rozwijanej, na przykład WebSphere Information Integrator OmniFind Edition.
3. Kliknij odsyłacz Support (Wsparcie) w lewej części wyświetlonej strony.
4. W sekcji Learn (Informacje) wybierz żądany odsyłacz. Jeśli dla wybranego produktu jest dostępne Centrum informacyjne, można wybrać odsyłacz do Centrum informacyjnego. Przykład można znaleźć na rysunku Rys. 1 na stronie 292.

## 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](#)

*Rysunek 1. Przykład odsyłaczy do dokumentacji produktu w serwisie WWW wsparcia produktów WebSphere Information Integration*

---

## Komentarze na temat dokumentacji

Prosimy o nadsyłanie wszelkich komentarzy na temat przedstawionych tutaj informacji lub innych dokumentów dotyczących produktów IBM WebSphere Information Integration.

Opinie klientów pomagają firmie IBM dostarczać informacje dobrej jakości. Prosimy o nadsyłanie wszelkich komentarzy na temat tych informacji lub innych dokumentów dotyczących produktów WebSphere Information Integration. Uwagi można przekazywać w dowolny z poniższych sposobów:

1. Komentarze można wysłać za pomocą elektronicznego formularza komentarzy czytelników dostępnego pod adresem [www.ibm.com/software/awdtools/rcf/](http://www.ibm.com/software/awdtools/rcf/).
2. Komentarze można wysłać pocztą elektroniczną na adres [comments@us.ibm.com](mailto:comments@us.ibm.com). Należy podać nazwę produktu, numer wersji produktu oraz nazwę i numer części podręcznika (jeśli ma zastosowanie). Jeżeli komentarz dotyczy konkretnego fragmentu tekstu, prosimy podać jego położenie (na przykład tytuł, numer tabeli lub numer strony).





---

## Kontakt z firmą IBM

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

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

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

Aby zlokalizować biuro firmy IBM w danym kraju lub regionie, należy skorzystać z informacji umieszczonych na stronie IBM Directory of Worldwide Contacts pod adresem [www.ibm.com/planetwide](http://www.ibm.com/planetwide).



---

## Znaki towarowe

Ta sekcja zawiera spis znaków towarowych firmy IBM i niektórych znaków towarowych firm trzecich.

Więcej informacji na temat znaków towarowych IBM zawiera strona WWW <http://www.ibm.com/legal/copytrade.shtml>.

Poniższe nazwy są znakami towarowymi lub zastrzeżonymi znakami towarowymi innych firm:

Java i wszystkie pokrewne znaki towarowe Java są znakami towarowymi lub zastrzeżonymi znakami towarowymi firmy Sun Microsystems, Inc. w Stanach Zjednoczonych i/lub innych krajach.

Microsoft, Windows, Windows NT oraz logo Windows są znakami towarowymi firmy Microsoft Corporation w Stanach Zjednoczonych i/lub innych krajach.

Intel, logo Intel Inside, MMX i Pentium są znakami towarowymi firmy Intel Corporation w Stanach Zjednoczonych i/lub innych krajach.

UNIX jest zastrzeżonym znakiem towarowym The Open Group w Stanach Zjednoczonych i innych krajach.

Linux jest znakiem towarowym Linusa Torvaldsa w Stanach Zjednoczonych i/lub innych krajach.

Nazwy innych firm, produktów i usług mogą być znakami towarowymi lub znakami usług firm trzecich.



---

## Uwagi

Niniejsza publikacja została przygotowana z myślą o produktach i usługach oferowanych w Stanach Zjednoczonych. IBM może nie oferować w innych krajach produktów, usług lub opcji, omawianych w tej publikacji. Informacje o produktach i usługach dostępnych w danym kraju można uzyskać od lokalnego przedstawiciela firmy IBM. Jakakolwiek wzmianka na temat produktu, programu lub usługi firmy IBM nie oznacza, że może być zastosowany jedynie ten produkt, ten program lub ta usługa firmy IBM. Zamiast nich można zastosować ich odpowiednik funkcjonalny, pod warunkiem, że nie narusza to praw własności intelektualnej firmy IBM. Jednakże cała odpowiedzialność za ocenę przydatności i sprawdzenie działania produktu, programu lub usługi pochodzących od producenta innego niż IBM spoczywa na użytkowniku.

IBM może posiadać patenty lub złożone wnioski patentowe na towary i usługi, o których mowa w niniejszej publikacji. Używanie tego dokumentu nie daje żadnych praw do tych patentów. Wnioski o przyznanie licencji można zgłaszać na piśmie pod adresem:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive Armonk, NY  
10504-1785 U.S.A.

Zapytania dotyczące zestawów znaków dwubajtowych (DBCS) należy kierować do lokalnych działów własności intelektualnej firmy IBM (IBM Intellectual Property Department) lub wysłać je na piśmie na adres:

IBM World Trade Asia Corporation  
Licensing  
2-31 Roppongi 3-chome, Minato-ku  
Tokyo 106-0032, Japan

Poniższy akapit nie obowiązuje w Wielkiej Brytanii, a także w innych krajach, w których jego treść pozostaje w sprzeczności z przepisami prawa miejscowego: FIRMA INTERNATIONAL BUSINESS MACHINES CORPORATION DOSTARCZA TĘ PUBLIKACJĘ W TAKIM STANIE, W JAKIM SIĘ ZNAJDUJE "AS IS" BEZ UDZIELANIA JAKIKOLWIEK GWARANCJI (W TYM TAKŻE RĘKOJMI), WYRAŻNYCH LUB DOMNIEMANYCH, A W SZCZEGÓLNOŚCI DOMNIEMANYCH GWARANCJI PRZYDATNOŚCI HANDLOWEJ ORAZ PRZYDATNOŚCI DO OKREŚLONEGO CELU LUB GWARANCJI, ŻE PUBLIKACJA NIE NARUSZA PRAW OSÓB TRZECICH. Ustawodawstwa niektórych krajów nie dopuszczają zastrzeżeń dotyczących gwarancji wyraźnych lub domniemanych w odniesieniu do pewnych transakcji; w takiej sytuacji powyższe zdanie nie ma zastosowania.

Informacje zawarte w niniejszej publikacji mogą zawierać nieścisłości techniczne lub błędy drukarskie. Informacje te są okresowo aktualizowane, a zmiany te zostaną ujęte w kolejnych wydaniach tej publikacji. IBM zastrzega sobie prawo do wprowadzania ulepszeń i/lub zmian w produktach i/lub programach opisanych w tej publikacji w dowolnym czasie, bez wcześniejszego powiadomienia.

Jakiekolwiek wzmianki na temat stron internetowych nie należących do firmy IBM zostały podane jedynie dla wygody użytkownika i nie oznaczają, że firma IBM w jakikolwiek sposób

firmuje te strony. Materiały dostępne na tych stronach nie są częścią materiałów opracowanych dla tego produktu IBM, a użytkownik korzysta z nich na własną odpowiedzialność.

IBM ma prawo do korzystania i rozpowszechniania informacji przysłanych przez użytkownika w dowolny sposób, jaki uzna za właściwy, bez żadnych zobowiązań wobec ich autora.

Licencjobiorcy tego programu, którzy chcieliby uzyskać informacje na temat programu w celu: (i) wdrożenia wymiany informacji między niezależnie utworzonymi programami i innymi programami (łącznie z tym opisywanym) oraz (ii) wykorzystywania wymienianych informacji, powinni skontaktować się z:

IBM Corporation J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003 U.S.A.

Informacje takie mogą być udostępnione, o ile spełnione zostaną odpowiednie warunki, w tym, w niektórych przypadkach, uiszczenie odpowiedniej opłaty.

Licencjonowany program opisany w tym dokumencie oraz wszystkie inne licencjonowane materiały dostępne dla tego programu są dostarczane przez IBM na warunkach określonych w Umowie IBM z Klientem, Międzynarodowej Umowie Licencyjnej IBM na Program lub w innych podobnych umowach zawartych między IBM i użytkownikami.

Wszelkie dane dotyczące wydajności zostały zebrane w kontrolowanym środowisku. W związku z tym rezultaty uzyskane w innych środowiskach operacyjnych mogą się znacząco różnić. Niektóre pomiary mogły być dokonywane na systemach będących w fazie rozwoju i nie ma gwarancji, że pomiary te wykonane na ogólnie dostępnych systemach dadzą takie same wyniki. Niektóre z pomiarów mogły być estymowane przez ekstrapolację. Rzeczywiste wyniki mogą być inne. Użytkownicy powinni we własnym zakresie sprawdzić odpowiednie dane dla ich środowiska.

Informacje dotyczące produktów innych firm zostały uzyskane od dostawców tych produktów z opublikowanych przez nich zapowiedzi lub innych powszechnie dostępnych źródeł. Firma IBM nie testowała tych produktów i nie może potwierdzić dokładności pomiarów wydajności, kompatybilności ani żadnych innych danych związanych z tymi produktami. Pytania dotyczące możliwości produktów innych firm należy kierować do dostawców tych produktów.

Jakiegokolwiek wzmianki na temat kierunków rozwoju firmy IBM mogą ulec zmianie lub anulowaniu bez uprzedzenia i dotyczą jedynie ogólnych celów i założeń.

Publikacja ta zawiera przykładowe dane i raporty używane w codziennej działalności biznesowej. W celu kompleksowego zilustrowania tej działalności podane przykłady zawierają nazwy osób, firm i ich produktów. Wszystkie te nazwiska/nazwy są fikcyjne i jakakolwiek ich zbieżność z prawdziwymi nazwiskami/nazwami jest całkowicie przypadkowa.

#### LICENCJA NA PRAWA AUTORSKIE:

Niniejsza publikacja zawiera przykładowe aplikacje w kodzie źródłowym, ilustrujące techniki programowania w różnych systemach operacyjnych. Użytkownik może kopiować, modyfikować i rozpowszechniać te programy przykładowe w dowolnej formie bez uiszczenia opłat, w celu rozbudowy, użytkowania, handlowym lub w celu rozpowszechniania aplikacji zgodnych z aplikacyjnym interfejsem programowym dla tego systemu operacyjnego, dla

którego napisane były programy przykładowe. Programy przykładowe nie zostały gruntownie przetestowane. Firma IBM nie może zatem gwarantować lub sugerować niezawodności, użyteczności i funkcjonalności tych programów. Użytkownik może kopiować, modyfikować i rozpowszechniać te programy przykładowe w dowolnej formie bez uiszczania opłat, w celu rozbudowy, użytkowania, handlowym lub w celu rozpowszechniania aplikacji zgodnych z aplikacyjnym interfejsem IBM.

Każda kopia lub dowolna część programów przykładowych, albo też dowolna praca pochodna, musi zawierać poniższą informację o prawach autorskich:

Outside In<sup>®</sup> Viewer Technology, ©1992-2005 Stellant, Chicago, IL., Inc. Wszelkie prawa zastrzeżone.

Materiały licencjonowane IBM XSLT Processor - własność firmy IBM ©Copyright IBM Corp., 1999-2005. Wszelkie prawa zastrzeżone.





# Indeks

## A

- adapter dla wyszukiwania korporacyjnego 200
- administrator kolekcji
  - konfiguracja roli 178
  - opis 176
- administrator wyszukiwania korporacyjnego
  - konfiguracja roli 178
  - opis 176
  - zmiana hasła na pojedynczym serwerze 6
  - zmiana hasła na wielu serwerach 7
- adnotatory 99
- agenty użytkownika 57
- aktywne serwisy WWW, monitorowanie 219, 221
- alerty
  - czasy odpowiedzi wyszukiwania 236
  - indeksowane dokumenty 236
  - konfiguracja serwera SMTP 239
  - limity indeksu 213
  - na poziomie kolekcji 213, 236
  - na poziomie systemu 237
  - opcje poczty elektronicznej 236, 237
  - opis 235
  - otrzymywanie wiadomości e-mail 240
  - przeszukane dokumenty 236
  - wolna pamięć na serwerach 237
- alerty o wolnej pamięci 237
- analiza globalna
  - analiza tekstu zakotwiczenia 175, 181
  - wykrywanie duplikatów dokumentów 175, 181
- Analiza składni, opis strony 1
- analiza tekstu
  - mechanizmy analizy tekstu 101, 102
  - odzworowywanie elementów XML 103
  - wspólne struktury analizy 104, 105
- analiza tekstu zakotwiczenia
  - analiza globalna 181
  - indeksowanie dokumentów 182
  - opis 175
  - zabezpieczenia kolekcji 181
- analizatory składni
  - analiza terminów złożonych 106
  - biały znak 108
  - monitorowanie 227
  - opis 87
  - podział na ciągi n-znakowe 108
  - rodzime wyszukiwanie XML 107
  - status systemu 227
  - typy analizowanych dokumentów 109
  - typy dokumentów dla sesji Stellent 112, 114
  - typy dokumentów dla usług analizatora składni 110, 111
  - uruchamianie 227
  - wątki 106
  - zatrzymywanie 227
  - znaki nowego wiersza 108
- aplikacja ESSearchServer
  - wyłączanie zabezpieczenia 179

- aplikacje korporacyjne
  - aplikacja ESSearchServer 179
  - ESAdmin, aplikacja 179
  - ESPACTServer.ear, plik 74
  - ESSearchAdapter.ear, plik 200
  - ESSearchApplication, aplikacja 179
- aplikacje wyszukujące
  - identyfikatory aplikacji 183
  - niestandardowe 155
  - powiązanie z kolekcjami 156
  - powiązanie ze źródłami zewnętrznymi 172
  - przykład 155, 156
  - zabezpieczenia na poziomie kolekcji 183

## B

- budowanie indeksów
  - równoległe 120
  - status systemu 229
  - uruchamianie 228
  - współbieżne 120
  - zatrzymywanie 228, 229

## C

- Centrum wyszukiwania dla produktu WebSphere Portal
  - adapter dla wyszukiwania korporacyjnego 195, 200
  - opis 195
  - portlet rejestrujący 195, 202
- Centrum wyszukiwania portalu WebSphere Portal
  - adapter dla wyszukiwania korporacyjnego 195, 200
  - opis 195
  - portlet rejestrujący 195, 202
- chiński
  - analizowanie znaków nowego wiersza 108
  - podział na ciągi n-znakowe 108
- config.properties, plik edycja 166
- opisy właściwości 158
- cookie, przeszukiwanie sieci WWW
  - format 69
  - konfigurowanie 70
  - opis 69
- cookies.ini, plik
  - format 69
  - konfigurowanie 70
  - opis 69
- crawl.rules, plik 70
- czasy odpowiedzi wyszukiwania
  - alerty o 236
  - monitorowanie 229
- często zadawane zapytania, monitorowanie 229

## D

- Default, aplikacja wyszukująca 166
- dokumentacja 277
- dokumenty HTML, wyszukiwanie 98
- dokumenty usunięte
  - konfiguracja pliku dziennika 232
  - opis 231
  - raporty na temat 233
- dokumenty XML
  - przeszukiwanie 96
  - rodzime wyszukiwanie XML 107
- domeny Lotus Domino 190, 191
- domeny Windows 188
- drzewa kategorii
  - migrowanie z portalu WebSphere Portal 203, 205
  - opis 92
- dyrektywy blokujące indeksowanie
  - konfigurowanie 73
  - opis 72
- dyrektywy blokujące wykorzystywanie odsyłaczy
  - konfigurowanie 73
  - opis 72
- Dziennik, opis strony 1

## E

- edycja
  - aplikacje nasłuchiwanie danych 77
  - kolekcje 16
  - przestrzeń przeszukiwania 23
  - właściwości aplikacji wyszukującej 158, 166
  - właściwości przeszukiwacza 22
- elementy XML
  - odzworowywanie na pola wyszukiwania 96
  - odzworowywanie na wspólne struktury analizy 103
  - przeszukiwanie 95, 103
- es.cfg, plik 6, 7
- ES\_INSTALL\_ROOT, opis 6, 7
- ES\_NODE\_ROOT, opis 6, 7
- es\_special\_field.default\_field 152
- es\_special\_field.regular\_text, pole 152
- ESAdmin, aplikacja
  - logowanie 5
  - wyłączanie zabezpieczenia 179
- esbackup.bat, skrypt 246, 247
- esbackup.sh, skrypt 246, 247
- eschangepw, skrypt 6, 7
- escrm.sh, skrypt 37
- escrm.vbs, skrypt 39
- escrmnote.sh, skrypt 46
- escrvbr.sh, skrypt 26
- escrvbr.vbs, skrypt 27
- ESPACTServer.ear, plik 74
- esrestore.bat, skrypt 246, 247
- esrestore.sh, skrypt 246, 247

ESSearchAdapter.ear, plik 200  
ESSearchAdapterPortlet.war, plik 202  
ESSearchApplication, aplikacja  
  config.properties, plik 158, 166  
  restartowanie 166, 168  
  włączanie zabezpieczeń 168  
  wyłączanie zabezpieczenia 179  
ESSearchPortlet.war, plik 198

## F

firewalle, przeszukiwanie dokumentów  
  serwera Exchange 42  
followindex.rules, plik  
  konfigurowanie 73  
  opis 72  
fragmenty XML, rodzime wyszukiwanie  
  XML 107

## G

global.rules, plik 70  
globalna przestrzeń przeszukiwania sieci  
  WWW 70  
głębokość ścieżki URL 58

## H

hasło administratora  
  zmiana na pojedynczym serwerze 6  
  zmiana na wielu serwerach 7  
hasło, administrator wyszukiwania  
  korporacyjnego 6, 7  
historia czasu odpowiedzi,  
  monitorowanie 229

## I

identyfikator kolekcji, reguły składni 14  
identyfikatory aplikacji 183  
identyfikatory sesji, wyszukiwanie  
  korporacyjne 249  
identyfikatory URI  
  formaty w wyszukiwaniu  
  korporacyjnym 80  
  reguły kategorii 89, 93  
  szybkie odsyłacze 141  
  usuwanie z indeksu 130  
  wpływanie na wyniki statyczne 146, 147  
  wyświetlanie szczegółów o 216  
  zasiegi 126, 127  
  zwinienie w wynikach wyszukiwania 128,  
  129  
Indeks, opis strony 1  
indeksy  
  alerty o 236  
  formaty URI 80  
  monitorowanie 228, 229  
  opis 117  
  planowanie 118  
  równoległe operacje budowania 120  
  tekst zakotwiczenia 182  
  usuwanie identyfikatorów URI 121, 130  
  usuwanie z kolejki 229  
  włączanie harmonogramu 119, 228

indeksy (*kontynuacja*)  
  współbieżne operacje budowania 120  
  wyłączanie harmonogramu 119, 228  
  zasiegi 121, 126  
  zmiana harmonogramu 119  
  znaki zastępcze 121, 124, 126  
  zwinione identyfikatory URI 121, 128,  
  129  
interfejsy API  
  wyszukiwanie i indeksowanie 155  
IOCP, konfiguracja przeszukiwacza 51

## J

japoński  
  analizowanie znaków nowego  
  wiersza 108  
  podział na ciągi n-znakowe 108  
JavaScript, obsługa przez przeszukiwacze sieci  
  WWW 58  
JDBC, źródła zewnętrzne  
  edycja 171  
  konfigurowanie 171  
  sterowniki JDBC 171  
  usuwanie 171

## K

kategorie  
  drzewa kategorii 92  
  formaty URI 80  
  migrowanie z portalu WebSphere  
  Portal 203, 205  
  oparte na modelach 91, 93  
  oparte na regułach 89, 93  
  opis 88  
  przeszukiwanie 88  
  tworzenie 93  
  typ klasyfikacji 93  
  zagnieżdżone podkategorie 92  
kategorie oparte na modelach  
  opis 91  
  wybór typu klasyfikacji 93  
kategorie oparte na regułach  
  opis 89  
  tworzenie 93  
  wybór typu klasyfikacji 93  
klasy kryteriów  
  konfiguracja 150, 151  
  odwzorowywanie pól na 150  
  opis 148  
  wartości domyślne 152  
  wykrywanie duplikatów  
  dokumentów 148  
zapytania o niskiej znajdawalności 148,  
  152  
zapytania o wysokiej  
  znajdawalności 148, 152  
kody powrotu HTTP  
  odebrane przez przeszukiwacze sieci  
  WWW 222  
  raport o przeszukiwaczu sieci WWW 222  
kody powrotu, wyszukiwanie  
  korporacyjne 249  
kolejka indeksów 229

kolekcje  
  analizowanie 87  
  domyślne ustawienia migracji 207  
  edycja 16  
  MigrationWizard.log, plik 205  
  migrowanie z portalu WebSphere  
  Portal 203, 205  
  monitorowanie 215  
  pomijanie kontroli dostępu na poziomie  
  dokumentu 192  
  powiązanie z aplikacjami  
  wyszukującymi 156  
  przeszukiwanie 121  
  robocze 14  
  serwery wyszukiwania 133  
  sposoby tworzenia 13  
  status systemu 215  
  stowarzyszenie 13  
  szacowanie wielkości 14  
  szacowanie zasobów 213  
  tworzenie przy użyciu Kreator  
  kolekcji 14  
  tworzenie przy użyciu widoku  
  Kolekcje 14  
  usuwanie 17  
  zabezpieczenia 181  
  zabezpieczenia duplikatów  
  dokumentów 181  
  zabezpieczenia przy użyciu identyfikatora  
  aplikacji 183  
  zabezpieczenia tekstu zakotwiczenia 181  
kolekcje stowarzyszone 13  
Kolekcje, widok  
  opis 1  
  tworzenie kolekcji 14  
komenda esadmin 249  
komendy, wyszukiwanie korporacyjne 249  
komunikaty o błędach  
  konfiguracja serwera SMTP 239  
  otrzymywanie wiadomości e-mail 238,  
  240  
  wyświetlanie plików dziennika 242  
konektor Java dla programu DB2 Content  
  Manager 37, 39  
konfiguracja opcji Directory Assistance 55  
konfiguracja serwera SMTP 239  
konfiguracja serwera zaufanego 191  
konfiguracja użytkownika programu Domino,  
  przeszukiwacze QuickPlace 54  
konsola administracyjna  
  interfejs 1  
  logowanie 5  
  podsumowanie zadań 1  
kopia zapasowa systemu 245, 246  
koreański  
  analiza terminów złożonych 106  
  podział na ciągi n-znakowe 108  
Kreator kolekcji 14  
Kreator migracji  
  domyślne ustawienia kolekcji 207  
  domyślne ustawienia  
  przeszukiwaczy 207  
kolekcje 205  
  opis 203  
  plik dziennika 208  
  systematyki oparte na modelach 203  
  systematyki oparte na regułach 205

Kreator migracji (*kontynuacja*)  
uruchamianie 203, 205

## L

LDAP, źródła zewnętrzne  
edycja 171  
konfigurowanie 171  
usuwanie 171  
logowanie się do konsoli administracyjnej 5

## M

maksymalny odstęp czasu przed ponownym przeszukiwaniem 63  
Mechanizm wyszukiwania środowiska Portal  
opis 195  
przekazywanie tokenów  
zabezpieczeń 197  
zabezpieczenia na poziomie dokumentu 197  
mechanizmy analizy tekstu  
dodawanie do systemu 101  
odzworowanie wyników analizy 104, 105  
odzworowywanie elementów XML 103  
opis 99  
powiązanie z kolekcjami 102  
migrowanie  
kolekcje 205  
systematyka oparta na modelach 203  
systematyka oparta na regułach 205  
minimalny odstęp czasu przed ponownym przeszukiwaniem 63  
moduł bos.iocp.rte 51  
moduł portu zakończenia we/wy, konfiguracja przeszukiwacza 51  
monitor  
konfiguracja roli 178  
opis 176  
Monitor, opis widoku 1  
monitorowanie  
Aktywne serwisy przeszukiwacza sieci WWW 221  
analizatory składni 227  
często zadawane zapytania 229  
dokumenty usunięte 233  
historia czasu odpowiedzi 229  
kolekcje 215  
nasłuchiwanie danych 231  
ostatnie zapytania 229  
pliki dzienników 242  
przeszukiwacze 217  
przeszukiwacze sieci WWW 219  
serwery wyszukiwania 229  
Szczegółowe informacje o wątkach przeszukiwacza sieci WWW 220  
szczegóły dotyczące URI 216  
Szybkość przeszukiwania przeszukiwacza sieci WWW 221  
wyszukiwanie korporacyjne 213

## N

nasłuchiwanie danych  
konfigurowanie 77  
monitorowanie 231

nasłuchiwanie danych (*kontynuacja*)  
restartowanie 77, 231  
niestandardowa analiza tekstu  
mechanizmy analizy tekstu 101, 102  
odzworowywanie elementów XML 103  
odzworowywanie wspólnych struktur analizy 104, 105  
odzworowywanie wyników analizy na indeks 104  
odzworowywanie wyników analizy na table JDBC 105  
opis 99  
numer portu, wyszukiwanie korporacyjne 11

## O

obsługa lingwistyczna  
niestandardowa analiza tekstu 99  
rodzime wyszukiwanie XML 107  
słowniki kryteriów 144  
słowniki słów ignorowanych 137  
słowniki synonimów 134  
wyszukiwanie semantyczne 99, 107  
ocena dynamiczna 143  
ocena statyczna  
opis 144  
w migrowanych kolekcjach 205  
włączanie dla kolekcji 14  
ocena wyników wyszukiwania  
dynamiczne 143  
klasy kryteriów 148, 150, 151, 152  
opis 143  
słowniki kryteriów 146  
statyczna 144  
wartościowanie danych tekstowych 143  
wzorce URI 146, 147  
odstępy czasu przed ponownym przeszukiwaniem przy użyciu przeszukiwaczy sieci WWW 63  
odświeżanie indeksów 117  
odświeżanie indeksu  
opis 117  
planowanie 118, 119  
odtworzenie  
pliki systemowe na nowych serwerach 247  
z kopii zapasowej 246  
odtworzenie systemu 245, 246, 247  
odtworzenie wyszukiwania korporacyjnego 245  
odzworowanie  
elementy XML, wspólne struktury analizy 103  
pola wyszukiwania HTML 98  
pola wyszukiwania XML 96  
pól na klasy kryteriów 150  
wspólne struktury analizy na indeks 104  
wspólne struktury analizy na table JDBC 105  
ograniczanie przestrzeni przeszukiwania sieci WWW 58  
opcje treści dokumentu 121  
opcje wyszukiwania  
dla wyników wyszukiwania 121  
treść dokumentu 121  
wyszukiwanie dla pól 121  
wyszukiwanie parametryczne 121  
opcje wyszukiwania (*kontynuacja*)  
wyszukiwanie pełnotekstowe 121  
operator  
konfiguracja roli 178  
opis 176  
ostatnie zapytania, monitorowanie 229  
ostatnio przeszukane adresy URL, monitorowanie 219

## P

pamięć podręczna wyszukiwania  
konfigurowanie 134  
opis 134  
parserTypes.cfg, plik 110  
PDF, dokumentacja 277  
planowanie  
budowanie indeksów 118, 119  
przeszukiwacze 19  
plik CCLServer\_date.log 11  
pliki dzienników  
dla śledzenia dokumentów 232  
filtrowanie 242  
konfiguracja serwera SMTP 239  
Kreator migracji 208  
maksymalna wielkość 238  
monitorowanie 242  
opcje poczty elektronicznej 240  
opis 235  
położenie domyślne 235  
poziomy istotności 238  
rotacja 238  
wyświetlanie 242  
pliki EAR  
aplikacja ESSearchServer 179  
ESAdmin, aplikacja 179  
ESPACServer.ear 74  
ESSearchAdapter.ear 200  
ESSearchApplication, aplikacja 179  
początkowe adresy URL dla przeszukiwaczy sieci WWW 58, 64  
podstawowe uwierzytelnianie HTTP 66  
podsumowania  
dostosowywanie 139, 140  
dynamiczne 139  
podsumowania dokumentów  
dostosowywanie 139  
edycja właściwości dla 140  
podsumowanie zadań, konsola administracyjna 1  
podsumowywanie dynamiczne 139  
podział na ciągi n-znakowe 108  
pola wyszukiwania HTML  
odzworowywanie elementów 98  
opis 98  
tworzenie 98  
pola wyszukiwania XML  
odzworowywanie elementów 95, 96, 103  
opis 95, 103  
tworzenie 96  
pola, odzworowywanie na klasy kryteriów 150  
ponowne przeszukiwanie adresów URL w najbliższym możliwym terminie 64  
Portal WebSphere Portal  
domyślne ustawienia migracji 207

- Portal WebSphere Portal *(kontynuacja)*
  - integracja z wyszukiwaniem korporacyjnym 195
  - migracja drzewa kategorii 203
  - migracja kolekcji 203
  - migracja systematyki 203
  - portlet Wyszukiwanie, wdrażanie 198
  - portlet zarządzania systematykami 203
  - systematyki oparte na modelach 203
  - Wyszukiwanie i przeglądanie, portlet 199
- portlet Wyszukiwanie i przeglądanie, konfiguracja 199
- portlet Wyszukiwanie, wdrażanie 198
- portlet zarządzania systematykami 203
- portlety
  - ESSearchAdapterPortlet.war, plik obsługiwane wersje portalu WebSphere Portal 195
  - opis 195
  - portlet zarządzania systematykami 203
  - rejestracja 202
  - Wyszukiwanie 198
  - wyszukiwanie i przeglądanie 199
  - wyszukiwanie korporacyjne 195
- powiadomienia pocztą elektroniczną dotycząca komunikatów 240
- dotyczące alertów 240
- konfiguracja serwera SMTP 239
- prawa dostępu
  - obsługa mechanizmu wyszukiwania środowiska Portal 197
  - opis 178
  - profile użytkowników 186
  - sprawdzanie poprawności bieżących 185
  - wyłączanie dla kolekcji 192
  - wymagania dla produktu Lotus Domino 190, 191
  - wymagania dla systemów plików Windows 188
  - zabezpieczenia na poziomie dokumentu 184, 197
- profile użytkowników
  - konfiguracja 187
  - opis 186
- program WebSphere II Event Publisher Edition, konfiguracja przeszukiwacza DB2 30
- program WebSphere MQ, konfiguracja przeszukiwacza DB2 32
- program WebSphere MQ, konfiguracja serwera przeszukiwacza 34, 35
- protokół DIIOP, konfiguracja przeszukiwacza 50
- protokół NRPC, konfiguracja przeszukiwacza 46, 48
- przestrzeń przeszukiwania
  - alerty o 236
  - edycja 23
  - konfiguracja przeszukiwacza sieci WWW 58
- przeszukiwacze
  - aplikacje nasłuchiwania danych Content Edition 24, 25
  - DB2 28
  - DB2 Content Manager 36
  - Domino Document Manager 40
  - domyślne ustawienia migracji 207
  - przeszukiwacze *(kontynuacja)*
    - edycja przestrzeni przeszukiwania 23
    - edycja właściwości przeszukiwacza Exchange Server 41, 42
    - formaty URI 80
    - łączenie typów przeszukiwaczy 19
    - monitorowanie 217
    - NNTP 43
    - Notes 43, 45
    - obsługa zewnętrznych 77
    - planowanie 19
    - Portal WebSphere Portal 73, 74, 75
    - przegląd konfiguracji 19
    - QuickPlace 52
    - sieć WWW 56
    - status systemu 217
    - system plików UNIX 55
    - system plików Windows 76
    - tworzenie 21
    - usuwanie 23
    - wartości początkowe 21
    - wartości podstawowe 19
    - włączanie zabezpieczeń na poziomie dokumentu 19
    - wtyczki 78
    - zabezpieczenia na poziomie dokumentu 184
  - przeszukiwacze Content Edition
    - formaty URI 80
    - konfiguracja 24
    - konfigurowanie w systemie operacyjnym AIX 26
    - konfigurowanie w systemie operacyjnym Linux 26
    - konfigurowanie w systemie Windows 27
    - konfigurowanie w Środowisku Operacyjnym Solaris 26
    - tryb bezpośredni 25
    - tryb serwera 25
  - przeszukiwacze DB2
    - formaty URI 80
    - instalacja programu WebSphere MQ 34
    - instalacja programu WebSphere MQ w systemie Windows 35
    - konfiguracja 28
    - konfiguracja programu WebSphere II Event Publisher Edition 30
    - konfiguracja programu WebSphere MQ 32
    - konfiguracja publikowania zdarzeń 34, 35
    - publikowanie zdarzeń 28
  - przeszukiwacze DB2 Content Manager
    - formaty URI 80
    - konfiguracja 36
    - konfigurowanie w systemie operacyjnym AIX 37
    - konfigurowanie w systemie operacyjnym Linux 37
    - konfigurowanie w systemie Windows 39
    - konfigurowanie w Środowisku Operacyjnym Solaris 37
  - przeszukiwacze Domino Document Manager
    - formaty URI 80
    - konfiguracja 40
    - konfiguracja IOCP 51
    - konfiguracja protokołu DIIOP 50
  - przeszukiwacze Domino Document Manager *(kontynuacja)*
    - konfigurowanie w systemie operacyjnym AIX 46
    - konfigurowanie w systemie operacyjnym Linux 46
    - konfigurowanie w systemie Windows 48
    - konfigurowanie w Środowisku Operacyjnym Solaris 46
    - NRPC, protokół 46, 48
  - przeszukiwacze NNTP, konfigurowanie 43
  - przeszukiwacze Notes
    - formaty URI 80
    - konfiguracja 43
    - konfiguracja IOCP 51
    - konfiguracja protokołu DIIOP 50
    - konfiguracja zabezpieczeń na poziomie dokumentu 190
    - konfigurowanie w systemie operacyjnym AIX 46
    - konfigurowanie w systemie operacyjnym Linux 46
    - konfigurowanie w systemie Windows 48
    - konfigurowanie w Środowisku Operacyjnym Solaris 46
    - NRPC, protokół 46, 48
    - reguły odwzorowywania pól 45
    - sprawdzanie poprawności bieżących referencji 191
    - wskazówki na temat używania 45
    - zaufany serwer Lotus Domino 191
  - Przeszukiwacze QuickPlace
    - formaty URI 80
    - konfiguracja 52
    - konfiguracja IOCP 51
    - konfiguracja opcji Directory Assistance 55
    - konfiguracja protokołu DIIOP 50
    - konfiguracja użytkownika programu Domino 54
    - konfigurowanie w systemie operacyjnym AIX 46
    - konfigurowanie w systemie operacyjnym Linux 46
    - konfigurowanie w systemie Windows 48
    - konfigurowanie w Środowisku Operacyjnym Solaris 46
    - NRPC, protokół 46, 48
    - zabezpieczenia użytkownika lokalnego 54
  - Przeszukiwacze serwera Exchange Server
    - formaty URI 80
    - konfiguracja 41
    - zabezpieczone dokumenty 42
  - przeszukiwacze sieci WWW
    - agenty użytkownika 57
    - aktywne serwisy 219, 221
    - cookie, format 69
    - cookie, konfiguracja informacji 70
    - dyrektywy blokujące indeksowanie 72, 73
    - dyrektywy blokujące wykorzystywanie odsyłaczy 72, 73
    - followindex.rules, plik 72, 73
    - globalna przestrzeń przeszukiwania 70
    - historia przeszukiwacza 219
    - historia serwisu 219

- przeszukiwacze sieci WWW (*kontynuacja*)
  - informacje cookie 69
  - JavaScript, obsługa 58
  - konfiguracja 56
  - monitorowanie 219
  - odstępy czasu przed ponownym przeszukiwaniem 63
  - ograniczanie przestrzeni przeszukiwania 58
  - ostatnio przeszukane adresy URL 219
  - początkowe adresy URL 58, 64
  - przeszukiwanie adresów URL w najbliższym możliwym terminie 64
  - reguły przeszukiwania 58
  - serwery proxy 68
  - serwisy WWW zabezpieczone hasłem 66, 67
  - status adresu URL 219
  - status systemu 219
  - strony błędów programowych 64
  - szczegóły wątków 219, 220
  - szybkość przeszukiwania 221
  - tworzenie raportów o 222
- przeszukiwacze systemu plików UNIX
  - formaty URI 80
  - konfiguracja 55
- przeszukiwacze systemu plików Windows
  - formaty URI 80
  - konfiguracja 76
  - konfiguracja zabezpieczeń na poziomie dokumentu 188
- przeszukiwacze WebSphere Portal
  - formaty URI 80
  - konfiguracja 73
  - kopiowanie adresów URL serwisów 75
  - wdrażanie aplikacji korporacyjnej 74
  - wdrażanie pliku ESPACServer.ear 74
- przeszukiwanie
  - dokumenty HTML 98
  - dokumenty XML 95, 96, 103
  - kategorie 88
  - kolekcje 121
  - szybkie odsyłacze 141
  - źródła zewnętrzne 121
- przeszukiwanie adresów URL w najbliższym możliwym terminie 64
- Przeszukiwanie, opis strony 1
- przetwarzanie tekstu
  - adnotatory 99
  - mechanizmy analizy tekstu 99
  - wspólne struktury analizy 99
- przykładowa aplikacja wyszukująca
  - config.properties, plik 158, 166
  - funkcje wyszukiwania 155, 156
  - instalacja domyślna 166
  - opis 156
  - uruchamianie 167
  - uzyskiwanie dostępu 167
  - WebSphere, zabezpieczenia globalne 168
  - włączanie zabezpieczeń 168
  - wyłączanie zabezpieczenia 179
- publikowanie zdarzeń
  - konfiguracja przeszukiwacza DB2 30, 32
  - konfigurowanie w systemie operacyjnym AIX 34
  - konfigurowanie w systemie operacyjnym Linux 34

- publikowanie zdarzeń (*kontynuacja*)
  - konfigurowanie w systemie Windows 35
  - konfigurowanie w Środowisku Operacyjnym Solaris 34
  - opis 28

## R

- raporty z historii przeszukiwacza
  - opis 219
  - raport o kodach powrotu HTTP 222
  - raport o serwisie 222
  - tworzenie 222
- raporty z historii serwisu
  - opis 219
  - tworzenie 222
- reguły domenowe dla przeszukiwaczy sieci WWW 58
- reguły dotyczące adresów dla przeszukiwaczy sieci WWW 58
- reguły dotyczące adresów IP dla przeszukiwaczy sieci WWW 58
- reguły dotyczące przedrostków dla przeszukiwaczy sieci WWW 58
- reguły kategorii
  - konfigurowanie 93
  - treść dokumentu 89, 93
  - wzorce URI 89, 93
- reorganizacja indeksu
  - opis 117
  - planowanie 118, 119
- reorganizowanie indeksów 117
- Robots Exclusion, protokół 57
- robots.txt, plik 57
- rodzime wyszukiwanie XML 107
- role administracyjne
  - administrator kolekcji 176, 178
  - administrator wyszukiwania korporacyjnego 176, 178
  - konfigurowanie 178
  - monitor 176, 178
  - operator 176, 178
  - opis 176
- rozszerzenia plików
  - obsługiwane przez analizatory składni kolekcji 110, 111
  - obsługiwane przez sesje Stellent 112, 114
  - wykluczanie z przestrzeni przeszukiwania w sieci WWW 58
- równoległe operacje budowania indeksu 120
- runtime-generic.properties, plik 140

## S

- serwery analizatorów składni
  - konfiguracja wątku 106
  - uruchamianie 209
  - zatrzymywanie 211
- serwery indeksowania
  - uruchamianie 209
  - zatrzymywanie 211
- serwery proxy 68
- serwery proxy HTTP 68
- serwery przeszukiwaczy
  - uruchamianie 209, 217
  - zatrzymywanie 211, 217

- serwery wyszukiwania
  - często zadawane zapytania 229
  - historia czasu odpowiedzi 229
  - monitorowanie 229
  - opis 133
  - ostatnie zapytania 229
  - pamięć podręczna wyszukiwania 134
  - powiązanie słowników kryteriów 146
  - powiązanie słowników słów ignorowanych 138
  - powiązanie słowników synonimów 136
  - słowniki kryteriów 144
  - słowniki słów ignorowanych 137
  - słowniki synonimów 134, 136
  - status systemu 229
  - uruchamianie 209, 229
  - zatrzymywanie 211, 229
- serwisy WWW zabezpieczone hasłem 66
  - podstawowe uwierzytelnianie HTTP 66
  - uwierzytelnianie oparte na formularzu 67
- SIAPI (interfejs API wyszukiwania i indeksowania) 155
- składnia zapytania XML, rodzima 107
- skrypt escrdb2.sh 34
- skrypt escrdb2.vbs 35
- skrypt escrnote.vbs 48
- skrypty
  - esbackup.bat 246, 247
  - esbackup.sh 246, 247
  - escrm.sh 37
  - escrm.vbs 39
  - escrdb2.sh 34
  - escrdb2.vbs 35
  - escrnote.sh 46
  - escrnote.vbs 48
  - escrvbr.sh 26
  - escrvbr.vbs 27
  - esrestore.bat 246, 247
  - esrestore.sh 246, 247
  - startccl 246
- skrypty odtwarzania
  - opis 245
  - uruchamianie 246, 247
- skrypty tworzenia kopii zapasowych
  - opis 245
  - uruchamianie 246, 247
- słowa kluczowe w szybkich odsyłaczach 141
- słowniki kryteriów
  - dodawanie do systemu 146
  - opis 144
  - powiązanie z kolekcją 146
- słowniki słów ignorowanych
  - dodawanie do systemu 138
  - opis 137
  - powiązanie z kolekcją 138
- słowniki synonimów
  - dodawanie do systemu 136
  - opis 134
  - powiązanie z kolekcją 136
- sprawdzanie poprawności bieżących referencji 185, 188, 190, 191
- sprawdzanie poprawności zapytań 185
- sprawdzanie zasobów systemu 214
- startccl, skrypt 246
- status systemu
  - analizatory składni 227
  - budowanie indeksów 229

- status systemu (*kontynuacja*)
  - kolekcje 215
  - przeszukiwacze 217
  - przeszukiwacze sieci WWW 219
  - serwery wyszukiwania 229
- Stellent, sesje
  - domyślne typy dokumentów 114
  - powiązanie typów dokumentów 112
  - typy analizowanych dokumentów 109
- stellent.properties, plik 112
- stellenttypes.cfg, plik 112
- strony błędów programowych, przeszukiwacze sieci WWW 64
- synonyms.xml, plik 203
- system operacyjny AIX
  - konfiguracja przeszukiwacza Content Edition 26
  - konfiguracja przeszukiwacza DB2 Content Manager 37
  - konfiguracja przeszukiwacza Domino Document Manager 46
  - konfiguracja przeszukiwacza Notes 46
  - konfiguracja przeszukiwacza QuickPlace 46
  - konfiguracja publikowania zdarzeń 34
- system operacyjny Linux
  - konfiguracja przeszukiwacza Content Edition 26
  - konfiguracja przeszukiwacza DB2 Content Manager 37
  - konfiguracja przeszukiwacza Domino Document Manager 46
  - konfiguracja przeszukiwacza Notes 46
  - konfiguracja przeszukiwacza QuickPlace 46
  - konfiguracja publikowania zdarzeń 34
- Środowisko Operacyjne Solaris
  - konfiguracja publikowania zdarzeń 34
- System, opis widoku 1
- systematyka oparta na modelach, migrowanie z portalu WebSphere Portal 203
- systematyka oparta na regułach, migrowanie z portalu WebSphere Portal 203, 205
- systematyki, migrowanie z portalu WebSphere Portal 203, 205
- szacowanie zasobów systemowych 213
- szczegóły dotyczące URI
  - dokumenty usunięte 233
  - monitorowanie 216
- szczegóły wątków, monitorowanie 219
- szybkie odsyłacze
  - formaty URI 80
  - opis 141
  - przeszukiwanie 141
  - tworzenie 141
- szybkość przeszukiwania, monitorowanie 221

## Ś

- śledzenie dokumentu
  - konfiguracja pliku dziennika 232
  - opis 231
  - raporty 233
  - włączanie 232
  - wyłączanie 232

- Środowisko Operacyjne Solaris
  - konfiguracja przeszukiwacza Content Edition 26
  - konfiguracja przeszukiwacza DB2 Content Manager 37
  - konfiguracja przeszukiwacza Domino Document Manager 46
  - konfiguracja przeszukiwacza Notes 46
  - konfiguracja przeszukiwacza QuickPlace 46

## T

- terminy złożone, analizowanie 106
- titles.xml, plik 203
- token zabezpieczeń
  - konfiguracja przeszukiwacza 184
  - przetwarzanie mechanizmu wyszukiwania środowiska Portal 197
  - wyłączanie dla kolekcji 192
  - zabezpieczenia na poziomie dokumentu 184, 197
- treenodes.xml, plik 203
- tryb bezpośredni, repozytoria Content Edition 25
- tryb serwera, repozytoria Content Edition 25
  - tworzenie
    - kategorie oparte na regułach 93
    - kolekcje 14
    - pola wyszukiwania HTML 98
    - pola wyszukiwania XML 96
    - przeszukiwacze 21
    - raporty o przeszukiwaczach sieci WWW 222
    - szybkie odsyłacze 141
    - zasięgi 127
  - tworzenie kopii zapasowej wyszukiwania korporacyjnego 245, 246
  - typ klasyfikacji
    - oparte na modelach 91
    - oparte na regułach 89
    - wybór 14, 93
  - typy dokumentów
    - analizowanie 109
    - sesje Stellant 112, 114
    - usługi analizatora składni 110, 111
  - typy MIME, włączanie do przestrzeni przeszukiwania w sieci WWW 58
  - typy przeszukiwaczy
    - łączenie w kolekcji 19
    - wartości podstawowe 19
  - typy źródeł danych
    - bazy danych DB2 28
    - bazy danych Domino Document Manager 40
    - bazy danych Notes 43, 45
    - bazy danych QuickPlace 52
    - foldery publiczne serwera Exchange Server 41
    - grupy dyskusyjne NNTP 43
    - relacyjne bazy danych 28
    - repozytoria Content Edition 24, 25
    - serwisy WebSphere Portal 73
    - serwisy WWW 56
    - systemy plików UNIX 55
    - systemy plików Windows 76

- typy źródeł danych (*kontynuacja*)
  - typy elementów DB2 Content Manager 36

## U

- UIMA
  - dodawanie mechanizmów analizy tekstu do systemu 101
  - odzworowywanie elementów XML 103
  - odzworowywanie wyników analizy na indeks 104
  - odzworowywanie wyników analizy na tabele JDBC 105
  - opis 99
  - powiązanie z kolekcjami 102
  - wspólne struktury analizy 104, 105
- ułatwienia dostępu 279
- uruchamianie
  - analizatory składni 227
  - budowanie indeksów 228
  - Kreator migracji 203, 205
  - nasłuchiwanie danych 231
  - przykładowa aplikacja wyszukująca 167
  - serwery przeszukiwaczy 217
  - serwery wyszukiwania 229
  - serwery wyszukiwania korporacyjnego 209
- usuwanie
  - indeksy z kolejki 229
  - kolekcje 17
  - przeszukiwacze 23
- usuwanie identyfikatorów URI z indeksu 130
- uwierzytelnianie
  - opis 178
  - wyłączanie dla aplikacji korporacyjnych 179
- uwierzytelnianie oparte na formularzu 66, 67

## V

- vbr\_access\_services.jar, plik 26, 27

## W

- wartościowanie danych tekstowych 143
- ważność dokumentu
  - klasy kryteriów 148, 152
  - słowniki kryteriów 146
  - statyczna 144
  - w migrowanych kolekcjach 205
  - włączanie dla kolekcji 14
  - wzorce URI 146, 147
- wątki
  - analizator składni 106
  - przeszukiwacz sieci WWW 220
- wdrażanie
  - adapter dla wyszukiwania korporacyjnego 200
  - ESPACServer.ear, plik 74
  - ESSearchAdapter.ear, plik 200
  - ESSearchAdapterPortlet.war, plik 202
  - ESSearchPortlet.war, plik 198
  - portlet rejestrujący 202
  - portlet Wyszukiwanie 198
  - Wyszukiwanie i przeglądanie, portlet 199

- WebSphere II OmniFind Edition 279
    - identyfikatory sesji 249
    - integracja z portalem WebSphere Portal 195
    - kody powrotu 249
    - komendy 249
    - konfiguracja numeru portu 11
    - ułatwienia dostępu 279
    - zarządzanie tożsamościami 187
    - zmiana hasła na pojedynczym serwerze 6
    - zmiana hasła na wielu serwerach 7
  - WebSphere, zabezpieczenia globalne
    - właściwości aplikacji wyszukującej 166
    - wyłączanie 179
  - Windows, system operacyjny
    - konfiguracja przeszukiwacza 48
    - konfiguracja przeszukiwacza Content Edition 27
    - konfiguracja przeszukiwacza DB2 Content Manager 39
    - konfiguracja publikowania zdarzeń 35
  - właściwości przeszukiwacza
    - edycja 22
  - włączanie harmonogramów operacji na indeksach 119
  - WpsMigratorLog.log, plik 208
  - wspólne struktury analizy
    - odzworowywanie elementów XML 103
    - odzworowywanie na indeks 104
    - odzworowywanie na tabelę JDBC 105
    - opis 99
  - współbieżne operacje budowania indeksu 120
  - współczynniki kryteriów
    - dla klas kryteriów 150, 152
    - dla słowników kryteriów 144
    - dla wzorców URI 146, 147
    - konfiguracja klasy kryteriów 148, 151
  - wtyczki dla przeszukiwaczy 78
  - wtyczki przeszukiwaczy 78
  - wykrywanie duplikatów dokumentów
    - analiza globalna 181
    - konfiguracja klasy kryteriów 148
    - opis 175
    - włączanie zabezpieczeń 181
  - wyłączanie harmonogramów operacji na indeksach 119
  - wyniki wyszukiwania
    - dostosowywanie podsumowań 139, 140
    - grupowanie 128, 129
    - konfiguracja klasy kryteriów 148, 150, 151, 152
    - konfiguracja wzorca URI 146
    - ocena 147
    - ocena dynamiczna 143
    - ocena statyczna 144
    - opis 143
    - podsumowania 139, 140
    - podsumowywanie dynamiczne 139
    - rozwijanie znaków zastępczych 126
    - wartościowanie danych tekstowych 143
    - znaki zastępcze 124
    - zwijanie 128, 129
  - wyszukiwanie dla pól 121
  - wyszukiwanie i indeksowanie, interfejs API 155
  - wyszukiwanie korporacyjne
    - analizatory składni 87
    - formaty URI 80
    - identyfikatory sesji 249
    - integracja z portalem WebSphere Portal 195
    - kody powrotu 249
    - komendy 249
    - konfiguracja numeru portu 11
    - monitorowanie 213
    - odtworzenie plików systemowych 247
    - odtworzenie z kopii zapasowej 246
    - pliki dzienników 235
    - role administracyjne 176
    - serwery indeksowania 117
    - serwery przeszukiwaczy 19
    - serwery wyszukiwania 133
    - skrypty odtwarzania 245
    - tworzenie kopii zapasowych 245
    - uruchamianie serwerów 209
    - zabezpieczenia 175
    - zabezpieczenia na poziomie dokumentu 184
    - zabezpieczenia na poziomie kolekcji 181
    - zatrzymywanie serwerów 209, 211
  - wyszukiwanie parametryczne 121
  - wyszukiwanie pełnotekstowe 121
  - wyszukiwanie semantyczne 99, 103, 107
  - Wyszukiwanie, opis strony 1
  - wyświetlanie
    - pliki dzienników 242
    - szczegóły dotyczące URI 216
- ## X
- XPath, rodzime wyszukiwanie XML 107
- ## Z
- zabezpieczenia
    - analiza tekstu zakotwiczenia 181
    - dokumenty Lotus Domino 190, 191
    - domeny Windows 188
    - identyfikatory aplikacji wyszukującej 183
    - na poziomie dokumentu 184, 185, 187, 192
    - na poziomie kolekcji 181, 192
    - opis 175
    - porównywanie kontroli dostępu na poziomie dokumentu 192
    - prawa dostępu 178
    - profile użytkowników 186
    - przykładowa aplikacja wyszukująca 168
    - role administracyjne 178
    - uwierzytelnianie 178, 179
    - WebSphere, zabezpieczenia globalne 179
    - włączanie dla kolekcji 14, 175
    - wtyczki przeszukiwaczy 78
    - wykrywanie duplikatów dokumentów 181
    - wyłączanie dla aplikacji korporacyjnej 179
    - zarządzanie tożsamościami 187
  - zabezpieczenia na poziomie dokumentu dla dokumentów Lotus Domino 190
  - dla systemów plików Windows 188
  - dokumenty Lotus Domino 191
  - konfiguracja przeszukiwacza 19
  - kontrola dostępu w indeksie 184
  - obsługa mechanizmu wyszukiwania środowiska Portal 197
  - opis 175, 184
  - profile użytkowników 186
  - sprawdzanie poprawności bieżących referencji 185
  - sprawdzanie poprawności w czasie rzeczywistym 185
  - token zabezpieczeń 184
  - wtyczki przeszukiwaczy 78
  - zarządzanie tożsamościami 187
  - zabezpieczenia na poziomie kolekcji
    - analiza tekstu zakotwiczenia 181
    - identyfikatory aplikacji 183
    - opis 175, 181
    - włączanie 14
    - wykrywanie duplikatów dokumentów 181
  - zabezpieczenia użytkownika lokalnego, przeszukiwacze QuickPlace 54
  - Zabezpieczenia, opis widoku 1
  - zapytania o niskiej znajdawalności
    - domyślne współczynniki kryteriów 152
    - opis 148
  - zapytania o wysokiej znajdawalności
    - domyślne współczynniki kryteriów 152
    - opis 148
  - zarządzanie tożsamościami
    - konfiguracja 187
    - profile użytkowników 186
  - zasiegi
    - formaty URI 80
    - opis 126
    - przeszukiwanie 126
    - tworzenie 127
    - wzorce URI 126, 127
  - zasoby systemowe
    - sprawdzanie 214
    - szacowanie 213
  - zatrzymywanie
    - analizatory składni 227
    - budowanie indeksów 228, 229
    - serwery przeszukiwaczy 217
    - serwery wyszukiwania 229
    - serwery wyszukiwania korporacyjnego 209, 211
    - zaufane serwery Lotus Domino 191
  - zewnętrzne przeszukiwacze
    - aplikacje nasłuchiwanie danych 77
    - konfigurowanie 77
  - znajdowanie dokumentacji dotyczącej wyszukiwania korporacyjnego 277
  - znaki zastępcze
    - rozwinięcie indeksu 124, 126
    - rozwinięcie zapytania 124, 126
  - zwinięte identyfikatory URI
    - konfigurowanie 129
    - opis 128
  - zwinięte wyniki wyszukiwania
    - konfigurowanie 129
    - opis 128

## Ź

- źródła zewnętrzne
  - konfigurowanie 171
  - opis 171
  - powiązanie z aplikacjami
    - wyszukującymi 172
  - przeszukiwanie 121
  - zabezpieczenia przy użyciu identyfikatora aplikacji 183





**IBM**



**Java**<sup>™</sup>  
**COMPATIBLE**

SC85-0162-02

