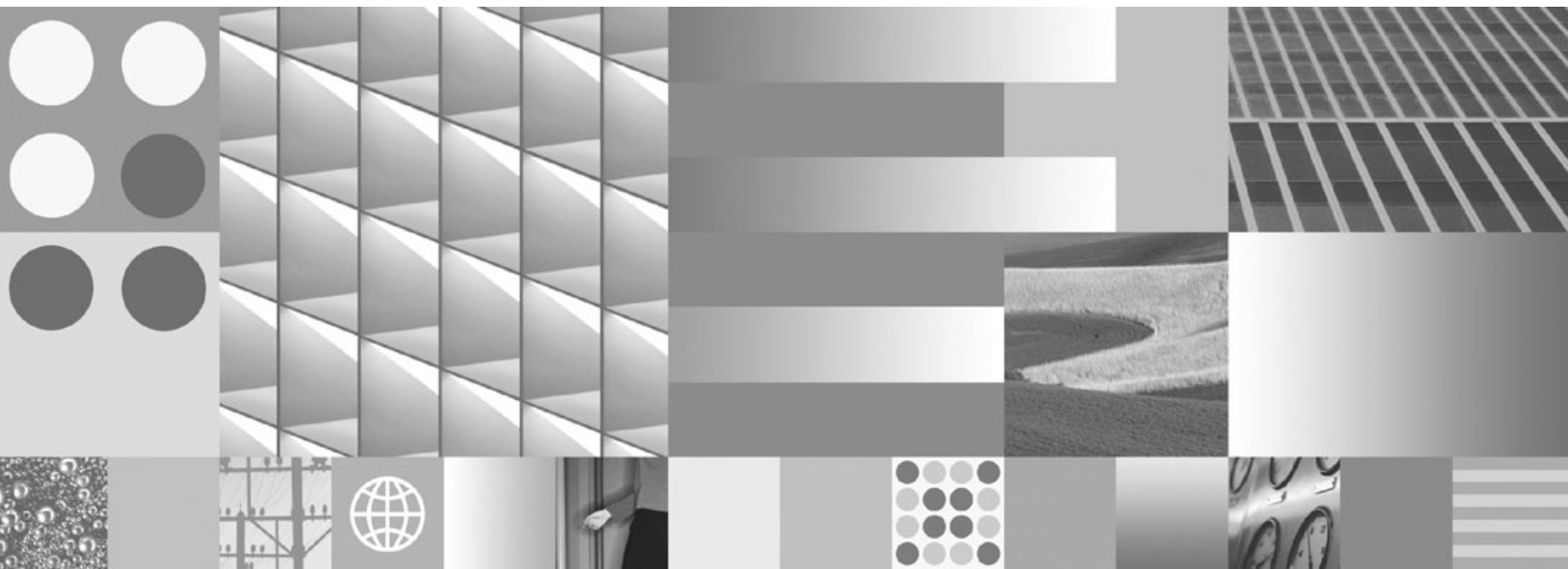


エンタープライズ・サーチの管理





エンタープライズ・サーチの管理

**お願い**

本書および本書で紹介する製品をご使用になる前に、463 ページの『特記事項および商標』に記載されている情報をお読みください。

本書は、IBM OmniFind Enterprise Edition のバージョン 8、リリース 5、モディフィケーション 0 (製品番号 5724-C74)、および新しい版で明記されていない限り、以降のすべてのリリースおよびモディフィケーションに適用されます。

IBM は、お客様が提供するいかなる情報も、お客様に対してなんら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。

IBM 発行のマニュアルに関する情報のページ

<http://www.ibm.com/jp/manuals/>

こちらから、日本語版および英語版のオンライン・ライブラリーをご利用いただけます。また、マニュアルに関するご意見やご感想を、上記ページよりお送りください。今後の参考にさせていただきます。

(URL は、変更になる場合があります)

お客様の環境によっては、資料中の円記号がバックスラッシュと表示されたり、バックスラッシュが円記号と表示されたりする場合があります。

原 典： SC18-9283-04  
IBM OmniFind Enterprise Edition  
Version 8.5  
Administering Enterprise Search

発 行： 日本アイ・ビー・エム株式会社

担 当： ナショナル・ランゲージ・サポート

# 目次

ibm.com および関連リソース . . . . .	vii
-----------------------------	-----

## エンタープライズ・サーチとは? . . . . . 1

エンタープライズ・サーチでサポートされているデータ・ソース・タイプ . . . . .	2
エンタープライズ・サーチ・コンポーネントの概要 . . . . .	3
エンタープライズ・サーチ・クローラー . . . . .	4
エンタープライズ・サーチ・パーサー . . . . .	5
エンタープライズ・サーチ索引 . . . . .	7
エンタープライズ・サーチ用検索サーバー . . . . .	9
エンタープライズ・サーチ管理コンソール . . . . .	10
エンタープライズ・サーチ・システムのモニター . . . . .	10
エンタープライズ・サーチ用ログ・ファイル . . . . .	11
エンタープライズ・サーチのカスタマイズ . . . . .	11
エンタープライズ・サーチ用サンプル検索アプリケーション . . . . .	12
エンタープライズ・サーチのデータ・フロー . . . . .	13

## エンタープライズ・サーチ・システムの管理 . . . . . 17

管理コンソールへのログイン . . . . .	21
シングル・サーバー構成のエンタープライズ・サーチ管理者パスワードの変更 . . . . .	22
マルチサーバー構成のエンタープライズ・サーチ管理者パスワードの変更 . . . . .	24
エンタープライズ・サーチに使用される TCP ポート番号 . . . . .	26
エンタープライズ・サーチ・システムのポート番号の変更 . . . . .	27
エンタープライズ・サーチ・サーバーのホスト名または IP アドレスの変更 . . . . .	29
二重 IP アドレスのサポートの構成 . . . . .	30
IPv6 プロトコルのサポートの使用可能化 . . . . .	31

## エンタープライズ・サーチ・コレクション 35

コレクション・ウィザードを使用したコレクションの作成 . . . . .	35
「コレクション」ビューを使用したコレクションの作成 . . . . .	37
コレクションの編集 . . . . .	38
コレクションの削除 . . . . .	40
コレクション ID の判別 . . . . .	41

## クローラーの管理 . . . . . 43

クローラーの作成 . . . . .	45
クローラーのプロパティの編集 . . . . .	46
クローラー・スペースの編集 . . . . .	47
クローラーの削除 . . . . .	47
クローラー・スケジュール . . . . .	48
Content Edition クローラー . . . . .	49

Content Edition リポジトリへの直接モード・アクセス . . . . .	50
WebSphere II Content Edition リポジトリへのサーバー・モード・アクセス . . . . .	52
DB2 クローラー . . . . .	55
UNIX 上での DB2 クローラー用のクローラー・サーバーの構成 . . . . .	57
Windows 上での DB2 クローラー用のクローラー・サーバーの構成 . . . . .	58
DB2 クローラーのための WebSphere Information Integrator Event Publisher Edition の構成 . . . . .	59
DB2 クローラーのための WebSphere MQ の構成 . . . . .	62
クラシック・データ・ソース・サーバーでの DB2 データベースのクローラー . . . . .	64
DB2 Content Manager クローラー . . . . .	65
UNIX での DB2 Content Manager クローラー用クローラー・サーバーの構成 . . . . .	66
Windows での DB2 Content Manager クローラー用クローラー・サーバーの構成 . . . . .	68
Lotus Domino Document Manager クローラー . . . . .	70
Exchange Server クローラー . . . . .	73
JDBC データベース・クローラー . . . . .	74
JDBC データベースの関係マップ . . . . .	76
複数の構造化された JDBC データベース表のクローラー . . . . .	78
NNTTP クローラー . . . . .	82
Lotus Notes クローラー . . . . .	83
Lotus Domino データベースのクローラーに関するヒント . . . . .	86
Lotus Domino ソースをクローラーするための UNIX でのクローラー・サーバーの構成 . . . . .	87
Lotus Domino ソースをクローラーするための Windows でのクローラー・サーバーの構成 . . . . .	89
DIHOP プロトコルを使用するサーバーの構成 . . . . .	91
Lotus Domino ソースをクローラーするための AIX での I/O 完了ポートの構成 . . . . .	92
QuickPlace クローラー . . . . .	94
シード・リスト・クローラー . . . . .	98
UNIX ファイル・システム・クローラー . . . . .	100
Web クローラー . . . . .	101
ユーザー・エージェントの構成 . . . . .	102
Web クローラーによるロボット排他プロトコルの使用方法 . . . . .	103
JavaScript のサポート . . . . .	105
Web クローラー・スペースを制限する規則 . . . . .	106
Web クローラーとの URL 接続のテスト . . . . .	111
Web クローラーにおける再クローリング間隔の設定 . . . . .	111
Web クローラーにおける URL アクセスのオプション . . . . .	112
Web クローラーがソフト・エラーのページを処理する方法 . . . . .	112

セキュア Web サイトのクロールのサポート . . . . .	114
プロキシ・サーバーによってサービスされてい る Web サイト . . . . .	116
Cookie の管理 . . . . .	117
グローバル Web クロール・スペースの構成 . . . . .	119
リンクをたどらない、索引付けしないディレクテ ィブ . . . . .	121
Web ページでのリンクをたどらないディレクテ ィブおよび索引付けをしないディレクティブのオ ーバーライド . . . . .	121
クロールされた文書の日付として Web クローラ ーが使用する日付の構成 . . . . .	122
Web Content Management クローラー . . . . .	123
WebSphere Portal クローラー . . . . .	126
クロールする URL を WebSphere Portal からコ ピーする . . . . .	127
Windows ファイル・システム・クローラー . . . . .	129
データ・リスナー・アプリケーションのサポートの 構成 . . . . .	131
カスタム・クローラー・プラグイン . . . . .	132
アーカイブ・ファイルのクロールのサポート . . . . .	134
エンタープライズ・サーチ索引における URI フォ ーマット . . . . .	135
<b>パーサーの管理 . . . . . 147</b>	
カテゴリに関する作業 . . . . .	148
ルール・ベースのカテゴリ . . . . .	148
カテゴリ・ツリー . . . . .	151
カテゴリ化タイプの選択 . . . . .	151
カテゴリの構成 . . . . .	152
XML 検索フィールドの作業 . . . . .	154
XML 検索フィールド . . . . .	154
XML エlementを検索フィールドにマップ . . . . .	155
HTML 検索フィールドの作業 . . . . .	158
HTML 検索フィールド . . . . .	158
検索フィールドへの HTML メタデータ・エレメ ントのマッピング . . . . .	159
カスタム・テキスト処理 . . . . .	160
システムへのテキスト分析エンジンの追加 . . . . .	162
テキスト分析エンジンとコレクションとの関連付 け . . . . .	163
XML エlementの共通分析構造へのマッピング . . . . .	164
共通分析構造の索引へのマッピング . . . . .	165
共通分析構造のリレーショナル・データベースへ のマッピング . . . . .	167
パーサー・サービス用スレッドの構成 . . . . .	168
複合語の拡張分析の使用可能化 . . . . .	168
ネイティブ XML 検索のサポートの使用可能化 . . . . .	169
<b>文書フォーマットの検出 . . . . . 171</b>	
デフォルトでサポートされる文書タイプ . . . . .	172
コレクション・パーサーと Stellent パーサーに関連 付けられた文書タイプ . . . . .	173
文書タイプとコレクション・パーサーとの関連付け . . . . .	174
デフォルトのコレクション・パーサー・サービ ス規則 . . . . .	176

不明な文書タイプの構文解析 . . . . .	176
一部の HTML タグの置換規則の変更 . . . . .	177
デフォルトの HTML 置換規則 . . . . .	178
文書タイプと Stellent パーサーとの関連付け . . . . .	179
Stellent パーサーのデフォルトの解析規則 . . . . .	182

## 言語とコード・ページのサポート . . . . . 187

言語自動検出 . . . . .	188
コード・ページ自動検出 . . . . .	189
中国語、日本語、韓国語の文書の言語分析 . . . . .	190
N-gram セグメンテーション . . . . .	190
テキストからの空白文字の除去 . . . . .	190

## 索引の管理 . . . . . 193

索引作成のスケジューリング . . . . .	194
索引スケジュールの変更 . . . . .	195
索引スケジュールの使用可能および使用不可設定 . . . . .	195
索引の並行作成の構成 . . . . .	196
変更を検出した場合にのみ索引を作成 . . . . .	198
索引作成の停止 . . . . .	199
索引の検索可能ビューに影響するオプション . . . . .	200
文書の検索に対する索引付きオプション . . . . .	200
重複文書の検出 . . . . .	204
照会内のワイルドカード文字 . . . . .	205
有効範囲 . . . . .	209
有効範囲の構成 . . . . .	210
縮小表示した URI . . . . .	211
検索結果内の URI の縮小表示 . . . . .	212
索引から URI を除去 . . . . .	213

## 検索サーバーの管理 . . . . . 215

検索キャッシュ . . . . .	216
検索キャッシュの構成 . . . . .	216
カスタム同義語辞書 . . . . .	216
システムへの同義語辞書の追加 . . . . .	218
同義語辞書とコレクションとの関連付け . . . . .	219
カスタムのストップワード辞書 . . . . .	219
システムへのストップワード辞書の追加 . . . . .	221
ストップワード辞書とコレクションとの関連付け . . . . .	221
カスタム辞書の再デプロイ . . . . .	222
動的要約 . . . . .	223
管理コンソールの文書サマリーのカスタマイズ . . . . .	224
プロパティの編集による文書サマリーのカスタ マイズ . . . . .	224
クイック・リンクに関する作業 . . . . .	225
クイック・リンク . . . . .	226
クイック・リンクの構成 . . . . .	226

## 文書ランキング . . . . . 229

テキスト・ベースのスコアリング . . . . .	229
静的ランキング . . . . .	230
静的文書ランキングのデフォルト値の復元 . . . . .	231
カスタム・ランキング調整ワード辞書 . . . . .	233
ランキング調整ワード辞書のシステムへの追加 . . . . .	234
ランキング調整ワード辞書とコレクションとの関 連付け . . . . .	235

URI パターンに基づいた文書ランキング	236
URI パターンに一致する文書のスコアの操作	236
ランキング調整クラスに基づいた文書ランキング	237
ランキング調整クラスへのフィールドのマッピング	239
ランキング調整クラスのランキング調整要因の構成	240
ランキング調整クラスのデフォルト値	241

## エンタープライズ・サーチの検索アプリケーション . . . . . **243**

コレクションとの検索アプリケーションの関連付け	244
サンプル検索アプリケーションの機能	244
検索アプリケーションのプロパティ	246
サンプル検索アプリケーション・プロパティの編集	264
検索アプリケーションのカスタマイズ	266
サンプル検索アプリケーションの複製	268
上位結果の分析	270
検索アプリケーションのアクセス	273
検索サーバーがセキュア (SSL) 要求のみを受け入れるようにする構成	274
検索サーバーがプロキシ・サーバーを介して要求を受け入れるようにする構成	276

## 外部ソースのサポート . . . . . **279**

システムへの外部ソースの追加	279
外部ソースと検索アプリケーションの関連付け	282

## エンタープライズ・サーチのセキュリティ . . . . . **285**

インストールのセキュリティ	287
認証とアクセス制御	288
管理役割	288
管理ユーザーの構成	289
コレクション・レベルのセキュリティ	290
重複文書の分析とコレクション・セキュリティ	290
検索アプリケーション ID	291
文書レベルのセキュリティ	292
検索結果の事前フィルターと事後フィルター	292
保管されたセキュリティ・トークンによる検証	293
照会処理中の現行のクレデンシャルの検証	295
アンカー・テキスト分析	302
エンタープライズ・サーチのセキュリティの使用可能化	304
WebSphere Application Server におけるグローバル・セキュリティと LDAP ユーザー・レジストリーの構成	305
シングル・サーバー・エンタープライズ・サーチ・システムに対するセキュリティの使用可能化	307
マルチサーバー・エンタープライズ・サーチ・システムに対するセキュリティの使用可能化	308
セキュリティをサポートするためのクローラーのセットアップ要件	310

セキュア Exchange Server 文書へのアクセスの検証	314
Lotus Domino 文書に対する文書レベルのセキュリティの施行	314
Windows ファイル・システム文書に対する文書レベルのセキュリティの施行	318
エンタープライズ・サーチのセキュリティの使用不可化	322
WebSphere Application Server 内でのエンタープライズ・アプリケーションのセキュリティの使用不可化	322
文書レベルのセキュリティを無効にする	324
縮小表示の検索結果に対するセキュリティの使用不可化	325

## エンタープライズ・サーチ・システムの開始と停止 . . . . . **327**

エンタープライズ・サーチ・システムの開始	327
エンタープライズ・サーチ・システムの停止	330
開始または停止対象のコンポーネントの制御	331
スタンドアロン・モードでの検索サーバーの管理	332

## エンタープライズ・サーチ・アクティビティのモニター . . . . . **335**

単一コレクション内の文書数の見積もり	335
コレクションのモニター	336
URI についての詳細表示	336
クローラーのモニター	338
Web クローラーのアクティビティ詳細の表示	339
Web クローラーのスレッドの詳細	340
Web クローラー・アクティブ・サイト	341
Web クローラーのクローリング速度	342
Web クローラー・レポートの作成	343
Web クローラーに戻される HTTP 状況コード	344
パーサーのモニター	348
コレクションの索引アクティビティのモニター	350
エンタープライズ・サーチの索引キューのモニター	351
検索サーバーのモニター	351
照会統計の計算方法の変更	353
データ・リスナーのモニター	354
文書トラッキング	355
文書トラッキングに対するログ・ファイルの構成	355
ドロップされた文書に関するレポートの表示	356
ドロップされた文書に関するログ・ファイルの表示	357

## ログ・ファイルおよびアラート . . . . . **359**

アラート	359
コレクション・レベル・アラートの構成	360
システム・レベル・アラートの構成	362
ログ・ファイルの構成	363
SMTP サーバー情報の構成	364
ログ・メッセージについての E メール受信	365
照会ログのサイズの変更	367
ログ・ファイルの表示	368

## エンタープライズ・サーチ・システムの バックアップとリストア . . . . . 371

- エンタープライズ・サーチ・システムのバックアップ . . . . . 372
- エンタープライズ・サーチ・システムのリストア . . . . . 373
- コレクション構成のエクスポートとインポート . . . . . 374

## Lotus Notes バージョン 8 との統合 379

- エンタープライズ・サーチ・プラグイン更新サイトの作成 . . . . . 379
- Lotus Notes バージョン 8 クライアントへのエンタープライズ・サーチ・プラグインのインストール . . . . . 379

## WebSphere Portal との統合 . . . . . 381

- エンタープライズ・サーチの WebSphere Portal との統合のセットアップ・スクリプト . . . . . 382
- WebSphere Portal バージョン 5.1 でのエンタープライズ・サーチのセットアップ . . . . . 384
  - WebSphere Portal バージョン 5.1 検索バーによるエンタープライズ・サーチ使用の構成 . . . . . 387
  - WebSphere Portal バージョン 5.1 からのエンタープライズ・サーチの除去 . . . . . 388
- WebSphere Portal バージョン 6 でのエンタープライズ・サーチのセットアップ . . . . . 389
  - WebSphere Portal バージョン 6 検索センターのエンタープライズ・サーチに対する構成 . . . . . 392
  - WebSphere Portal バージョン 6 検索バーによるエンタープライズ・サーチ使用の構成 . . . . . 394
  - Lotus Quickr 用のエンタープライズ・サーチ・ポートレットのセットアップ . . . . . 396
  - WebSphere Portal バージョン 6 からのエンタープライズ・サーチの除去 . . . . . 397

- WebSphere Portal クラスター化システムとのエンタープライズ・サーチの統合 . . . . . 398

- WebSphere Portal クラスター化システムでのエンタープライズ・サーチのセットアップ . . . . . 399

- WebSphere Portal クラスター化システムからのエンタープライズ・サーチの除去 . . . . . 403

## WebSphere Portal からエンタープライズ・サーチへの移行 . . . . . 405

- WebSphere Portal からのコレクションの移行 . . . . . 405
- 移行されたコレクションの設定 . . . . . 407
- 移行ウィザードのログ・ファイル . . . . . 408

## エンタープライズ・サーチ・コマンド、 戻りコード、およびセッション ID . . . . . 409

## エンタープライズ・サーチにおける大/小 文字の区別 . . . . . 443

## エンタープライズ・サーチの資料 . . . . . 445

## アクセシビリティ機能 . . . . . 447

## エンタープライズ・サーチの用語集 . . . . . 449

## 特記事項および商標 . . . . . 463

- 特記事項 . . . . . 463
- 商標 . . . . . 465

## 索引 . . . . . 467



---

## ibm.com および関連リソース

製品サポートおよび文書は、ibm.com® より入手できます。

### サポートおよび支援

製品サポートを Web 上で受けられます。

#### IBM® OmniFind™ Enterprise Edition

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

#### IBM OmniFind Discovery Edition

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

#### IBM OmniFind Yahoo! Edition

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

### インフォメーション・センター

Web ブラウザーによる Eclipse ベースのインフォメーション・センターで製品資料を表示できます。インフォメーション・センターは、<http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/> にアクセスしてご覧ください。

### PDF 資料

ご使用の OS で Adobe® Acrobat Reader を使用して、オンラインで PDF 資料を閲覧できます。ご使用の OS に Acrobat Reader がインストールされていない場合は、Adobe の Web サイト <http://www.adobe.com> からダウンロードできます。

PDF 資料の Web サイトを以下に示します。

製品	Web サイト・アドレス
OmniFind Enterprise Edition バージョン 8.5	<a href="http://www.ibm.com/support/docview.wss?rs=63&amp;uid=swg27010938">http://www.ibm.com/support/docview.wss?rs=63&amp;uid=swg27010938</a>
OmniFind Discovery Edition バージョン 8.4	<a href="http://www.ibm.com/support/docview.wss?rs=3035&amp;uid=swg27008552">http://www.ibm.com/support/docview.wss?rs=3035&amp;uid=swg27008552</a>
OmniFind Yahoo! Edition バージョン 8.4	<a href="http://www.ibm.com/support/docview.wss?rs=3193&amp;uid=swg27008932">http://www.ibm.com/support/docview.wss?rs=3193&amp;uid=swg27008932</a>



---

## エンタープライズ・サーチとは?

エンタープライズ・サーチ・システムは、1回の照会により、構造化されているかどうかを問わず任意の数のデータ・ソースを検索できる幅広い機能を持っています。照会に対する応答が速く、結果セットは大規模なテキスト分析に基づいて統合およびランク付けされているため、必要な情報を簡単に見つけることができるだけでなく、文書コンテンツから意味を取り出すことができます。

エンタープライズ・サーチ・コンポーネントは、IBM OmniFind Enterprise Edition と共にインストールされ、企業の全域から情報を収集します。ユーザーは、Web ブラウザーに照会を入力するだけで、ローカルおよびリモート・データベース、コラボレーション・システム、コンテンツ・マネージメント・システム、ファイル・システム、および内外の Web サイトを同時に検索することができます。

エンタープライズ・サーチ・システムは、既存のシステムとシームレスに統合できるように設計されており、多種多様なソースからデータを収集して、高速に検索するためにデータに索引をつけるのに必要な流れを処理します。また、エンタープライズ・サーチでは、言語分析その他の分析をデータに適用し、有用性の高い検索結果を提供することができます。異なるタイプのリポジトリを検索する場合でも、異なるインターフェースの使い方を習得する必要もありません。

エンタープライズ・サーチ索引に含めることを希望しないデータ・ソースを検索するサポートを追加できます。エンタープライズ・サーチのフェデレーテッド・サーチ機能を使用して、索引付けされたデータ・ソースを検索するときに同時に、これらの外部ソースを検索できます。

### 検索品質

探している情報をユーザーが確実に検出するために、OmniFind Enterprise Edition は IBM Unstructured Information Management Architecture (UIMA) をサポートします。UIMA は、テキスト分析法の共通かつ標準インターフェースを定義するオープン・フレームワークです。大規模なセマンティック分析により、エンタープライズ・サーチは、概念、潜在的意味、関係、事実、および非構造化テキストではしばしば隠されている他の関連データを識別できます。分析中に抽出される情報は、検索結果の品質を向上させるために使用したり、ビジネス・インテリジェンスやデータ・マイニングなど、他のアプリケーションの品質を向上させるために使用したりすることができます。

### セキュリティ

セキュリティは、エンタープライズ・サーチに不可欠な要素です。システム管理は、権限のあるユーザーしか操作できないようになっています。IBM WebSphere® Application Server で使用可能なセキュリティ・メカニズムを使用して、管理役割を構成して、さまざまな管理機能に対するアクセス権を、どのユーザーが持つかを制御できます。

また、データ収集時に、セキュリティー・トークンを収集データに関連付けるオプションを指定することもできます。検索アプリケーションでセキュリティーを使用可能にすれば、索引内の文書と一緒に保管されている、これらのトークンを使用して、アクセス制御を施行し、適切なクレデンシャルを持つユーザーのみがデータを照会し、検索結果を表示できるようにすることができます。

特定のタイプのデータ・ソースの場合、オプションを構成して、照会処理中に、現行のアクセス制御によってユーザーのログイン・クレデンシャルを妥当性検査することができます。この追加のセキュリティーの層は、ネイティブ・データ・ソースによって、リアルタイムでユーザーの権限を確実に妥当性検査するようにします。このオプションによって、文書およびそのセキュリティー・トークンが索引付けされた後、ユーザーのクレデンシャルが変更になったインスタンスに対してプロテクトすることができます。

## 製品チュートリアル


オンライン・チュートリアルが <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html> で利用できます。このチュートリアルには、インストールおよび構成手順の説明のほか、さまざまなタイプのデータ・ソースを検索する方法が示されています。また、当該製品のアプリケーション・プログラミング・インターフェースを使用してエンタープライズ・サーチを拡張する方法についての説明もあります。このチュートリアルは OmniFind Enterprise Edition の旧バージョンに対応するものですが、概念や手順の多くは適用可能です。

### 関連概念

3 ページの『エンタープライズ・サーチ・コンポーネントの概要』

13 ページの『エンタープライズ・サーチのデータ・フロー』

エンタープライズ・サーチ・セキュリティー

 カスタム・テキスト分析統合

 テキスト分析処理で使用される基本概念

### 関連資料

『エンタープライズ・サーチでサポートされているデータ・ソース・タイプ』

---

## エンタープライズ・サーチでサポートされているデータ・ソース・タイプ

データ・ソースを検索する場合に、さまざまなデータ・ソース・タイプが検索できるようにあらかじめサポート設定されています。

IBM OmniFind Enterprise Edition をインストールすると、次の各タイプのデータ・ソースからのデータの収集を開始できます。

- IBM DB2<sup>®</sup> Content Manager 項目タイプ (文書、リソース、および項目)
- IBM DB2 データベース
- IBM Domino<sup>®</sup> Document Manager (以前の Domino.Doc<sup>®</sup>) データベース
- IBM Lotus Notes<sup>®</sup> データベース
- IBM Lotus<sup>®</sup> QuickPlace<sup>®</sup> データベース
- IBM Lotus Quickr<sup>™</sup> コンテンツ・ライブラリー

- IBM WebSphere Information Integrator Content Edition リポジトリ (Documentum、FileNet® Panagon Content Services、FileNet P8 Content Manager、Hummingbird® Document Management (DM)、Microsoft® SharePoint、OpenText Livelink Enterprise Server、および WebSphere Portal Document Manager (PDM) を含む)
- 多くのデータベース・システム・タイプ (IBM DB2 for z/OS®、IBM Informix®、Microsoft SQL Server、Oracle、および Sybase を含む) の IBM WebSphere Information Integrator ニックネーム表
- IBM WebSphere Portal サイト
- IBM Workplace Web Content Management™ サイト
- Microsoft Exchange Server パブリック・フォルダー
- Microsoft SQL Server データベース
- Microsoft Windows® ファイル・システム
- NNTP ニュースグループ
- Oracle データベース
- UNIX® ファイル・システム
- インターネット上およびイントラネット内の Web サイト

これらのソースをエンタープライズ・サーチ索引に追加せずに、以下のタイプの外部ソースを検索するサポートを追加できます。

- Java™ Database Connectivity (JDBC) プロトコルをサポートするデータベース (DB2 および Oracle データベース・システムのみ)。検索を可能にする表ごとに、別個の外部ソースが作成されます。
- Lightweight Directory Access Protocol (LDAP) サーバー。各 LDAP サーバーごとに 1 つの外部ソースが作成されます。

サポートされるデータ・ソース・タイプとサポートされる製品バージョンについての最新情報は、OmniFind Enterprise Edition サポート Web サイトの「システム要件」のページを参照してください。

### 関連概念

外部ソースのサポート

13 ページの『エンタープライズ・サーチのデータ・フロー』

---

## エンタープライズ・サーチ・コンポーネントの概要

エンタープライズ・サーチでは、そのコンポーネントによって企業全体のデータが収集され、その情報が分析、解析、カテゴリ化されます。さらに、ユーザーが検索に使用する索引が作成されます。

エンタープライズ・サーチ・コレクションは、ユーザーが 1 回の照会で検索できるデータ・ソースの集合を表します。コレクションの作成時には、コレクションに入れるソースを指定し、ユーザーが索引付けされたデータを検索するときのオプションを構成します。

コレクションは複数作成することができ、各コレクションにはさまざまな種類のデータ・ソースのデータを入れることができます。例えば、IBM DB2 Universal

Database™、IBM Lotus Notes、および IBM DB2 Content Manager データベースの文書を含むコレクションを作成することができます。ユーザーがこのコレクションを検索すると、検索結果にはこれらの各データ・ソースの文書が含まれることとなります。

フェデレーテッド検索のサポートにより、1回の照会で複数のコレクションを検索することができます。場合によっては、検索結果にエンタープライズ・サーチ・システム内のすべてのコレクションと外部ソースの文書が組み込まれる可能性があります。

コレクションの作成および管理には、次のようなアクティビティーがあります。

#### データの収集

クローラー・コンポーネントによって、連続的に、または指定したスケジュールに従ってデータ・ソースから文書が収集されます。頻繁にクロールを行うことにより、ユーザーが常に最新の情報にアクセスできるようになります。

#### データの分析

パーサー・コンポーネントによって、文書からテキストが抽出され、クローラーがクロールした各文書に対して言語分析その他の分析が行われます。この詳細なコンテンツ分析により、検索結果の品質が向上します。

#### データの索引作成

索引コンポーネントが、定期的にスケジュールに従って実行され、新規の文書、変更された文書に関する情報が索引に追加されます。索引コンポーネントでは、コレクション内の文書に対してグローバル分析を行い、検索結果の品質を高めます。

#### データの検索

検索コンポーネントによって、索引が検索され、検索アプリケーションと連携して検索要求を処理し、処理された検索結果を戻します。

他の OmniFind Enterprise Edition コンポーネントによって、セキュリティ設定の指定、システム・アクティビティーのモニター、さらに問題が発生した場合のトラブルシューティングが行えるようになります。また、この製品では、独自の検索アプリケーションを作成するときのテンプレートとして使用できる検索アプリケーションの実用サンプルも提供されています。

#### 関連概念

クローラーの管理

パーサーの管理

索引の管理

検索サーバーの管理

13 ページの『エンタープライズ・サーチのデータ・フロー』

## エンタープライズ・サーチ・クローラー

エンタープライズ・サーチ・クローラーは、データ・ソースから文書を収集し、これらの文書が分析、索引作成、および検索できるようにします。

OmniFind Enterprise Edition で提供されているクローラー・コンポーネントには、次のような機能があります。

- クローラーを構成すると、検出 プロセスは、Lotus Notes データベースのすべてのビューとフォルダー名、または UNIX サーバーの全ファイル・システム名など、クローラーに使用することができるソースに関する情報を探し出します。
- クロールするソースを選択し、クローラーを開始すると、クローラー・コンポーネントはソースからデータを収集し、それらのデータが分析および索引作成できるようにします。

1 つのコレクションには複数のクローラーを入れることができ、各クローラーは特定タイプのデータ・ソースからデータを収集するように設計されています。例えば、ファイル・システム、Lotus Notes® データベース、およびリレーショナル・データベースのデータを結合するために、1 つのコレクションの中に 3 つのクローラーを作成することができます。また、同一タイプのクローラーを複数作成し、各クローラーがクロールしているデータの変更頻度に従って、それぞれ異なるクローラー・スケジュールを設定することもできます。

Web クローラーは、開始すると継続的に実行されます。クロール対象の URL を指定すると、クローラーは、定期的にそれらの対象に戻って、新規または変更されたデータがないかどうかをチェックします。他のタイプのクローラーは、手動で開始および停止することもできますし、クローラー・スケジュールを設定することもできます。クローラーのスケジュールを設定する場合、初回実行の日時、および新規および変更された文書をクロールするためにデータ・ソースにクローラーがアクセスする頻度を指定します。

クローラーのプロパティは、特定のクローラーがクロールするときの動作を制御する規則の集合です。例えば、クローラーのシステム・リソースの使用を制御する規則を指定します。クロール対象として適格なデータ・ソースの集合は、クローラーのクロール・スペースを構成します。クローラーを作成後、クローラーのプロパティは、いつでも編集でき、クローラーのデータ収集方法を変更することができます。また、クロール・スペースを編集して、クローラー・スケジュールの変更、新規ソースの追加、または今後検索しないソースの除去を行うこともできます。

#### 関連概念

クローラーの管理

#### 関連タスク

クローラーのモニター

## エンタープライズ・サーチ・パーサー

エンタープライズ・サーチ・パーサーは、クローラーによって収集された文書を分析し、文書の索引作成の準備を行います。

OmniFind Enterprise Edition で提供されているパーサー・コンポーネントは、文書コンテンツおよび文書のメタデータを分析します。分析結果は、索引作成コンポーネントがアクセスできるようにデータ・ストアに格納します。パーサーは、以下のようなタスクを実行します。

- さまざまなフォーマットによる文書からテキストを抽出する。例えば、パーサーは、XML および HTML 文書のタグ内のテキストを抽出します。パーサーはま

た、Oracle (以前の Stellent) Outside In Content Access for IBM OmniFind Enterprise Edition を使用して、Microsoft Word 文書や Adobe Acrobat PDF 文書などのバイナリー・フォーマットからもテキストを抽出します。

- 各文書の文字セット・エンコードを検出する。パーサーは、言語分析を行う前処理として、検出したこの情報を使用してすべてテキストをユニコードに変換します。
- 各文書のソース言語を検出する。
- テキストを抽出し、トークンを追加してデータの検索機能を拡張する。このプロセスでは、パーサーは、以下のタスクを実行します。
  - 大文字の正規化、およびドイツ語のウムラウトなどの発音区別符号の正規化などの文字の正規化。
  - 段落、文、単語、および空白文字の構造分析。パーサーは、このような言語分析により、複合語を分解し、辞書検索および同義語検索ができるようにするためのトークンの指定を行います。
- コレクションに管理者が指定した解析規則を適用する。パーサーの構成時、次の構文解析アクティビティを構成できます。

#### **XML 文書および HTML 文書のフィールド・マッピング規則**

このオプションを使用して、ユーザーは XML 文書および HTML 文書内の構造化されたコンテンツ、および構造化されていないコンテンツを区別して検索することができます。XML エlementまたは HTML メタデータ・Elementをエンタープライズ・サーチ索引の検索フィールドにマップすると、ユーザーは照会でフィールド名を指定して、XML 文書および HTML 文書の特定の部分を検索することができます。特定フィールドを検索する照会では、文書の全コンテンツを検索するフリー・テキスト照会より精度の高い検索結果が得られます。

#### **カテゴリー**

このオプションで、ユーザーは、文書が属するカテゴリーを基準にして文書を検索できます。また、ユーザーは検索結果でカテゴリーを選択して、同じカテゴリーに属する文書だけを参照することもできます。

コレクションの作成時に、カテゴリー化を使用するかどうかを選択します。「ルール・ベース」のカテゴリーを使用すると、文書は管理者が定義した規則に従ってカテゴリーに関連付けられます。ルール・ベースのカテゴリーは、新規作成したエンタープライズ・サーチ・コレクションと IBM WebSphere Portal から移行したコレクションで構成することができます。

#### **カスタム・テキスト分析**

アプリケーション開発者は、ユーザーが検索するデータに複雑な言語分析を実行するために、カスタム分析プログラムを作成することができます。こうしたプログラムをエンタープライズ・サーチ・システムに埋め込み、これを使用してコレクションのコンテンツに注釈を付けることができます。注釈に索引付けをすることで、コレクションにセマンティック検索を行うことができます。

例えば、ユーザーは、互いに隣接して出現する照会条件や、同一文に出現する照会条件を検索したり、あるいは、照会条件間の関係を検索すること



ができます。例えば、ユーザーは、Smith という名前の IBM 技術者ではなく、Smith という名前の IBM 営業担当者についての文書が必要かもしれません。

#### **N-gram セグメンテーションのサポート**

中国語、日本語、韓国語で書かれた文書の検索性能を高めるためには、字句解析の N-gram セグメンテーション方式を使用可能にします。この分析方法では、ワードの区切りに空白文字を使用しません。コレクション作成後に、セグメンテーション方式を変更することはできません。

#### **ネイティブ XML 検索での XML 文書検索サポート**

ネイティブ XML 検索は、XML マークアップを検索することで、より正確な検索結果を提供できます。例えば、照会で、あるワードが特定の XML エlement に出現しなければならないと指定することができます。

#### **フィールドの相対的重要度スコアをランキング調整するクラス**

フィールドをランキング調整クラスにマップすると、検索結果でどのように文書をランク付けるかに影響を与えることができます。例えば、タイトル・フィールドのスコアをランキング調整したいとします。照会条件がタイトルに出現した場合、タイトルに条件が入っている文書が、検索結果でより高くランク付けされます。

#### **関連概念**

カテゴリーに関する作業

XML 検索フィールド

HTML 検索フィールド



カスタム・テキスト分析統合



エンタープライズ・サーチに組み込まれたテキスト分析

#### **関連タスク**

パーサーのモニター

## **エンタープライズ・サーチ索引**

エンタープライズ・サーチの索引作成コンポーネントは、定期的なスケジュールに従って実行され、新規の文書、変更された文書に関する情報を索引に追加します。

ユーザーが、常に検索対象のソースの最新情報にアクセスできるようにするため、索引の作成には、次の 2 つの段階があります。

#### **主索引の作成**

主索引の作成では、構成が最適な編成になるように索引全体が再作成されず。索引作成プロセスによって、クローラーが収集し、パーサーが分析したすべてのデータが読み込まれます。

#### **差分索引の作成**

差分索引の作成では、最後の主索引の作成後にクロールされた情報が、索引に追加されます。

コレクションの索引オプションを構成するときに、主索引と差分索引の作成スケジュールを指定できます。索引作成の頻度は、システム・リソースや、索引付けされるソースが静的コンテンツと動的コンテンツのどちらを含むかによって異なります。

新しい情報が確実に利用されるようにするため、差分索引作成が頻繁に行われるようにスケジュールしてください。主索引作成を定期的に行うようスケジュールすることにより、すべての新規の情報が統合され、新規のコンテンツが分析され、索引のパフォーマンスが最適化されます。

また、索引作成プロセスは、スケジューリングせずに開始することもできます。例えば、特定の解析規則を変更し、その変更を使用している検索アプリケーションで有効になるようにしたい場合、スケジュールされた時に開始する索引作成を待たずにデータを再クロールし、解析した後、主索引作成を開始することができます。

リソースの使用を制御するには、いくつかのコレクションが索引付けプロセスを共用し、索引作成要求を同時にサブミットするかを制御します。同時に索引作成を行うことで、大規模な主索引作成によって、他のコレクションの差分索引作成が妨げられることがないようにできます。索引作成は、リソースを集中的に使用するプロセスになりかねないので、大規模なシステムの場合、システム負荷をモニターして、主索引および差分索引作成スケジュールを調整してください。

索引作成時、索引プロセスでは、グローバル文書分析を行います。この分析では、文書の重複を識別し、文書のリンク構造を分析し、アンカー・テキスト (ハイパーテキスト・リンクでターゲット・ページを記述するテキスト) に対する特殊処理を行うアルゴリズムが適用されます。

以下の索引付けアクティビティのオプションが指定できます。

- ユーザーがワイルドカード文字を指定できるように、照会条件を索引に拡張するサポートを構築するか、照会処理中に照会条件を拡張することを指定します。これを判断するときは、リソースの使用量と照会応答時間のトレードオフを考慮する必要があります。
- 有効範囲を構成できます。有効範囲を設定すると、ユーザーがコレクションで表示できる内容を制限することができます。例えば、技術サポート部の文書の URI を含む有効範囲を作成したり、人事部の文書の URI を含む別の有効範囲を作成したりすることができます。検索アプリケーションで有効範囲をサポートすると、ユーザーはコレクション中の該当するサブセットだけを対象として文書を検索、リトリブすることができます。
- 同じ URI 接頭部をもつ検索結果文書を縮小表示するオプションが指定できます。また、URI 接頭部が異なる文書を検索結果にまとめて縮小表示できるようにグループ名を指定することもできます。
- 索引を作成後、ユーザーに検索させたくない URI を削除することができます。

#### 関連概念

索引の管理

照会内のワイルドカード文字

有効範囲

縮小表示された URI

URI パターンに基づいた文書ランキング

### 関連タスク

索引作成のスケジューリング

索引の並行作成の構成

索引から URI を除去

コレクションの索引アクティビティのモニター

エンタープライズ・サーチの索引キューのモニター

## エンタープライズ・サーチ用検索サーバー

エンタープライズ・サーチ用の検索サーバーは、検索アプリケーションと連携して、照会の処理、索引の検索、および検索結果を戻す処理を行います。

エンタープライズ・サーチ用検索サーバーは、OmniFind Enterprise Edition のインストール時にインストールされます。コレクションに検索サーバーを設定すると、コレクションを検索する方法のオプションを指定できます。

- 検索キャッシュを構成して、頻繁に要求される検索結果を保持することができます。検索キャッシュを利用すると、検索効率が高まります。
- コレクション内の文書検索のデフォルト言語を指定できます。
- アプリケーション開発者がカスタム辞書を作成する場合、それら辞書をコレクションに関連付けることができます。
  - ユーザーが同義語辞書を使用するコレクションを照会するとき、照会条件の同義語を含む文書が、検索結果に含まれます。
  - ユーザーがストップワード辞書を使用するコレクションを照会するとき、ストップワードは、照会が処理される前に、照会から除去されます。
  - ユーザーがランキング調整ワード辞書を使用するコレクションを照会するとき、ランキング調整ワードを含む文書の重要度は、辞書のワードと関連付けられているランキング調整要因に従って調整されます。
- 特定の文書が特定の照会に関連付けられることが事前にわかっている場合、クイック・リンクを構成することができます。クイック・リンクでは、特定の URI と特定のキーワードおよび句を関連付けます。照会にクイック・リンク定義で指定したキーワードまたは句がある場合、関連付けられた URI が自動的に検索結果に戻されます。

マルチサーバー構成では、障害保護がサーバー・レベルではなく、コレクション・レベルで使用可能です。1 つの検索サーバー上のコレクションがいかなる理由にしろ使用不可になった場合、その後そのコレクションの照会は、自動的に別の検索サーバーにルーティングされます。

### 関連概念

エンタープライズ・サーチの検索アプリケーション

検索キャッシュ

カスタム同義語辞書

カスタムのストップワード辞書

カスタムのランキング調整ワード辞書

クイック・リンク

## 関連タスク

検索サーバーのモニター

# エンタープライズ・サーチ管理コンソール

エンタープライズ・サーチ管理コンソールは、ブラウザで稼働します。すなわち、管理ユーザーは、いつでもどこからでもこのコンソールにアクセスできるようになります。同時に、セキュリティ機構によって、管理機能にアクセスすることを許可されたユーザーしかコンソールにアクセスできないことが保証されています。

エンタープライズ・サーチの管理コンソールは、OmniFind Enterprise Edition のインストール時に、検索サーバーにインストールされます。

管理コンソールには、さまざまな基本的管理用タスクの実行を助けるウィザードが付属しています。例えば、コレクション・ウィザードを使用すれば、簡単にコレクションを作成でき、ドラフト・モードで作業をする手間が省けます。クローラー・ウィザードは、データ・ソース・タイプによって異なり、ユーザーが検索できるようにしたいソースを選択する作業を支援します。

他の管理タスクについては、管理する個別の項目を選択します。例えば、コレクションを編集する場合、索引スケジュールを変更するには「索引」ページを選択し、XML 文書の構文解析規則を変更するには「解析」ページを選択します。

## 関連概念

エンタープライズ・サーチ・システムの管理

管理役割

## 関連タスク

管理コンソールへのログイン

# エンタープライズ・サーチ・システムのモニター

エンタープライズ・サーチ管理コンソールを使用すれば、システム・アクティビティをモニターし、必要に応じてその運用を調整することができます。

OmniFind Enterprise Edition をインストールし、最少でも 1 つのコレクションを作成すると、主要な各アクティビティ（クローラー、解析、索引作成、検索）の詳細な統計データを表示することができます。統計データには、平均応答時間や、クローラーまたは索引作成セッションでいくつの文書がクローラーまたは索引作成されたかなどの進行状況データがあります。

ほとんどのアクティビティは、管理者が開始および停止することができます。例えば、アクティビティを一時停止して、構成の変更または問題のトラブルシューティングを行い、アクティビティを進められる状況になった時点で処理を再開することができます。

また、アラートを設定することもでき、これによりモニターしている事象が発生したときはいつでも、特定のアクティビティに関する E メールを受信できるようになります。例えば、検索応答時間が指定したしきい値を超過した場合にアラートを受信することができます。

文書がエンタープライズ・サーチ・システムから除去された場合、文書をトラッキングして、いつ、どこで、なぜ、その文書が除去されたのかを判別できます。例えば、パーサーがその文書を構文解析できなかったか、または管理者が索引から文書を除去したことが考えられます。

#### 関連概念

エンタープライズ・サーチ・アクティビティのモニター  
エンタープライズ・サーチ・システムの開始と停止

## エンタープライズ・サーチ用ログ・ファイル

ログ・ファイルは、個別のコレクションおよびシステム・レベル・セッションに対して作成されます。

特定のエンタープライズ・サーチ・コレクションまたはシステムに対してログ・オプションを構成する場合、ログに記録するメッセージのタイプ (エラー・メッセージ、警告メッセージなど) を指定します。また、新しいメッセージのスペースを確保するためにシステムが古くなったログ・ファイルを循環使用する頻度を指定します。特定のメッセージ (アラートを含む)、あるいはエラーが発生したときのすべてのエラー・メッセージについて、E メールを受信するオプションを選択できます。

ログ・ファイルを表示する場合、表示するログ・ファイルを選択します。ファイル名を見れば、ファイル作成日時、メッセージを発行したコンポーネントがわかります。目視フィルターを指定することもできます。例えば、エラー・メッセージだけを表示したり、特定エンタープライズ・サーチ・セッションから出されたメッセージだけを表示することができます。

#### 関連概念

ログ・ファイルとアラート  
アラート

 エンタープライズ・サーチのメッセージ

#### 関連タスク

ログ・ファイルの構成  
SMTP サーバー情報の構成  
ログ・メッセージについての Eメールの受信  
ログ・ファイルの表示

## エンタープライズ・サーチのカスタマイズ

エンタープライズ・サーチ用のアプリケーション・プログラミング・インターフェース (API) を使用すれば、カスタム検索アプリケーション、コレクションのコンテンツを更新するカスタム・アプリケーション、テキスト分析用のカスタム・プログラム、および同義語、ストップワード、ランキング調整ワード用の辞書を作成することができます。

OmniFind Enterprise Edition をインストールすると、エンタープライズ・サーチ・コレクションの拡張に使用可能な、次の一連の API が使用できるようになります。

## 検索および索引 API (SI-API)

この API を使用して、カスタム検索アプリケーションおよびカスタム管理インターフェースを作成します。

## クローラー・プラグイン

プラグイン API を使用して、クロール中の文書にメタデータを追加したり、あるいは組織の業務を強化するセキュリティー・トークンとセキュリティー規則を関連付けたりします。

言語分析用カスタム・プログラムをエンタープライズ・サーチ・コレクションと統合することで情報の検索性を高めることができます。カスタム・テキスト分析エンジンをシステムに追加すると、そのエンジンをコレクションに関連付けることができます。ユーザーがコレクションを照会したときに、カスタム・プログラムで索引に構築したワードの関連付けが有効に働きます。例えば、ユーザーは単に用語そのものだけではなく、概念および用語と用語の間の関係を検索できます。

また、頭字語、省略語、業界固有の用語などを反映したカスタム辞書を統合することで、情報をより多く取得できるようになります。辞書をシステムに追加した後、それらの辞書をコレクションに関連付けることができます。ユーザーがコレクションを照会したとき、次のような仕方です利点がもたらされます。

- 照会に同義語として定義されたワードが含まれる場合、照会条件の同義語を含む文書が、検索結果に含まれることとなります。
- 照会にストップワードが含まれる場合、そのストップワードは照会から除外されることになり、無関係な文書は検索結果に戻されません。
- 照会にランキング調整ワードが含まれる場合、ランキング調整ワードを含む文書は、検索結果において、辞書のワードと関連付けられたランキング調整値に従って調整されて、ランク付けされます。

### 関連概念

エンタープライズ・サーチの検索アプリケーション

カスタム同義語辞書

カスタムのストップワード辞書

カスタムのランキング調整ワード辞書

 [検索および索引 API の概要](#)

 [クローラー・プラグイン](#)

## エンタープライズ・サーチ用サンプル検索アプリケーション

エンタープライズ・サーチ用のサンプル検索アプリケーションは、カスタム検索アプリケーションを作成するためのテンプレートとして使用することができます。

サンプル検索アプリケーションは、OmniFind Enterprise Edition のインストール時に、併せてインストールされます。サンプル検索アプリケーションによって、エンタープライズ・サーチで使用可能な検索および取り出し機能のほとんどを実際に試すことができます。また、このアプリケーションは実用的なサンプルで、これによってエンタープライズ・サーチ・システム内のすべてのアクティブ・コレクション

および外部ソースを検索できます。コレクションまたは外部ソースをユーザーに提供する前に、サンプル・アプリケーションを使用して、新規のコレクションおよび外部ソースをテストできます。

サンプル検索アプリケーションは、一度に複数のコレクションおよび外部ソースの検索を可能にして、フェデレーテッド検索の対応も実際に試すことができます。

特定のクローラー・タイプについては、ユーザーが検索アプリケーションにアクセスするときにエンタープライズ・サーチの ID 管理コンポーネントを使用して現行のクレデンシャルを検証できます。検索対象のドメインがシングル・サインオン (SSO) セキュリティーによって保護されていれば、検索セッション中のユーザーの検証に SSO メカニズムを使用できます。そうでない場合、ID 管理コンポーネントは、ユーザーのクレデンシャルを暗号化してプロファイルの中に保管し、検索結果から禁止文書を除外するためにそのクレデンシャルを使用することができます。

サンプル検索アプリケーションをカスタマイズするには、検索アプリケーション・カスタマイザーを使用できます。これは、加えた変更の影響を確認できるグラフィカル・ユーザー・インターフェースです。また、アプリケーションの構成ファイルを編集して検索アプリケーションをカスタマイズすることもできます。

カスタム検索アプリケーションを作成するには、エンタープライズ・サーチ用検索および索引 API を使用します。

#### 関連概念

エンタープライズ・サーチの検索アプリケーション

サンプル検索アプリケーションの機能

 [検索および索引 API の概要](#)

#### 関連タスク

検索アプリケーションのアクセス

サンプル検索アプリケーション・プロパティの編集

検索アプリケーションのカスタマイズ

---

## エンタープライズ・サーチのデータ・フロー

OmniFind Enterprise Edition と共にインストールされるエンタープライズ・サーチ・コンポーネントは、互いに連携してシステム内のデータ・フローを確保しています。

クローラーは、企業全体のデータ・ソースから文書を収集します。パーサーは、クロール済みの文書から有用な情報を抽出し、さらに、例えば文書をカテゴリーに関連付け、検索要求にある用語と文書の適合性を判断するときに役立つトークンを生成します。そして、索引コンポーネントは、効率的な検索が行われるようにデータを保管します。

Web ブラウザーと検索アプリケーションを使用することにより、ユーザーは索引処理済みのコレクションおよび外部ソースを検索します。検索アプリケーションは、ユーザーがブラウザーでクリックして利用できる結果リストを表示することができます。

ます。また、より高度に使用して、さまざまなデータ・ソースの情報に基づいて動的に生成されたコンテンツを戻すこともできます。

例えば、カタログ検索アプリケーションを使用すれば、検索要求を満たす結果表示をカスタマイズすることができます。1回の照会で、例えば IBM DB2 Content Manager および Lotus Notes リポジトリにある文書の組み合わせなど、異なるタイプのデータ・ソースから文書を検索することができます。

管理者は、収集対象のデータを指定し、それらのデータのクロール、解析、索引作成、および検索に関する設定を行います。システム・アクティビティをモニターすることにより、管理者は、必要な調整を行ってデータのスループットを最適にします。

次のダイアグラムは、エンタープライズ・サーチ・システムにおけるデータの流れを示しています。



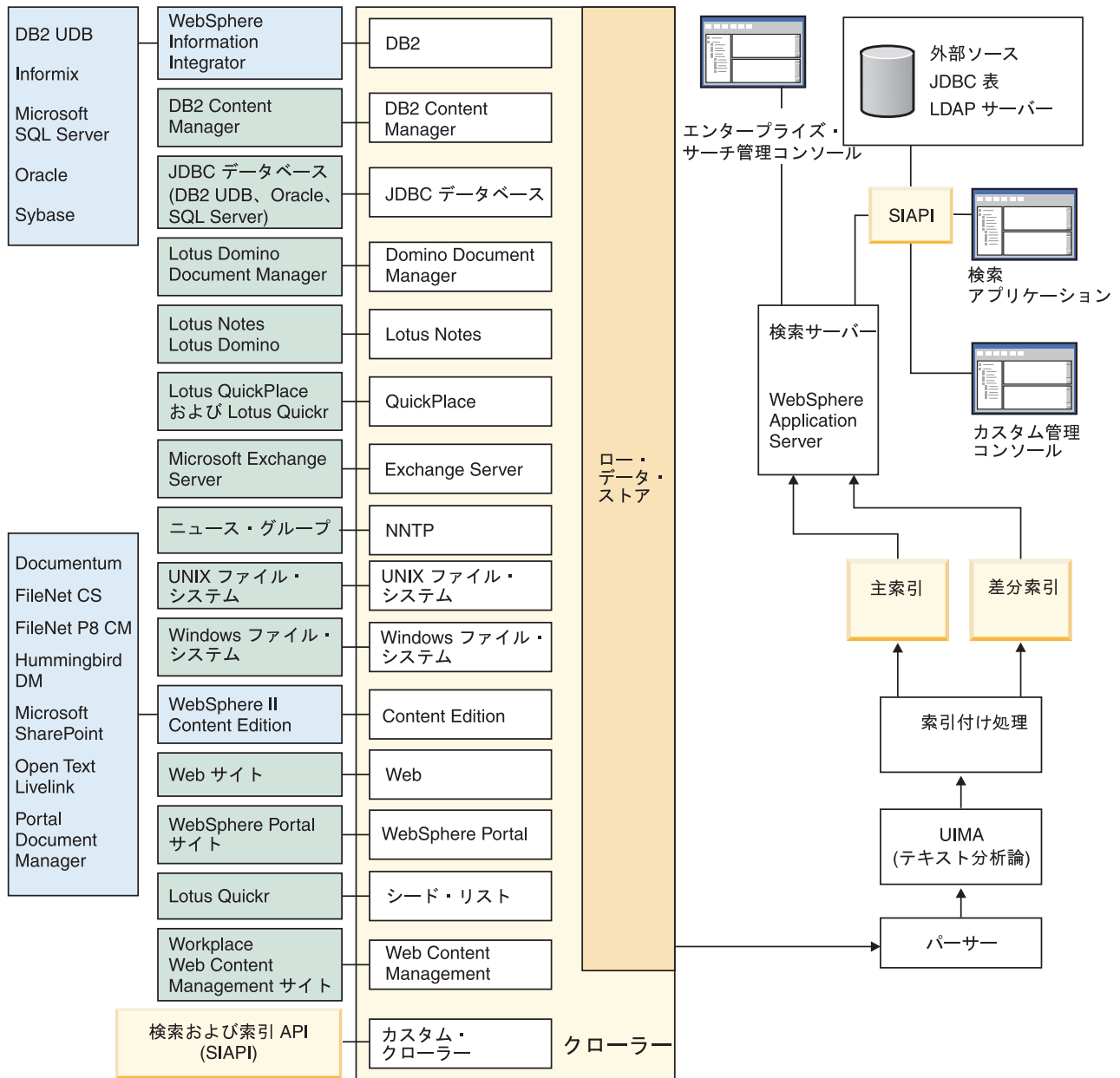


図1. エンタープライズ・サーチ・システムにおけるデータ・フロー

### 関連概念

- 1 ページの『エンタープライズ・サーチとは?』
- 3 ページの『エンタープライズ・サーチ・コンポーネントの概要』

### 関連資料

- 2 ページの『エンタープライズ・サーチでサポートされているデータ・ソース・タイプ』



---

## エンタープライズ・サーチ・システムの管理

コレクションおよび外部ソースの作成と管理、コンポーネントの開始と停止、システム・アクティビティおよびログ・ファイルのモニター、管理ユーザーの構成、検索アプリケーションのコレクションおよび外部ソースへの関連付け、セキュリティー施行情報の指定には、エンタープライズ・サーチ管理コンソールを使用します。

**ヒント:** オンライン・チュートリアルが <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html> で利用できます。このチュートリアルには、インストールおよび構成手順の説明のほか、さまざまなタイプのデータ・ソースを検索する方法が示されています。また、当該製品のアプリケーション・プログラミング・インターフェイスを使用してエンタープライズ・サーチを拡張する方法についての説明もあります。このチュートリアルは OmniFind Enterprise Edition の旧バージョンに対応するものですが、概念や手順の多くは適用可能です。


小規模、中規模、および大規模組織においてクローラーを構成し、セキュリティーを有効にする方法の詳細例については、IBM Redbook 「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」を参照してください。

### 「コレクション」ビュー

以下のガイドラインに従って、最初のコレクションを作成し、システムを管理します。

#### ログイン

エンタープライズ・サーチ管理コンソールにログインします。「コレクション」ビューは、コレクションの作成および管理のエントリー・ポイントです。

**ヒント:** 管理コンソールを使用する場合は、ツールバーの「ヘルプ」または「このページのヘルプ」をクリックします。メッセージに関する詳細説明および修正アクションがある場合は、メッセージの最後にある「 詳細」アイコンをクリックして、詳細を見ることができます。


#### コレクションの作成

以下のいずれかの方法を選択して、空の新規コレクションを作成します。

- コレクション・ウィザードを使用してコレクションを作成するには、「**コレクション・ウィザード**」をクリックし、ウィザードの指示に従ってください。
- 「コレクション」ビューを使用してコレクションを作成するには、「**コレクションの作成**」をクリックし、「コレクションの作成」ページの各フィールドに必要なデータを入力して「**OK**」をクリックします。

#### コレクションの構成

新規コレクションは空なので、編集してコンテンツを追加する必要があります。またデータをクロールするときやデータを検索可能にする場合のオプションを指定します。

作成した新規のコレクションに対して「 編集」をクリックし、続いてそのコレクションのオプションを指定するページを選択します。

**重要:** 直前に表示されたページに移動したり、管理コンソールの情報を最新表示したりするには、エンタープライズ・サーチ管理コンソールの「前へ」および「リフレッシュ」をクリックしてください。Web ブラウザーで「戻る」または「最新表示」をクリックすると、矛盾した結果やデータ損失の恐れが発生する可能性があります。

- 「一般」ページでは、コレクション全体に適用するオプションを指定します。
  - 一般オプションを編集して、コレクションの名前や説明を変更したり、あるいは、コレクションの見積もりサイズを変更することができます。
  - コレクション ID や検索結果で文書をランキングする静的ランキング方法などのコレクションに関する変更不可能な情報を表示することができます。
  - コレクション作成時にそのコレクションのセキュリティーを有効に設定している場合、文書レベルのセキュリティー制御を有効にしたり、無効にしたりすることができます。
- 「クロール」ページでは、クローラーを必ず 1 つは構成します。

1 つのコレクションには、さまざまな種類のデータ・ソースのデータを入れることができます。コレクションに組み込む各データ・ソースのタイプごとに、1 つ以上のクローラーを構成する必要があります。クローラーを作成する場合、クロール対象のデータ・タイプに固有のウィザードがあり、これを使用すると簡単にクローラーを構成できます。
- 「構文解析」ページでは、クロール対象データの検索が効率的に行われるようにするための構文解析方法のオプションを構成します。
  - ネイティブ XML 検索で検索できるように XML 文書を構文解析するかどうかを指定することができます。
  - 文書をカテゴリーに関連付けることができます。こうすることでユーザーはコレクションのサブセットを検索したり、文書が含まれるカテゴリーを基準にして検索結果の文書を参照することができます。
  - XML エlementと HTML メタデータ・Elementを索引の検索フィールドにマップすると、ユーザーが照会にフィールド名を指定して文書の特定部分を検索することができますようになります。
  - エンタープライズ・サーチ・システムにカスタム・テキスト分析エンジンを追加した場合、コレクションで使用するものを選択し、テキスト処理オプションを指定して情報のリトリブ可能性を高め、セマンティック検索をサポートすることができます。
  - フィールドをランキング調整クラスに関連付けて、照会条件に一致するフィールドの、検索結果におけるランクを調整することができます。
- 「索引」ページでは、索引の作成のスケジュールを設定します。スケジュールは、ユーザーが常に最新の情報にアクセスできるようにするため、索

引作成が頻繁に実行されるように設定してください。オプションで次のようなアクティビティーを実行することもできます。

- ユーザーが照会条件にワイルドカード文字を指定できるようにします。
- 有効範囲を設定し、ユーザーが索引のすべての文書ではなくコレクションの限定された一部分を検索できるようにします。
- 検索結果を縮小表示して、同じソースの文書が検索結果に縮小表示されるようにします。
- ランキング調整要因を割り当てて、URI パターンに一致する文書の、検索結果におけるランクを調整します。
- 索引から URI を除去します。例えば、コレクション作成後に、特定の文書をユーザーが表示できないようにする必要がある場合に行います。
- 「検索」ページでは、コレクションの文書を検索するためのオプションを指定します。
  - 検索結果用のキャッシュ・スペースを確保し、またコレクションのデフォルト言語を変更できます。
  - エンタープライズ・サーチ・システムに、同義語、ストップワード、ランキング調整ワードに関するカスタム辞書を追加した場合、コレクションの検索時に使用する辞書を選択できます。
  - 検索結果に表示する文書サマリーの長さを指定できます。
  - 照会に特定のキーワードまたは句が含まれているときに必ず特定の URI が自動的に表示されるように、クイック・リンクを構成することができます。
- 「ログ」ページでは、以下のアクティビティーを実行することができます。
  - ログに記録するメッセージのタイプと、ログ・ファイルのリサイクル頻度を指定します。
  - コレクション・アクティビティーに関するアラートを受信する場合のオプションを指定します。例えば、検索の平均応答時間が指定した限度を超える場合にアラート通知されるように設定することができます。
  - また、特定のメッセージあるいは特定タイプのメッセージがログに記録された場合に、必ず E メールを受信するようにするオプションを指定できます。
  - 文書がエンタープライズ・サーチ・システムからドロップされた時間、場所、理由を判別するためのロギング情報のオプションを指定します。

## コンポーネントの始動

クローリング対象のデータ・ソース、およびデータ収集、検索に関するオプションを指定すると、コレクション作成プロセスを開始することができます。コンポーネントを開始する順序は重要です。クローラーはデータが構文解析および索引付けされる

前にデータをクローलする必要があり、主索引は検索サーバーが検索要求を処理する前に作成される必要があります。

## 「外部ソース」ビュー

データ・ソースをクローलまたは索引付けせずに検索する場合は、ツールバーの「外部ソース」をクリックして、データ・ソースを検索可能にするオプションを指定できます。ご使用の Java Database Connectivity (JDBC) データベースおよび Lightweight Directory Access Protocol サーバーをエンタープライズ・サーチでアクセスできるように設定する情報を指定する必要があります。外部ソースと検索アプリケーションを関連付けると、クローल、構文解析、索引付けされたデータを使用してコレクションを検索できると同時に、関連付けた外部ソースを検索することができます。

## 「システム」ビュー

エンタープライズ・サーチ管理者の役割を担当している場合、ツールバーの「システム」をクリックして、以下のアクティビティーを行うことができます。コレクション管理者、オペレーター、およびモニター担当者は、エンタープライズ・サーチ管理者によってアクセス権限が付与されている場合にのみ、このビューにアクセスできます。

- カスタム・テキスト分析エンジンをシステムに追加する。
- システムに、同義語、ストップワード、ランキング調整ワードに関するカスタム辞書を追加する。
- 並行して索引を作成できるコレクションの数を指定する。また、単一のコレクションに対する主索引作成を差分索引作成と並行して実行できるかどうかを指定する。
- システム・レベルのイベントに対するアラートを構成する。
- システム・レベル・セッションで生成されるロギング・メッセージのオプションを指定する。
- メール・サーバーに関する情報を指定して、エンタープライズ・サーチ・アクティビティーに関する E メールを受信できるようにする。

## 「セキュリティ」ビュー

エンタープライズ・サーチ管理者の役割を担当している場合、「セキュリティ」をクリックしてセキュリティ・オプションを指定することができます。コレクション管理者、オペレーター、およびモニター担当者は、このビューにはアクセスできません。


WebSphere Application Server でセキュリティを有効に設定している場合は、「セキュリティ」ビューを使用して管理役割を構成することができます。管理役割を構成することで、システム管理を実施できるユーザーを増やすと同時に、それぞれのユーザーがアクセスできる対象を、特定の機能およびコレクションに制限することができます。

また、「セキュリティ」ビューを使用して、ID 管理オプションの構成も行います。例えば、照会処理中にユーザーの現行のクレデンシャルを検証するのに使用できるプロファイルに、ユーザー・クレデンシャルを保管するためのオプションを指

定できます。検索されるソースがシングル・サインオン (SSO) セキュリティーで保護される場合、SSO 認証方式を使用して照会処理中にユーザーの現行のクレデンシャルを検証するためのオプションも指定できます。

独自の検索アプリケーションを作成するまで、サンプル検索アプリケーションを使用して、あらゆるコレクションおよび外部ソースの検索ができます。カスタム検索アプリケーションの作成後は、「セキュリティ」ビューを使用して、アプリケーションを、アプリケーションで検索可能なコレクションおよび外部ソースに関連付けてください。

## 「モニター」ビュー

 **モニター** をクリックすると、システムまたはコレクション・コンポーネントをいつでもモニターできます。管理役割で権限が与えられていれば、モニターするだけでなく、コンポーネント・プロセスの開始および停止もできます。

### 関連概念

43 ページの『クローラーの管理』

335 ページの『エンタープライズ・サーチ・アクティビティのモニター』

### 関連タスク

327 ページの『エンタープライズ・サーチ・システムの開始』

332 ページの『スタンドアロン・モードでの検索サーバーの管理』

330 ページの『エンタープライズ・サーチ・システムの停止』

35 ページの『コレクション・ウィザードを使用したコレクションの作成』

37 ページの『「コレクション」ビューを使用したコレクションの作成』

---

## 管理コンソールへのログイン

エンタープライズ・サーチ・システムを管理するには、Web ブラウザーで URL を指定した後で、管理コンソールにログインします。

### 始める前に

ログインに使用するユーザー ID は、エンタープライズ・サーチ管理コンソールへのアクセスを認可されている必要があります。

- WebSphere Application Server でグローバル・セキュリティを使用可能にしていない場合、管理コンソールにアクセスできるのは、OmniFind Enterprise Edition のインストール時に指定されたエンタープライズ・サーチ管理者だけです。
- WebSphere Application Server でグローバル・セキュリティを使用可能に設定している場合、エンタープライズ・サーチ管理コンソールを使用して、管理役割を構成できます。構成するユーザー ID は、WebSphere Application Server ユーザー・レジストリー内に存在していなければなりません。管理役割を構成すると、より多くのユーザーが管理コンソールにログインするのを許可することになりますが、各管理ユーザーがアクセス可能な機能およびコレクションは制御することができます。

### 手順

エンタープライズ・サーチ管理コンソールにログインするには、次のようにします。

1. Web ブラウザーで、管理コンソールの URL を入力します。次に例を示します。

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

*SearchServer.com* は、エンタープライズ・サーチ用の検索サーバーのホスト名です。

ご使用の Web サーバー構成によっては、ポート番号を指定する必要がある場合もあります。次に例を示します。

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

2. 「ようこそ」ページで、ユーザー ID とパスワードを入力し、「ログイン」をクリックします。

システムおよびコレクションの管理の入り口点である「コレクション」ビューが表示されます。管理役割を使用する場合、実行できるアクションと表示されるコレクションは、使用する管理役割によって決まります。

一定の時間、セッションが非アクティブな状態が続くと、システムによって自動的にログアウトされます。システムの管理を続けるには、もう一度ログインしてください。

コレクションの管理の終了後、「ログアウト」をクリックすると、コンソールをログアウトできます。その後、別の ID とパスワードを使ってログインすることもできますし、Web ブラウザーを閉じて管理コンソールを終了することもできます。

#### 関連概念

288 ページの『管理役割』

---

## シングル・サーバー構成のエンタープライズ・サーチ管理者パスワードの変更

エンタープライズ・サーチ管理者のパスワードは、暗号化されたフォーマットで保管されます。パスワードを変更するには、**eschangepw** スクリプトを使用してください。

#### 制約事項

パスワードには、以下の特殊文字を使用できます。

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

AIX<sup>®</sup>、Linux<sup>®</sup>、Solaris システムで、特殊文字を含むパスワードを指定する場合は、パスワード全体を単一引用符で囲んでください。

例: 'mypwd@\$%'

Windows システムで、特殊文字を含むパスワードを指定する場合は、パスワード全体を二重引用符で囲んでください。



例: "my?+!pwd"

## このタスクについて

エンタープライズ・サーチ管理者 ID の初期パスワードは、OmniFind Enterprise Edition のインストール時に指定されます。

このパスワードを変更するためには、**eschangepw** スクリプトを実行して、エンタープライズ・サーチ・システム全体に変更を周知させる必要があります。インストール・プログラムでは、**eschangepw** スクリプトで使用できる、次の 2 つの環境変数を生成します。

### ES\_INSTALL\_ROOT

エンタープライズ・サーチのインストール・ディレクトリー

### ES\_NODE\_ROOT

エンタープライズ・サーチのデータ・ディレクトリー。エンタープライズ・サーチ管理者 ID のパスワードは、このディレクトリーの **es.cfg** ファイルに保管されます。

## 手順

シングル・サーバー構成のエンタープライズ・サーチ管理者パスワードを変更するには、以下のようにします。

1. エンタープライズ・サーチ管理者でログインし、エンタープライズ・サーチ・システムを停止します。

```
esadmin system stopall
```

**重要:** システムが停止しているときには、検索要求を実行できません。

2. オペレーティング・システムのコマンド (AIX、Linux、または Solaris の場合)、またはパスワード変更機能 (Windows の場合) を使用して、エンタープライズ・サーチ管理者のユーザー ID のシステム・パスワードを変更します。
3. 以下のスクリプトを実行します。ここで、*newValue* は、ステップ 2 で指定したパスワードです。

### AIX、Linux、または Solaris

```
eschangepw.sh newValue
```

### Windows

```
eschangepw newValue
```

4. エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```

## 関連資料

409 ページの『エンタープライズ・サーチ・コマンド、戻りコード、およびセッション ID』

## マルチサーバー構成のエンタープライズ・サーチ管理者パスワードの変更

エンタープライズ・サーチ管理者のパスワードは、暗号化されたフォーマットで保管されます。このパスワードを変更するには、ご使用のエンタープライズ・サーチ・システムのすべてのコンピューターで **eschangepw** スクリプトを実行して変更してください。

### 制約事項

パスワードには、以下の特殊文字を使用できません。

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

AIX、Linux、Solaris システムで、特殊文字を含むパスワードを指定する場合は、パスワード全体を単一引用符で囲ってください。

例: `mypwd@\$%`

Windows システムで、特殊文字を含むパスワードを指定する場合は、パスワード全体を二重引用符で囲ってください。

例: "my?+!pwd"

### このタスクについて

エンタープライズ・サーチ管理者 ID のパスワードは、OmniFind Enterprise Edition のインストール時に初期設定されるもので、すべてのエンタープライズ・サーチ・サーバーで同じでなければなりません。

パスワードを変更し、その変更をエンタープライズ・サーチ・システム全体に周知させるには、エンタープライズ・サーチに使用する各コンピューターで **eschangepw** スクリプトを実行する必要があります。以下の手順では、すべてのサーバーでパスワードを変更するための順序を示しています。この順序に従う必要はありませんが、各サーバー・タイプに必要なステップを必ず完了してください。

インストール・プログラムでは、**eschangepw** スクリプトで使用できる、次の 2 つの環境変数を生成します。

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

エンタープライズ・サーチのインストール・ディレクトリー

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

エンタープライズ・サーチのデータ・ディレクトリー。エンタープライズ・サーチ管理者 ID のパスワードは、このディレクトリーの **es.cfg** ファイルに保管されます。

### 手順

マルチサーバー構成のエンタープライズ・サーチ管理者パスワードを変更するには、以下のようにします。

1. エンタープライズ・サーチ索引サーバーで、エンタープライズ・サーチ管理者としてログインし、エンタープライズ・サーチ・システムを停止します。

```
esadmin system stopall
```

**重要:** システムが停止しているときには、検索要求を実行できません。

- a. オペレーティング・システムのコマンド (AIX、Linux、または Solaris の場合)、またはパスワード変更機能 (Windows の場合) を使用して、エンタープライズ・サーチ管理者のユーザー ID のシステム・パスワードを変更します。
- b. 以下のスクリプトを実行します。ここで、*newValue* は、ステップ 1a で指定したパスワードです。

**AIX、Linux、または Solaris**

```
eschangepw.sh newValue
```

**Windows**

```
eschangepw newValue
```

2. エンタープライズ・サーチ・システム内の他のコンピューターで以下のステップを行います。

- a. エンタープライズ・サーチ管理者としてログインします。
- b. エンタープライズ・サーチの共通通信層 (CCL) を停止します。

**AIX、Linux、または Solaris**

```
stopccl.sh
```

**Windows コマンド・プロンプト**

```
stopccl
```

**Windows サービス管理ツール**

- 1) 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」 により、Windows サービスを起動します。
  - 2) 「**IBM OmniFind Enterprise Edition**」 チャートの追加を右クリックして、「停止」をクリックします。
- c. オペレーティング・システムのコマンド (AIX、Linux、または Solaris の場合)、またはパスワード変更機能 (Windows の場合) を使用して、エンタープライズ・サーチ管理者のユーザー ID のシステム・パスワードを変更します。このパスワードは、ステップ 1a で指定したパスワードと一致している必要があります。
  - d. 以下のスクリプトを実行します。ここで、*newValue* は、ステップ 1a で指定したパスワードです。

**AIX、Linux、または Solaris**

```
eschangepw.sh newValue
```

**Windows**

```
eschangepw newValue
```

- e. CCL を再始動します。

**AIX、Linux、または Solaris**

```
startccl.sh -bg
```

**Windows コマンド・プロンプト**

```
startccl
```

**Windows サービス管理ツール**

バックグラウンドで CCL を開始するには、次のようにします。

- 1) 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」 により、Windows サービスを起動します。
  - 2) 「IBM OmniFind Enterprise Edition」チャートの追加を右クリックして、「プロパティ」チャートの追加をクリックします。
  - 3) 「ログオン」タブをクリックします。
  - 4) 新規パスワードの値を指定してパスワードを変更してから、「OK」をクリックします。
  - 5) 「IBM OmniFind Enterprise Edition」チャートの追加を右クリックして、「開始」をクリックします。
3. エンタープライズ・サーチ索引サーバーで、エンタープライズ・サーチ管理者としてログインし、エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```

#### 関連資料



暗号化された管理者パスワードをすべてのサーバーで同じに設定

409 ページの『エンタープライズ・サーチ・コマンド、戻りコード、およびセッション ID』

## エンタープライズ・サーチに使用される TCP ポート番号

リソースの構成や他のアプリケーションへのポート番号の割り当てでポートの競合が発生しないように、エンタープライズ・サーチ・システムで使用されるデフォルトのポート番号を確認します。

ファイアウォールを構成した場合、特定のポート番号へのアクセスを明示的に使用可能にする必要があります。また、すべてのエンタープライズ・サーチ・サーバーがファイアウォールの内部にあるようにしてください。

表 1. エンタープライズ・サーチのポート番号の構成

ポート名	ポート番号	構成場所
共通通信層 (CCL)	6002	すべてのエンタープライズ・サーチ・サーバー上の ES_NODE_ROOT/nodeinfo/es.cfg および ES_NODE_ROOT/master_config/nodes.ini
検索サーバーの HTTP	80	検索サーバーの HTTP_SERVER_ROOT/conf/http.conf
WebSphere Application Server バージョン 6 管理コンソール	9060	検索サーバー
WebSphere Application Server バージョン 5.1 管理コンソール	9090	検索サーバー
DB2 クローラー	6000、6001、6002、60003、50000	クローラー・サーバー
インフォメーション・センター	8888	検索サーバー

表 1. エンタープライズ・サーチのポート番号の構成 (続き)

ポート名	ポート番号	構成場所
CCL、ファイル転送 (ESFTP)、および索引コピー用の匿名または動的ポート	49152 から 65535	すべてのエンタープライズ・サーチ・サーバー
Apache Derby Network Server	1527	クローラー・サーバー
カスタム通信	8890	クローラー・サーバー
DB2 サーバーへのリモート・クライアント接続 (OmniFind Enterprise Edition バージョン 8.4 より前のリリースでのみ使用)	50000	クローラー・サーバー
WebSphere Information Integrator Content Edition FastObjects データベース	6001 (代替として 6002)	WebSphere Information Integrator Content Edition 管理コンソール
WebSphere Information Integrator Content Edition リモート・メソッド呼び出し (RMI) プロキシ・コネクター	1251 (RMI ポート)	WebSphere Information Integrator Content Edition 管理コンソール

## エンタープライズ・サーチ・システムのポート番号の変更

エンタープライズ・サーチ・システムが通信用に使用するポート番号が、別の製品が使用するポート番号と競合する場合、エンタープライズ・サーチのポート番号を変更する必要があります。

### このタスクについて

エンタープライズ・サーチ・システムのポート番号は、OmniFind Enterprise Edition のインストール時に指定されます。マルチサーバー構成では、すべてのサーバーで同じポート番号が指定されます。

ポート番号が使用不能の場合 (例えば、そのポート番号が同じサーバー上の別の製品に割り当てられている場合など)、競合が原因となって、CCLServer\_date.log ファイルに以下のエラー・メッセージが出力されます。ここで、date は、ログ・ファイルの作成日を示しています。

```
FFQ00273W 内部警告が発生しました。例外メッセージ: {0}
at java.net.PlainSocketImpl.socketBind(Native Method)
at java.net.PlainSocketImpl.bind(PlainSocketImpl.java:357)
at java.net.ServerSocket.bind(ServerSocket.java:341)
at java.net.ServerSocket.<init>(ServerSocket.java:208)
at java.net.ServerSocket.<init>(ServerSocket.java:120)
```

### 手順

エンタープライズ・サーチが使用するポート番号を変更するには、次のようにします。

1. ポート番号を変更する必要があるコンピューターに移動し、エンタープライズ・サーチ管理者でログインして、エンタープライズ・サーチ・システムを停止します。

```
esadmin system stopall
```

**重要:** システムが停止しているときには、検索要求を実行できません。

2. ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルを編集し、以下のプロパティーを付けて新しいポート番号値を指定し、ファイルを保存してクローズします。

```
CCLPort=new_port_number
```

3. エンタープライズ・サーチの共通通信層 (CCL) を再始動します。

**AIX、Linux、または Solaris**

```
startccl.sh
```

**Windows コマンド・プロンプト**

```
startccl
```

**Windows サービス管理ツール**

バックグラウンドで CCL を開始するには、次のようにします。

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」により、Windows サービスを起動します。
  - b. 「**IBM OmniFind Enterprise Edition**」チャートの追加を右クリックして、「開始」をクリックします。
4. 索引サーバーに移動し、エンタープライズ・サーチ管理者でログインして CCL を停止します。

**AIX、Linux、または Solaris**

```
stopccl.sh
```

**Windows コマンド・プロンプト**

```
stopccl
```

**Windows サービス管理ツール**

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」により、Windows サービスを起動します。
  - b. 「**IBM OmniFind Enterprise Edition**」チャートの追加を右クリックして、「停止」をクリックします。
5. 次のようにして、ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルを編集します。
    - a. 以下のプロパティーを見つけます。ここで、*computer\_name* は、ステップ 2 でポート番号を変更したコンピューターの名前です。 *nodeN* プロパティー内の *N* は、エンタープライズ・サーチ・サーバーを識別する番号です。

```
nodeN.destination=computer_name
```
    - b. 以下のサブプロパティーを見つけ、ステップ 2 でサーバーに指定したのと同じポート番号を指定し、ファイルを保存してクローズします。

```
nodeN.port=new_port_number
```
  6. エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```

**関連資料**

409 ページの『エンタープライズ・サーチ・コマンド、戻りコード、およびセッション ID』

26 ページの『エンタープライズ・サーチに使用される TCP ポート番号』

---

## エンタープライズ・サーチ・サーバーのホスト名または IP アドレスの変更

エンタープライズ・サーチ・サーバーが使用する構成済みのホスト名および IP アドレスを変更することができます。

例えば、各エンタープライズ・サーチ・サーバーにネットワーク・インターフェース・カード (NIC) が複数あるとして、索引サーバーが低速ネットワークを使用するように構成されていることが分かった場合などに、IP アドレスを変更するとします。構成ファイル中の設定を変更すれば、索引サーバーにより高速なネットワークを使用させることができます。

**ヒント:** 構成ファイルを編集しない場合は、OmniFind Enterprise Edition を再インストールして、インストール・プログラムの実行時に新しいホスト名または IP アドレスを指定することができます。

### 手順:

ホスト名または IP アドレスを変更するには、次のようにします。

1. エンタープライズ・サーチ管理者としてログインします。複数サーバー構成の場合は、どれか 1 つのエンタープライズ・サーチ・サーバーにログインします。
2. 次のコマンドを入力してエンタープライズ・サーチ・セッションを停止します。

```
esadmin system stopall
```

3. 共通通信層 (CCL) を停止します。複数サーバー構成の場合は、次のいずれかの方法で各エンタープライズ・サーチ・サーバーの CCL を停止します。

#### AIX、Linux、または Solaris

```
stopccl.sh
```

#### Windows コマンド・プロンプト

```
stopccl
```

#### Windows サービス管理ツール

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」の順にクリックして、Windows サービスを開始します。
  - b. 「**IBM OmniFind Enterprise Edition**」を右クリックし、「停止」をクリックします。
4. 各エンタープライズ・サーチ・サーバーで以下のステップを実行します。
    - a. ES\_INSTALL\_ROOT/configurations/ccl.properties ファイルを編集し、このサーバーで使用したいホスト名または IP アドレスを **es\_server\_hostName** パラメーターで指定します。
    - b. ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルを編集し、このサーバーで使用したいホスト名または IP アドレスを **LocalHostName** パラメーターで指定します。

- c. ES\_NODE\_ROOT/master\_config/nodes.ini ファイルを編集し、ファイル中のすべての **nodeN.destination** パラメーターを、このサーバーに使用したいホスト名または IP アドレスに置き換えます。
  - d. ES\_NODE\_ROOT/config/nodes.ini ファイルを編集し、ファイル中のすべての **nodeN.destination** パラメーターを、このサーバーに使用したいホスト名または IP アドレスに置き換えます。
5. CCL を再始動します。複数サーバー構成の場合は、次のいずれかの方法で各エンタープライズ・サーチ・サーバーの CCL を再始動します。

#### AIX、Linux、または Solaris

```
startccl.sh -bg
```

#### Windows コマンド・プロンプト

```
startccl
```

#### Windows サービス管理ツール

バックグラウンドで CCL を開始するには、次のようにします。

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」の順にクリックして、Windows サービスを開始します。
  - b. 「**IBM OmniFind Enterprise Edition**」を右クリックし、「開始」をクリックします。
6. 次のコマンドを入力してエンタープライズ・サーチ・セッションを再開します。

```
esadmin system startall
```

---

## 二重 IP アドレスのサポートの構成

エンタープライズ・サーチをインストールするサーバーが二重 IP アドレスをサポートするように構成されている場合はエンタープライズ・サーチ・サーバーをその環境で実行するように手動で構成する必要があります。

例えば、デスクトップ管理者は、Microsoft SQL Server など、ある種の製品のネットワーク要件をサポートする仮想ネットワークを作成するために、Microsoft Loopback Adapter をインストールします。

#### 手順:

二重 IP アドレスをサポートするサーバーでエンタープライズ・サーチ・システムを実行できるように構成するには、次のようにします。

1. エンタープライズ・サーチ管理者としてログインします。複数サーバー構成の場合は、どれか 1 つのエンタープライズ・サーチ・サーバーにログインします。
2. 次のコマンドを入力してエンタープライズ・サーチ・セッションを停止します。

```
esadmin system stopall
```

3. 共通通信層 (CCL) を停止します。複数サーバー構成の場合は、次のいずれかの方法で各エンタープライズ・サーチ・サーバーの CCL を停止します。

#### AIX、Linux、または Solaris

```
stopccl.sh
```



### Windows コマンド・プロンプト

```
stopccl
```

### Windows サービス管理ツール

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」の順にクリックして、Windows サービスを開始します。
  - b. 「IBM OmniFind Enterprise Edition」を右クリックし、「停止」をクリックします。
4. 二重 IP アドレスをサポートする各エンタープライズ・サーチ・サーバーで、ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルを編集し、**LocalIPAddress** というパラメーターを追加します。値には、DNS が解決できる IP アドレスを割り当てます。
  5. CCL を再始動します。複数サーバー構成の場合は、次のいずれかの方法で各エンタープライズ・サーチ・サーバーの CCL を再始動します。

### AIX、Linux、または Solaris

```
startccl.sh -bg
```

### Windows コマンド・プロンプト

```
startccl
```

### Windows サービス管理ツール

バックグラウンドで CCL を開始するには、次のようにします。

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」の順にクリックして、Windows サービスを開始します。
  - b. 「IBM OmniFind Enterprise Edition」を右クリックし、「開始」をクリックします。
6. 次のコマンドを入力してエンタープライズ・サーチ・セッションを再開します。

```
esadmin system startall
```

---

## IPv6 プロトコルのサポートの使用可能化

IP バージョン 6 (IPv6) プロトコルに従ったアドレスおよび URL のサポートを使用可能にするには、IPv6 ソケット・アドレスのみを使用するように、エンタープライズ・サーチ・システムに指示する環境変数を定義する必要があります。

### 始める前に

1. %ES\_CFG% 内の **LocalHostName** プロパティおよび **LocalIPAddress** プロパティの値 (指定されている場合) が、ローカル・マシンのホスト名か有効な IPv6 アドレスのいずれかであることを確認してください。必要であれば、この値を手動で変更して、%ES\_CFG% を保存できます。
2. %ES\_CFG% 内の **DerbyServerHostName** プロパティの値がホスト名であることを確認します。この値は、IPv4 アドレスや IPv6 アドレスであってはなりません。このプロパティを変更する場合は、%ES\_CFG% ファイルを保存します。
3. 索引サーバー上の ES\_NODE\_ROOT/master\_config/nodes.ini ファイル内の **destination** および **serverhost** プロパティの値が、有効な IPv6 アドレスか有効なホスト名のいずれかであることを確認します。この値は、IPv4 アドレスであってはなりません。このファイルを変更する場合は、変更内容を保存します。

## 制約事項

IPv6 プロトコルは、Windows 2003 サーバーにインストールされたエンタープライズ・サーチ・システムでのみサポートされます。IPv6 アドレスのサポートを使用可能にすると、エンタープライズ・サーチはソケット通信に IPv4 アドレスを使用しなくなります。

### サポートされるブラウザ

Internet Explorer と Mozilla Firefox ブラウザーの IPv6 アドレスの処理は異なります。

#### Mozilla Firefox

エンタープライズ・サーチ管理コンソールまたは検索アプリケーションを実行する場合、URL に IPv6 アドレスかホスト名を指定できます。次に例を示します。

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

#### Internet Explorer

エンタープライズ・サーチ管理コンソールまたは検索アプリケーションを実行する場合、URL に IPv6 アドレスを指定できません。次の形式を使用し、DNS または `c:\windows\system32\etc\hosts` ファイル内の IPv6 アドレスにホスト名がマップされるようにする必要があります。

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

URL に IPv6 アドレスを指定すると、メッセージ「Invalid syntax error」が表示されます。この制約の詳細については、<http://support.microsoft.com/kb/325414> を参照してください。

### IPv4 データ・ソースのサポート

IPv4 サーバーで稼働するデータ・ソースは、エンタープライズ・サーチ・クローラーでサポートされます。Web Crawler を構成して、IPv6 アドレス・フォーマットで開始 URL を指定する場合は、必ず URL を大括弧で囲ってください。次に例を示します。

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

### 手順

IPv6 プロトコルのサポートを使用可能にするには、次の手順で行います。

1. エンタープライズ・サーチ索引サーバーで、エンタープライズ・サーチ管理者としてログインし、エンタープライズ・サーチ・システムを停止します。

```
esadmin system stopall
```

**重要:** システムが停止しているときには、検索要求を実行できません。

2. すべてのエンタープライズ・サーチ・サーバーの共通通信層 (CCL) サーバーを停止します。

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」 により、Windows サービスを起動します。
  - b. 「IBM OmniFind Enterprise Edition」を右クリックし、「停止」をクリックします。
3. システム環境変数に **ES\_IPV6=1** を定義します。エンタープライズ・サーチのマルチサーバー・インストールの場合は、このステップを検索サーバー上で実行します。
  4. WebSphere Application Server の Java 仮想マシン (JVM) カスタム・プロパティを更新します。エンタープライズ・サーチのマルチサーバー・インストールの場合は、これらのステップを検索サーバー上で実行します。
    - a. 実行中の場合は、WebSphere Application Server の ESSearchServer アプリケーション・サーバーを停止します。
    - b. WebSphere Application Server 管理コンソールを開いて、Java 仮想マシンのカスタム・プロパティ・パネルにナビゲートします。「サーバー」 → 「アプリケーション・サーバー」 → 「ESSearchServer」 → 「Java およびプロセス管理 (Java and Process Management)」 → 「プロセス定義」 → 「Java 仮想マシン」 → 「カスタム・プロパティ」を選択します。
    - c. 次のプロパティを構成します。
 

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

カスタム・プロパティがまだリストされていない場合は、新規プロパティを作成して、「名前」フィールドにプロパティ名、「値」フィールドに有効な値を入力します。
    - d. WebSphere Application Server バージョン 6.1 を稼働している場合、**LocalIPAddress** プロパティを ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルに追加して、その検索サーバーの IPv6 アドレスを組み込みます。
    - e. ESSearchServer アプリケーションを再始動します。
  5. WebSphere Application Server バージョン 6.1 を稼働している場合は、次のようになります。エンタープライズ・サーチのマルチサーバー・インストールの場合は、これらのステップを検索サーバー上で実行します。
    - a. IBM HTTP Server 用の httpd.conf ファイルを編集します。
    - b. 次の行の先頭から # 文字を除去して、命令をアンコメントします。これにより、IBM HTTP Server がポート 80 で IPv6 接続を listen します。
 

```
# Listen [::]:80
```
    - c. IBM HTTP Server を再始動します。
  6. すべてのエンタープライズ・サーチ・サーバーで CCL を再開します。
    - a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」 により、Windows サービスを起動します。
    - b. 「IBM OmniFind Enterprise Edition」を右クリックし、「開始」をクリックします。
  7. エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```

8. ES\_NODE\_ROOT/node/logs/cc10.log ファイルを検査して、CCL サーバーがエラーなしで始動したことを確認します。IPv6 が正しく使用可能に設定されると、以下のようなメッセージがログに記録されます。

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

9. 検索および管理アプリケーションが正しく動作していることを確認するには、次のようにします。
  - a. いずれかのコレクションについての検索サーバーを始動し、サンプル検索アプリケーションを開いて照会を実行します。このステップで、索引に検索する文書がなくても、照会の実行依頼処理が機能していることが検証されます。
  - b. 管理コンソールを開いて、問題なしにログインできることを確認します。

#### IPv6 プロトコルのサポートを使用不可にする

1. 始める前にセクションを参照し、IPv6 アドレスのすべてのインスタンスについて、その値をホスト名または IPv4 アドレスで置換します。
2. IPv6 プロトコルのサポートを使用可能にする手順に従います。ただし、次を除きます。
  - ステップ 3 (33 ページ) の **ES\_IPV6** 環境変数を除去します。
  - ステップ 4 (33 ページ) の、WebSphere Application Server の **ESSearchServer** アプリケーションに設定されたカスタム JVM プロパティを除去します。
3. ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルに IPv6 アドレスが含まれている場合は、ファイルから **LocalIPAddress** プロパティを除去します。
4. ステップ 9 の手順に従って、検索および管理アプリケーションを検証します。

---

## エンタープライズ・サーチ・コレクション

エンタープライズ・サーチ・コレクションには、ユーザーが 1 回の照会で検索できる全ソースが格納されています。フェデレーションによって、ユーザーは複数のコレクションを 1 回の照会で検索することができます。

コレクションを作成する場合、コレクション全体に適用されるオプションを指定します。コレクションは、コンテンツを追加するまでは空の状態です。

コレクションをエンタープライズ・サーチ・システムに追加する方法には、次の 2 つの方法があります。

- エンタープライズ・サーチ管理コンソールの操作に不慣れな場合、また、さまざまなコレクション・コンポーネントが連携してどのように機能するかが、まだよくわからない場合は、コレクション・ウィザードを使用してコレクションを作成してください。コレクション・ウィザードを使用すれば、作成タスクを順序良く実行することができ、ドラフト・コレクションで作業する手間が省けます。
- 管理コンソールの操作に慣れている場合は、「コレクション」ビューで管理を行う特定のページを選択しながらコレクションを作成する方法もあります。

コレクションの作成後は、「コレクション」ビューのコントロールを使用して、コレクション、エンタープライズ・サーチ・システム、およびセキュリティー・オプションの編集とモニターをすることができます。

### コレクション・フェデレーション

フェデレーションのサポートが検索アプリケーションに組み込まれて構築されている場合、ユーザーは同時に複数のコレクションを検索できます。また、フェデレーションにより、コレクションのサイズ制限 (コレクションあたり 2,000 万 (20,000,000) 文書) を拡大することができます。例えば、それぞれ 2,000 万 (20,000,000) 件の文書を含む 2 つのコレクションを検索できます。

検索の質は、マージされて最終的な結果セットを構成する前の、個別のコレクションで生成されたスコアに依存します。結果は、2 つの別々の検索を実行してから、その結果をマージしてランキングした場合と同じになります。

#### 関連タスク

336 ページの『コレクションのモニター』

---

## コレクション・ウィザードを使用したコレクションの作成

エンタープライズ・サーチにまだ慣れていなければ、ウィザードのガイドに従いながら、コレクションを作成できます。ウィザードは、処理の各ステップの詳細を提供し、途中で設定を保存できます。

### 始める前に

コレクションを作成するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

コレクションにコンテンツを追加したり、コレクション内のコンテンツの構文解析、索引付け、または検索方法のオプションを指定したりするには、エンタープライズ・サーチ管理者であるか、そのコレクションのコレクション管理者でなければなりません。

### 制約事項

コレクション・ウィザードを使用して次のクローラー・タイプを作成できます。

- Content Edition
- DB2
- DB2 Content Manager
- Exchange Server
- Notes
- UNIX ファイル・システム
- Web
- Windows ファイル・システム

### このタスクについて

コレクションの作成途中で、それをドラフト状態として保存できます。それがドラフト状態にある間、コレクションを管理する権限を持つ管理者は、それを変更することができます。例えば、Lotus Notes ソースに詳しいコレクション管理者が Lotus Notes クローラーを構成します。その後、ドラフト状態のそのコレクションを、UNIX システムに詳しいコレクション管理者が編集して UNIX ファイル・システムクローラーを構成します。


### 手順

コレクション・ウィザードを使用してコレクションを作成するには、次のようにします。

1. 「**コレクション**」をクリックして「**コレクション**」ビューを開きます。
2. 「**コレクション・ウィザード**」をクリックします。
3. ウィザードの指示に従って、空のコレクションを作成し、それにコンテンツを追加します。

コレクションについての一般情報を構成し、少なくとも 1 つのクローラーを作成しなければなりません。その他の構成オプションについては、デフォルト値を受け入れることも、あるいは、新規コレクションのオプションを指定することもできます。

4. コレクションの作成が完了する前にコレクションを保存するには、「**ドラフトとして保存**」をクリックします。

そのコレクションが他のドラフト・コレクションと一緒に「**コレクション**」ビューにリストされます。コレクションのセキュリティーを有効にしてある場合、「 **コレクション・セキュリティー有効**」アイコンがコレクション名の横に表示されます。

5. 作成途中のコレクションに戻るには、「**コレクション**」ビューで、「**ウィザードに戻る**」をクリックします。

## 6. 「完了」をクリックして、コレクションを作成します。

新規コレクションが他のコレクションと一緒に「コレクション」ビューにリストされます。

コレクションを作成したあと、コレクションのクローल、構文解析、索引付け、および検索のプロセスを開始しなければなりません。コレクションを検索できる検索アプリケーションにコレクションを関連付ける準備ができるまで、サンプル検索アプリケーション (名前は Default) を使用して、新規コレクションを検索できます。

---

## 「コレクション」ビューを使用したコレクションの作成

「コレクション」ビューを使用して、空のコレクションを作成します。その後、このコレクションを編集して、コレクションにデータを追加したり、コレクションを検索可能にするオプションを指定できます。

### 始める前に

コレクションを作成するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

コレクションにコンテンツを追加したり、コレクション内のコンテンツの構文解析、索引付け、または検索方法のオプションを指定したりするには、エンタープライズ・サーチ管理者であるか、そのコレクションのコレクション管理者でなければなりません。

### このタスクについて

新規コレクションに指定できる値の情報は、コレクションを作成している間に「ヘルプ」をクリックすると表示されます。

### 手順

「コレクション」ビューからコレクションを作成するには、次のようにします。


1. 「コレクション」ビューで、「**コレクションの作成**」をクリックします。
2. 「コレクションの作成」ページで、次のフィールドに情報を入力するか、選択を行います。
  - **コレクション名**。コレクションのコンテンツまたは目的の記述名を指定します。
  - **コレクション・セキュリティ**。コレクションのセキュリティを有効にするかどうかを指定します。コレクションを作成した後、この値を変更することはできません。コレクション・セキュリティが有効であれば、あとで、文書レベルのアクセス制御を実行するためのオプションを指定できます。
  - **文書の重要度 (静的ランキング・モデル)**。検索結果での文書のランク付けに使用される静的ランキング要素を割り当てるための方法を指定します。コレクションの作成後、この値を変更することはできません。
  - **カテゴリー化タイプ**。文書を、所属するカテゴリーで検索できるようにするかどうかを指定します。
  - **使用する言語**。コレクション内の文書検索のデフォルト言語を指定します。

3. 次のフィールドについては、デフォルト値を受け入れることも、あるいは、このコレクションで使用するオプションを指定することもできます。

- **説明。** デフォルトでは、説明は作成されません。
- **見積もり文書数。** コレクションのデフォルトの見積もり文書数は 1,000,000 です。システムは、この値を使用して、コレクションのメモリーとディスク・リソースを見積もります。コレクションのサイズを制限することはありません。
- **コレクション・データの場所。** コレクション関連ファイルのデフォルトの場所は、索引サーバー上にあります。コレクションの作成後、この値を変更することはできません。
- **コレクション ID。** デフォルトのコレクション ID は、コレクション名に基づきます。コレクションの作成後、この値を変更することはできません。カスタムのコレクション ID を指定すると、検索アプリケーションは、システムが作成する意味の分かりにくい ID の代わりに、この ID でコレクションを呼びます。
- **N-gram セグメンテーション。** デフォルトのセグメンテーション方式は、Unicode ベースの、空白文字セグメンテーションです。このオプションは、コレクションに中国語、日本語、または韓国語の文書が含まれており、ワードを区切るのにパーサーに N-gram セグメンテーションを使用させたい場合にのみ、選択します。コレクションの作成後、この値を変更することはできません。

エンタープライズ・サーチ・コレクションの完全な N-gram 構文解析およびトークン化のサポートを構成する方法の詳細については、<http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088> を参照してください。

4. 「OK」をクリックします。

「コレクション」ビューに、新規コレクションが、エンタープライズ・サーチ・システム内の他のコレクションと一緒にリストされます。コレクションのセキュリティを有効にしてある場合、「 コレクション・セキュリティ有効」アイコンがコレクション名の横に表示されます。

コレクションは、コンテンツを追加するまでは空の状態です。新規コレクションにコンテンツを追加するには、「コレクション」ビューでコレクションを選択し、それを編集し、少なくとも 1 つのクローラーを作成します。そして、データの構文解析、索引付け、および検索の方法を指定します。

そして、コレクションのクローラー、構文解析、索引付け、および検索のプロセスを開始しなければなりません。カスタム検索アプリケーションを使用する準備を終えるまでは、新規コレクションの検索にサンプル検索アプリケーションを使用できません。

---

## コレクションの編集

コレクションに入れる文書に関する情報を指定するために、コレクションを編集します。

始める前に




コレクションを編集するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

コレクションの編集では、データ・ソースのクロール、文書の構文解析、索引の作成、索引付きコンテンツの検索、およびエラー・メッセージのログ記録のオプションを指定します。コレクションを作成するときに、コレクションを編集して、それにコンテンツを追加しなければなりません。あとでコレクションを編集して、コンテンツを更新したり、情報のクロール、構文解析、索引付け、検索、またはログの方法を変更したりできます。

### 手順

コレクションを編集するには、次のようにします。

1. ツールバーの中の「**コレクション**」をクリックして「コレクション」ビューを開きます。
2. コレクションのリストで編集したいコレクションを指して、「 **編集**」をクリックします。
3. 次のページで変更を行います。

#### クロール

クロールするデータ・ソースを指定し、コンテンツのクロール方法のオプションを指定します。すべてのコレクションは、それぞれ少なくとも 1 つのクロールを含む必要があり、1 つのコレクションに、複数のタイプのデータ・ソースからのデータを含めることができます。コレクションに含めるデータ・ソースのタイプごとに、少なくとも 1 つのクロールを構成しなければなりません。

#### 構文解析

クロールされた文書の構文解析および分析方法のオプションを指定します。カテゴリを構成することで、ユーザーがコレクションのサブセットを検索できるようにすることができます。また、XML 文書および HTML 文書の特定の部分をユーザーが検索できるようにする規則を構成できます。カスタム・テキスト分析エンジンをエンタープライズ・サーチ・システムに追加する場合、その 1 つをこのコレクション内のコンテンツの分析と注釈付けに使用するよう選択できます。フィールドをランキング調整クラスに関連付けて、検索結果における文書のランキング調整を制御することもできます。

#### 索引

索引全体を作成するスケジュールと、新規および変更されたコンテンツで索引を更新するスケジュールを指定します。照会でワイルドカード文字を使用するためのオプション、コレクションのビューを一定範囲の URI に制限するためのオプション、同じ Web サイトからの検索結果を縮小表示するためのオプション、索引から URI を除去するためのオプションも構成できます。

#### 検索

検索キャッシュの構成、検索言語の選択といった、コレクションの検索処理オプションを指定します。また、クイック・リンクを構成することもできます。これは、ユーザーが照会の中に特定の語または句を含めたときは、必ず事前に決められた URI が戻されるようにする機能です。

エンタープライズ・サーチ・システムにカスタム同義語辞書を追加した場合、それらの辞書をこのコレクションの検索に使用するよう選択できます。

- ログ** ログに記録するメッセージのタイプと、ログ・ファイルの作成と循環のオプションを指定します。特定のイベントが発生した時に通知されるようにアラートを構成でき、特定のメッセージまたは特定のタイプのメッセージがログに記録されたら必ず E メールを受信するオプションを指定できます。また、エンタープライズ・サーチ・システムからドロップされた文書に関する情報をログに記録するオプションを指定することもできます。
- 一般** コレクションに関する一般情報を指定し、変更できない設定を確認します。コレクションの作成時にそのコレクションに対してセキュリティーを有効にすると、文書レベルのセキュリティーのオプションを構成することができます。

---

## コレクションの削除

コレクションを削除すると、そのコレクションに関するすべての情報がエンタープライズ・サーチ・システムから除去されます。

### 始める前に

コレクションを削除するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

コレクションを削除するには、その前に、そのコレクションに関連付けられたすべてのプロセスを停止しなければなりません。


### このタスクについて

コレクションの削除は、処理に時間がかかることがあります。コレクションの削除を確認すると、システムは、システムにあるコレクションに関連するすべてのデータを削除します。

**ヒント:** プロセスがバックグラウンドでまだ実行されているにもかかわらず、要求された操作がタイムアウトになったというメッセージが表示される場合があります。タスクが完了したかどうかを判断するには、管理コンソールで「リフレッシュ」をクリックします (Web ブラウザーの「リフレッシュ」はクリックしないでください)。削除処理は、コレクション名がコレクション・リストに表示されなくなったときに完了します。

### 手順

コレクションを削除するには、次のようにします。

1. 「**コレクション**」をクリックして「**コレクション**」ビューを開きます。
2. コレクションのリストで、削除するコレクションを指して、「 **削除**」をクリックします。

---

## コレクション ID の判別

多くの管理用タスクでは、コレクション ID を認識する必要があります。

### 始める前に

コレクション ID を表示するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### このタスクについて

コレクションを作成する場合は、ID の値を指定することも、システムから ID を割り当てすることもできます。コレクションが作成されてからコレクション ID を判別するには、管理コンソールを使用するか構成ファイルを表示します。

### 手順

1. 管理コンソールを使用してコレクション ID を判別するには、次のようにします。
  - a. 「コレクション」をクリックして「コレクション」ビューを開きます。
  - b. 「一般」ページで、「コレクション設定の表示」をクリックします。

「コレクション設定(Collection Settings)」ページに、コレクション ID、コレクション・データが保管された場所の完全修飾パス、このコレクション内の文書で使用された静的ランキング・モデルが表示されます。

2. 構成ファイルを表示してコレクション ID を判別するには、次のようにします。
  - a. `ES_NODE_ROOT/master_config/collections.ini` ファイルを開きます。見やすくするために、このファイルをソートします。次の出力例では、コレクション ID は `coll` です。

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



---

## クローラーの管理

コレクションに入れるデータ・タイプに応じて、さまざまなタイプのクローラーを構成できます。1 つのコレクションに任意の数のクローラーを入れることができます。

**ヒント:** オンライン・チュートリアルが <http://www.ibm.com/developerworks/edu/dm-dw-dm-0503buehler-i.html> で利用できます。このチュートリアルには、インストールおよび構成手順の説明のほか、さまざまなタイプのデータ・ソースを検索する方法が示されています。また、当該製品のアプリケーション・プログラミング・インターフェースを使用してエンタープライズ・サーチを拡張する方法についての説明もあります。このチュートリアルは OmniFind Enterprise Edition の旧バージョンに対応するものですが、概念や手順の多くは適用可能です。

小規模、中規模、および大規模組織においてクローラーを構成し、セキュリティーを有効にする方法の詳細例については、IBM Redbook 「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」を参照してください。

### クローラーの構成

クローラーの作成、編集、および削除には、エンタープライズ・サーチ管理コンソールを使用します。クローラーの構成は、通常、クロール対象のデータ・タイプに精通した担当者が行います。例えば、Lotus Notes データ・ソースをクロールするクローラーをセットアップする場合、コレクション管理者には Lotus Notes 管理者がなるか、または、クロール対象のデータベースを十分理解している担当者と緊密な協力を得ながら作業を行います。

クローラーを作成する場合、クロール対象のデータ・タイプ専用のウィザードがあり、これを使用するとクローラーによるシステム・リソースの使用を制御するプロパティーを指定できます。また、このウィザードによって検索対象のソースを選択することもできます。

既存のクローラーは、いつでも変更できます。クローラー・プロパティーの編集またはクロール・スペースの一部変更を必要に応じて行うことができます。クローラー・ウィザードを使用すれば、これらの変更を行うこともできます。

### 新規のクローラーにベースとなる値を移植する

クローラーは、システム・デフォルト値を使用するか、または同じタイプの別のクローラーに指定されている値をコピーして作成することができます。既存のクローラーを新規クローラーのベースとして使用すると、類似したプロパティーを持つ複数のクローラーを手早く作成でき、それらのクローラーを構成して異なるソースをクロールするようにしたり、また、例えば、異なるクロール・スケジュールで運用したりすることができます。

同一のクローラーをコピーすれば、同じクロール規則に従う複数のクローラーにクロールの負荷を分散することができます。例えば、異なる Lotus Notes サーバーで同じプロパティーおよびフィールド・クロール規則を使用するために、Lotus Notes

クローラーをコピーすることが考えられます。この場合、各クローラーで異なるのは、クロール対象のデータベースと文書レベルのセキュリティー設定だけとなります。

## 単一コレクション内で異なるクローラー・タイプを組み合わせる

エンタープライズ・サーチ・クローラーは、特定タイプのデータ・ソースから情報を収集するように設計されています。コレクションに対してクローラーを構成する場合、さまざまな異なるデータ・ソース・タイプをどのように組み合わせればユーザーが簡単にエンタープライズ・データを検索できるか、その組み合わせを決定する必要があります。例えば、ユーザーが Microsoft Windows ファイル・システムと Microsoft Exchange Server のパブリック・フォルダーを 1 回の照会で検索できるようにするには、Windows ファイル・システム・クローラーと Exchange Server クローラーを含むコレクションを作成します。

1 つのコレクションに複数のタイプのクローラーを組み合わせる場合、必ずすべてのクローラーが同一の静的ランキング方式を使用できることを確認してください。(静的ランキング方式は、コレクション作成時に指定します。) 例えば、Web ソース (ランキング係数として文書リンクを使用) と NNTP ソース (ランキング係数として、通常、文書日付を使用) を組み合わせると、検索結果の品質が低下する可能性があります。

## 文書レベルのセキュリティーの構成

コレクションの作成時に、そのセキュリティーを有効にすると、文書レベルのセキュリティー・オプションを構成することができます。各クローラーは、クロール対象の文書にセキュリティー・トークンを関連付けることができます。クローラーの構成時に文書レベルのセキュリティーを使用することを指定すると、そのクローラーは指定したセキュリティー・トークンを各文書に関連付けます。そして、これらのトークンは文書の索引に追加されます。

カスタム検索アプリケーションでセキュリティーを有効にした場合、そのアプリケーションはクローラーが文書に関連付けたセキュリティー・トークンを使用してユーザーを認証します。この機能を使えば、コレクション内の一部の文書をアクセス制限し、他の文書はすべてのユーザーが検索できるように設定することができます。例えば、あるコレクションで、Microsoft Exchange Server のパブリック・フォルダー内の全文書をすべてのユーザーがアクセスできるようにし、Lotus Notes データベース文書にアクセスできるのは特定のユーザー ID のユーザーだけに限定するように設定できます。

規則を Java クラスでエンコードすることにより、セキュリティー・トークンの値の判別にカスタム・ビジネス・ルールを適用することができます。クローラー・プロパティーを構成するときに、文書をクロールする際にクローラーに使用させたいプラグインの名前を指定します。プラグインが追加するセキュリティー・トークンは、索引に保管され、文書へのアクセス制御に使用されます。

特定のタイプのクローラーを構成する場合、さらにセキュリティー制御を追加指定することができます。例えば、照会処理中にユーザーの検証を行うように指定することができます。このオプションを設定すると、ユーザーのクレデンシャルは検索対象のデータ・ソースで維持されている現在のアクセス制御リストと比較されま

す。この現行のクレデンシャルに関する検証は、エンタープライズ・サーチ索引に保管されているセキュリティー・トークンに基づく検証の代わりに、あるいはその検証に加えて行うことができます。

#### 関連概念

292 ページの『文書レベルのセキュリティー』

#### 関連タスク

338 ページの『クローラーのモニター』

#### 関連資料

310 ページの『セキュリティーをサポートするためのクローラーのセットアップ要件』

---

## クローラーの作成

クローラーを作成するとき、作成するクローラーのタイプを指定します。コレクションに含めるデータについて、指定する情報をウィザードがガイドします。

### 始める前に

クローラーを作成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのクローラーを所有するコレクションのコレクション管理者でなければなりません。

### このタスクについて

1 つのコレクションに対して、クローラーを少なくとも 1 つ作成しなければなりません。作成するクローラーのタイプは、コレクションに含めるデータのタイプによって異なります。作成するクローラーのタイプに対応するウィザードが、クローラーのオプションの指定をガイドします。ウィザードは、例えば、クローラーがシステム・リソースをどのように使用するかについてのオプションを指定する手助けをします。また、ウィザードは、コレクションに含めるデータ・ソースの選択の手助けもします。

### 手順

クローラーを作成するには、次のようにします。

1. コレクションを編集して「クロール」ページを選択し、「クローラーの作成」をクリックします。
2. クローラー・タイプと、クローラーの基本値を選択します。
  - a. Web サイト、Lotus Notes データベース、または UNIX ファイル・システムなど、クロールするデータのタイプをサポートするクローラーのタイプを選択します。

クローラー・タイプを選択したあと、その作成方法についてのオプションが表示されます。

- b. クローラーの基本値を選択します。

#### 「システム・デフォルト値を新規クローラーに使用する」

クローラー初期設定値に、インストールのデフォルト値を設定します。

このオプションを選択した場合、「次へ」をクリックすると、新規クローラーの構成が開始されます。

#### 「既存のクローラーの値を新規クローラーにコピーする」

クローラー初期設定値に、このタイプの別のクローラーに構成されている値を設定します。

このオプションを選択すると、このクローラー・タイプに一致するクローラーのリストが表示されます。新規クローラーとして使用するクローラーを選択し、「次へ」をクリックして、新規クローラーの構成を開始します。

作成しているクローラー・タイプ用のウィザードが開きます。ウィザードのプロンプトに従って、クローラーを作成します。ウィザードのどのページでも、「ヘルプ」をクリックすると、そのタイプのクローラーに指定できるオプションの説明が表示されます。

新規クローラーは、そのコレクションに属する他のクローラーと一緒に「クローラー」ページにリストされます。クローラーを変更する必要があるときはいつでも、クローラーのプロパティおよびクローラー・スペースを編集するオプションをクリックできます。

---

## クローラーのプロパティの編集

クローラーについての情報と、クローラーがデータをクローリングする方法を変更できます。例えば、クローラーがシステム・リソースを使用する方法を変更できます。

### 始める前に


クローラー・プロパティを編集するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのクローラーを所有するコレクションのコレクション管理者でなければなりません。

### このタスクについて

クローラーのプロパティを編集するときに、「ヘルプ」をクリックすると、変更できるタイプを知ることができます。編集できるプロパティは、クローラー・タイプによって異なります。

### 手順

クローラーのプロパティを編集するには、次のようにします。

1. コレクションを編集して「クローリング」ページを選択し、編集するクローラーを指して、「 クローラーのプロパティ」をクリックします。
2. クローラーのプロパティを変更して、「OK」をクリックします。
3. 変更を有効にするために、クローラーを停止し、再開します。(クローラーの説明のみを変更した場合、クローラーを再始動する必要はありません。)



---

## クローラー・スペースの編集

クローラーがクローリングするデータ・ソースについての情報を変更できます。例えば、データ・ソースの追加、データ・ソースの除去、クローラー・スケジュールの変更、および、特定のデータ・ソース内の文書のクローリング規則の変更を行うことができます。

### 始める前に


クローラー・スペースを編集するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのクローラーを所有するコレクションのコレクション管理者でなければなりません。

### このタスクについて

管理するクローラーのタイプで行える変更については、クローラー・スペースを編集している間に「ヘルプ」をクリックすると説明が表示されます。

### 手順

クローラー・スペースを編集するには、次のようにします。

1. コレクションを編集して「クローリング」ページを選択し、編集するクローラーを指して、「 クローラー・スペース」をクリックします。
2. 変更するオプションを選択して、クローラー・スペースを変更します。

選択できるオプションは、クローラー・タイプによって異なります。データ・ソースをコレクションに追加するといった一部のオプションについては、クローラー・タイプに応じたウィザードが開き、クローラー・スペースの変更を手助けします。

3. 変更を有効にするために、クローラーを停止し、再開します。

---

## クローラーの削除

クローラーを削除すると、そのクローラーに関するすべての情報がエンタープライズ・サーチ・システムから除去されます。前にそのクローラーでクローリングされた情報は、次の主索引作成が行われるまで索引内にそのまま残ります。

### 始める前に

クローラーを削除するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのクローラーを所有するコレクションのコレクション管理者でなければなりません。

### このタスクについて


クローラーの削除は、時間のかかるプロセスです。クローラーの削除を確認すると、システムは、システムにあるクローラーに関連するすべてのデータを削除します。

**ヒント:** このタスクは完了するのに時間がかかるため、プロセスがバックグラウンドでまだ実行中であっても、要求された操作のタイムアウトに関するメッセージが

表示されます。タスクが完了したかどうかを確認するには、時々、管理コンソールで「リフレッシュ」をクリックします (Web ブラウザーの「リフレッシュ」をクリックしないでください)。クローラー名がクローラーのリストに表示されなくなったら、削除プロセスは終了です。

### 手順

クローラーを削除するには、次のようにします。

1. コレクションを編集して「クロール」ページを選択します。
2. 削除するクローラーを指して「 削除」をクリックします。

---

## クローラー・スケジュール

Web ソース用に作成したクローラーは、連続的に実行されます。Web クローラーを開始した後は、通常、クローラーの構成を変更しない限り、停止する必要はありません。それ以外のすべてのクローラー・タイプについては、クローラーの構成時にクロール・スケジュールを指定します。

データ・ソース・タイプによっては、1 つのスケジュールで、クローラーがクロール・スペース内のすべてのデータ・ソースにアクセスする日時が管理されます。また、他のデータ・ソース・タイプの場合は、特定のデータ・ソースに対して複数の異なるスケジュールを指定できます。例えば、Lotus Notes データベースについて、クローラーがクロールする各データベースに異なるクロール・スケジュールを指定できます。

スケジュールを構成する際に、行われるクロールのタイプを指定します。クロール・スペース内の全文書の完全クロールをスケジュールしたり、クロール・スペースに対するすべての更新 (新規文書、変更された文書、削除された文書) を対象とするクロールをスケジュールしたり、新規および変更された文書のみを対象とするクロールをスケジュールすることができます。完全クロールは、最も時間がかかります。削除された文書を除去するクロールは、削除された文書を無視するクロールに比べて時間がかかります。

クローラーのクロール・スペースを編集する際に、2 番目のクロール・スケジュールを指定することができます。例えば、毎週土曜日の夜にクロール・スペース内のすべての文書をクロールするスケジュールを構成し、新規および変更された文書のクロールをもっと頻繁に実行する 2 番目のスケジュールを構成することができます。

複数のクローラー・スケジュールを作成すれば、クローラーのターゲット・ソースへのアクセス日時をより適切に管理できます。例えば、異なるタイム・ゾーンにある複数のデータベースをクロールする場合、ユーザーが日中の仕事を終了していると思われる時間にクローラーがクロールするように、タイム・ゾーンに合わせて複数のスケジュールを設定できます。

---

## Content Edition クローラー

IBM WebSphere Information Integrator Content Edition リポジトリをエンタープライズ・サーチ・コレクションに入れる場合は、Content Edition クローラーを構成する必要があります。

Content Edition クローラーを使用して、次のリポジトリ・タイプをクロールすることができます。

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

クローラーを構成する場合、クローラーがクロール・スペースのすべてのリポジトリをクロールするときのオプションを指定します。また、各リポジトリでクロールする項目クラスも選択します。

### ヒント:

コネクタおよびセキュアな Content Edition クローラーを構成する方法の詳細例については、IBM Redbook「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある中規模組織向けのシナリオを参照してください。

### クローラー・サーバーの構成

リポジトリをクロールするための準備は、クロールされるデータへの接続に直接モードとサーバー・モードのどちらを使用するかによって異なります。直接モードを使用する場合は、WebSphere Information Integrator Content Edition でコネクタを構成する必要があります。サーバー・モードを使用する場合は、クローラー・サーバーでスクリプトを実行してください。このスクリプトは、OmniFind Enterprise Edition で提供されており、これによって Content Edition クローラーが WebSphere Information Integrator Content Edition サーバーと通信できるようになります。

サーバー・モードを使用する場合、Content Edition クローラーを作成する前に、ご使用の環境に応じたタスクを実行してください。

- 52 ページの『UNIX での WebSphere II Content Edition 用クローラー・サーバーの構成』
- 53 ページの『Windows での WebSphere II Content Edition 用クローラー・サーバーの構成』

WebSphere Information Integrator Content Edition リポジトリを検索するようにエンタープライズ・サーチ・システムを構成する方法の詳細手順については、IBM developerWorks の記事「*Search WebSphere Portal Document Manager using WebSphere Information Integrator OmniFind Edition*」(URL: <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-0606lee/>) を参照してください。

## 構成の概要

Content Edition クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内のすべてのリポジトリをクロールする方法を制御します。
- クローラーがリポジトリへのアクセスに直接モードまたはサーバー・モードのいずれを使用するかを指定します。サーバー・モードの場合は、クローラーが Web アプリケーション・サーバーにアクセスするための情報を指定する必要があります。
- クロールするリポジトリを選択します。
- 選択されたリポジトリのコンテンツにクローラーがアクセスできるようにするためのユーザー ID とパスワードを指定します。
- リポジトリをクロールするスケジュールを設定します。
- 各リポジトリでクロールする項目クラスを選択します。
- 項目クラスのプロパティを検索可能にするときのオプションを指定します。例えば、特定タイプの文書をクロール・スペースから除外したり、特定リポジトリの特定バージョンをクロールすると指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、アクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

Documentum、FileNet Panagon Content Services、Hummingbird DM、Portal Document Manager、SharePoint 項目クラスの場合は、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。このタイプの現行クレデンシャルの検証は、他のリポジトリ・タイプでは使用できません。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

## Content Edition リポジトリへの直接モード・アクセス

直接モードで WebSphere Information Integrator Content Edition のリポジトリにアクセスできるように Content Edition クローラーを構成できます。

このタスクについて

直接モードでは、クローラーは、OmniFind Enterprise Edition のインストール時にクローラー・サーバーにインストールされる WebSphere Information Integrator Content Edition コネクタを使用します。クローラーは、コンテンツ統合 API を使用して、クロール対象のリポジトリに直接接続します。コンテンツ統合サーバーが直接モードで稼働する場合、すべてのコンテンツ統合サーバー機能が使用できるとは限りません。コンテンツ統合サーバーの直接モードでの稼働、サーバー・モードで稼働するコンテンツ統合サーバーとの機能の相違については、WebSphere Information Integrator Content Edition の資料を参照してください。

以下の手順では、直接モード・アクセスのセットアップに必要なステップを簡単に説明します。詳細手順については、IBM developerWorks の記事「*Search WebSphere Portal Document Manager using WebSphere Information Integrator OmniFind Edition*」(URL: <http://www-128.ibm.com/developerworks/db2/library/techarticle/dm-06061ee/>) を参照してください。

## 手順

クローラーがリポジトリに直接モードでアクセスするようにシステムを構成するには、次のようにします。

1. `iice_install_root/bin/config.sh` ファイル (UNIX) または `iice_install_root\bin\config.bat` ファイル (Microsoft Windows) の `VBR_HOME` および `JAVA_HOME` 環境変数で正しいディレクトリが指定されているか確認します。
2. WebSphere Information Integrator Content Edition 管理コンソールが直接モードで実行されるように構成するために、`-Dvbr.as.operationMode=direct` の Java システム・プロパティを `iice_install_root/bin/Admin.sh` ファイル (UNIX) または `iice_install_root\bin\Admin.bat` ファイル (Windows) に追加します。例えば、次のようにします。

### Admin.sh ファイル

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

### Admin.bat ファイル

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

3. WebSphere Information Integrator Content Edition 管理コンソールを直接モードで開始して、OmniFind Enterprise Edition クローラー・サーバーのコネクタを構成します。
4. エンタープライズ・サーチ管理コンソールを使用して Content Edition クローラーを構成する際に、直接モードのオプションを選択します。

## WebSphere II Content Edition リポジトリへのサーバー・モード・アクセス

Content Edition クローラーを構成して、サーバー・モードでリポジトリにアクセスすることができます。

サーバー・モードでは、クローラーがデータのアクセスに使用する WebSphere Information Integrator Content Edition コネクタが WebSphere Application Server にエンタープライズ・アプリケーションとしてインストールされ、クローラーはサーバーを介してリポジトリにアクセスします。この方法により、J2EE アプリケーション・サーバー環境の利点を活用することができます。

サーバー・モードで WebSphere Information Integrator Content Edition リポジトリにアクセスできるようにクローラーを構成する前に、クローラー・サーバーでスクリプトを実行する必要があります。このスクリプトは、OmniFind Enterprise Edition で提供されており、これによって、Content Edition クローラーがサーバー上のリポジトリにアクセスできるようになります。

ご使用の環境に応じたタスクを実行してください。

- 『UNIX での WebSphere II Content Edition 用クローラー・サーバーの構成』
- 53 ページの『Windows での WebSphere II Content Edition 用クローラー・サーバーの構成』

### UNIX での WebSphere II Content Edition 用クローラー・サーバーの構成

IBM AIX、Linux、または Solaris オペレーティング環境で稼働するコンピューターに OmniFind Enterprise Edition をインストールし、Content Edition クローラーがリポジトリのアクセスにサーバー・モードを使用するように構成する場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトによって、Content Edition クローラーは、WebSphere Information Integrator Content Edition リポジトリにアクセスできるようになります。

#### このタスクについて

Content Edition クローラーは、Java クライアントとして WebSphere Information Integrator Content Edition の Java ライブラリーを使用します。サーバー・モードで、これらの Java ライブラリーは、WebSphere Application Server の EJB 関連 Java ライブラリーを必要とします。Content Edition クローラーが Java ライブラリーと動作するようにするため、WebSphere Application Server をインストールした後、OmniFind Enterprise Edition によって提供されているセットアップ・スクリプトをクローラー・サーバー上で実行してください。

WebSphere Information Integrator Content Edition は、OmniFind Enterprise Edition のインストール時に、クローラー・サーバーにインストールされます。Content Edition クローラーをサーバー・モードで使用できるようにするには、`vbr_access_services.jar` ファイルを WebSphere Information Integrator Content Edition サーバーからクローラー・サーバーにコピーする必要があります。

#### 手順

クローラー・サーバーが WebSphere Information Integrator Content Edition リポジトリをクローリングできるように構成するには、次のようにします。

1. OmniFind Enterprise Edition がマルチサーバー構成でインストールされる場合、WebSphere Application Server Java ライブラリーをインストールおよびバインドします。
2. クローラー・サーバーで Content Edition クローラー用のセットアップ・スクリプトを実行します。
  - a. エンタープライズ・サーチ管理者としてログインします。
  - b. 以下のスクリプト (これは、`$ES_INSTALL_ROOT/bin` ディレクトリーにインストールされています) を開始し、プロンプトに応えます。

```
escrvbr.sh
```

3. クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

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

4. `vbr_access_services.jar` ファイルを WebSphere Information Integrator Content Edition サーバーからクローラー・サーバーにコピーします。

**コピー元:**

`vbr_access_services.jar` ファイルは、次のデフォルトの場所にありません。

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

*was\_install\_root* は WebSphere Application Server インストール・ディレクトリー、*server\_name* はサーバーに指定した名前、*application\_name* は WebSphere Application Server で WebSphere Information Integrator Content Edition アプリケーションに指定した名前です。

**コピー先:**

クローラー・サーバー上のターゲット・ディレクトリーは `ice_install_root/lib` です。ここで、*ice\_install\_root* は、クローラー・サーバー上の WebSphere Information Integrator Content Edition インストール・ディレクトリーです。

## Windows での WebSphere II Content Edition 用クローラー・サーバーの構成

Microsoft Windows コンピューターに OmniFind Enterprise Edition をインストールし、Content Edition クローラーがリポジトリのアクセスにサーバー・モードを使用するように構成する場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトによって、Content Edition クローラーは、WebSphere Information Integrator Content Edition リポジトリにアクセスできるようになります。

### このタスクについて

Content Edition クローラーは、Java クライアントとして WebSphere Information Integrator Content Edition の Java ライブラリーを使用します。サーバー・モードで、これらの Java ライブラリーは、WebSphere Application Server の EJB 関連 Java ライブラリーを必要とします。Content Edition クローラーが Java ライブラリーと動作するようにするため、WebSphere Application Server をインストールした後、OmniFind Enterprise Edition によって提供されているセットアップ・スクリプトをクローラー・サーバー上で実行してください。

WebSphere Information Integrator Content Edition は、OmniFind Enterprise Edition のインストール時に、クローラー・サーバーにインストールされます。Content Edition クローラーをサーバー・モードで使用できるようにするには、`vbr_access_services.jar` ファイルを WebSphere Information Integrator Content Edition サーバーからクローラー・サーバーにコピーする必要があります。

### 手順

クローラー・サーバーが WebSphere Information Integrator Content Edition リポジトリをクローラーできるように構成するには、次のようにします。

1. OmniFind Enterprise Edition がマルチサーバー構成でインストールされる場合、WebSphere Application Server Java ライブラリーをインストールおよびバインドします。
2. クローラー・サーバーで Content Edition クローラー用のセットアップ・スクリプトを実行します。
  - a. エンタープライズ・サーチ管理者 ID でログインします (このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです)。
  - b. 以下のスクリプト (これは、`%ES_INSTALL_ROOT%bin` ディレクトリーにインストールされています) を開始し、プロンプトに応答します。

```
escrvbr.vbs
```

3. クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

```
esadmin system stopall
```

```
esadmin system startall
```

4. `vbr_access_services.jar` ファイルを WebSphere Information Integrator Content Edition サーバーからクローラー・サーバーにコピーします。

#### コピー元:

`vbr_access_services.jar` ファイルは、次のデフォルトの場所にありません。

```
was_install_root%installedApps%server_name%application_name
```

*was\_install\_root* は WebSphere Application Server インストール・ディレクトリー、*server\_name* はサーバーに指定した名前、*application\_name* は WebSphere Application Server で WebSphere Information Integrator Content Edition アプリケーションに指定した名前です。

#### コピー先:

クローラー・サーバー上のターゲット・ディレクトリーは



`iice_install_root%lib` です。ここで、`iice_install_root` は、クローラー・サーバー上の WebSphere Information Integrator Content Edition インストール・ディレクトリーです。

---

## DB2 クローラー

DB2 クローラーを使用して、IBM DB2 データベースをコレクションに組み込むことができます。

IBM WebSphere Information Integrator を使用して次のデータベース・システム・タイプをフェデレートし、それに対してニックネーム表を作成すれば、DB2 クローラーにより、ニックネームで表をクロールすることができます。

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

クローラーは、クロール対象の各データベース・サーバーごとに独立したクローラーを構成する必要があります。クローラーを構成する場合、クローラーが同じサーバー上のすべてのデータベースをクロールするときのオプションを指定します。また、各データベースでクロールする特定の表も選択します。

クロール用に選択する表は、データベース表、ニックネーム表、ビューのいずれかでなければなりません。DB2 クローラーは結合表をサポートしません。

### ヒント:

セキュアな DB2 クローラーを構成する方法の詳細例については、IBM Redbook「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある大規模組織向けのシナリオを参照してください。

### クローラー・サーバーの構成

データベース表をクロールするには、その前に、クローラー・サーバーに DB2 Administration Client をインストールする必要があります。その後、クローラー・サーバーでスクリプトを実行してください。このスクリプトは、OmniFind Enterprise Edition で提供されており、これによって、DB2 クローラーがデータベース・サーバーと通信できるようになります。

エンタープライズ・サーチ管理コンソールを使用して、DB2 クローラーを構成する前に、ご使用の環境に応じたタスクを実行する必要があります。

- 57 ページの『UNIX 上での DB2 クローラー用のクローラー・サーバーの構成』
- 58 ページの『Windows 上での DB2 クローラー用のクローラー・サーバーの構成』

## イベント・パブリッシング

WebSphere Information Integrator Event Publisher Edition を使用している場合に、クローラーデータベースをパブリッシング・キュー・マップに関連付けると、DB2 クローラーは、マップを使用してデータベース表への更新をクローラーできます。

パブリッシング・キュー・マップは、データベース表に対する更新が行われると、XML メッセージを受け取る WebSphere MQ キューを識別します。クローラーは、表が更新されると、発行されたイベントに関する情報のキューを listen し、クローラー・スペースを更新します (クローラーは、最初に表をクローラーする場合、すべての文書をクローラーします)。

イベント・パブリッシングにより、新規文書および変更された文書は、クローラーがクローラー・スケジュールに従ってクローラーする文書よりも速く検索されるようになります。

一部またはすべての表がイベント・パブリッシングを使用するように構成されている場合は、クローラーの構成時に、クローラーが WebSphere MQ およびパブリッシング・キュー・マップにアクセスできるように指定することができます。

また、クローラーされるサーバーで WebSphere MQ および WebSphere Information Integrator Event Publisher Edition が構成されており、クローラー・サーバーで WebSphere MQ クライアント・モジュールが構成されていることを確認する必要があります。イベント・パブリッシングを DB2 クローラーで使用するには、以下のタスクを実行します。

- 62 ページの『DB2 クローラーのための WebSphere MQ の構成』
- 59 ページの『DB2 クローラーのための WebSphere Information Integrator Event Publisher Edition の構成』

### 構成の概要

DB2 クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティは、特定のデータベース・サーバーにあるすべてのデータベースをクローラーがクローラーする方法を制御します。
- クローラーするデータベースのタイプに関する情報を指定します。

ローカル・データベース・サーバーでカタログされていないリモート・データベースをクローラーする予定がある場合、まず、リモート・サーバーで DB2 Administration Server を始動してから、DB2 クローラーを使用してそれらのデータベースをクローラーします。また、クローラーの構成時にそのリモート・データベースのホスト名とポートを指定する必要があります。

- クローラーするデータベースを指定します。

- アクセス制御をしているデータベースにクローラーがアクセスできるようにするためのユーザー ID とパスワードを指定します。
- データベースをクロールするスケジュールを設定します。
- 各データベースでクロールする表を選択します。

**重要:** 検出プロセスのパフォーマンスを最適化するため (そしてクローラー構成プロセスがタイムアウトになるのを防止するため)、すべての表をクロールする選択は、データベース中の表の数が少ない場合か、各表に含まれる列の数が少ない場合のみ行うようにしてください。現時点では一部の表についてクロールする選択をした場合、後からクロール・スペースを編集して、さらに多くの表をコレクションに追加できます。

- イベント・パブリッシング・キューで表に対する更新が行われた際にクロールする表を選択し、クローラーがイベント・パブリッシング・キューにアクセスするための情報を指定します。
- 特定の表の列を検索可能にするときのオプションを指定します。例えば、特定の列をパラメトリック照会で使用できるようにしたり、検索結果に戻すことのできる列を指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

## UNIX 上での DB2 クローラー用のクローラー・サーバーの構成

IBM AIX、Linux、または Solaris オペレーティング環境で稼働するコンピューターに OmniFind Enterprise Edition をインストールした場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトにより、DB2 クローラーがデータベース・サーバーと通信できるようになります。イベント・パブリッシングを使用する場合、このスクリプトによってクローラーは WebSphere MQ キュー・マネージャーおよびキューへのアクセスも可能になります。

### このタスクについて

DB2 クローラーがデータベース表をクロールできるようにするには、OmniFind Enterprise Edition によって提供されているセットアップ・スクリプト `escrdb2.sh` をクローラー・サーバーで実行します。

スクリプトを実行する前に、クローラー・サーバーに DB2 Administration Client がインストールされていることを確認してください。

イベント・パブリッシングを使用する場合、DB2 クローラーが WebSphere MQ キュー・マネージャーおよびキューにアクセスできるように、クローラー・サーバーに Java Messaging 用の WebSphere MQ 5.3 モジュールをインストールする必要があります。

あります。 WebSphere MQ モジュールをインストールした後に、escrdb2.sh セットアップ・スクリプトを実行してください。

### 手順

DB2 クローラーによるクロールをサポートするようクローラー・サーバーを構成するには、次のようにします。

1. オプション: イベント・パブリッシングを使用する予定の場合、Java Messaging 用の WebSphere MQ 5.3 モジュールをクローラー・サーバーにインストールします。
  - a. root ユーザーとしてログインし、以下のコマンドを入力します。

```
export LD_ASSUME_KERNEL=2.4.19
```

- b. WebSphere MQ CD を挿入します。

- c. Java Messaging 用の MQ モジュールが入っているディレクトリーに移動します。

- d. これらのモジュールをインストールするため、以下のコマンドを入力します。

```
rpm -i MQSeriesJava-5.3.0-1.i386.rpm
```

2. クローラー・サーバーで DB2 クローラー用のセットアップ・スクリプトを実行します。
  - a. エンタープライズ・サーチ管理者としてログインします (このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです)。
  - b. 以下のスクリプト (これは、\$ES\_INSTALL\_ROOT/bin ディレクトリーにインストールされています) を開始し、プロンプトに応答します。

```
escrdb2.sh
```

3. クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

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

## Windows 上での DB2 クローラー用のクローラー・サーバーの構成

Microsoft Windows コンピューターに OmniFind Enterprise Edition をインストールした場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトにより、DB2 クローラーがデータベース・サーバーと通信できるようになります。イベント・パブリッシングを使用する場合、このスクリプトによってクローラーは WebSphere MQ キュー・マネージャーおよびキューへのアクセスも可能になります。

### このタスクについて

DB2 クローラーがデータベース表をクロールできるようにするには、OmniFind Enterprise Edition によって提供されているセットアップ・スクリプト escrdb2.vbs をクローラー・サーバーで実行します。

スクリプトを実行する前に、クローラー・サーバーに DB2 Administration Client がインストールされていることを確認してください。

イベント・パブリッシングを使用する場合、DB2 クローラーが WebSphere MQ キュー・マネージャーおよびキューにアクセスできるように、クローラー・サーバーに Java Messaging 用の WebSphere MQ 5.3 モジュールをインストールする必要があります。WebSphere MQ モジュールをインストールした後に、escrdb2.vbs セットアップ・スクリプトを実行してください。

### 手順

DB2 クローラーによるクロールをサポートするようクローラー・サーバーを構成するには、次のようにします。

1. オプション: イベント・パブリッシングを使用する予定の場合、Java Messaging 用の WebSphere MQ 5.3 モジュールをクローラー・サーバーにインストールします。
  - a. WebSphere MQ CD を挿入します。
  - b. WebSphere MQ インストーラーを開始します。
  - c. 「製品フィーチャーの選択」ウィンドウで、インストール・オプションに「**Java Messaging**」を選択します。
2. クローラー・サーバーで DB2 クローラー用のセットアップ・スクリプトを実行します。
  - a. エンタープライズ・サーチ管理者 ID でログインします (このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです)。
  - b. 以下のスクリプト (これは、%ES\_INSTALL\_ROOT%\bin ディレクトリーにインストールされています) を開始し、プロンプトに応答します。

```
escrdb2.vbs
```

3. クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

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

## DB2 クローラーのための WebSphere Information Integrator Event Publisher Edition の構成

イベント・パブリッシングを使用するように DB2 クローラーを構成する前に、クロール対象のサーバー上に IBM WebSphere Information Integrator Event Publisher Edition が構成されていることを確認してください。

### このタスクについて

DB2 クローラーと共に使用するために WebSphere Information Integrator Event Publisher Edition を構成する際には、以下のガイドラインに従ってください。

- ソース表内の変更された列と未変更の列の両方がパブリッシングに選択されなければなりません。
- ソース表内の削除された行がパブリッシングに選択されなければなりません。

- 1 つのイベント・パブリッシング・キューを複数のデータベースで共用することはできません。
- 単一のデータベースは複数のキュー・マップおよびキューを持つことができます。
- 1 つの表は、1 つのパブリッシング・キュー・マップに関連付けられた 1 つの XML パブリケーションを持っているべきです。(1 つの表が、1 つのパブリッシング・キュー・マップに関連付けられた複数の XML パブリケーションを持つべきではありません。各 XML パブリケーションがそれぞれ異なる 1 つのパブリッシング・キュー・マップに関連付けられている場合は、1 つの表が複数の XML パブリケーションを持つことができます。)

## 手順

DB2 クローラーがイベント・パブリッシング・キューにパブリッシュされる表更新にアクセスできるようにデータベース・サーバーを構成するには、次のステップを実行します (これらのステップについての不明点は、WebSphere Information Integrator Publisher Edition 資料を参照してください)。

1. クロールされるデータベース・サーバー上に WebSphere Information Integrator Event Publisher Edition をインストールします。
2. レプリケーション・センターのランチパッドを開始します。

### AIX、Linux、または Solaris

db2rc

### Windows コマンド・プロンプト

「スタート」 → 「IBM DB2 Replication Center」をクリックします。

3. Q キャプチャー制御テーブルを作成します。
  - a. ランチパッドのビューとして「**イベント発行**」を選択し、「**Q キャプチャー・コントロール表の作成**」を選択して「**次へ**」をクリックします。
  - b. 「**Q キャプチャー・サーバー**」フィールドで、選択可能なデータベース・サーバーのリストから、Q キャプチャー・サーバーとして使用するサーバーを選択し、「**OK**」をクリックします。
  - c. 選択した Q キャプチャー・サーバーへのアクセスを認可されているユーザー ID とパスワードを指定します。Q キャプチャーのスキーマを変更するか、デフォルトのスキーマ名を受け入れ、「**次へ**」をクリックします。
  - d. このデータベース・サーバー上に WebSphere MQ を構成したときに指定した、キュー・マネージャー、管理キュー、および再始動キューの名前を指定し、「**次へ**」をクリックします。
  - e. 「**完了**」をクリックします。メッセージと SQL スクリプトを示すページが表示された後、「**クローズ**」をクリックします。
  - f. 処理オプションに「**今すぐ実行する**」を選択し、「**OK**」をクリックします。SQL スクリプトが完了したことを示すメッセージが表示された後、「**クローズ**」をクリックします。
4. XML パブリケーションを作成します。
  - a. レプリケーション・センターのランチパッドで、ランチパッドのビューとして「**イベント発行**」を選択し、「**XML 発行の作成**」を選択して「**次へ**」をクリックします。

- b. 「開始」 ページで、「次へ」をクリックします。
  - c. 「サーバーとキュー・マップ」 ページで、Q キャプチャー・サーバーと Q キャプチャー・スキーマが正しいことを確認し、「パブリッシング・キュー・マップ」 フィールドの横のオプションをクリックし、パブリッシング・キュー・マップを作成するため「新規」をクリックします。
  - d. 「一般」 ページで、キュー・マップの名前を入力します。
  - e. 「プロパティ」 ページで、送信キューの名前 (例えば、このサーバー上に WebSphere MQ を構成したときに指定したデータ・キューの名前) を指定し、メッセージ内容のタイプに「行操作」または「トランザクション」を選択し、ハートビート・メッセージ送信のチェック・ボックスと JMS メッセージ・ヘッダー追加のチェック・ボックスをクリアし、「OK」をクリックします。
  - f. メッセージと SQL スクリプトを示すページが表示された後、「クローズ」をクリックします。
  - g. 処理オプションに「今すぐ実行する」を選択し、「OK」をクリックします。SQL スクリプトが完了したことを示すメッセージが表示された後、「クローズ」をクリックします。
  - h. 「パブリッシング・キュー・マップの選択」 ページで、作成したキュー・マップを選択して「OK」をクリックします。
  - i. 「サーバーとキュー・マップ」 ページで、キュー・マップ名が正しいことを確認して「次へ」をクリックします。
  - j. 「ソース表」 ページで、「追加」をクリックし、「すべてリトリブ」をクリックし、イベント・パブリッシングに使用可能にしたい表を選択して「OK」をクリックし、「次へ」をクリックします。
  - k. 「列と行」 ページで、DB2 クローラーがクロールする対象にする列 (またはすべての列) を選択し、キー列を選択します。クロールする行 (またはすべての行) を選択したページで、ソース表の削除をパブリッシュするオプションを選択します。これらのオプションの構成を完了した後、「次へ」をクリックします。
  - l. 「メッセージ内容」 ページで、列データには変更された列と未変更の列の両方を組み込むオプションを選択し、新規データ値のみのオプションを選択します。XML パブリケーションを自動的に開始するよう指定するチェック・ボックスが選択されていることを確認し、「次へ」をクリックします。
  - m. 「XML 発行の検討および完了」 ページで「次へ」をクリックします。
  - n. 「サマリー」 ページで、「完了」 ページをクリックします。メッセージと SQL スクリプトを示すページが表示された後、「クローズ」をクリックします。
  - o. 処理オプションに「今すぐ実行する」を選択し、「OK」をクリックします。SQL スクリプトが完了したことを示すメッセージが表示された後、「クローズ」をクリックします。
5. Q キャプチャー・サーバーを開始します。
- a. レプリケーション・センターのランチパッドをクローズし、レプリケーション・センターを開始します。
  - b. オブジェクト・ツリーで、「Q レプリケーション」 → 「定義」 → 「Q キャプチャー・サーバー」を順にクリックします。

- c. 構成した Q キャプチャー・サーバーのアイコンを右クリックし、「データベースを Q レプリケーション可能にする」を選択します。
- d. 警告メッセージが表示された後、「OK」をクリックします。
- e. DB2 メッセージを示すページが表示された後、「クローズ」をクリックします。
- f. オブジェクト・ツリーで、Q キャプチャー・サーバーのアイコンを右クリックし、「Q キャプチャー・プログラムの開始」を選択します。
- g. 処理オプションに、「今すぐ実行する」を選択し、システム名、DB2 ユーザーのユーザー ID とパスワード、ログを保管するディレクトリーのパス、および DB2 インスタンス名を指定し、「OK」をクリックします。
- h. 要求が実行されたことを示すメッセージが表示された後、「クローズ」をクリックします。
- i. オブジェクト・ツリーで、Q キャプチャー・サーバーのアイコンを右クリックし、「状況のチェック」を選択します。

Q キャプチャー・サーバーの状況が表示されます。エラーが発生した場合、サーバーがダウンしたと推定されることを示す状況メッセージが表示されます。ログを検討し、エラーの原因を判別するには、コマンド行で以下のコマンドを入力します。

```
asncap Capture_Server=capture server name LOGSTDOUT=y
```

## DB2 クローラーのための WebSphere MQ の構成

イベント・パブリッシングを使用するように DB2 クローラーを構成する前に、クローラーが listen するサーバー上に IBM WebSphere MQ が構成されていることを確認してください。

### 始める前に

DB2、WebSphere Information Integrator Event Publisher Edition、および WebSphere MQ がターゲット・データベース・サーバーにインストール済みであることを確認してください。

### 制約事項

ターゲット・データベース・サーバーが Linux コンピューターにインストールされている場合、すべての DB2 ユーザー、WebSphere MQ ユーザー、および OmniFind Enterprise Edition ユーザーは、以下の環境変数を設定する必要があります。

```
export LD_ASSUME_KERNEL=2.4.19
```

この環境変数は、LinuxThread スレッド化インプリメンテーションのエクスポートが、インストールが実行されるシェル、WebSphere MQ 制御コマンドが発行されるシェル、または WebSphere MQ アプリケーションが実行される任意のシェルから行われることを可能にします。WebSphere MQ は、この環境変数がエクスポートされる必要があります。



イベント・パブリッシングを使用する DB2 クローラーは、クライアント接続で WebSphere MQ キューに接続します。クライアント接続できるようにするには、WebSphere MQ 管理者でログインし、次のコマンドを実行してキュー・マネージャー CCSID を 819 に設定します。

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

### このタスクについて

DB2 クローラーは、WebSphere MQ サーバーへのクライアント接続モードをサポートします。このクローラーは、イベント・パブリッシング・キューにパブリッシュされる XML メッセージを listen します。クローラーは、複数のキューを経由して移送される XML メッセージを listen できません。

WebSphere MQ を構成すると、それ以降、DB2 クローラーはキュー・マネージャー名、キュー名、サーバー・ホスト名、サーバー・ポート番号、およびサーバー・チャンネル名を使用して、パブリッシング・キューから XML メッセージを取得します。クローラーは、メッセージを構文解析し、更新された表についての情報でクローラー・スペースを更新します。

### 手順

DB2 クローラーがイベント・パブリッシング・キューを listen できるようにデータベース・サーバーを構成するには、次のステップを実行します (これらのステップについての不明点は、WebSphere MQ 資料を参照してください)。

1. WebSphere MQ 管理者役割としてログインし、キュー・マネージャーおよびキューを作成するため、以下のコマンドを入力します。
  - a. コマンド行で、以下のコマンドを入力します。

```
crtmqm QM1
```

- b. Setup completed メッセージが表示されたら、以下のコマンドを入力します。

```
strmqm QM1
```

- c. 'QM1' started メッセージが表示されたら、以下のコマンドを入力します。

```
runmqsc QM1
```

- d. Starting MQSC for queue manager QM1 メッセージが表示されたら、以下のコマンドを入力して管理キューを作成します。

```
DEFINE QLOCAL(*ASN.QM1.ADMINQ*)
```

- e. WebSphere MQ queue created メッセージが表示されたら、以下のコマンドを入力して再始動キューを作成します。

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

- f. WebSphere MQ queue created メッセージが表示されたら、以下のコマンドを入力してデータ・キューを作成します。

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

- g. WebSphere MQ queue created メッセージが表示されたら、以下のコマンドを入力して終了します。

end

2. 以下のコマンドを入力してデータベース・サーバー上で MQ Listener を開始します (MQ Listener は、イベント・パブリッシングを使用する DB2 クローラー作成時に実行中でなければなりません)。この例では、サーバーのポート番号は 1414 であり、デフォルトのチャンネル SYSTEM.DEF.SVRCONN が使用されています。

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

3. 以下のコマンドを入力して、イベント・パブリッシングのために Message Queuing Interface (MQI) を使用してキュー・マネージャーおよびキューにアクセスすることを、DB2 ユーザーに認可します (この例では、ユーザー ID は db2inst1 です)。

```
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. DB2 クローラーを作成し、イベント・パブリッシングでこのクローラーを実行するのに使用されるユーザー ID に対して、以下のコマンドを入力します。これらのコマンドは、イベント・パブリッシングのために Message Queuing Interface (MQI) を経由してそのユーザー ID がキュー・マネージャーおよびキューにアクセスすることを認可します。この例では、ユーザー ID は esuser です。

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

## クラシック・データ・ソース・サーバーでの DB2 データベースのクローラー

DB2 クローラーは、WebSphere Information Integrator Classic Federation を介してクラシック・データ・ソース・サーバー上の DB2 データベースをクローラーできます。

### このタスクについて

クラシック・データ・ソース・サーバー上の DB2 データベースをクローラーするには、WebSphere Information Integrator Classic Federation の ODBC ラッパーを使用して、非クラシック・データ・ソース・サーバー上の DB2 データベースとフェデレートする必要があります。

### 手順

データベースをフェデレートするには、以下のようになります。

1. WebSphere Information Integrator Classic Federation をクラシック・データ・ソース・サーバーにインストールします。

- クラシック・データ・ソース・サーバー上のデータベースとフェデレートするデータベースがある非クラシック・データ・ソース・サーバーに、 WebSphere Information Integrator Classic Federation クライアント・モジュールをインストールします。
- WebSphere Information Integrator Classic Federation の ODBC ドライバーがクラシック・データ・ソース・サーバーに接続するように構成します。
- フェデレートするデータベースに接続し、クラシック・データ・ソース・サーバー上のデータベースとフェデレートする ODBC ラッパーを作成します。
- DB2 クローラーが非クラシック・データ・ソース・サーバー上のフェデレートするデータベースをクロールするように構成します。 これにより、クラシック・データ・ソース・サーバー上のデータベースが、 WebSphere Information Integrator Classic Federation を介してクロール可能になります。

---

## DB2 Content Manager クローラー

IBM DB2 Content Manager 項目タイプをエンタープライズ・サーチ・コレクションに入れる場合は、 DB2 Content Manager クローラーを構成する必要があります。

### クローラー・サーバーの構成

DB2 Content Manager サーバーをクロールするには、あらかじめクローラー・サーバーで特定のスクリプトを実行する必要があります。このスクリプトは、OmniFind Enterprise Edition で提供されており、これによって DB2 Content Manager クローラーが DB2 Content Manager サーバーと通信できるようになります。

エンタープライズ・サーチ管理コンソールを使用して、DB2 Content Manager クローラーを構成する前に、ご使用の環境に応じたタスクを実行する必要があります。

- 66 ページの『UNIX での DB2 Content Manager クローラー用クローラー・サーバーの構成』
- 68 ページの『Windows での DB2 Content Manager クローラー用クローラー・サーバーの構成』

### ヒント:

セキュアな DB2 Content Manager クローラーを構成する方法の詳細例については、IBM Redbook「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある大規模組織向けのシナリオを参照してください。

### 構成の概要

DB2 Content Manager クローラーを使用して、任意の数の DB2 Content Manager サーバーをクロールすることができます。クローラーを構成する際に、クローラーがクロール・スペースのすべての DB2 Content Manager サーバーをクロールするためのオプションを指定します。また、各サーバーでクロールする特定の項目タイプも選択します。

DB2 Content Manager クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ

ズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内のすべての DB2 Content Manager サーバーのあらゆる項目タイプをクロールする方法を制御します。
- クロールする DB2 Content Manager サーバーを選択します。
- クローラーが DB2 Content Manager サーバーのコンテンツにアクセスできるようにするためのユーザー ID とパスワードを指定します。
- サーバーをクロールするスケジュールを設定します。
- 各 DB2 Content Manager サーバーでクロールする項目タイプを選択します。
- 一部の項目タイプの属性を検索可能にするときのオプションを指定します。例えば、特定タイプの文書をクロール・スペースから除外したり、検索結果に戻すことのできる属性を指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

## UNIX での DB2 Content Manager クローラー用クローラー・サーバーの構成

IBM AIX、Linux、または Solaris オペレーティング環境で稼働するコンピュータに OmniFind Enterprise Edition をインストールした場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトにより、DB2 Content Manager クローラーが IBM DB2 Content Manager サーバーと通信できるようになります。

### このタスクについて

DB2 Content Manager クローラーは、DB2 Content Manager バージョン 8 用の Java コネクタを使用して、DB2 Content Manager サーバーにアクセスします。このコネクタは、クローラー・サーバーに次の製品の 1 つをインストールすると、インストールされます。

- IBM DB2 Information Integrator for Content for AIX、Solaris、Linux バージョン 8.3

- IBM DB2 Information Integrator for Content for AIX、Solaris バージョン 8.2
- IBM DB2 Content Manager Toolkit for Linux バージョン 8.2

DB2 Content Manager クローラーが DB2 Content Manager と連動して機能できるようにするには、コネクタをインストールした後で、OmniFind Enterprise Edition が提供するセットアップ・スクリプトをクローラー・サーバー上で実行します。

## 手順

クローラー・サーバーが DB2 Content Manager サーバーをクロールできるように構成するには、次のようにします。

- クローラー・サーバーに DB2 Content Manager バージョン 8 の Java コネクタをインストールします。
  - クローラー・サーバーに、root ユーザーとしてログインします。
 

```
su - root
```
  - db2profile ファイルを実行します。次に例を示します。
 

```
./home/db2inst/sql1lib/db2profile
```
  - JAVAHOME 環境変数をエクスポートします。次に例を示します。
 

```
export JAVAHOME=/usr/IBMJava2-141
```
  - Java ディレクトリーを PATH 環境変数に追加します。
 

```
export PATH=$PATH:$JAVAHOME/bin
```
  - DB2 Information Integrator for Content インストール CD を挿入し、インストール・ウィザードを実行します。
  - 「コンポーネントの選択」ウィンドウで、次の操作を行います (DB2 Information Integrator for Content バージョン 8.3 を使用して作業している場合、「コンポーネントの選択」ウィンドウには、カスタム・インストールのオプションが表示されます。)
    - 「コンポーネント」リストから「ローカル・コネクタ」を選択し、「サブコンポーネント」リストから「Content Manager V8 コネクタ」を選択します。
    - 「コンポーネント」リストから「コネクタ・ツールキットとサンプル」を選択し、「サブコンポーネント」リストから「Content Manager V8 コネクタ」を選択します。
  - DB2 Content Manager ライブラリーのデータベース名、ユーザー名、およびパスワードを指定し、残りのウィンドウでは、デフォルトの設定値を受け入れます。
- クローラー・サーバーに、DB2 管理グループの中のユーザー ID でログインします。
- リモート DB2 Content Manager ライブラリー・サーバー・データベースをカタログし、クローラー・サーバーが DB2 Content Manager サーバーに接続できることを確認します。

```
db2 catalog tcpip node ノード名 remote ホスト名 server port
db2 catalog database データベース名 as 別名 at node ノード名
```

ここで、それぞれ次のことを表します。

ノード名

DB2 Content Manager サーバーのホストのショート名 (例えば `ibmes`)。

ホスト名

DB2 Content Manager サーバーのホストの完全修飾名 (例えば `ibmes.ibm.com`)。

*port*

DB2 Content Manager サーバーのポート番号。

データベース名

DB2 Content Manager データベースの名前 (例えば `ICMNLSDB`)。

*alias*

DB2 Content Manager データベースの別名 (例えば `CMSVR`)

- オプション: `root` ユーザーとしてログインし、データベース接続をテストします。

```
. Information_Integrator_for_Content_install_directory/bin/cmbenv81.sh
cd Information_Integrator_for_Content_install_directory/samples/java/icm
javac *.java
java SConnectDisconnectICM ICM データベース名 CM 管理者 ID CM 管理者パスワード
```

- クローラー・サーバーで DB2 Content Manager クローラー用のセットアップ・スクリプトを実行します。
  - 次のようにして、`ES_INSTALL_ROOT/bin` ディレクトリーに移動します。

```
cd $ES_INSTALL_ROOT/bin
```

- 以下のスクリプトを開始し、プロンプトに応答します。

```
escrcm.sh
```

- クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

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

## Windows での DB2 Content Manager クローラー用クローラー・サーバーの構成

Microsoft Windows コンピューターに OmniFind Enterprise Edition をインストールした場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトにより、DB2 Content Manager クローラーが IBM DB2 Content Manager サーバーと通信できるようになります。

### このタスクについて

DB2 Content Manager クローラーは、DB2 Content Manager バージョン 8 用の Java コネクターを使用して、DB2 Content Manager サーバーにアクセスします。このコネクターは、クローラー・サーバーに IBM DB2 Information Integrator for Content for Windows バージョン 8.2 またはバージョン 8.3 をインストールする

と、インストールされます。DB2 Content Manager クローラーが DB2 Content Manager と連動して機能できるようにするには、コネクタをインストールした後で、OmniFind Enterprise Edition が提供するセットアップ・スクリプトをクローラー・サーバー上で実行します。

## 手順

クローラー・サーバーが DB2 Content Manager サーバーをクロールできるように構成するには、次のようにします。

1. クローラー・サーバーに DB2 Content Manager バージョン 8 の Java コネクタをインストールします。

a. DB2 Information Integrator for Content インストール CD を挿入します。インストール・プログラムが自動的に始動します。

DB2 Content Manager Enterprise Information Portal インストール・ウィザードが開きます。

b. 「コンポーネントの選択」ウィンドウで、次の操作を行います (DB2 Information Integrator for Content バージョン 8.3 を使用して作業している場合、「コンポーネントの選択」ウィンドウには、カスタム・インストールのオプションが表示されます。)

1) 「コンポーネント」リストから「ローカル・コネクタ」を選択し、「サブコンポーネント」リストから「Content Manager V8 コネクタ」を選択します。

2) 「コンポーネント」リストから「コネクタ・ツールキットとサンプル」を選択し、「サブコンポーネント」リストから「Content Manager V8 コネクタ」を選択します。

c. DB2 Content Manager ライブラリーのデータベース名、ユーザー名、およびパスワードを指定し、残りのウィンドウでは、デフォルトの設定値を受け入れます。

2. リモート DB2 Content Manager ライブラリー・サーバー・データベースをカタログし、クローラー・サーバーが DB2 Content Manager サーバーに接続できることを確認します。クローラー・サーバー上のコマンド・プロンプトで、次のコマンドを入力します。

```
db2 catalog tcpip node ノード名 remote ホスト名 server port  
db2 catalog database データベース名 as 別名 at node ノード名
```

ここで、それぞれ次のことを表します。

ノード名

DB2 Content Manager サーバーのホストのショート名 (例えば `ibmes`)。

ホスト名

DB2 Content Manager サーバーのホストの完全修飾名 (例えば `ibmes.ibm.com`)。

ポート

DB2 Content Manager サーバーのポート番号。

データベース名

DB2 Content Manager データベースの名前 (例えば `ICMNLSDDB`)。

### 別名

DB2 Content Manager データベースの別名 (例えば CMSVR)

- オプション: データベース接続をテストするため、コマンド・プロンプトをオープンし、以下のコマンドを入力します。

```
cmbenv81.bat
cd Information_Integrator_for_Content_install_directory%samples%java%icm
javac *.java
java SConnectDisconnectICM ICM データベース名 CM 管理者 ID CM 管理者パスワード
```

- クローラー・サーバーで DB2 Content Manager クローラー用のセットアップ・スクリプトを実行します。

- 次のようにして、ES\_INSTALL\_ROOT%bin ディレクトリーに移動します。

```
cd %ES_INSTALL_ROOT%bin
```

- 以下のスクリプトを開始し、プロンプトに応答します。

```
escrcm.vbs
```

- クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

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

---

## Lotus Domino Document Manager クローラー

Domino Document Manager ライブラリーおよびキャビネットをエンタープライズ・サーチ・コレクションに入れる場合は、Domino Document Manager クローラーを構成する必要があります。

### クローラー・サーバーの構成

クロールする Lotus Domino Document Manager サーバーが Lotus Notes リモート・プロシージャー・コール (NRPC) プロトコルを使用している場合は、クローラー・サーバーでセットアップ・スクリプトを実行する必要があります。このスクリプトは、OmniFind Enterprise Edition で提供されており、これによって、Lotus Domino Document Manager クローラーが NRPC を使用するサーバーと通信できるようになります。

クロールする Lotus Domino Document Manager サーバーが Lotus Domino Internet Inter-ORB Protocol (IIOP) を使用している場合、クローラー・サーバーでセットアップ・スクリプトを実行する必要はありません。ただし、Lotus Domino Document Manager クローラーが Lotus Domino Document Manager サーバーにアクセスできるようにサーバーを構成する必要があります。

**重要:** Domino Document Manager サーバーが IIOP を使用し、クローラーとサーバーの間の伝送を暗号化するために HTTPS または IIOP over SSL を使用するように構成する場合は、TrustedCerts.class ファイル (例えば c:%certs または /data/certs) を Domino Document Manager サーバーからクローラー・サーバーにコピーする必要があります。2 サーバー構成または 4 サーバー構成では、



TrustedCerts.class ファイルを、検索コンポーネントがインストールされているサーバーにコピーする必要もあります。ファイルがクローラー・サーバーと検索サーバーとで同じ場所にあるようにしてください。クローラーを構成するときに、TrustedCerts.class ファイルのディレクトリー・パスを指定します。

OmniFind Enterprise Edition が IBM AIX システムにインストールされている場合は、I/O 完了ポート・モジュールがクローラー・サーバーにインストールされており、使用可能であることを確認してください。

エンタープライズ・サーチ管理コンソールを使用して Lotus Domino Document Manager クローラーを構成する前に、ご使用の環境に応じたタスクを実行する必要があります。

- 87 ページの『Lotus Domino ソースをクロールするための UNIX でのクローラー・サーバーの構成』
- 89 ページの『Lotus Domino ソースをクロールするための Windows でのクローラー・サーバーの構成』
- 91 ページの『DIHOP プロトコルを使用するサーバーの構成』
- 92 ページの『Lotus Domino ソースをクロールするための AIX での I/O 完了ポートの構成』

## 文書レベルのセキュリティー

コレクションのセキュリティーが有効で、クロールするサーバーが NRPC プロトコルを使用している場合は、クローラー・サーバーで Lotus Domino トラステッド・サーバーを構成する必要があります。トラステッド・サーバーは、文書レベルのアクセス制御を行うために使用されます。ユーザーに対してコレクションを検索可能にする前に、以下のタスクを行います。

- Lotus Domino トラステッド・サーバーがユーザーのクレデンシャルを検証するように構成します。
- WebSphere Application Server でグローバル・セキュリティーを有効にし、検索アプリケーションがセキュリティーを使用するように構成します。このステップにより、ユーザーが検索アプリケーションを使用しようとしたときに、ログイン・クレデンシャルが検証されるようになります。検索サーバーは、クレデンシャルを使用して、Lotus Domino 文書に対する各ユーザーのアクセス権限を検査します。

## 構成の概要

Lotus Domino Document Manager クローラーを使用して、任意の数の Lotus Domino Document Manager ライブラリーをクロールすることができます。クローラーを作成する際に、単一の Lotus Domino Document Manager サーバーからクロールするライブラリーを選択します。後でクロール・スペースを編集する際に、同じクロール・スペースに別の Domino Document Manager サーバーの文書を追加できます。クローラーを作成または編集する際に、クロール対象として選択したライブラリー内のすべてのキャビネットをクロールするか、または、特定のキャビネットをクロールするかを指定できます。

Lotus Domino Document Manager クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタ

ープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内のすべての文書をクロールする方法を制御します。
- Domino Document Manager サーバーおよび通信プロトコルを識別します。
- サーバーが DIIOIP プロトコルを使用するように構成されている場合、クローラーが Lotus Domino オブジェクトに接続する方法を指定できます。例えば、通信の暗号化に HTTPS または Secure Sockets Layer (SSL) を使用するオプションを指定できます。
- クロールするライブラリーを選択します。
- ライブラリーをクロールするスケジュールを設定します。
- クロールする文書を選択します。クローラーは、ライブラリー内のすべてのキャビネットをクロールするか、またはユーザーが選択したキャビネット内のある特定の文書のみをクロールすることができます。
- さまざまなライブラリーおよびキャビネットのフィールドを検索可能にするオプションを指定します。例えば、特定タイプの文書をクロール・スペースから除外したり、添付ファイル検索のオプションを指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

#### 関連概念

86 ページの『Lotus Domino データベースのクロールに関するヒント』

314 ページの『Lotus Domino 文書に対する文書レベルのセキュリティーの施行』

#### 関連タスク

87 ページの『Lotus Domino ソースをクロールするための UNIX でのクローラー・サーバーの構成』

89 ページの『Lotus Domino ソースをクロールするための Windows でのクローラー・サーバーの構成』

91 ページの『DIIOIP プロトコルを使用するサーバーの構成』

92 ページの『Lotus Domino ソースをクロールするための AIX での I/O 完了ポートの構成』

## Exchange Server クローラー

Microsoft Exchange Server のパブリック・フォルダーをエンタープライズ・サーチ・コレクションに入れる場合は、Exchange Server クローラーを構成する必要があります。

Exchange Server クローラーを使用して、Exchange Server のパブリック・フォルダー・サーバー上の任意の数のフォルダーおよびサブフォルダーをクロールすることができます。クローラーの作成時に、パブリック・フォルダー・サーバーでクロールする対象コンテンツを選択します。クロール対象については、後からクロール・スペースを編集して、別のパブリック・フォルダー・サーバーのコンテンツを追加することができます。

Exchange Server クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内の全サーバーの全サブフォルダーをクロールする方法を制御します。
- クロールする Exchange Server パブリック・フォルダー・サーバーに関する情報を指定します。

クローラーがサーバーのコンテンツにアクセスできるようにするためのユーザー ID とパスワードを指定する必要があります。サーバーが Secure Sockets Layer (SSL) プロトコルを使用している場合、クローラーがクローラー・サーバーの鍵ストア・ファイルにアクセスできるようにするためのオプションを指定することができます。

- パブリック・フォルダー・サーバーをクロールするスケジュールを設定します。
- クロールするフォルダーおよびサブフォルダーを選択します。
- サブフォルダー内の文書を検索可能にするときのオプションを指定します。例えば、特定タイプの文書をクロール・スペースから除外することができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

### 関連タスク

314 ページの『セキュア Exchange Server 文書へのアクセスの検証』

## JDBC データベース・クローラー

JDBC データベース・クローラーは、Java Database Connectivity (JDBC) プロトコルでアクセスできるデータベースをエンタープライズ・サーチ・コレクションに入れる場合に使用します。

クローラーは、クロール対象の各データベース・システム・タイプごとに別のクローラーを構成する必要があります。クローラーの作成時に、1 つのデータベースをクロールするためのオプションを指定します。後で、同じタイプのデータベースをクロール・スペースに追加できます。

データベース表の各行は文書と見なされ、データベース列の値は解析されて、検索可能フィールドとして索引付けされます。クローラーのプロパティを構成する際にクローラーにプラグインを関連付けることで、複数の構造化されたテーブルをクロールするようクローラーを構成することができます。このプラグインがあれば、同じキー・フィールドを含む、リレーショナル・データベースの複数の表の行を連結することができます。連結された行は 1 つの文書と見なされます。ユーザーがデータベースを検索すると、文書が検索結果に表示される際、結合された表のデータは追加フィールドとして表示されます。

### サポートされるデータベース・システムとドライバー

JDBC プロトコルを使用してデータベース内の表をクロールするには、適切な JDBC ドライバーがクローラー・サーバーに存在しなければなりません。JDBC データベース・クローラーは、次のデータベース・システムとタイプ 4 JDBC ドライバーをサポートします。

データベース・システム	JDBC タイプ 4 ドライバー名	標準の JDBC ドライバー・クラス・パス
IBM DB2 Universal Databaseバージョン 8.2 ならびに IBM DB2 Enterprise Server Editionバージョン 9.1 for Linux、UNIX、および Windows	com.ibm.db2.jcc.DB2Driver	<i>db2_install_root</i> /java/db2jcc.jar <i>db2_install_root</i> /java/db2jcc_license_cu.jar
Oracle 9i および 10g	Oracle.jdbc.driver.OracleDriver	<i>oracle_home</i> /jdbc/lib/ojdbc14.jar
Microsoft SQL Server 2000	com.microsoft.jdbc.sqlserver.SQLServerDriver	<i>mssql_jdbc_home</i> /lib/mssqlserver.jar <i>mssql_jdbc_home</i> /lib/msbase.jar <i>mssql_jdbc_home</i> /lib/msutil.jar
Microsoft SQL Server 2005	com.microsoft.sqlserver.jdbc.SQLServerDriver (SQL Server 2005 の JDBC ドライバーは AIX システムでサポートされません。)	<i>install_dir</i> /sqljdbc_1.0/loc/sqljdbc.jar ここで、 <i>loc</i> は、 <i>install_dir</i> /sqljdbc_1.0/enu/sqljdbc.jar

など、ご使用のロケールを表します。

## JDBC データベース・クローラーと DB2 クローラー

現在 DB2 クローラーを使用している場合には、それを引き続き使用したほうがよいかもしれません。DB2 クローラー用に保管されたデータは、JDBC データベース・クローラーに移行できません。

以下の状況では、JDBC データベース・クローラーでなく、DB2 クローラーを使用してください。

- JDBC タイプ 2 ドライバーで DB2 データベースをクロールする。
- DB2 データベースとフェデレートされた Oracle および SQL Server データベースをクロールする。DB2 クローラーを使用すると、これらのすべてのデータベース・タイプにニックネームでアクセスできます。
- DB2 for z/OS、DB2 for iSeries、Informix、Sybase、VSAM、IMS、CA-Datcom、または Software AG Adabas データベースをクロールする。これらのタイプのデータベースを DB2 データベースとフェデレートし、DB2 クローラーでニックネームによりアクセスします。
- イベント・パブリッシングを使用して、データベースへの更新が発行されたときに、エンタープライズ・サーチ索引を更新する。

### 構成の概要

JDBC データベース・クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内のすべてのデータベースをクロールする方法を制御します。
- クロールするデータベースのタイプを指定します。
- クロールするデータベースを選択し、必要な場合、クローラーのデータベース・アクセスを可能にするユーザー ID とパスワードを指定します。
- データベースをクロールするスケジュールを設定します。
- クロールする表を選択します。

**重要:** ディスカバリー・プロセスのパフォーマンスを最適化するため、またクローラー構成プロセスがタイムアウトになるのを防止するため、すべての表をクロールする選択は、データベース中の表の数が少ない場合か、各表に含まれる列の数が少ない場合のみ行うようにしてください。現時点では一部の表についてクロールする選択をした場合、後からクロール・スペースを編集して、さらに多くの表をコレクションに追加できます。

- 特定の表の列を検索可能にするときのオプションを指定します。例えば、特定の列をパラメトリック照会で使用できるようにしたり、検索結果に戻すことのできる列を指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引

内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクローリングに必要な情報をクローラーに指定する方法が表示されます。

## JDBC データベースの関係マップ

複数の構造化された JDBC データベース表をクローリングするプラグインの規則を作成する際には、ルート表および親表と子表の結合に関する情報を指定します。

エンタープライズ・サーチ用のプラグインがあれば、JDBC データベース・クローラーは複数の構造化されたテーブルを結合できます。このプラグインを作成するには、ES\_INSTALL\_ROOT/default\_config/crawler\_rdb\_plugin.xml ファイルで規則を指定します。このプラグインを使用するようにクローラーを構成すると、同じキー・フィールドを含む表の行は、連結されて 1 つの文書と見なされます。ユーザーがデータベースを検索すると、文書が検索結果に表示される際、結合された表のデータは追加フィールドとして表示されます。

### キー列による表の結合

次の図は、複数の表の関係マップがどのようにして作成されるかを示しています。JDBC データベース・クローラーが、データベース内のルート表をスキャンします。表内のいくつかの列が、この表と他の表とを結合するために使用できるキー・フィールドになっています。したがって、結合された表内の列をキーとして使用することで、さらに別の表を結合することができます。これらの複数の表の行は、クローリング・スペース内で 1 つの文書と見なされます。列の値は文書のメタデータと見なされます。表の関係でいえば、ルート表が親で、第 1 レベルで結合される表が子となります。また、第 1 レベルの子表は、第 2 レベルで結合される表の親にもなることができます。

この例では、ルート (親) 表の「キー 1」列と「キー 2」列が、ルート表と、やはり「キー 1」および「キー 2」列を持つ下位表との結合を可能にするキー・フィールドになっています。この第 1 レベルで結合された一方の表にはキー・フィールド (「キー 3」および「キー 4」) があり、これらのキー・フィールドによって、この表はさらに別の表と結合することができます。

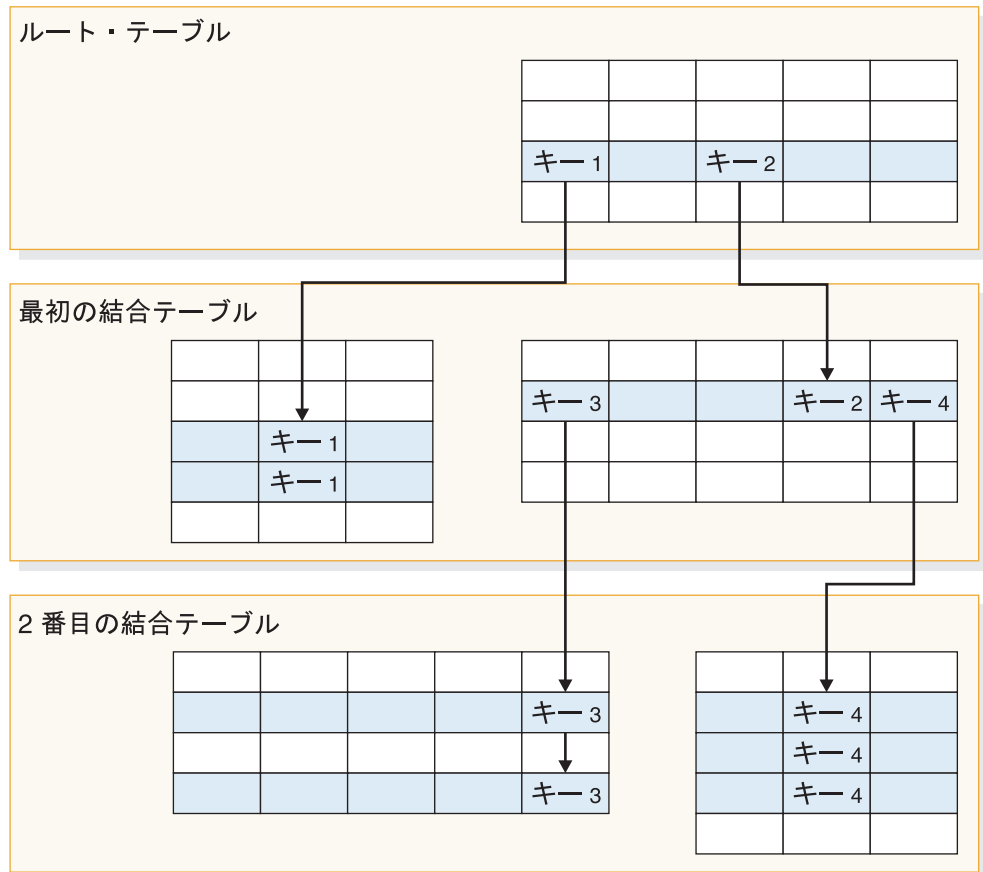


図2. キー・フィールドによって結合される JDBC データベース表

## 検索結果の表示

次の図は、検索結果において、複数の構造化されたテーブルのデータがどのようにして 1 つの文書として表示されるかを示しています。プラグインがないと、EMPLOYEE 表を検索したユーザーには検索結果に現れるルート表の行が表示され、EMPLOYEE 表列 (ID、名前、および事業所) の値のみが表示されます。

プラグインがあると、クローラーは「事業所」列をキーとして使用することで、EMPLOYEE 表と OFFICE 表とを結合することができます。OFFICE 表の「国」列は、OFFICE 表と COUNTRY 表とを結合するキーとなります。以上の表が結合されると、EMPLOYEE 表を検索したユーザーには、OFFICE および COUNTRY 表の列の値が検索結果で追加フィールドとして表示されます。

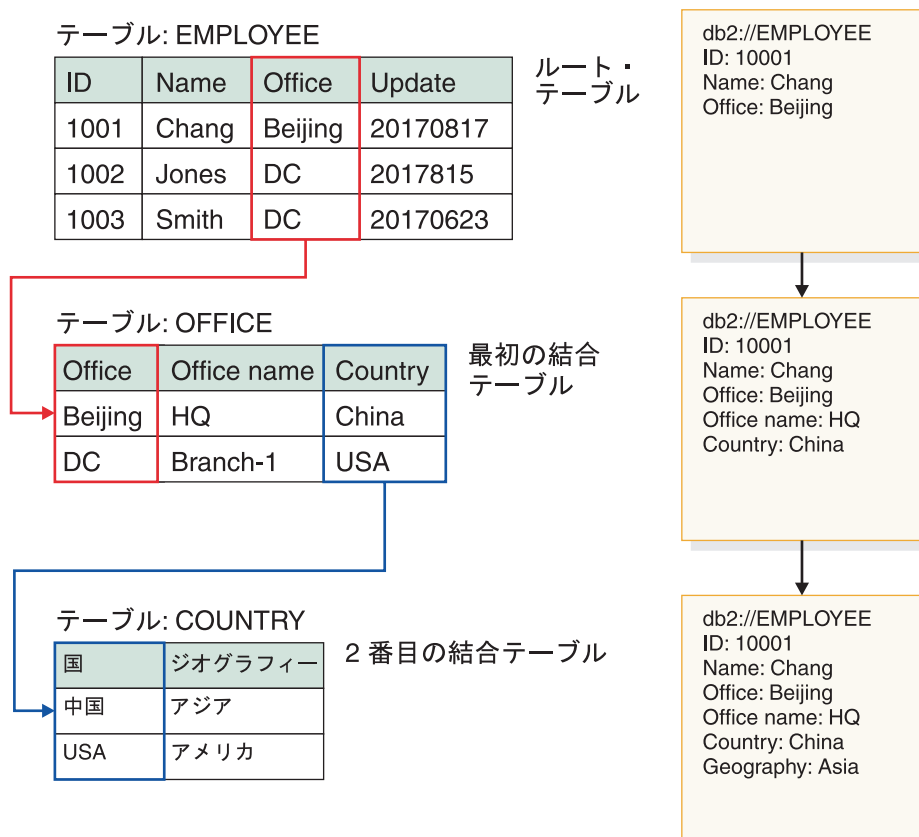


図3. 結合された JDBC 表から取り出されて検索結果に表示される値

## 複数の構造化された JDBC データベース表のクロール

同じキー・フィールドを持つ複数の構造化されたテーブルを結合するように JDBC データベース・クローラーを構成することができます。

### 始める前に

このタスクを実行するには、エンタープライズ・サーチ管理者としてログインする権限が必要です。

### このタスクについて

JDBC データベース・クローラーのクローラー・プロパティを構成するとき、キー・フィールドによって相互に関連付けられている複数の構造化されたテーブルをクロールするプラグインを指定することができます。このプラグインがない場合は、データベース表の行は個々の文書と同様に扱われ、データベース列の値は個々のフィールドとして検索可能です。このプラグインがある場合は、同じキー・フィールドを含む、リレーショナル・データベースの複数の表の行は、連結されて1つの文書と見なされます。クローラーは、連結された表からリトリブしたデータを、データベース表の元の行のメタデータに追加します。ユーザーがデータベースを検索すると、文書が検索結果に表示されたとき、この追加データは追加フィールドとして表示されます。



## 制約事項

### クローラーできないデータ・タイプ

クローラーは、結合した表のフィールドのうち、以下のバイナリー・データ・タイプを含むフィールドをクローラーできません。

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

### クローラー・スペースの範囲に関する制限事項

結合する表は、同じリレーショナル・データベース内の表でなければなりません。異なるデータベースの表を結合することはできません。

データベースの表が他の表と結合されるように構成されている場合、この設定は、同一コレクション内の、プラグインを使用できるようになっている、すべてのクローラーに共通します。ただし、複数のコレクションを作成して、個々のクローラーを、さまざまなルート表をクローラーし、さまざまな表を結合するように構成することができます。

### 他のプラグインの使用に関する制限事項

複数の構造化されたテーブルをクローラーするプラグインを使用するクローラーを構成した場合は、そのクローラーに別のプラグインを関連付けることはできません。例えば、業務およびセキュリティ規則を適用するカスタム・プラグインを指定することはできません。同じクローラーに複数のプラグインを関連付けることはできません。

### 表、行、フィールド、およびキーの数に関する制限事項

1 つのデータベースで結合できる表の最大数は 5 つです。これらの表に含まれる行の合計数は 100 万行未満でなければなりません。1 つの表から読み取れるフィールドの最大数は 10 個です。表の結合には一対のキーが使用されます。つまり、複数のキーを使用して表を結合することはできません。

### 結合された表の変更に対するクローラーの実行

クローラーとクローラーの間でルート表の行が変更されず、またクローラーがフル・クローラーを実行するように構成されていない場合は、クローラーは未変更の行を無視します。ルート表が変更されなくても、ルート表に結合されている表の行が変更された場合は、次のいずれかのアクションを実行して、その変更が検出されてクローラーされるようにする必要があります。

- ターゲット・データベース内のルート表にタイム・スタンプ・フィールドを設けます。ルート表の行が変更されたときや、結合されているいずれかの下位表の行が変更されたときに更新されるタイム・スタンプ・フィールドを設けるように、ターゲット・データベースを構成します。JDBC データベース・クローラーをセットアップするときは、クローラーが表の変更の有無を決定する際に使用するフィールドとして、このタイム・スタンプ・フィールドを必ず指定してください。
- クローラー・スケジュールの構成時にクローラーがフル・クローラーを実行するように指定します。このオプションを使用すれば、変更の有無にかかわらず毎回すべての表がクローラーされます。

## 手順

複数の構造化されたテーブルをクローリングする JDBC データベース・クローラーをセットアップするには、次のようにします。


1. エンタープライズ・サーチ管理者としてクローラー・サーバーでログインし、ES\_INSTALL\_ROOT/default\_config/crawler\_rdb\_plugin.xml ファイルをコピーして ES\_NODE\_ROOT/master\_config/crawler\_rdb\_plugin.xml ファイルを作成します。
2. UTF-8 エンコード方式をサポートするテキスト・エディターを使用して ES\_NODE\_ROOT/master\_config/crawler\_rdb\_plugin.xml ファイルを編集します。
  - a. `<Server DBURL="jdbc:db2://db_server_url:50000/SAMPLE">` エレメントを編集し、`jdbc:db2://db_server_url:50000/SAMPLE` を、クローリングする JDBC データベースの URL で置き換えます。クローラーを構成する際には、クローリングするデータベースに必ずこれと同じ URL を指定してください。
  - b. クローリングするデータベースが DB2 データベースでない場合は、`<JDBCdriver>com.ibm.db2.jcc.DB2Driver</JDBCdriver>` エレメントを編集し、`com.ibm.db2.jcc.DB2Driver` を適切な JDBC ドライバーで置き換えます。クローラーを構成する際には、クローリングするデータベースに必ずこれと同じドライバーを指定してください。
  - c. `<User>username</User>` エレメントを編集し、`username` を、クローリングするデータベースにアクセスする権限があるユーザー ID で置き換えます。
  - d. `<Password Encryption="True">encrypted_password</Password>` エレメントを編集し、`encrypted_password` を、指定したユーザー ID の暗号化パスワードで置き換えます。暗号化パスワードは、ES\_NODE\_ROOT/master\_config/col\_collection\_name.JDBC\_crawler\_name/jdbccrawler.xml ファイルからコピーして、ここに貼り付けることができます。パスワードを暗号化する必要がない場合は、`Encryption="True"` を `Encryption="False"` で置き換え、さらに `encrypted_password` をプレーン・テキスト・パスワードで置き換えます。
  - e. `<Delimiters Use="True">` エレメントをそのままにしておくと、列内の複数の用語は、`<Delimiter>` エレメントに定義された文字 (,) で区切られます。表ごとの用語の組は、`<SecondDelimiter>` エレメントに定義された文字 (;) で区切られます。`<Delimiters Use="True">` を `<Delimiters Use="False">` に設定すると、区切り文字は使用されず、同じフィールド名を持つ複数のメタデータ・フィールドは文書メタデータとして追加されます。
  - f. `<Delimiters Use="True">` エレメントを使用する場合は、`<Delimiter>,</Delimiter>` および `<SecondDelimiter>;</SecondDelimiter>` エレメントを編集して、値区切り記号として使用する文字を指定します。
  - g. `<RelationMap Root="DB2INST1.TABLE_0">` エレメントを編集し、`DB2INST1.TABLE_0` を、クローリングするルート表の名前で置き換えます。
  - h. `<Relation Parent="DB2INST1.TABLE_0" ParentAlias="T0" ParentKey="ID" Child="DB2INST1.TABLE_1" ChildAlias="T1" ChildKey="ID"/>` エレメントを編集します。
    - `Parent="DB2INST1.TABLE_0"` を、表間の関係でいえば親である表の名前で置き換えます。

- ParentKey= "T0" を親表の別名で置き換えます。この別名は、crawler\_rdb\_plugin.xml ファイル内で固有にし、かつ重複しないようにしてください。
- ParentKey="ID" を、表間の関係のキー・フィールドとして使用される列の名前で置き換えます。
- Child="DB2INST1.TABLE\_1" ChildAlias="T1" ChildKey="ID" を、クロールする子表に関する情報で置き換えます。

この構造は表の結合方法を定義するものです。例えば、次の関係マップは、DB2INST1.TABLE\_A という名前のルート表をクロールするように指定するものです。DB2INST1.TABLE\_B および DB2INST1.TABLE\_C という表が、DB2INST1.TABLE\_A.ID=DB2INST1.TABLE\_B.ID AND DB2INST1.TABLE\_B.ID=DB2INST1.TABLE\_C.ID という条件のもとで結合されます。

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

- ステップ 2h (80 ページ) を繰り返して、ルート表からの表を結合するすべての関係について <Relation> エレメントを作成します。
- <Target TableAlias="T1"> エレメントを編集し、TableAlias="T1" を、ステップ 2h (80 ページ) で定義した ChildAlias 値で置き換えます。
- <Field Name="ID" FieldName="ID\_1" Enabling="True" Searchable="True" FieldSearchable="True" IsContent="True"/> エレメントを編集します。
  - Name="ID" を、クロールする文書内の列の名前で置き換えます。
  - FieldName="ID\_1" を、クロールする文書内のメタデータ・フィールドの名前で置き換えます。この値は、エンタープライズ・サーチ管理コンソールおよび検索結果で列の表示名として使用されます。
  - Enabling="True" を "False" で置き換えます (この列を文書メタデータに組み込まない場合)。
  - Searchable="True" を "False" で置き換えて、ユーザーがこの列をフリー・テキスト照会で検索しないようにします。
  - FieldSearchable="True" を "False" で置き換えて、ユーザーがこの列を列名で検索しないようにします。
  - IsContent="True" を "False" で置き換えて、列に検索可能コンテンツが含まれていないことを示します。Searchable="True" と IsContent="True" を指定した場合は、列の値が重複文書の検出に使用され、また検索結果で動的文書サマリーの一部となります。
- ステップ 2k を繰り返して、クロールするすべての列について <Field> エレメントを作成します。
- ステップ 2j と 2k を繰り返して、関係マップ (<RelationMap>) で参照されるすべての子表について <Target> および <Field> エレメントを作成します。
- ステップ 2g (80 ページ) から 2m を繰り返して、複数のルート表について複数の関係マップを作成します。

- o. ステップ 2a (80 ページ) から 2n (81 ページ) を繰り返して、別のデータベース用の関係マップを構成します。
- 3. 次のようにして、クローラーがプラグインを使用するように構成します。
  - a. エンタープライズ・サーチ管理コンソールを開き、コレクションを編集し、「クロール」ページを選択します。
  - b. JDBC データベース・クローラーを作成するか、変更したいクローラーを探して、「 クローラーのプロパティ」をクリックします。
  - c. 「プラグイン・クラス名」フィールドで、複数の構造化されたテーブルをクロールするプラグインの名前を入力します。

```
com.ibm.es.plugin.rdb.RDBPlugin
```

- d. 「プラグイン・クラス・パス」フィールドで、プラグインの絶対パスとプラグインが使用する JDBC ドライバーを入力します。例えば、Windows システムの場合は、DB2 データベース用の JDBC ドライバーのパスは次のようになります。
- ```
C:%Program Files%IBM%es%lib%plugin_rdb.jar;C:%Program Files%IBM%SQLLIB%java%db2jcc.jar;C:%Program Files%IBM%SQLLIB%java%db2jcc_license_cu.jar;
```
- e. 「次へ」をクリックしてクローラーの作成を続行するか、「OK」をクリックして行った変更を保存します。
  - 4. crawler\_rdb\_plugin.xml ファイルをシステム構成にデプロイするには、エンタープライズ・サーチ・システムを再始動します。

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

---

## NNTP クローラー

NNTP ニュースグループの記事をエンタープライズ・サーチ・コレクションに入れる場合は、NNTP クローラーを構成する必要があります。

NNTP クローラーを使用して、任意の数の NNTP サーバーをクロールすることができます。クローラーの構成時に、1 つの NNTP サーバーからクロールするニュースグループを選択します。後でクロール・スペースを編集して、クローラーにクロールさせる別の NNTP サーバーを追加できます。

クロールするニュースグループを特定する際には、組み込むグループを選択することや、クロール・スペースから除外するグループを選択することが可能です。この設計により、特定サーバーの大多数のニュースグループをクローラーのクロール対象として容易に許可することができ、また、ユーザーに検索してほしくない少数のニュースグループをクロール禁止にすることができます。

例えば、特定の NNTP サーバーのすべてのニュースグループを含める規則を指定して、そのサーバーのニュースグループのうち、private というストリングが名前に含まれているニュースグループは除外する規則を指定できます。

NNTP クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内の全ニュースグループをクロールする方法を制御します。
- クロールする NNTP サーバーを特定します。サーバーがパスワード保護されている場合には、そのサーバー上のニュースグループのアクセスにクローラーが使用するユーザー ID とパスワードを指定してください。
- サーバーをクロールするスケジュールを設定します。
- クロール対象のニュースグループを指定し、さらにクロール・スペースから除外する一定のニュースグループのパターンを指定します。
- クローラーが、クロールする記事の言語とコード・ページを自動的に検出するか、または特定の言語とコード・ページを使用するかを指定します。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

---

## Lotus Notes クローラー

IBM Lotus Notes データベースをエンタープライズ・サーチ・コレクションに入れる場合は、Lotus Notes クローラーを構成する必要があります。

**ヒント:**

セキュアな Notes クローラーを構成する方法の詳細例については、IBM Redbook「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある大規模組織向けのシナリオを参照してください。

### クローラー・サーバーの構成

クロールする Lotus Notes サーバーが Lotus Notes リモート・プロシージャ・コール (NRPC) プロトコルを使用している場合は、クローラー・サーバーでセットアップ・スクリプトを実行する必要があります。このスクリプトは、OmniFind Enterprise Edition で提供されており、これによって、Lotus Notes クローラーが NRPC を使用するサーバーと通信できるようになります。

クロールする Lotus Notes サーバーが Lotus Domino Internet Inter-ORB Protocol (DIOP) を使用している場合、クローラー・サーバーでセットアップ・スクリプト

を実行する必要はありません。ただし、Lotus Notes クローラーが Lotus Notes サーバーにアクセスできるようにサーバーを構成する必要があります。

**重要:** Lotus Notes サーバーが DIHOP を使用し、クローラーとサーバーの間の伝送を暗号化するために HTTPS または DIHOP over SSL を使用するように構成する場合は、TrustedCerts.class ファイル (例えば c:%certs または /data/certs) を Lotus Notes サーバーからクローラー・サーバーにコピーする必要があります。2 サーバー構成または 4 サーバー構成では、TrustedCerts.class ファイルを、検索コンポーネントがインストールされているサーバーにコピーする必要があります。ファイルがクローラー・サーバーと検索サーバーとで同じ場所にあるようにしてください。クローラーを構成するときに、TrustedCerts.class ファイルのディレクトリー・パスを指定します。

OmniFind Enterprise Edition が IBM AIX システムにインストールされている場合は、I/O 完了ポート・モジュールがクローラー・サーバーにインストールされており、使用可能であることを確認してください。

エンタープライズ・サーチ管理コンソールを使用して Lotus Notes クローラーを構成する前に、ご使用の環境に応じたタスクを実行する必要があります。

- 87 ページの『Lotus Domino ソースをクロールするための UNIX でのクローラー・サーバーの構成』
- 89 ページの『Lotus Domino ソースをクロールするための Windows でのクローラー・サーバーの構成』
- 91 ページの『DIHOP プロトコルを使用するサーバーの構成』
- 92 ページの『Lotus Domino ソースをクロールするための AIX での I/O 完了ポートの構成』

## 文書レベルのセキュリティー

コレクションのセキュリティーが有効で、クロールするサーバーが NRPC プロトコルを使用している場合は、クローラー・サーバーで Lotus Domino トラステッド・サーバーを構成する必要があります。トラステッド・サーバーは、文書レベルのアクセス制御を行うために使用されます。ユーザーに対してコレクションを検索可能にする前に、以下のタスクを行います。

- Lotus Domino トラステッド・サーバーがユーザーのクレデンシャルを検証するように構成します。
- WebSphere Application Server でグローバル・セキュリティーを有効にし、検索アプリケーションがセキュリティーを使用するように構成します。このステップにより、ユーザーが検索アプリケーションを使用しようとしたときに、ログイン・クレデンシャルが検証されるようになります。検索サーバーは、クレデンシャルを使用して、Lotus Domino 文書に対する各ユーザーのアクセス権限を検査しません。

## 構成の概要

Lotus Notes クローラーを使用して、任意の数の標準 Lotus Notes データベース (.nsf ファイル) をクロールすることができます。クローラーを作成するときに、単一の Lotus Notes サーバーからクロールするデータベースまたはディレクトリーを選択します。後でクロール・スペースを編集して、同じクロール・スペースに別の

Lotus Notes サーバーから文書を追加できます。クローラーを作成または編集するときに、サーバー上のすべてのデータベースまたはディレクトリーをクロールするか、あるいは、特定のデータベース、ビューまたはフォルダーをクロールするかを指定できます。

Lotus Notes クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティーを指定します。クローラーのプロパティーで、クローラーがクロール・スペース内のすべての文書をクロールする方法を制御します。
- Lotus Notes サーバーのホスト名、ポート、および通信プロトコルを指定します。
- サーバーが DIOP プロトコルを使用するように構成されている場合、クローラーが Lotus Domino オブジェクトに接続する方法を指定できます。例えば、通信の暗号化に HTTPS または Secure Sockets Layer (SSL) を使用するオプションを指定できます。
- クロールするデータベースまたはディレクトリーを選択します。ディレクトリーをクロールするときは、データベースの組み込みまたは除外を行うパターンを指定できます。このパターンは、大きいディレクトリーをクロールするタスクを複数のクローラーに分割するのに役立ちます。
- データベースまたはディレクトリーをクロールするスケジュールを設定します。
- クロールする文書を選択します。ディレクトリー内のすべての文書、データベース内のすべての文書、またはデータベース内の選択したビューおよび選択したフォルダー内の文書をクロールすることができます。
- さまざまなデータベース、ビュー、およびフォルダーのフィールドを検索可能にするオプションを指定します。例えば、特定タイプの文書をクロール・スペースから除外したり、添付ファイル検索のオプションを指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

### 関連概念

86 ページの『Lotus Domino データベースのクロールに関するヒント』

314 ページの『Lotus Domino 文書に対する文書レベルのセキュリティの施行』

### 関連タスク

87 ページの『Lotus Domino ソースをクロールするための UNIX でのクローラー・サーバーの構成』

89 ページの『Lotus Domino ソースをクロールするための Windows でのクローラー・サーバーの構成』

91 ページの『DIOP プロトコルを使用するサーバーの構成』

92 ページの『Lotus Domino ソースをクロールするための AIX での I/O 完了ポートの構成』

315 ページの『ユーザーのクレデンシャルを検証するための Lotus Domino トラストレッド・サーバーの構成』

## Lotus Domino データベースのクロールに関するヒント

Lotus Notes クローラーを構成する前に、Lotus Domino データベースのクロールに関するガイドラインを確認してください。

- 標準テンプレートに基づく Lotus Notes データベース (ディスカッション・データベースなど) は、クロールするのに最適なタイプのデータベースです。
- Lotus Notes クローラーは、以下のフィールド・マッピング規則を適用します。
  - Lotus Domino 標準テンプレートの主フィールド名が最初に登録されます。
  - マッピング規則表に指定されている Lotus Notes フィールドの値は、検索結果で文書サマリーとして使用されます。
  - マッピング規則表に指定されていない Lotus Notes フィールドの値は、文書サマリーで使用されません。
  - Title フィールドにマップされる Lotus Notes フィールドの値は、検索結果で文書タイトルとして使用されます。
  - 以下の表のフィールドは、デフォルトで、以下の検索フィールド名にマップされます。

表 2. デフォルトのフィールド・マッピング規則

| Lotus Notes データベースのフィールド名 | 検索フィールド名     |
|---------------------------|--------------|
| Title                     | Title        |
| EventTitle                | Title        |
| Subject                   | Title        |
| Body                      | Body         |
| Mission                   | Body         |
| From                      | Creator      |
| Author                    | Creator      |
| Keywords                  | Categories   |
| Categories                | Categories   |
| TeamRoomName              | Organization |
| TeamName                  | Organization |
| Department                | Organization |



- Lotus Notes クローラーは、computed for display フィールドを除くすべてのタイプのフィールドをクローリングすることができます。
- Lotus Notes フォーム上の静的テキストおよびイメージは、クローリングされません。
- クローラーを構成する際に、「クローリング」の「すべて」チェック・ボックスを選択すると、すべてのフィールドをクローリングし、クローリングされるフィールド・データが最大限になります（「以下を除くすべてのフィールドをクローリング」フィールドを使用すると、クローリングされるフィールドを制限できます）。

不要なフィールドのクローリングを最小限にとどめるには、検索フィールドにマップされているフィールドを除くすべてのフィールドについて「クローリング」チェック・ボックスのチェックを外します。

#### 関連概念

70 ページの『Lotus Domino Document Manager クローラー』

83 ページの『Lotus Notes クローラー』

94 ページの『QuickPlace クローラー』

## Lotus Domino ソースをクローリングするための UNIX でのクローラー・サーバーの構成

IBM AIX、Linux、または Solaris オペレーティング環境で稼働するコンピューターに OmniFind Enterprise Edition をインストールし、Lotus Notes リモート・プロシージャー・コール (NRPC) プロトコルを使用するサーバーをクローリングする予定の場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトの実行によって、Lotus Notes クローラー、QuickPlace クローラー、および Lotus Domino Document Manager クローラーがデータベース・サーバーと通信できるようになります。

#### 制約事項

Domino Server は、NRPC プロトコルを使用するよう構成された Notes クローラー、QuickPlace クローラー、または Domino Document Manager クローラーとは、同じコンピューター上で同時に稼働できません。Lotus Domino Server が稼働中にこれらのクローラーのいずれかを開始しようとする、エラーが発生し、クローラーは停止します。

#### このタスクについて

NRPC プロトコルを使用するクローラーは、Domino ライブラリーをクライアントとして使用します。これらのライブラリーは、クローラー・サーバーに Lotus Domino Server をインストールすることによってインストールされます。クローラーが Lotus Domino ライブラリーと動作するようにするため、Lotus Domino ライブラリーをインストールした後、OmniFind Enterprise Edition によって提供されているセットアップ・スクリプトをクローラー・サーバー上で実行してください。

#### 手順

Lotus Notes サーバー、Lotus QuickPlace サーバー、および Domino Document Manager サーバーをクローリングできるようにクローラー・サーバーを構成するには、次のようにします。

1. クローラー・サーバー上にユーザー `server` とグループ `notes` を作成します。
  - a. `root` ユーザーとしてログインします。

```
su - root
```

- b. ユーザーを追加します。

```
useradd server
```

- c. このユーザーのパスワードを追加します。

```
passwd server
```

パスワードを変更するようプロンプトが出されます。

2. クローラー・サーバーに Lotus Domino Server をインストールします。
  - a. Lotus Domino Server CD を挿入し、マウントします。(CD がない場合、イメージをダウンロードできます。)
  - b. ご使用のオペレーティング・システムに応じたフォルダーに移動します。

```
AIX: cd /mnt/cdrom/aix
```

```
Linux: cd /mnt/cdrom/linux
```

```
Solaris: cd /mnt/cdrom/solaris
```

- c. インストール・プログラムを開始します。

```
./install
```

- d. プロンプトに回答し、デフォルト値を受け入れるか、希望するインストール設定値 (インストール・ディレクトリーおよびデータ・ディレクトリーのパスなど) を指定します。

Domino Server のインストールで不明な点があれば、Domino 資料を参照してください。

- e. エンタープライズ・サーチ管理者 ID に、`home/server` ディレクトリーにアクセスする権限があることを確認してください。この管理者 ID は、OmniFind Enterprise Edition のインストール時に指定されます。

3. クローラー・サーバーで、OmniFind Enterprise Edition により提供されるセットアップ・スクリプトを実行します。

- a. エンタープライズ・サーチ管理者としてログインします (このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです)。
  - b. 以下のスクリプトを開始します。これは、`$ES_INSTALL_ROOT/bin` ディレクトリーにインストールされています。

```
escrnote.sh
```

- c. プロンプトに回答します。

- 次のプロンプトに対して、Domino Server がデフォルトのディレクトリーにインストールされている場合は Y、そうでない場合は N と応答します。

```
Lotus Notes ディレクトリー・パス /opt/lotus/notes/latest/linux  
が見つかりました。
```

```
これは正しい Lotus Notes ディレクトリー・パスですか?
```

AIX の場合のデフォルトのパスは /opt/lotus/notes/latest/ibmpow です。  
Linux の場合のデフォルトのパスは /opt/lotus/notes/latest/linux です。  
Solaris の場合のデフォルトのパスは /opt/lotus/notes/latest/sunspa です。

- Domino Server がクローラー・サーバー上のデフォルトのディレクトリーにインストールされていない場合、次のプロンプトに対して Domino がインストールされている場所を指定します。

Lotus Notes ディレクトリーのパスを入力してください

例えば、Linux コンピューターで /opt/lotus/notes/latest/linux のように指定します。

- 次のプロンプトに対して、Domino Server データ・ディレクトリーがデフォルトのディレクトリーにインストールされている場合は Y、そうでない場合は N と応答します。

Lotus Notes データ・ディレクトリー・パス /local/notesdata が見つかりました。  
これは正しい Lotus Notes データ・ディレクトリー・パスですか?

デフォルトのパスは /local/notesdata です。

- Domino Server データ・ディレクトリーがクローラー・サーバー上のデフォルトの場所にデプロイされていない場合、次のプロンプトに対して、Domino データ・パスを指定します。

Lotus Notes データ・ディレクトリーのパスを入力してください。

4. クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

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

#### 関連概念

70 ページの『Lotus Domino Document Manager クローラー』

83 ページの『Lotus Notes クローラー』

94 ページの『QuickPlace クローラー』

## Lotus Domino ソースをクロールするための Windows でのクローラー・サーバーの構成

Microsoft Windows コンピューターに OmniFind Enterprise Edition をインストールし、Lotus Notes リモート・プロシージャ・コール (NRPC) プロトコルを使用するサーバーをクロールする予定の場合、クローラー・サーバーを構成するためのスクリプトを実行する必要があります。このスクリプトの実行によって、Lotus Notes クローラー、QuickPlace クローラー、および Lotus Domino Document Manager クローラーがデータベース・サーバーと通信できるようになります。

#### 制約事項

Lotus Domino Server および Lotus Notes クライアントは、NRPC プロトコルを使用するよう構成された Notes クローラー、QuickPlace クローラー、または Domino Document Manager クローラーとは、同じコンピューター上で同時に稼働できません。Lotus Domino Server が稼働中にこれらのクローラーのいずれかを開始しようとすると、エラーが発生し、クローラーは停止します。

## このタスクについて

NRPC プロトコルを使用するクローラーは、Lotus Domino クライアント・ライブラリーを使用します。これらのライブラリーは、クローラー・サーバーに Lotus Domino Server をインストールすることによってインストールされます。クローラーが Lotus Domino ライブラリーと動作するようにするため、Lotus Domino ライブラリーをインストールした後、OmniFind Enterprise Edition によって提供されているセットアップ・スクリプトをクローラー・サーバー上で実行してください。

## 手順

Lotus Notes サーバー、Lotus QuickPlace サーバー、および Domino Document Manager サーバーをクロールできるようにクローラー・サーバーを構成するには、次のようにします。

1. クローラー・サーバー上で、管理者グループのメンバーであるユーザー ID でログインします。このユーザー ID に Lotus Notes をインストールする権限があることを確認してください。
2. Lotus Notes をインストールします。
  - a. Lotus Domino Server CD を挿入します。(CD がない場合、イメージをダウンロードできます。)
  - b. インストール・プログラム `setup.exe` を開始します。
  - c. プロンプトに回答し、デフォルト値を受け入れるか、希望するインストール設定値 (インストール・ディレクトリーおよびデータ・ディレクトリーのパスなど) を指定します。

不明な点があれば、Lotus Domino 資料を参照してください。

- d. エンタープライズ・サーチ管理者 ID に、Domino データ・ディレクトリーにアクセスする権限があることを確認してください。この管理者 ID は、OmniFind Enterprise Edition のインストール時に指定されます。
3. クローラー・サーバーで、OmniFind Enterprise Edition により提供されるセットアップ・スクリプトを実行します。
    - a. エンタープライズ・サーチ管理者 ID でログインします (このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです)。
    - b. 以下のスクリプトを開始します。これは、`%ES_INSTALL_ROOT%\bin` ディレクトリーにインストールされています。

`escrnote.vbs`

- c. プロンプトに回答します。
  - 次のプロンプトに対して、Lotus Notes がデフォルトのディレクトリーにインストールされている場合は Y、そうでない場合は N と応答します。  
Lotus Notes ディレクトリー・パス `c:\%lotus%\notes` が見つかりました。  
これは正しい Lotus Notes ディレクトリー・パスですか?

Windows コンピューターでの標準的なインストール・パスは、`c:\%lotus%\notes` または `c:\%lotus%\domino` です。

- Lotus Notes がクローラー・サーバー上のデフォルトのディレクトリーにインストールされていない場合、次のプロンプトに対して Lotus Notes がインストールされている場所を指定します。

Lotus Notes ディレクトリーのパスを入力してください

- 次のプロンプトに対して、Lotus Notes データ・ディレクトリーがデフォルトの場所にデプロイされている場合は Y、そうでない場合は N と応答します。

Lotus Notes データ・ディレクトリー・パス c:%1lotus%notes%data  
が見つかりました。  
これは正しい Lotus Notes データ・ディレクトリー・パスですか?

Windows コンピューターでの標準的なパスは、c:%1lotus%notes%data または c:%1lotus%domino%data です。

- Lotus Notes データ・ディレクトリーがクローラー・サーバー上のデフォルトの場所にデプロイされていない場合、次のプロンプトに対して、データ・ディレクトリー・パスを指定します。

Lotus Notes データ・ディレクトリーのパスを入力してください。

4. クローラー・サーバーでエンタープライズ・サーチ・システムを一度停止してから再始動します。

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

#### 関連概念

70 ページの『Lotus Domino Document Manager クローラー』

83 ページの『Lotus Notes クローラー』

94 ページの『QuickPlace クローラー』

## DIIOP プロトコルを使用するサーバーの構成

Lotus Domino Internet Inter-ORB Protocol (DIIOP) を使用するサーバーをクローリングするには、Lotus Notes クローラー、QuickPlace クローラー、および Lotus Domino Document Manager クローラーがこのプロトコルを使用できるようにサーバーを構成する必要があります。

#### 始める前に

クローリングしたいサーバーは、DIIOP タスクおよび HTTP タスクを実行中でなければなりません。

#### 手順

DIIOP プロトコルを使用するサーバーを構成するには、次のようにします。

1. サーバー文書を構成します。
  - a. クローリングする Lotus Notes、Lotus QuickPlace、または Lotus Domino Document Manager サーバー上の server 文書を開きます。この文書は、Domino ディレクトリーに格納されています。
  - b. 「構成」ページで、「サーバー」セクションを拡張します。

- c. 「セキュリティ」ページの「プログラマビリティ制約事項」領域の以下のフィールドに、ご使用の環境に適したセキュリティ制約事項を指定します。
- 制限された Lotus Script/Java エージェントの実行
  - 制限された Java/Javascript/COM の実行
  - 制限されない Java/Javascript/COM の実行

例えば、アスタリスク (\*) を指定して Lotus Script/Java エージェントによる無制限のアクセスを許可し、Java/Javascript/COM 制約事項に対しては Domino ディレクトリーに登録済みのユーザー名を指定します。

**重要:** DIOP プロトコルを使用するこのサーバーをクロールするために構成するクローラーは、これらのフィールドに指定するユーザー名を使用できなければなりません。

- d. 「インターネット・プロトコル」ページを開き、次に「HTTP」ページを開き、「HTTP クライアントにデータベースの参照を許可する」オプションを「はい」に設定します。
2. ユーザー文書を構成します。
- a. クロールしたい Lotus Notes サーバー、Lotus QuickPlace サーバー、または Domino Document Manager サーバー上の user 文書を開きます。この文書は、Domino ディレクトリーに格納されています。
  - b. 「基本」ページの「インターネット・パスワード」フィールドに、パスワードを指定します。

エンタープライズ・サーチ管理コンソールを使用してこのサーバーのクロールについてのオプションを構成するときに、このユーザー ID とパスワードを、クロールするサーバーを指定するページで指定します。クローラーは、これらのクレデンシャルを使用してサーバーにアクセスします。

3. サーバーで DIOP タスクを再始動します。

#### 関連概念

70 ページの『Lotus Domino Document Manager クローラー』

83 ページの『Lotus Notes クローラー』

94 ページの『QuickPlace クローラー』

## Lotus Domino ソースをクロールするための AIX での I/O 完了ポートの構成

Lotus Notes クローラー、QuickPlace クローラー、または Lotus Domino Document Manager クローラーを IBM AIX システム上で使用するには、その前に、I/O 完了ポート (I/O completion port (IOCP)) モジュールをインストールし、このモジュールをクローラーが使用できるよう構成する必要があります。

#### このタスクについて

IOCP モジュールがないと、クローラーを作成しようとしたときに検出処理が失敗します。以下のエラー・メッセージが表示されます。

FFQM0105E サーバーからエラーを受け取りました -  
メッセージ: FFQG0024E 予期しない例外がキャッチされました: discover

以下のメッセージ (ENOEXEC エラーを含んでいます) が、\$ES\_NODE\_ROOT/logs/system\_yyyymmdd.log ファイルに書き込まれます (読みやすくするために、一部のメッセージ・テキストは複数行に分けて記述してあります)。

```
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 例外がキャッチされました。詳細 '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)
```

## 手順

IOCP モジュールをインストールし、それがクローラー・サーバー上に正しくインストールされたことを確認するには、次のようにします。

1. AIX 製品 CD から IOCP モジュール (bos.iocp.rte) をクローラー・サーバーにインストールします。

IOCP モジュールのインストールを完了したら、Lotus Notes クローラー、QuickPlace クローラー、または Lotus Domino Document Manager クローラーを作成する前に、このモジュールのソフトウェア修正を適用します。以下のリンクで説明を参照してください。

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

2. IOCP モジュールがクローラー・サーバーにインストールされたことを確認するため、以下のコマンドを入力します。

```
$ ls1pp -l bos.iocp.rte
```

ls1pp コマンドからの出力は、次の例に似たものです。

| 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
```

- IOCP ポート状況が **Available** であることを確認するため、以下のコマンドを入力します。

```
$ lsdev -Cc iocp
```

lsdev コマンドからの出力は、次の例と一致するはずです。

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

- IOCP ポート状況が **Defined** である場合、次のようにして状況を **Available** に変更します。
  - root としてクローラー・サーバーにログインし、以下のコマンドを発行します。

```
# smit iocp
```
  - 「I/O 完了ポートの特性の変更/表示」を選択し、「システム再始動時に構成される状態」を「定義済み」から「使用可能」に変更します。
  - クローラー・サーバーをリブートします。
  - もう一度 lsdev コマンドを入力し、IOCP ポート状況が **Available** に変わったことを確認します。

#### 関連概念

70 ページの『Lotus Domino Document Manager クローラー』

83 ページの『Lotus Notes クローラー』

『QuickPlace クローラー』

---

## QuickPlace クローラー

Lotus QuickPlace のプレースおよびルームをエンタープライズ・サーチ・コレクションに入れる場合は、QuickPlace クローラーを構成する必要があります。

また、QuickPlace クローラーを使用して、Lotus Domino 用の Lotus Quickr サービスで管理するプレースをクローリングすることもできます。WebSphere Portal 用の Lotus Quickr サービスを使用する場合は、シード・リスト クローラーを使用して、Lotus Quickr ライブラリーに保管されている文書をエンタープライズ・サーチ・コレクションに追加します。

#### ヒント:

セキュアな QuickPlace クローラーを構成する方法の詳細例については、IBM Redbook「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある小規模組織向けのシナリオを参照してください。

#### クローラー・サーバーの構成

クローリングしようとしているサーバーが Lotus Notes リモート・プロシージャ・コール (NRPC) プロトコルを使用している場合は、クローラー・サーバーでセットアップ・スクリプトを実行する必要があります。このスクリプトは、OmniFind



Enterprise Edition で提供されており、これによって、Lotus QuickPlace クローラーが NRPC を使用するサーバーと通信できるようになります。

クロールしようとしているサーバーが Domino Internet Inter-ORB Protocol (DIIOP) を使用している場合は、クローラー・サーバーでセットアップ・スクリプトを実行する必要はありません。ただし、QuickPlace クローラーがアクセスできるようにターゲット・サーバーを構成する必要があります。

クロールしようとしているサーバーが Lightweight Directory Access Protocol (LDAP) サーバーを使用している場合は、DIIOP プロトコルを使用するようにターゲット・サーバーを構成する必要があります (QuickPlace クローラーは NRPC プロトコルを使用して LDAP データをクロールできません)。また、Directory Assistance データベースを構成し、LDAP サーバーを 2 次 Lotus Domino Server として使用するようターゲット・サーバーを構成する必要があります。

**重要:** ターゲット・サーバーが DIIOP を使用し、クローラーとサーバーの間の伝送を暗号化するために HTTPS または DIIOP over SSL を使用するよう構成する場合は、TrustedCerts.class ファイル (例えば c:\certs または /data/certs) をターゲット・サーバーからクローラー・サーバーにコピーする必要があります。2 サーバー構成または 4 サーバー構成では、TrustedCerts.class ファイルを、検索コンポーネントがインストールされているサーバーにコピーする必要があります。ファイルがクローラー・サーバーと検索サーバーとで同じ場所にあることを確認してください。クローラーを構成するときに、TrustedCerts.class ファイルのディレクトリー・パスを指定します。

クローラーを構成し、クローラーが使用するユーザー ID を指定する際には、Domino ドメイン内のすべての QuickPlace プレースにアクセスする十分な権限を持つ ID を指定する必要があります。これを確実にするために、ユーザー ID を QuickPlaceAdministratorsSUGroup という名前の予約済みグループに割り当ててください。

OmniFind Enterprise Edition が IBM AIX システムにインストールされている場合は、I/O 完了ポート・モジュールがクローラー・サーバーにインストールされており、使用可能であることを確認してください。

エンタープライズ・サーチ管理コンソールを使用して、QuickPlace クローラーを構成する前に、ご使用の環境に応じたタスクを実行する必要があります。

- 87 ページの『Lotus Domino ソースをクロールするための UNIX でのクローラー・サーバーの構成』
- 89 ページの『Lotus Domino ソースをクロールするための Windows でのクローラー・サーバーの構成』
- 91 ページの『DIIOP プロトコルを使用するサーバーの構成』
- 316 ページの『ローカル・ユーザー・セキュリティーを使用するための QuickPlace サーバーの構成』
- 318 ページの『QuickPlace での Directory Assistance の構成』
- 92 ページの『Lotus Domino ソースをクロールするための AIX での I/O 完了ポートの構成』

## 文書レベルのセキュリティー

コレクションのセキュリティーが有効で、クロールするサーバーが NRPC プロトコルを使用している場合は、クローラー・サーバーで Lotus Domino トラステッド・サーバーを構成する必要があります。トラステッド・サーバーは、文書レベルのアクセス制御を行うために使用されます。ユーザーに対してコレクションを検索可能にする前に、以下のタスクを行います。

- Lotus Domino トラステッド・サーバーがユーザーのクレデンシャルを検証するように構成します。
- WebSphere Application Server でグローバル・セキュリティーを有効にし、検索アプリケーションがセキュリティーを使用するように構成します。このステップにより、ユーザーが検索アプリケーションを使用しようとしたときに、ログイン・クレデンシャルが検証されるようになります。検索サーバーは、クレデンシャルを使用して、Lotus Domino 文書に対する各ユーザーのアクセス権限を検査します。

## 添付ファイルのクロール

Lotus QuickPlace で、Microsoft Office 文書をインポートおよび公開できます (インポートのときに選択するオプションは、「ページの呼び出し」、「Microsoft Word ページ」、「Microsoft Excel ページ」、「Microsoft PowerPoint ページ」、および「複数ページの呼び出し」)。QuickPlace クローラーは、次の場合にのみ、これらのタイプのインポートされた文書を添付ファイルとしてクロールできます。

- サーバーが DIOP プロトコルを使用している。
- クローラーのクロール・オプションの構成時に添付ファイル・クロールを使用可能にしている。
- 「\$FILE」フィールドまたはすべてのフィールドをクロールするようにクローラーを構成している。

## 構成の概要

QuickPlace クローラーを使用して、任意の数の QuickPlace プレースをクロールすることができます。クローラーを作成する際に、単一の QuickPlace サーバーからクロールするプレースを選択します。後でクロール・スペースを編集して、同じクロール・スペースに別の QuickPlace サーバーから文書を追加できます。クローラーを作成または編集するときは、サーバー上のすべてのプレースをクロールするか、それとも指定したプレースのみをクロールするかを指定できます。また、クロールするプレース内のすべてのルームをクロールするか、それとも指定したルームのみをクロールするかを指定することもできます。

**制約事項:** クローラーが使用するユーザー ID を指定する際には、Lotus Domino ドメイン内のすべての QuickPlace プレースにアクセスするための十分な権限を持った ID を必ず指定してください。これを行うには、QuickPlace サーバーを構成して、QuickPlaceAdministratorsSUGroup という予約グループにユーザー ID を割り当てます。このグループ名は埋め込みスペースを含まず、大/小文字を区別します。

QuickPlace クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内のすべての文書をクロールする方法を制御します。
- QuickPlace サーバーおよび通信プロトコルを識別します。
- サーバーが DIOP プロトコルを使用するように構成されている場合、クローラーが Lotus Domino オブジェクトに接続する方法を指定できます。例えば、通信の暗号化に HTTPS または Secure Sockets Layer (SSL) を使用するオプションを指定できます。
- サーバーに関連したユーザー・ディレクトリーに関する情報を指定します (クローラーはこの情報を必要とするので、ユーザーがコレクションを検索する際にアクセス制御が行われます)。
- クロールするプレースを選択します。
- プレースをクロールするスケジュールを設定します。
- クロールする文書を選択します。クローラーは、プレース内のすべてのルームをクロールできるか、またはユーザーが選択したルーム内の特定の文書のみをクロールすることができます。
- さまざまなプレースおよびルームのフィールドを検索可能にするオプションを指定します。例えば、特定タイプの文書をクロール・スペースから除外したり、添付ファイル検索のオプションを指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

#### 関連概念

86 ページの『Lotus Domino データベースのクロールに関するヒント』

314 ページの『Lotus Domino 文書に対する文書レベルのセキュリティーの施行』

#### 関連タスク

87 ページの『Lotus Domino ソースをクロールするための UNIX でのクローラー・サーバーの構成』

- 89 ページの『Lotus Domino ソースをクロールするための Windows でのクローラー・サーバーの構成』
- 91 ページの『DIOP プロトコルを使用するサーバーの構成』
- 92 ページの『Lotus Domino ソースをクロールするための AIX での I/O 完了ポートの構成』
- 315 ページの『ユーザーのクレデンシャルを検証するための Lotus Domino トラストッド・サーバーの構成』
- 316 ページの『ローカル・ユーザー・セキュリティーを使用するための QuickPlace サーバーの構成』
- 318 ページの『QuickPlace での Directory Assistance の構成』

---

## シード・リスト・クローラー

WebSphere Portal 用の IBM Lotus Quickr サービスを使用した場合は、シード・リスト・クローラーを使用して、Lotus Quickr ライブラリー内の文書をエンタープライズ・サーチ・コレクションに追加することができます。

Lotus Quickr ライブラリーは文書ファイルを収容するコンテナです。シード・リスト・クローラーは、ウィキペディアやブログといった Web ベース・コンテンツのクロールをサポートしません。

Lotus Domino 用の Lotus Quickr サービスを使用した場合は、QuickPlace クローラーを使用して、Lotus Quickr 文書をコレクションに追加してください。

### WebSphere Portal サーバーの構成

Lotus Quickr を WebSphere Portal バージョン 6 サーバーにインストールすれば、シード・リスト・クローラーを使用して、Lotus Quickr ライブラリー文書をクロールすることができます。これらの文書をクロールするためのオプションは、WebSphere Portal クローラーがクロールするポータル・サイトに指定するオプションとは別に構成できます。

シード・リスト・クローラーを作成する前に、WebSphere Portal でエンタープライズ・サーチをセットアップする手順に従う必要があります。エンタープライズ・サーチ環境をセットアップするには、検索サーバー上の OmniFind Enterprise Edition と一緒に提供されるスクリプト (AIX、Linux、または Solaris の場合は `wp6_install.sh`、Windows の場合は `wp6_install.bat`) を実行します。

ユーザー・エージェント・ストリングが、サーバーにアクセスするブラウザーまたはロボットを識別します。シード・リスト・クローラーは、Lotus Quickr サーバーをクロールする際、ユーザー・エージェント・ストリング `OmniFind SeedlistCrawler/1.0` を使用します。

### 構成の概要

シード・リスト・クローラーを使用して、Lotus Quickr 文書をいくつでもクロールすることができます。クローラーを構成する際には、クロールするサーバーを指定します。そうすると、クローラーは、そのサーバー上の Lotus Quickr ライブラリーにあるすべての文書をクロールします。

クローリングする文書は、同じ Lotus Quickr 管理者 ID およびパスワードによってアクセスできるものでなければなりません。異なるクレデンシャルを使用するサイトをクローリングするには、別の シード・リスト クローラーを構成する必要があります。

シード・リスト・クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティは、クローラーが Lotus Quickr 文書をどのようにクローリングするかを制御します。
- クローリングするサーバーとクローラーがサーバーに接続するための情報を指定します。

クローラーを作成または編集したら、クローラーがクローリング対象文書に接続できるかどうかをテストできます。クローラーを開始する前に、クローリングされる文書にクローラーがアクセスできるかどうかメッセージで通知されます。

- 文書レベルのセキュリティー・オプションを指定します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

**重要:** セキュアな Lotus Quickr 文書を検索するには、WebSphere Portal 内からエンタープライズ・サーチ用の検索ポートレットを使用して、検索を実行する必要があります。サンプル検索アプリケーション ESSearchApplication から実行された検索は、適切なクレデンシャルを持たず、文書へのユーザーのアクセス権限を検査できません。

- クローラーがプロキシ・サーバーと通信するための情報を指定します (ページの提供にプロキシ・サーバーが使用される場合)。
- WebSphere Portal サーバーおよび Lotus Quickr 文書を保護するのに別の製品 (IBMTivoli® Access Manager WebSEAL や CA SiteMinder SSO Agent for PeopleSoft など) を使用する場合は、クローラーがサーバー上の文書にアクセスできるようにするシングル・サインオン・クレデンシャルを指定します。
- クローラーが Secure Sockets Layer (SSL) プロトコルを使用してサーバーに接続するための鍵ストア・ファイルに関する情報を指定します。
- クローリングする文書の言語とコード・ページを指定します。
- Lotus Quickr 文書内のメタデータをクローリングおよび検索するためのオプションを指定します。
- Lotus Quickr 文書をクローリングするスケジュールを指定します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクローリングに必要な情報をクローラーに指定する方法が表示されます。

#### 関連概念

381 ページの『WebSphere Portal との統合』

#### 関連タスク

389 ページの『WebSphere Portal バージョン 6 でのエンタープライズ・サーチのセットアップ』

396 ページの『Lotus Quickr 用のエンタープライズ・サーチ・ポートレットのセットアップ』

---

## UNIX ファイル・システム・クローラー

AIX、Linux、または Solaris ファイル・システムに保管されている文書をエンタープライズ・サーチ・コレクションに入れる場合は、UNIX ファイル・システム・クローラーを構成する必要があります。

UNIX ファイル・システム・クローラーを使用して、任意の数のファイル・システムをクローリングすることができます。クローラーの構成時に、クローリングするローカルおよびリモートのディレクトリー、サブディレクトリーを選択します。

Windows コンピューターにクローラー・サーバーをインストールした場合、そのサーバーを使用して AIX、Linux、または Solaris ファイル・システム・ソースをクローリングすることはできません (UNIX ファイル・システム・クローラーは、選択可能なクローラー・タイプのリストに表示されません)。

UNIX ファイル・システム・クローラーは、エンタープライズ・サーチ管理者に指定された読み取り許可に従って文書をクローリングします。

UNIX ファイル・システム・クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクローリング・スペース内のすべてのサブディレクトリーをクローリングする方法を制御します。
- ファイル・システムをクローリングするスケジュールを設定します。
- クローラーがクローリングするサブディレクトリー、サブディレクトリーのレベルを選択します。
- サブディレクトリー内の文書を検索可能にするときのオプションを指定します。例えば、特定タイプの文書をクローリング・スペースから除外することができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引

内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクローリングに必要な情報をクローラーに指定する方法が表示されます。

---

## Web クローラー

Web サイトのページをエンタープライズ・サーチ・コレクションに入れる場合は、Web クローラーを構成する必要があります。

Web クローラーを使用すれば、任意の数の HTTP サーバーおよびセキュア HTTP (HTTPS) サーバーをクローリングすることができます。クローラーが Web サイトにアクセスし、サイト上のデータを読み取ります。その後、文書内のリンクをたどって他の文書をクローリングします。Web クローラーは、個々のページあるいはフレームセット (HTML フレームを使用して作成されたページ) のリンクをクローリングして、抜き出すことができます。

クローリングされたデータは、フォーマットもソースも様々です。一般的に使用されるフォーマットは数多くありますし、イントラネットまたはインターネットの様々なところから取得されます。一般的なフォーマットには、HTML、PDF、Microsoft Word、Lotus WordPro、Extensible Markup Language (XML) などがあります。

### ヒント:

Web クローラーを構成する方法の詳細例については、IBM Redbook「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある中規模組織向けのシナリオを参照してください。

Web クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。また、この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクローリング・スペース内の全 Web ページをクローリングする方法を制御します。
- Web サイトへのアクセスを許可または禁止するための規則を指定します。クローリング規則を指定する場合、規則をテストして、ユーザーがクローリング・スペースに入りたいサイトにクローラーがアクセスできるかどうか検証することができます。
- 特定のタイプのファイルを含めるためのオプション、および特定のファイル拡張子を持つファイルを除外するためのオプションを指定します。
- Web クローラーがソフト・エラーのページを処理するための規則を指定します。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引

内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

- パスワード保護された Web サイトをクロールする場合のオプションを指定します (クロール対象の Web サーバーは HTTP 基本認証、またはパスワード入力のプロンプトを出す HTML フォームを使用する必要があります)。
- プロキシ・サーバーによってサービスされている Web サイトをクロールするためのオプションを指定します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

## ユーザー・エージェントの構成

ロボット排除プロトコルを使用する Web サイトをクロールするには、その Web サイト上の robots.txt ファイルで、Web クローラー用に設定したユーザー・エージェント名が Web サイトへのアクセスを許可されるようにしてください。

エンタープライズ・サーチ・システムが開始されると、Web クローラーは、構成されたユーザー・エージェント名をロードします。クローラーが、まだ訪れたことのない (あるいは、しばらく訪れていない) Web サイトからページをダウンロードする前に、クローラーは、まず robots.txt というファイルをダウンロードしようとします。このファイルは、Web サイトのルート・ディレクトリーにあります。

robots.txt ファイルがない場合、その Web サイトは、制限なく自由にクロールできます。このファイルがある場合、それは、サイト (ディレクトリー) のどの領域がクローラーの立ち入り禁止になっているかを明示します。robots.txt ファイルは、ユーザー・エージェント名を特定してクローラーに許可を出します。

ロボット排除プロトコルは、任意のもですが、エンタープライズ・サーチ Web クローラーはこれに従うように試行します。

- robots.txt ファイルに、Web クローラー用に設定されたユーザー・エージェント名の項目が含まれている場合、Web クローラーはそのユーザー・エージェント上の制約に従います。
- ユーザー・エージェント名が robots.txt ファイルになく、最後の項目が User-agent: \* (任意のユーザー・エージェントを意味する) と指定されていて、その制約が Disallow: / (この Web サイトのルートを起点とするいかなるクロールも許可しないことを意味する) になっている場合、Web クローラーはそのサイトのクロールができません。
- ユーザー・エージェント名が robots.txt ファイルになく、最後の項目が User-agent: \* と指定されていて、その制約が Allow: / になっている場合は、Web クローラーはそのサイトをクロールすることができます。

Web サイト管理者は、明示的にアクセスを認可されていないすべてのクローラーに対してアクセスを禁止する最後の項目を指定することがよくあります。新規に Web クローラーを構成するとき、クロール対象の Web サイトの一部でロボット排除プロトコルを使用していることがわかっている場合は、Web サイト管理者に依頼して、クローラーの項目をその robots.txt ファイルに追加してもらってください。



Web クローラーのプロパティと、対象とする Web サイトに含まれるすべての robots.txt ファイルに、同じユーザー・エージェント名が指定されていることを確認してください。

クロール予定の Web サイトのいずれもロボット排他プロトコルを使用していない場合は、ユーザー・エージェント・プロパティに指定する値は通常、何でもかまいません。ただし、一部のアプリケーション・サーバー、JSP、およびサーブレットでは、ユーザー・エージェント名に対応するように調整しています。例えば、ブラウザの非互換性に対処するために異なる応答が用意されています。このような場合は、ロボット排他プロトコルの使用に関係なく Web クローラーに指定したユーザー・エージェント名が重要になります。このタイプのサイトをクロールする必要がある場合は、Web サイト管理者に相談して Web クローラーのアクセスが許可されるようにしてください。

## Web クローラーによるロボット排他プロトコルの使用方法

Web クローラーはロボット排他プロトコルに従い、サーバーの robots.txt ファイル内の規則がクロールを許可しない場合には Web サイトをクロールしないようにします。

ダウンロードの成功とは、クローラーが robots.txt ファイルを Web サーバーからリトリブできる、または robots.txt ファイルが存在しないことを確認できる場合をさします。クローラーが規則を入手できない、または robots.txt ファイルが存在することを確認できない場合、ダウンロードは失敗と見なされます。

ダウンロードが成功しても、クローラーがクロール許可を持つわけではありません。これは、robots.txt ファイルの規則でクロールが許可されない可能性があるためです。ダウンロードが失敗すると、クローラーが規則を判別できないため、クロールは一時的に禁止されます。

robots.txt ファイルのダウンロードを試行する際にクローラーが行うステップは、以下のとおりです。

1. クローラーは、新しいサイトを検出すると、サーバーの IP アドレスを入手しようとしています。この試行が失敗した場合、クロールはできません。
2. 少なくとも 1 つの IP アドレスが使用可能な場合、クローラーは **HTTP (または HTTPS) GET** を使用して robots.txt ファイルのダウンロードを試行します。
3. ソケット接続がタイムアウトになるか、中断したか、または別の下位のエラー (SSL 証明書の問題など) が発生した場合、クローラーは問題をログに記録し、ターゲット・サーバーに対して認識されるすべての IP アドレスについて試行を繰り返します。
4. クローラーがすべてのアドレスを試行した後に接続されていない場合、クローラーは 2 秒待機してから、すべてのアドレスをもう一度試行します。
5. 接続されて HTTP ヘッダーが交換されると、戻り状況が検査されます。状況コードが 500 以上の場合、クローラーはこれを正しくない接続と解釈し、他の IP アドレスの試行を続行します。その他の状況の場合、クローラーは代替 IP アドレスの試行を停止し、状況コードに応じて進みます。

クローラーは、500 未満の HTTP 状況コードを受け取るか、またはすべての IP アドレスを 2 回試行した後、次のように進みます。

1. 500 未満の HTTP 状況を受け取らなかった場合、サイトは当面の間、不適格と見なされます。
2. HTTP 状況 400、404、または 410 を受け取った場合、サイトは規則なしのクローラーに適格と見なされます。
3. HTTP 状況 200 から 299 を受け取った場合、以下の条件により、次のアクションが方向付けされます。
  - コンテンツが切り捨てられた場合、サイトは一時不適格と見なされます。
  - コンテンツがエラーなしで解析された場合、サイトは、検出された規則によるクローラーに適格と見なされます。
  - コンテンツの解析でエラーが発生した場合、サイトは、規則なしのクローラーに適格と見なされます。
4. その他の HTTP 状況が戻された場合、サイトは一時不適格と見なされます。

クローラーは、サイトの robots.txt ファイルのダウンロードを試行するとき、ロボット日付と呼ばれる、そのサイトの永続タイム・スタンプを更新します。robots.txt 情報が使用不可のためにサイトが不適格である場合、永続ロボット失敗カウントが 1 つ増えます。

再試行間隔に達すると、クローラーは、失敗したサイトの robots.txt 情報のリトリブを再試行します。連続失敗回数が失敗の最大許容回数に達した場合、クローラーは、そのサイトの robots.txt ファイルをリトリブする試行を停止し、そのサイトをクローラーに不適格と見なします。

サイトがクローラーに適格となる (robots.txt ファイルの規則の検査が正常終了する) と、失敗カウントはゼロに設定されます。クローラーは、規則検査間隔が経過するまで、ダウンロードの結果を使用します。その時点で、サイトは、再び適格性を判断される必要があります。

#### ヒント:

- サーバーがコンテンツを戻すが構文エラーを含む場合、サーバーが 1994 年版以外のロボット・プロトコルを使用する場合、またはコンテンツにロボット規則以外のもの (ソフト・エラー・ページなど) が含まれる場合、クローラーは、サーバーが適用可能な規則ファイルを持たない場合と同じように振る舞い、サイトをクローラーします。コレクション管理者はサイト・コンテンツまたはデフォルトのサーバーの振る舞いを制御しないため、このアクションは、通常、正しいものです。Web サーバー管理者が、サイトがクローラーされることを望まず、適合規則ファイルをインストールしたくない場合、コレクション管理者は、サイトのドメイン、IP アドレス、または HTTP 接頭部をクローラーの規則で指定することによって、Web クローラーからサイトを遮断することができます。
- サーバーが 302 状況コードまたはその他のリダイレクト・コードを戻した場合、クローラーは、サイトに使用すべき robots.txt ファイルがあるが、そのファイルが適合した場所 (サイト・ルート) にないことを意味するものとして、そのコードを解釈します。Web サーバー管理者は、Web クローラーがファイル内の規則に従うことができるように、そのファイルを正しい場所に移動する必要があります。

- 証明書の問題 (例えば、証明書が期限切れである、認証局が信頼されない、証明書が自己署名だがクローラーは自己署名の証明書を受諾するよう構成されていない) がある場合、クローラーは、そのサイトとの接続の失敗として問題を解釈し、サイトを不適格と見なします。同じ問題により、そのサイトの他のページをクロールできなくなる可能性があります。サイトをクロール可能にするには、コレクション管理者が、自己署名証明書を有効にするか、サイトの権限をトラステッド鍵ストア・ファイルに追加するか、または Web サーバー管理者に最新の証明書の取得を依頼する必要があります。
- Web クローラーは、HTTP 基本認証 (HTTP 基本プロキシ認証を含む) を使用するように構成されている可能性があります。適切に構成されている場合には、robots.txt ファイルのダウンロードにも認証が必要です。状況コード 403、407、または他の認証に関連した応答は、許可の問題を示し、クローラーはサイトを不適格と見なします。(HTTP 基本認証のみがサポートされます。)
- サイトの robots.txt ファイルがロボット・ページの最大長を超えた場合、コレクション管理者は構成済みの最大値 (デフォルト値の 100 万バイトで十分なはずです) を上げることができます。

問題のトラブルシューティングに役立てるために、Web クローラーのモニター時にサイト・レポートを要求することができます。robots.txt ファイルの内容を表示する (これにより、サイトへの Web クローラーのアクセスが規則で禁止されているかを確認する) オプション、クローラーが robots.txt ファイルのダウンロードを最後に試行した日時 (再試行間隔が経過するまでクローラーは再試行しません) を確認するオプション、robots.txt ファイルのダウンロードをクローラーが何回連続して試行して失敗したかを確認するオプションを選択します。これらのサイト・レポート・オプションの詳細や、結果の解釈方法については、Web クローラーのモニター中に「ヘルプ」をクリックしてください。

ロボット排他プロトコルについて詳しくは、URL: <http://www.robotstxt.org/wc/exclusion.html> を参照してください。

## JavaScript のサポート

エンタープライズ・サーチの Web クローラーが、Web 文書の JavaScript™ 部分に含まれているリンク (URL) を検出する場合があります。

Web クローラーは、相対リンクと絶対リンクの両方を検出できます。HTML 文書に BASE エlementが含まれている場合、クローラーはそのElementを使用して相対リンクを解決します。そうでない場合は、クローラーは、文書固有の URL を使用します。

JavaScript のサポートはリンク抽出に限定されています。クローラーは JavaScript の構文解析、DOM (Document Object Model) の構築、および、JavaScript ステートメントの解釈や実行は行いません。クローラーは JavaScript ステートメントの URL である可能性の高いストリングを文書コンテンツ (JavaScript 部分を含むがそれに限定しない) で検索します。これは 2 つのことを意味しています。

- 一部の URL は検出されても、厳しい HTML パーサーでは無視されます。クローラーは、構文的に有効でない URL はすべて拒否しますが、スキャンのステップで戻される有効な URL の中には、検索にあまり重要でないものがあります。

- 人がブラウザでページを表示してそのブラウザが JavaScript を実行するときなどに、JavaScript によって生成される文書コンテンツは、Web クローラーで検出できず、したがって、索引付けもされません。

Web クローラーは HTML ファイル内の JavaScript を解析しないため、JavaScript 内の URL はクロールされません。Web クローラーが JavaScript 内の URL をクロールできるようにするには、次のいずれかのアクションを実行します。

- エンタープライズ・サーチ管理コンソールで、Web クローラーを編集し、「Web クロール・スペース」ページで、クローラーがコレクションに URL を追加するための開始点として使用する URL (開始 URL) のリストに URL を追加します。行った変更を有効にするには、Web クローラーを再始動します (フル・クロールを開始する必要はありません)。
- アンカー・タグ (<a href="..">) を使用して、URL を HTML ファイル内でハイパーテキスト・リンクとして指定します。

## Web クロール・スペースを制限する規則

ユーザーが、検索を許可されている Web サイトのみにアクセスするように、Web クローラーのクロール対象を制限する規則を指定します。

Web クローラーが Web ページをクロールするときに他のページへのリンクを検出すると、それらのリンクを次にクロールされるようにキューに入れます。このクロールと検出は、時間とメモリー・リソースが許す限り繰り返して行われます。Web クローラーを構成する場合、クローラーがクロールを開始する場所を指定します。Web クローラーは、この初期 URL (開始 URL といいます) から始めて、直接または間接的に接続されている Web 上のあらゆる文書に到達できます。

クロール・スペースを制限するには、Web クローラーを構成するときに特定の URL は完全にクロールし、関心がある領域外を示すリンクは無視するように設定します。デフォルトではクローラーは、検出したすべての URL を受け付けるようになっているため、コレクションに入れる URL を指定し、それ以外のページは除外する規則を指定する必要があります。

Web クローラーにクロール対象とクロールから除外する対象を指定するには、いくつかの方法があります。次のような指定を行うことができます。

- クローラーがクロールを開始する開始 URL のリスト
- 3 種類のクロール規則 (ドメイン、IP アドレス、URL 接頭部)
- クロール対象の文書を指定する MIME タイプのリスト
- 除外対象の文書を指定するファイル拡張子のリスト
- 1 つの URL パスに許容される最大ディレクトリー数

クロール規則は、次の書式で指定します。

```
action type target
```

action は forbid または allow、type はドメイン (domain)、IP アドレス (IP address)、または URL 接頭部 (HTTP または HTTPS) で、target は type の値によって異なります。パターンに一致するターゲットの指定には、ワイルドカード文字としてアスタリスク (\*) を指定できますが、その使用は限定的になります。

## ドメイン規則

ドメイン規則のターゲットは DNS ドメイン・ネームです。例えば、www.ibm.com ドメイン全体をクロール対象にする場合、次のように指定します。

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

ワイルドカード文字としてアスタリスクを指定できます。これにより、残りのパターンに一致するすべてのホスト名にこの規則が適用されることとなります。例えば、server で始まり ibm.com で終わるドメインをクロールしない場合、次のように指定できます。

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

ホスト名の突き合わせでは、具体的なドメイン・ネームを明示して指定した場合でも、ドメイン・ネームのパターンを指定した場合でも大/小文字は区別されます。例えば、「\*.user.ibm.com」と指定すると、「joe.user.ibm.com」および「mary.smith.user.ibm.com」は一致となりますが、「joe.user.IBM.com」は一致となりません。

ポート番号を指定しないドメイン規則は、そのドメイン上のすべてのポートに適用されます。以下の例では、sales ドメインのすべてのポートが許可されます。

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

ドメイン規則でポート番号を指定する場合、その規則はそのポートのみに適用されます。以下の例では、sales ドメインのポート 443 のみ許可されます。

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

## 接頭部規則

接頭部規則は、指定されたストリングで始まる URL のクロールを制御します。ターゲットは、単独の URL で、通常、1 つ以上アスタリスクを入れてパターンを指定します。例えば、よくある方法として、接頭部ストリングの末尾文字にアスタリスクを指定します。

接頭部規則を使用すれば、Web サイトのすべて、または一部をクロールすることができます。ディレクトリー・パスかパターンを指定し、ディレクトリー・ツリーのそれ以降のすべてを対象として、クロールを許可または禁止します。例えば、次の 2 つの規則では、sales.ibm.com の public ディレクトリーのすべてのコンテンツをクローラーのクロール対象とし、そのサイトの他のすべてのページについてはアクセスを禁止します。

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

接頭部規則を指定する場合、アスタリスクは複数指定が可能で、末尾に限らず接頭部ストリングの任意の位置に指定できます。例えば、次の規則では、sales.ibm.com サイトでディレクトリー名の末尾が fs の最上位ディレクトリーの文書を、クローラーがクロールするのを禁止します。(検索索引で有用な情報がないファイル・システム・マウントなどの場合が考えられます。)

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

## アドレス規則

アドレス規則は、ターゲットとして IP アドレスとネットマスクを指定することにより、ホストまたはネットワーク全体のクローラを制御します。次に例を示します。

**IPv4** allow address 9.0.0.0 255.0.0.0

**IPv6** エンタープライズ・サーチを Windows 2003 サーバーで実行するとき、エンタープライズ・サーチ・システムが IP バージョン 6 (IPv6) プロトコルを使用できるように設定した場合は、アドレスを大括弧で囲む必要があります。

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

ネットマスクを使用するとパターン・マッチングを指定できます。アドレス規則を候補となる IP アドレスに適用するには、規則の IP アドレスと候補 IP アドレスが、ネットマスクでゼロで隠されている部分を除き、同一でなければなりません。アドレス規則はパターンを定義し、ネットマスクはアドレス・パターン中の有意のビット部分を定義します。ネットマスク中のゼロは、ワイルドカードのような働きをし、アドレス内でゼロと同じビット位置にある値はすべて一致すると見なされません。

上記の例の allow 規則は、最初のオクテットが 9 で末尾の 3 つのオクテットが任意の値のすべての IP アドレスに適合します。

次の規則は、アドレス規則のリストの最後のアドレスとして組み込む有用な規則です。この規則は、ネットマスクの全ビットを有意でないビットとしているため、あらゆる IP アドレスが一致となります (この規則では、規則リストでこの前のすべての規則で許可されていないすべてのアドレスを禁止することになります)。

**IPv4** forbid address 0.0.0.0 0.0.0.0

**IPv6**

forbid address :: ::

**プロキシ・サーバーに対する制限:** プロキシ・サーバーによってサービスされている Web サイトをクローラする場合は、IP アドレス規則を指定しないでください。プロキシ・サーバーは、通常、ユーザー・エージェント (ブラウザまたはクローラ) が、Web サーバーがあるネットワークに対して直接アクセス権限を持っていない場合に使用されます。例えば、HTTP プロキシ・サーバーは、HTTP 要求をクローラから Web サーバーに中継し、それに対する応答をクローラに送信することができます。

Web クローラがプロキシ・サーバーを使用する場合、プロキシ・サーバーの IP アドレスは、クローラが別のホスト用に持っている唯一の IP アドレスです。IP アドレス規則により、クローラをあるサブネットの IP アドレスに制約している場合、この制約によって、ほとんどすべての URL が状況コード 760 (URL が Web スペースによって禁止されていることを示す) に分類されます。

## クローリング規則の順序

クローラーは、URL の検出およびクローリング処理中、クローリング規則を何度も適用します。規則の順序は重要ですが、それは同タイプの規則に限られます。あるアドレス規則が別のアドレス規則の前に来るか、後ろに来るかでは意味が違いますが、あるアドレス規則が別の接頭部規則の前に来るか、後ろに来るかでは意味の違いはありません。クローラーがこの 2 つの規則を同時に適用することはないからです。

同タイプの規則セット内については、クローラーは候補ドメイン、アドレス、または URL を最初に指定した規則から順次最後の規則まで適合する規則が見つかるまで 1 つずつテストします。適合する最初の規則に指定された処置が使用されます。

順序に依存しているため、多くのクローリング規則は以下のような典型的構成になります。

- ドメイン規則のセットは、通常、クローリング・スペースから単独のドメインを除外する `forbid` 規則を先に置きます。例えば、コレクション管理者は、有用な情報がないドメインをここで除外指定します。
- `forbid` 規則のリストの後には、通常、ワイルドカード文字を入れた一連の `allow` 規則を設定し、これによって、エンタープライズ・イントラネットを定義する上位レベルのドメイン・ネームが末尾に来る任意のドメイン (`*.ibm.com`、`*.lotus.com` など) をクローラーがアクセスできるようにします。

ドメイン規則のセットの最後に次のデフォルト規則を入れ、それより前の規則で許可されなかったドメインをクローリング対象から除外します。

```
forbid domain *
```

この最後の規則は、クローリング・スペースにインターネット全体が入らないようにするため、重要です。

- アドレス規則のセットは、通常、少数の `allow` 規則が先頭に置かれ、これによってクローラーがエンタープライズ・イントラネットにわたる上位レベル (クラス A、クラス B、またはクラス C) のネットワークをクローリングできるようにします。

アドレス規則のリストに最終規則を指定して企業ネットワークの外部の Web サイトをクローラーがクローリングしないようにする方法の例については、アドレス規則に関する前述の説明を参照してください。

- 通常、接頭部規則のセットは、ツリーまたはサブツリーで許可および禁止指定される領域が詳細にその都度指定されるため、この規則セットが一番大きなセットになります。よい方法としては、まず、より特定化された領域を許可または禁止し、それから逆に一般的なパターンでその他すべての領域を許可または禁止します。

接頭部セクションの最後は、通常の規則とは異なります。ドメイン規則およびアドレス規則について推奨される最終規則を使用すれば、URL の接頭部をテストするよりも効率的に、クローラーがエンタープライズ・ネットワークを超えてクローリングしないようにすることができます。

接頭部規則の場合、処置 (`forbid` または `allow`) 別に規則をまとめるほうが、クローラーはより効率的に処理を行うことができます。例えば、`allow` および `forbid`

規則を交互に短い指定で並べるよりも、片方の処置を規定する規則を長く並べ、それから他方の処置を規定する規則を同様に長く並べるようにします。allow 規則と forbid 規則を混在させるようにしても目的のクローल・スペースを作成することはできますが、allow 規則と forbid 規則をそれぞれひとまとめにしたほうが、クローラーのパフォーマンスは良くなります。

## ファイル拡張子、MIME タイプ、およびクロールの最大の深さ

これらのオプションを使用すれば、クロール・スペースの内容をさらに詳細に指定することができます。特定タイプの文書を文書のファイル拡張子に基づいて除外したり、文書の MIME タイプを基に特定タイプの文書を組み入れたりすることができます。クローラーにクロールさせる MIME タイプを指定する場合、Web 文書では MIME タイプが誤って設定されていることが多いので注意してください。

クロールの最大の深さは、サイト・ルートから数えた URL 内のスラッシュ数で表されます。このオプションを使用すると、クローラーが深さが無限の再帰的なファイル・システム構造にはまり込むのを防ぐことができます。クロールの深さは、クローラーがある文書から別の文書にリンクをトラバースする時のレベル数とは一致しません。

## 開始 URL

開始 URL は、クローラーがクロールを開始する URL のことを指し、クローラーが開始されるたびにこの URL がクロール対象となります。開始 URL が既に検出された場合は、その URL がクロール規則で許可された他の Web サイトよりも早くクロールまたは再クロールされることはありません。

開始 URL は Web クローラーが初めて始動され、クロール・スペースが空の場合に重要です。また、まだ検出されていなかった開始 URL をクロール・スペースの開始 URL のリストに追加するときにも重要となります。

開始 URL にはドメイン・ネームだけでなく完全修飾 URL を指定する必要があります。プロトコル、およびポートが 80 でない場合は、ポート番号も指定してください。

以下に、有効な開始 URL の例を示します。

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

以下の URL は無効な開始 URL です。

```
www.ibm.com
```

また、開始 URL をクロール規則に入れる必要があります。例えば、クロール規則で開始 URL をクロール対象として許可していなければ、クローラーは指定された開始 URL でクロールを開始することができなくなります。

**IPv6 アドレスのサーバー・ポート:** エンタープライズ・サーチを Windows 2003 サーバーで実行するとき、エンタープライズ・サーチ・システムが IP バージョン 6 (IPv6) プロトコルを使用できるように設定した場合は、開始 URL を大括弧で囲む必要があります。例えば、次のようにします。



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

### 関連タスク

31 ページの『IPv6 プロトコルのサポートの使用可能化』

## Web クローラーとの URL 接続のテスト

Web クローラーがクロールする URL を指定したら、クロール規則の構成をテストできます。

クロールするドメイン、HTTP 接頭部、または IP アドレスを指定したら、「テスト」をクリックできます。また、「URL のテスト」ページを選択して、指定した URL に加えて、開始 URL にクローラーが接続できるかをテストすることができます。

テスト結果により、クローラーのプロパティに指定されているユーザー・エージェント名を使用して、クローラーが URL にアクセスできるかどうかを示されます。また、テスト結果により、URL が排他規則によりクロールできないかどうかも示されます (例えば、クロール・スペースから除外されている拡張子に一致するファイル拡張子を持つ文書はクロールされません)。

サイトが少なくとも 1 度クロールされた後、URL をテストして、追加情報を入手できます。例えば、テスト・レポートにより、最新の HTTP 状況コード (URL のクロールが正常に行われたかどうかを示します)、URL の最終クロール日時、再クロールがスケジュールされている日時、ユーザー・エージェントが Web サーバーの現行の robots.txt ファイルを使用しているかどうか、などが示されます。

## Web クローラーにおける再クロール間隔の設定

Web クローラーが URL に再アクセスする頻度を調整するには、Web クローラーのプロパティのオプションを指定します。

エンタープライズ・サーチ・システムの他のほとんどのクローラー・タイプの場合、管理者が指定したスケジュールに従ってクロールが実行されます。これに対して、Web クローラーの場合、通常、開始されると連続して実行されます。前にクロールした URL に再アクセスする頻度を調節するには、最大および最小再クロール間隔を指定します。

エンタープライズ・サーチ管理コンソールを使用して Web クローラーを作成する、または Web クローラーのプロパティを編集するときに、拡張プロパティを構成するオプションを選択できます。最大および最小再クロール間隔は、この「Web クローラーの拡張プロパティ」ページで指定します。Web クローラーは、指定したこの値を使用して、データの再クロール間隔を計算します。

あるページが最初にクロールされると、クローラーはそのページがクロールされた日時と指定された最大と最小再クロール間隔の平均値を使用して、再クロール日付を設定します。そのページはその日付までは再クロールされません。その日付以後、そのページが再クロールされる実際の日時は、クローラー負荷とクロール・スペース内の新旧の URL のバランスによって異なります。

毎回、そのページが再クロールされるごとに、クローラーはその内容に変更がないか確認します。内容に変更がある場合、次の再クロール間隔は前回よりも短くなりますが、指定された最小再クロール間隔よりも短くなることはありません。内容に変更がない場合、次の再クロール間隔は前回よりも長くなりますが、指定された最大再クロール間隔よりも長くなることはありません。

## Web クローラーにおける URL アクセスのオプション

Web クローラーには、特定の URL をできる限り早くアクセスするように強制することができます。

特定の Web サイトの情報でクロール・スペースをリフレッシュする必要がある場合、クローラーをモニターし、「**アクセスまたは再アクセスする URL**」オプションを選択して、クロールまたは再クロールする必要があるページの URL または URL パターンを指定します。

例えば、企業の情報通信部門がイントラネットにある Web ページを追加する、または重要なポリシー変更があったため、それを反映するようにページを改訂する場合、その新規、または変更済みのページの URL を指定します。クローラーが実行中の場合、クローラーは指定された URL を次にクロールするようにキューに入れ、アクセスを待機しているページがないかその URL を確認します (通常 10 分ごと)。クローラーが実行されていない場合は、クローラーは指定された URL をキューに入れ、クローラーが次回開始されたときにクロールできるようにします。

クロール規則の中に、指定した URL のアクセスをクローラーに許可する規則が必ず含まれるようにしてください。クローラーは、指定した URL には通常より早くアクセスできますが、前提としてすべてのクロール対象の URL には、その URL のクロールを許可するクロール規則がなければなりません。

新規にクロールされたデータは、次に主索引作成が行われた時点で検索可能になります。

## Web クローラーがソフト・エラーのページを処理する方法

特定ページ要求の返答に標準的なエラー・コードを返したくない場合に、Web サイト管理者が作成したカスタム・ページを処理する Web クローラーを構成することができます。

HTTP サーバーが、クライアントが要求したページを戻すことができない場合、サーバーは、通常、ヘッダーから成る応答に状況コードを付けて戻します。状況コードで、何が問題であるかを示します (例えば、ファイルが見つからなかったことを表すエラー 404 など)。Web サイト管理者によっては、問題をより詳細に説明する特別なページを作成し、代わりにそのページを戻すように HTTP サーバーを構成します。こうしたカスタム・ページは、ソフト・エラー・ページと呼ばれます。

ソフト・エラー・ページは、Web クローラーの結果をゆがめることがあります。例えば、クローラーは、問題を表すヘッダーを受け取る代わりに、ソフト・エラー・ページと状況コード 200 (有効な HTML ページの正常なダウンロードを表す) を受け取ることがあります。しかし、ダウンロードされるこのソフト・エラー・ページは要求された URL とは関連がなく、その内容は、要求されたページの代わりに戻

されるたびに、ほとんど同じです。関連のない、ほとんど重複しているこれらのページが、索引および検索結果をゆがめます。

この状況に対応するために、Web クローラーの構成時にソフト・エラー・ページの処理オプションを指定することができます。Web クローラーは、ソフト・エラー・ページを返す各 Web サイトについて以下の情報を必要とします。

- ソフト・エラー・ページを使用するサイトの URL パターン。この URL パターンは、プロトコル (HTTP または HTTPS)、ホスト名、ポート番号 (標準でない場合)、およびパス名で構成されます。パターンには、ワイルドカード文字としてアスタリスク (\*) を使用することができ、非ワイルドカード文字が次に現れる位置までの 1 字以上の文字とマッチングさせることができます。指定するパターンは、大/小文字を区別します。
- HTML 文書の <TITLE> タグに対応するテキストのタイトル・パターン。このパターンの指定には、ワイルドカード文字としてアスタリスク (\*) を使用できます。指定するパターンは、大/小文字を区別します。
- HTML 文書のコンテンツに対応するテキストのコンテンツ・パターン。<BODY> タグがある場合、コンテンツは、単に <BODY> タグの内容だけではありません。コンテンツは、ファイルの HTTP ヘッダーの後にあるすべてです。このパターンの指定には、ワイルドカード文字としてアスタリスク (\*) を使用できます。指定するパターンは、大/小文字を区別します。
- 指定した URL、タイトル、およびコンテンツ・パターンと一致する文書に使用する状況コードを表す整数。

## 例

以下の構成は、Web クローラーに、`http://www.mysite.com/hr/*` という Web サイトから戻される、すべての有効な HTML ページ (状況コード 200) を、指定されたタイトルおよびコンテンツ・パターンと比較するように指示します。ページの <TITLE> タグが、「Sorry, the page」で始まっていて、文書のコンテンツに任意の文字 (\*) が含まれている場合には、クローラーはそのページを、状況コード 404 を戻す (ページが見つからなかった) 場合と同じ方法で扱います。

表 3. ソフト・エラー・ページの例

| URL パターン                                | タイトル・パターン                     | コンテンツ・パターン     | HTTP 状況コード |
|-----------------------------------------|-------------------------------|----------------|------------|
| <code>http://www.mysite.com/hr/*</code> | <code>Sorry, the page*</code> | <code>*</code> | 404        |

異なる状況コードを扱うために、同一の Web サイトについて複数の項目を作成できます。同一 Web サイトからの各状況コードごとに、Web クローラーの構成で固有の項目が必要です。

## ワイルドカード文字の使用

URL、タイトル、および、コンテンツ・パターンは、正規表現ではありません。アスタリスク文字は、非ワイルドカード文字が次に現れる位置までの任意の文字と一致します。次に例を示します。

\*404 は、任意の文字 404 と一致します。

404: \* は、404: 任意の文字と一致します。

http://\*.mysite.com/\* は、http://任意のホスト.mysite.com/任意のファイル名と一致します。

\* は、任意の文字 と一致します。

## パフォーマンスへの影響

ソフト・エラー・ページの処理オプションを構成すると、正常にクロールされたページをすべて検査する必要があるため、クローラー・プロセスの時間が長くなります。パターン・マッチングの検査と、ページまたは置き換えた状況コードを戻すかどうかの判断に、さらに処理時間が必要になります。

## セキュア Web サイトのクロールのサポート

エンタープライズ・サーチ管理コンソールにクレデンシャルを指定することで、アクセスにパスワードが必要な文書などの制限付きコンテンツに Web クローラーが、アクセスできるように設定することができます。

Web サーバーが HTTP 基本認証または HTML フォーム・ベースの認証を使用している場合、Web サイトへのアクセスを制限している場合、Web クローラーの構成でクレデンシャルを指定して、パスワード保護された Web サイト上のページをクロールできるようにします。また、Cookie ファイルを手動で構成するオプションも指定できます。

### HTTP 基本認証によって保護された Web サイト

Web サーバーが HTTP 基本認証を使用して Web サイトへのアクセスを制限している場合、Web クローラーがパスワード保護されたページにアクセスできるように認証のためのクレデンシャルを指定してください。

ユーザー (あるいはクライアント・アプリケーション) に Web サイト上のページへのアクセス権があるかどうかを判別するために、多くの Web サーバーが HTTP 基本認証と呼ばれるクライアント認証方式を採用し、ユーザーの身元を確認しています。通常、この操作は、対話式に行われます。

- HTTP ユーザー・エージェント (Web ブラウザーなど) が HTTP 基本認証で保護されているページを要求すると、Web サーバーは、要求側に要求ページへのアクセス権がないことを示す、状況コード 401 で応答します。
- また、Web サーバーは、ユーザーが制限付きコンテンツへのアクセス許可を得ているかどうかを検証するためのクレデンシャルを提示するよう要求側に求めます。
- Web ブラウザーは、ユーザー名、パスワード、その他のユーザーのクレデンシャルに必要な情報を要求するダイアログをユーザーに表示します。
- Web ブラウザーはクレデンシャルをエンコードして、保護されたページが要求されるたびにそれを組み込みます。
- クレデンシャルが有効であれば、Web サーバーは、戻りコード 200 と要求ページのコンテンツを戻します。
- それ以降の同じ Web サーバーからのページ要求には、通常、同じクレデンシャルが含まれています。これにより、許可ユーザーは、追加のクレデンシャルを指定せずに他の制限付きコンテンツにアクセスできるようになります。

ユーザーの身元が確認された後、Web サーバーと HTTP ユーザー・エージェントは、通常、*Cookie* というトークンを交換します。*Cookie* を使用すると、ユーザー・ログイン状況の認識を複数の HTTP 要求間で保持することができます。

Web クローラーは対話式で動作しないので、パスワード保護されたページのクロールを可能にするためのクレデンシャルを、クロールの開始前に指定しておく必要があります。Web クローラーを作成したり、クロール・スペースを編集する場合、クロール対象のセキュア Web サイトそれぞれについて情報を指定してください。

この情報を指定するには、HTTP 基本認証によって保護されている Web サイトまたは Web サーバーの管理者と緊密に作業する必要があります。管理者は、クロール対象の Web サイトのセキュリティー要件を提供してくれるはずです。これには、Web クローラーの識別の認証や、制限付きページをクロールするアクセス権がクローラーにあるかどうかの判別使用するあらゆる情報が含まれます。

コレクションの作成時にそのコレクションに対してセキュリティーを有効に設定している場合は、文書へのアクセスを制御するために、ユーザー ID、グループ ID、あるいはユーザー役割などのセキュリティー・トークンを、クローラー構成時に指定してください。Web クローラーは、このセキュリティー・トークンを、指定されたルート URL のファイル・システム・ツリーでクロールするあらゆる文書に関連付けます。このトークンは Web クロール・スペース全体に対して設定する文書レベルのセキュリティー・トークンとは別に使用されます。

URL の順序は、重要です。パスワード保護された Web サイトに関する情報を追加したら、それをクローラーに処理させたい順序に配置する必要があります。リスト内では、特定の URL を最初に置き、汎用的な URL が続くように配置してください。Web クローラーは、候補の URL を評価するとき、その候補の URL と一致した、リストの最初にある URL に対して指定された認証データを使用します。

## フォーム・ベースの認証によって保護された Web サイト

Web サーバーが HTML フォームを使用して Web サイトへのアクセスを制限する場合、Web クローラーがパスワード保護されたページにアクセスできるようにする認証のためのクレデンシャルを指定できます。

ユーザー (あるいはクライアント・アプリケーション) に Web サイト上のページへのアクセス権があるかどうかを判別するために、多くの Web サーバーで HTML フォームを使用して、ユーザーの身元を確認しています。通常、この操作は、対話式に行われます。

- HTTP ユーザー・エージェント (Web ブラウザーなど) がフォーム・ベースの認証で保護されているページを要求すると、Web サーバーは、要求にユーザーの身元を証明する *Cookie* が含まれているかどうかを確認します。
- *Cookie* がない場合は、Web サーバーはユーザーにプロンプトを出して、セキュリティー・データをフォームに入力するよう求めます。ユーザーがフォームを実行すると、Web サーバーは、必要な *Cookie* を戻し、それからパスワード保護されたページの要求処理が進められます。
- また、必要な *Cookie* を組み込んでいれば、その後の要求も処理されます。許可ユーザーは、要求のたびごとにフォームへの入力やクレデンシャルの指定を求められることなく、他の制限付きコンテンツにアクセスできます。

Web クローラーは対話式で動作しないので、パスワード保護されたページのクローリングを可能にするためのクレデンシャルを、クローリングの開始前に指定しておく必要があります。Web クローラーを作成したり、クローリング・スペースを編集する場合、クローリング対象のセキュア Web サイトそれぞれについて情報を指定してください。

指定したフィールドは、対話式ユーザーが Web ブラウザーにプロンプトで求められて入力するフィールド、および、ログイン実行に必要な非表示フィールドや静的フィールドと対応します。

この情報を指定するには、フォーム・ベースの認証によって保護されている Web サイトまたは Web サーバーの管理者と緊密に作業する必要があります。管理者は、クローリング対象の Web サイトのセキュリティー要件を提供してくれるはずです。これには、Web クローラーの識別の認証や、制限付きページをクローリングするアクセス権がクローラーにあるかどうかの判別に使用するあらゆる情報が含まれます。

URL パターンの順序は、重要です。パスワード保護された Web サイトに関する情報を追加したら、それをクローラーに処理させたい順序に配置する必要があります。リスト内では、特定の URL パターンを最初に置き、汎用的な URL パターンが続くように配置してください。Web クローラーは、候補の URL を評価するとき、その候補の URL と一致した、リストの最初にある URL パターンに対して指定されたフォーム・データを使用します。

### **セキュアな WebSphere Portal サイトをクローリングするプラグインの使用**

WebSphere Application Server でグローバル・セキュリティーが有効になっているとき、Web クローラーで WebSphere Portal セキュア・サイトをクローリングする場合は、フォーム・ベースの認証要求を処理するクローラー・プラグインを作成する必要があります。フォーム・ベースの認証、およびカスタム Web クローラー・プラグインに適合させることができるサンプル・プログラムについては、<http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani> を参照してください。

プラグインが必要なのは、Web クローラーを使用し、WebSphere Portal を介してサイト (Workplace Web Content Management サイトおよび Lotus Quickr サイトを含む) をクローリングする場合です。

## **プロキシ・サーバーによってサービスされている Web サイト**

Web クローラーがネットワークへの直接アクセスを許可されていない場合、クローラーが HTTP プロキシ・サーバーを使用してクローリング対象のコンテンツにアクセスするように構成することができます。

Web クローラーが稼働するコンピューター上で TCP/IP ネットワークへのアクセスが使用できない場合、あるいは、アクセスが特権のある処理に制限されている場合は、Web クローラーで HTTP プロキシ・サーバーを使用するように構成できます。HTTP プロキシは、指定されたホストの指定されたポートで HTTP 要求を listen するプロセスです。プロキシ・サーバーは要求を Web サーバーにつなぎ、Web サーバーからの応答を要求元のクライアント (Web クローラー) にリレーします。プロキシ・サーバーは、Web クローラーと同じコンピューター上で稼働することも、別のコンピューター上で稼働することもできます。

プロキシ以外のクローラでは、URL に対する要求は直接ホストに送信されます。プロキシのクローラでは、要求はプロキシ・サーバーに送信されます。

Web クローラーの作成や、クローラ・スペースの編集の際は、Web クローラーがプロキシ・サーバー・ドメインでページのクローラに使用するプロキシ・サーバーの情報を指定してください。プロキシ・サーバーをクローラ・スペースに追加する前に、プロキシ・サーバーによってサービス提供されるドメインの名前、プロキシ・サーバーのホスト名または IP アドレス、プロキシ・サーバーが使用するポート番号を入手してください。

プロキシ・サーバーが認証を必要とする場合は、そのプロキシ・サーバーによって提供されるページへのアクセスに Web クローラーが使用できるユーザー名とパスワードも入手します。Web クローラーは、RFC2616 (<http://rfc.net/rfc2616.html>) に記述された HTTP 基本プロキシ認証のみをサポートします。他のタイプの認証 (Windows NT LAN Manager (NTLM) など) はサポートされません。

プロキシ・サーバーを追加した後は、それを選択して、クローラーに処理させたい順序で配置する必要があります。リスト内では、特定のドメイン・ネームを最初に置き、汎用的なドメイン・ネームが続くように配置してください。Web クローラーは、候補 URL を評価するときに、候補 URL と一致して、かつ、リストの最初のドメインに指定されたプロキシ・サーバー・データを使用します。(どのプロキシ規則にも一致しない URL は、クローラーに直接アクセス可能であると想定されます。)

## Cookie の管理

通常、Cookie の管理は自動的に行われ、エンタープライズ・サーチ管理者が特別な処置をする必要はありません。必要な場合は、Web クローラ・セッションの Cookie を手動で指定することができます。

Cookie は、Web サーバーが HTTP 応答ヘッダーの一部としてユーザー・エージェントに戻す内部が見えないトークンです。これは、発行した Web サーバーにのみ意味があるもので、HTTP 要求の間の状態維持に使用されます。例えば、クライアントの認証の間、Web サーバーは Cookie を戻し、認証されたユーザーが既にログインされていることをサーバーが判断できるようにします。Cookie があることでユーザーは再度ログインを求められることなく、その Web サーバー上の他のページの要求を出すことができます。

Web クローラーは、Web サーバーから受け取った Cookie を保持し、そのクローラー・インスタンスが継続している間、それを使います。Cookie は、`cookies.ini` ファイルに保管されます。これは、クローラー・セッションが終了するたびに再度書き込まれます。Web クローラーが停止すると、期限切れになっていない Cookie はすべて保存されて次のセッションの開始時に再ロードされます。

手動で Cookie を指定する場合は、まず別ファイルに保管して、必要なときに `cookies.ini` ファイルの Cookie とマージしてください。クローラーは、期限切れになっていない Cookie を廃棄しませんが、問題が発生して Cookie コレクション全体の書き込みができない場合にも、手動で指定した Cookie を失わないようにするためです。クローラ・セッションが開始するまでにクローラーが自動的に保守している Cookie とユーザーの Cookie をマージする必要があります。

## Cookie フォーマット

エンタープライズ・サーチの `cookies.ini` ファイルとマージする予定の Cookie は特定のフォーマットである必要があります。

- 各 Cookie は単一行に入っていなければなりません。空白行とコメントは許可されていますが、`cookies.ini` ファイルに保存することはできません。
- 各 Cookie は次のフォーマットでなければなりません。

```
CookieN(cookie_length,URL_length)cookie_text,validation_URL
```

### Cookie

Cookie 項目の始まりを示す必要なキーワード。

Cookie キーワードには、空白を含むことはできず、また、最後に 0 か、1 か、2 のいずれかの 1 桁の数字が付いていなければなりません。この数字は、次の Cookie タイプを表します: バージョン -0 (Netscape)、バージョン -1 (RFC2109)、あるいはバージョン -2 (RFC2965)。ポート・リストは、RFC2965 Cookie ではサポートされていません。

### *cookie\_length*

関連した Cookie テキストの文字数。

### *URL\_length*

関連した検証 URL の文字数。

### *cookie\_text*

発信元の Web サーバーに送信される予定の Cookie のコンテンツ。このストリング (HTTP 応答ヘッダーにある Set-Cookie ディレクティブの右側に相当する部分) には、Cookie の名前と値の組、およびその Cookie と一緒に送信するその他のコンテンツ (パス、セキュリティ設定など) が入ります。このストリングの後ろには、区切り記号のコンマ (,) が続きます。

### *validation\_URL*

この Cookie が検出された URL。この URL は、Cookie の送信先の決定に使用されます (例えば、ドメイン名とパス名を指定することで決定します)。検証 URL は、発信元の Web サーバーの Cookie に関するセキュリティとプライバシーの制約事項を満たしている必要があります。

次の例では、読みやすくするために 2 行に分けて示します。Cookie を指定するときは、一行に入れてください:

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

## Web クローラーの Cookie の構成

Web クロール・セッションの Cookie を手動で指定し、Web クローラーが維持する Cookie とマージできます。

### 始める前に

Web クローラーが使用する Cookie を手動で構成するには、エンタープライズ・サーチ管理者でなければなりません。

### 手順

Web クローラーの Cookie を手動で構成するには、次のようにします。



1. エンタープライズ・サーチ管理コンソールから、Cookie を指定したいコレクションをモニターして、Web クローラーを停止します。
2. クローラー・サーバーで、エンタープライズ・サーチ管理者としてログインします。(このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。)
3. 構成したいクローラーのデータ・ディレクトリーに移動します。ここで、`crawler_session_ID` は、エンタープライズ・サーチ・システムによりクローラー・セッションに割り当てられたものです。次に例を示します。  
`ES_NODE_ROOT/data/col_56092.WEB_88534`
4. `cookies.ini` ファイルを編集し、手動で指定した Cookie 項目を、既にリストされている項目に追加して、そのファイルを保存して終了します。追加する Cookie が、既に存在しているものをオーバーライドしないようにします。
5. エンタープライズ・サーチ管理コンソールから、停止した Web クローラーを再始動します。

## グローバル Web クロール・スペースの構成

Web クローラーのためにグローバル・クロール・スペースを構成できます。これにより索引からの URL の除去をより制御できるようになります。

Web クローラーはそれぞれ、クロール対象の URL あるいは、クロール対象外の URL を定義するクロール・スペースで構成されます。クロール・スペースにある検出済み URL は、後でクロールするときのために (データベースに) 保存されます。クロール・スペースにない URL は、廃棄されます。クローラーが空のデータベースで始動した場合、クローラーが稼働中は、そのクロール・スペースの定義とデータベースは整合しています。

時々、クローラーは停止されて、そのクロール・スペースが、(例えば、ページのクロールを禁止する新規規則によって) 削減されます。クローラーが再始動されるとそのクロール・スペース定義とデータベースは、整合しなくなります。データベースには URL (クロールされるものもあれば、クロールされないものもある) が含まれ、これらの URL は小さくなった新しいクロール・スペースにはありません。

コレクションに Web クローラーが 1 つしかない場合、これらの URL 用の HTTP 状況コードを 760 (除外されることを示す) に変更し、現在除外されているページを索引から除去するように要求することで、その Web クローラーには整合性が戻ります。

複数の Web クローラー間でクロール・スペースを分割する場合 (例えば、一部のページは残りのものより頻繁にクロールされるようにするなど)、各 Web クローラーが独立したデータベース表 (初期は空) を保守し、それらはそれぞれ Web クロール・スペースの別の部分をクロールします。元のクローラーのクロール・スペースは削減されて、他のクローラーがクロールする予定の部分が除去された後に残った分のみになります。元のクローラーが、移動済みのページを索引から除去することで整合性を復元しようとするとう問題が発生します。移動済みのページは、他のクローラーによってクロールされているので、そのページは索引に残っている必要があります。

高水準のグローバル・クローल・スペースを構成することで、元のクローラーによってクローलされないが、索引からも削除されない URL を識別することができます。クローラーのクロール・スペースに既にある URL は、検出プロセスによって除外とマークされたままになり、再度クロールされるときに索引から除去されません。

グローバル・クロール・スペースは、`global.rules` という名前の構成ファイルに定義されます。このファイルは、クローラー構成ディレクトリに存在しなければなりません。( `global.rules` ファイルがあることで、グローバル・クロール・スペース機能が使用可能になります。 ) このファイルがある場合は、クローラーの初期化時に読み取られます。このファイルが存在しないと、クローラーは単一レベルのクロール・スペースで作動することになり、クロール・スペースの定義とデータベースの間の整合性を維持するために、必要に応じて索引から文書を除去します。

グローバル・クロール・スペースが存在する場合、クローラーは従来どおり URL が存在するかしないかを判定しますが、索引からの URL の除去を要求するのは URL がどの Web クロール・スペースにも存在しない場合のみです。

`global.rules` ファイルは、ローカルの `crawl.rules` ファイルと構文は同じですが、ドメイン名の規則のみを含むことができる点が異なります。この制約事項により、クロール・スペースは IP アドレスや HTTP 接頭部パターンではなく DNS ホスト名に基づいてのみ、複数クローラー間でパーティション化されるようになります。URL 接頭部やローカル・クロール・スペースの IP アドレスの規則で ( `crawl.rules` に定義されているように ) 除外される URL はグローバル・クロール・スペースによって影響を受けません。そのような URL はやはり除外されません。

グローバル・クロール・スペースは、索引から URL を除去できなくするためにのみ使用されます。この URL はローカルのドメイン規則によってクローラーのクロール・スペースから除去されます。以下の規則が以下の順序で適用されます。

1. ローカルの接頭部規則やアドレス規則によって、クローラーのデータベースから URL が除外される場合、その URL には状況コード 760 が割り当てられ、索引から除去されます。URL は再度クロールされません。
2. ローカルのドメイン規則によって、クローラーのデータベースから URL が除外されて、さらにグローバル・クロール・スペースがない場合に、URL には状況コード 760 が割り当てられ、索引から除去されます。URL は再度クロールされません。
3. ローカルのドメイン規則によって、クローラーのデータベースから URL が除外されたが、グローバル・クロール・スペースの規則によって明示的に許可されている場合、URL には状況コード 761 が割り当てられます。クローラーは、その URL を再度クロールしませんが、索引から URL が除去されることはありません。(他のクローラーのローカル・クロール・スペースにあるものと想定されません。)
4. ローカルのドメイン規則によってクローラーのデータベースから URL が除外されており、グローバル・クロール・スペースの規則によって明示的には許可されていない場合、その URL には状況コード 760 が割り当てられ、索引から除去されます。

グローバル・クローल・スペースは、ローカル・クローल・スペースにより既に除外されている URL の除去を防ぐためにのみ調べられるので、グローバル・クロール・スペースからのデフォルトの結果は、候補 URL に適用する規則がない場合に、その URL に対するクロールを禁止することです。

global.rules ファイルは、グローバル・クロール・スペースを共用するあらゆるクローラーの master\_config ディレクトリーに存在しなければなりません。

global.rules ファイルとローカルの crawl.rules ファイルのすべてのコピーを注意深く編集して、それらが相互に整合性を保っていることを確認してください。

## リンクをたどらない、索引付けしないディレクティブ

ページ上のリンクをたどるか、ページに索引付けをするかを制御する Web クローラー用のディレクティブを指定することで検索の質を上げることができます。

一部の Web ページには、リンクをたどらない、あるいは索引付けをしないディレクティブがあり、これにより Web クローラーなどのロボットがそのページで検出したリンクをたどらないように指示したり、あるいは、そのページのコンテンツを索引に含めないように指示したり、あるいは、そのどちらも行わないように指示します。

これらの設定を制御することで、クロールの質を向上させることができます。例えば、ディレクトリー・ページの中には、多数のリンクを含んでいるものの、他に有効なコンテンツが無いものがあります。こうしたページはクローリングしてリンクをたどる必要はありますが、ディレクトリー・ページそのものを索引付けする利点はありません。

また、クローラーが階層の下方面に行かないようにしたい場合がありますが、要求したリーフ・ページにリンクが組み込まれていて、しかもリンクをたどらないディレクティブが含まれていない場合があります。これらのページには自動生成されているものがあり、必要なディレクティブを挿入できる所有者がいません。

そうしたページのクローリングに規則を指定するために、followindex.rules という名前の構成ファイルを作成、編集します。このファイルに規則を指定するときは、以下のガイドラインに従ってください。

- 設定する規則では、URL 接頭部を指定する必要があります (IP アドレスあるいは DNS ホスト名で Web サイトを識別することはできません)。
- URL 接頭部には、類似した URL を持つ複数のサイトを許可したり禁止するためにワイルドカード文字としてアスタリスク (\*) を含めることができます。
- 順序は重要です (クローラーは、候補 URL と一致する最初の規則を適用します)。
- リンクをたどることや索引付けを明示的に許可および禁止する規則は、ターゲット文書内の設定を含む他の設定に優先します。

## Web ページでのリンクをたどらないディレクティブおよび索引付けをしないディレクティブのオーバーライド

構成ファイルに規則を指定して、Web クローラーがリンクをたどらないディレクティブまたは索引付けをしないディレクティブが含まれているページへのリンクまたは索引ページをたどるかどうかを制御できます。

## 始める前に

Web クローラーについてリンクをたどらないディレクティブおよび索引付けをしないディレクティブを指定するには、エンタープライズ・サーチ管理者でなければなりません。指定されたディレクティブは、クロールされるページに存在するディレクティブをオーバーライドします。

## 手順

リンクをたどらないディレクティブおよび索引付けをしないディレクティブをオーバーライドするには、次のようにします。

1. エンタープライズ・サーチ管理コンソールから、規則を構成したいコレクションをモニターして、Web クローラーを停止します。
2. クローラー・サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
3. 構成したいクローラーの構成ディレクトリーに移動します。ここで、*crawler\_session\_ID* は、エンタープライズ・サーチ・システムによりクローラー・セッションに割り当てられたものです。次に例を示します。

```
ES_NODE_ROOT/master_config/col_56092.WEB_88534
```

4. *followindex.rules* という名前のファイルを作成または編集します。
5. 以下の形式でクローラーの規則を入力します。ここで、*URLprefix* は、リンクのたどりまたは索引付けを許可または禁止したい Web サイトの先頭文字です。

```
forbid follow URLprefix  
allow follow URLprefix  
forbid index URLprefix  
allow index URLprefix
```

6. ファイルを保存して、終了します。
7. エンタープライズ・サーチ管理コンソールから、停止した Web クローラーを再始動します。

## クロールされた文書の日付として Web クローラーが使用する日付の構成

構成ファイルの中にオプションを指定して、Web クローラーがクロールされた文書の日付として使用する日付を制御できます。

## 始める前に

デフォルトでは、Web クローラーは、クロールされた文書の *Date* (日付) フィールドの値を、文書がクロールされた日に設定します。しかし、好みで、Web クローラーがこの日付を、Web サーバーによって戻された *Last-Modified* (最終変更日時) の値に設定するように構成できます。

*Last-Modified* のデータは、すべての文書で使用可能であるわけではありません。この値を使用するようにクローラーを構成し、Web サーバーが文書の *Last-Modified* データを戻さない場合、クローラーは、クロールされた文書の *Date* フィールドの値に、その文書がクロールされた日時を設定します。

表 4. Web クローラーがクロールされた文書の日付フィールドを構成する方法

| クローラー構成                           | Web サーバーによって戻されたデータ                       | クロールされた文書の Date メタデータ・フィールド      |
|-----------------------------------|-------------------------------------------|----------------------------------|
| デフォルトの構成                          | Web サーバーは、クロールされた文書の Last-Modified 値を戻す   | クロールされた日時                        |
|                                   | Web サーバーは、クロールされた文書の Last-Modified 値を戻さない | クロールされた日時                        |
| Last-Modified (最終変更日時) を使用するように構成 | Web サーバーは、クロールされた文書の Last-Modified 値を戻す   | Web サーバーによって戻された Last-Modified 値 |
|                                   | Web サーバーは、クロールされた文書の Last-Modified 値を戻さない | クロールされた日時                        |

Web クローラーがクロールされた文書にどちらの日付を使用するかを指定するには、エンタープライズ・サーチ管理者でなければなりません。

### 手順

クロールされた文書の日付を、Web サーバーによって戻された Last-Modified 日時にするように Web クローラーを構成するには、次のようにします。

1. クローラー・サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. 構成する Web クローラーの `crawl.properties` ファイルを編集します。ここで、`crawler_session_ID` は、エンタープライズ・サーチ・システムによってクローラー・セッションに割り当てられた ID です。

`ES_NODE_ROOT/master_config/crawler_session_ID/crawl.properties`

例:

```
/home/esadmin/master_config/col_00112.WEB_23344/crawl.properties
C:\Program Files\IBM\es\esadmin\master_config\col_55667.WEB_78899\crawl.properties
```

3. 次の行を追加して、ファイルを保存します。

```
which_date=LastModified
```

4. エンタープライズ・サーチ管理コンソールから、Web クローラーを再始動します。

Web サーバーがこの Web クローラーによってクロールされた文書の Last-Modified データを戻す場合、クローラーは、クロールされた文書の Date として戻された日時を使用します。

## Web Content Management クローラー

IBM Workplace Web Content Management 文書をエンタープライズ・サーチ・コレクションに入れる場合は、Web Content Management クローラーを構成する必要があります。

## WebSphere Portal サーバーの構成

Workplace Web Content Management を WebSphere Portal バージョン 6 サーバーにインストールした場合、Web Content Management クローラーを使用して Web Content Management サイトをクロールできます。これらのサイトをクロールするためのオプションは、WebSphere Portal クローラーによってクロールされるポータル・サイトに指定するオプションとは別に構成できます。

Web Content Management ・クローラーを作成する前に、WebSphere Portal でエンタープライズ・サーチをセットアップする手順に従う必要があります。エンタープライズ・サーチ環境をセットアップするには、検索サーバー上の OmniFind Enterprise Edition と一緒に提供されるスクリプト (AIX、Linux、または Solaris の場合は `wp6_install.sh`、Windows の場合は `wp6_install.bat`) を実行します。

### URL フォーマット

クロールする URL を指定する際には、次のフォーマットを使用してください。

```
http_protocol://ポータル・ホスト名:ポート番号/ポータル接頭部  
/WCM 検索シード・サブレット・パス/searchseed?site=WCM サイト名&lib=WCM ライブラリー名
```

次の例は、WebSphere Portal 上の Workplace Web Content Management のデフォルト・インストール・パスにあるサイトの URL を示しています。

```
http://portal.server.ibm.com:80/wps/wcmsearchseed/searchseed?  
site=SiteTest01&lib=Web+Content
```

サイト名またはライブラリー名にスペースが含まれる場合は、そのスペースを正符号 (+) 文字に置き換えてください。例えば、Web Content は Web+Content に置き換えます。

### 構成の概要

Web Content Management クローラーを使用して、任意の数の Web Content Management サイトをクロールすることができます。クローラーを構成する際に、クロールするサイトの URL を指定します。これにより、クローラーは、指定されたサイトからリンクされたページをダウンロードします。

クロールされるサイトは、同じ WebSphere Portal 管理者 ID およびパスワードによりアクセス可能でなければなりません。異なるクレデンシャルを使用するサイトをクロールするには、別の Web Content Management クローラーを構成する必要があります。

Web Content Management クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクロール・スペース内のすべての文書をクロールする方法を制御します。

- クロールするサイトの URL とクローラーがサイトに接続するための情報を指定します。

クローラーを作成または編集したら、クロールされる URL にクローラーが接続できるかをテストできます。クローラーを開始する前に、クロールされる文書にクローラーがアクセスできるかどうかメッセージで通知されます。

- 文書レベルのセキュリティー・オプションを指定します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

**重要:** セキュアな Web Content Management ページを検索するには、WebSphere Portal 内からエンタープライズ・サーチの検索ポートレットを使用して検索を実行する必要があります。サンプル検索アプリケーション `ESSearchApplication` から実行された検索は、適切なクレデンシャルを持たず、文書へのユーザーのアクセス権限を検査できません。

- Web Content Management サイトが文書の提供にプロキシ・サーバーを使用する場合に、クローラーがプロキシ・サーバーと通信するための情報を指定します。
- WebSphere Portal サーバーおよび Web サイトを保護するのに別の製品 (IBM Tivoli Access Manager WebSEAL や CA SiteMinder SSO Agent for PeopleSoft など) を使用する場合は、クローラーがサーバー上の文書にアクセスできるようにするシングル・サインオン・クレデンシャルを指定します。
- クローラーが Secure Sockets Layer (SSL) プロトコルを使用して Web Content Management サイトに接続するための鍵ストア・ファイルに関する情報を指定します。
- クロールする文書の言語とコード・ページを指定します。
- Web Content Management 文書内のメタデータをクロールおよび検索するためのオプションを指定します。
- Web Content Management サイトをクロールするスケジュールを指定します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

#### 関連概念

381 ページの『WebSphere Portal との統合』

#### 関連タスク

389 ページの『WebSphere Portal バージョン 6 でのエンタープライズ・サーチのセットアップ』

---

## WebSphere Portal クローラー

IBM WebSphere Portal サイトのページをエンタープライズ・サーチ・コレクションに入れるには、WebSphere Portal クローラーを構成する必要があります。

### WebSphere Portal サーバーの構成

WebSphere Portal クローラーを作成する前に、WebSphere Portal でエンタープライズ・サーチ環境をセットアップするスクリプトを実行する必要があります。

WebSphere Portal のバージョンごとに別のスクリプトが必要です。スクリプトは、OmniFind Enterprise Edition のインストール時に検索サーバーにインストールされません。

- WebSphere Portal バージョン 5.1.0 以降の場合、AIX、Linux、Solaris システムでは `wp5_install.sh` スクリプト、Windows システムでは `wp5_install.bat` スクリプトを実行します。
- WebSphere Portal バージョン 6 の場合、AIX、Linux、Solaris システムでは `wp6_install.sh` スクリプト、Windows システムでは `wp6_install.bat` スクリプトを実行します。

#### ヒント:

セキュアな WebSphere Portal クローラーを構成する方法の詳細例については、IBM Redbook 「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある中規模組織向けのシナリオを参照してください。

### 構成の概要

WebSphere Portal クローラーを使用して、単一の WebSphere Portal サイトをクロールすることができます。クローラーを構成する際に、クロールするポータル・サイトの URL を指定します。これにより、クローラーは、指定されたポータル URL からリンクされたポートレットとページをダウンロードします。別のポータル・サイトをクロールするには、別のクローラーを作成します。

WebSphere Portal クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがサイトの全ページをクロールする方法を制御します。
- クロールするポータル・サイトの URL とクローラーがサイトに接続するための情報を指定します。このようなタイプの URL は長く、エンコードされた非 ASCII 文字が含まれているので、URL を WebSphere Portal サーバーからコピーし、エンタープライズ・サーチ管理コンソールに貼り付けると便利です。

クローラーを作成または編集したら、クロールされる URL にクローラーが接続できるかをテストできます。クローラーを開始する前に、クロールされる文書にクローラーがアクセスできるかどうかが届きメッセージで通知されます。



- 文書レベルのセキュリティー・オプションを指定します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

**重要:** セキュアな WebSphere Portal ページを検索するには、 WebSphere Portal 内からエンタープライズ・サーチの検索ポートレットを使用して検索を実行する必要があります。サンプル検索アプリケーション `ESSearchApplication` から実行された検索は、適切なクレデンシャルを持たず、文書へのユーザーのアクセス権限を検査できません。

- WebSphere Portal サイトがページの提供にプロキシ・サーバーを使用する場合に、クローラーがプロキシ・サーバーと通信するための情報を指定します。
- WebSphere Portal サーバーおよび Web サイトを保護するのに別の製品 (IBMTivoli Access Manager WebSEAL や CA SiteMinder SSO Agent for PeopleSoft など) を使用する場合は、クローラーがサーバー上の文書にアクセスできるようにするシングル・サインオン・クレデンシャルを指定します。
- クローラーが Secure Sockets Layer (SSL) プロトコルを使用して WebSphere Portal サイトに接続するための鍵ストア・ファイルに関する情報を指定します。
- クロールする文書の言語とコード・ページを指定します。
- WebSphere Portal 文書内のメタデータをクロールおよび検索するためのオプションを指定します。
- WebSphere Portal 文書をクロールするスケジュールを指定します。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクロールに必要な情報をクローラーに指定する方法が表示されます。

#### 関連概念

381 ページの『WebSphere Portal との統合』

#### 関連タスク

『クロールする URL を WebSphere Portal からコピーする』

389 ページの『WebSphere Portal バージョン 6 でのエンタープライズ・サーチのセットアップ』

384 ページの『WebSphere Portal バージョン 5.1 でのエンタープライズ・サーチのセットアップ』

## クロールする URL を WebSphere Portal からコピーする

誤った URL の入力をできる限り少なくするために、 WebSphere Portal または Web Content Management クローラーの構成時に、クロールするサイトの URL を適切なフィールドにコピーして貼り付けることができます。

このタスクについて

WebSphere Portal または Web Content Management クローラーの作成時に、クローリングする WebSphere Portal サーバー上のサイトの URL を指定します。URL は長く、通常、エンコードされた非 ASCII 文字が含まれているので、この手順を使用して、URL を WebSphere Portal サーバーからコピーし、エンタープライズ・サーチ管理コンソールに貼り付けると便利です。

## 手順

クローラーがクローリングする URL を指定するには、次のようにします。

1. エンタープライズ・サーチ管理コンソールでクローリングする URL を指定する準備ができたなら、WebSphere Portal サーバーが始動していることを確認し、管理者として WebSphere Portal にログインします。
2. WebSphere Portal バージョン 5.1 を使用する場合は、WebSphere Portal サーバーで次のステップを完了します。
  - a. 右上隅にある「管理」をクリックします。
  - b. 左側のナビゲーション領域で「ポータル設定」をクリックし、「検索管理」をクリックします。
  - c. 「検索コレクションの管理」ページで、「検索コレクション」領域にある「PortalCollection」をクリックします。他のコレクションが選択可能な場合、別のコレクションを選択できます。
  - d. 「コレクション」領域の「コンテンツ・ソース」で、「コンテンツ・ソースの追加」をクリックします。
  - e. 「クローリング・ソース・タイプ」で、「ポータル・サイト」をクリックします。「この URL でリンクされる文書を収集」フィールドにサイトの URL が表示されます。
  - f. URL をクリップボードにコピーします。例えば、URL を強調表示して、Ctrl キーを押しながら Insert (挿入) キーを押します。
3. WebSphere Portal バージョン 6 を使用する場合は、WebSphere Portal サーバーで次のステップを完了します。
  - a. 左下隅にある「管理」をクリックします。
  - b. 左側のナビゲーション領域から「検索の管理」をクリックします。
  - c. 「検索の管理」ページで「検索コレクション」リンクをクリックします。
  - d. 「検索の管理」ページで、検索サービス・オプションから「デフォルトのポータル検索サービス」をクリックします。
  - e. 検索コレクション表で「ポータル・コンテンツ」コレクションをクリックします。
  - f. コンテンツ・ソース表で「ポータル・コンテンツのソース」の横の右端にある「編集」アイコン (鉛筆形アイコン) をクリックします。
  - g. 「コンテンツ・ソース・タイプ」に対して「ポータル・サイト」をクリックします。「この URL でリンクされる文書を収集」フィールドにサイトの URL が表示されます。
  - h. URL をクリップボードにコピーします。例えば、URL を強調表示して、Ctrl キーを押しながら Insert (挿入) キーを押します。
4. エンタープライズ・サーチ管理コンソールに戻って、コピーした URL をサイト URL フィールドに貼り付けます。

## Windows ファイル・システム・クローラー

Microsoft Windows ファイル・システムに格納された文書をエンタープライズ・サーチ・コレクションに入れる場合は、Windows ファイル・システム・クローラーを構成する必要があります。

Windows ファイル・システム・クローラーを使用して、任意の数の Windows ファイル・システムをクロールすることができます。クローラーの構成時に、クロールするローカルおよびリモートのディレクトリー、サブディレクトリーを選択します。

AIX、Linux、Solaris システムにクローラー・サーバーをインストールした場合、そのサーバーを使用して Windows ファイル・システム・ソースをクロールすることはできません (Windows ファイル・システム・クローラーは、選択可能なクローラー・タイプのリストに表示されません)。

### ヒント:

セキュアな Windows ファイル・システム クローラーを構成する方法の詳細例については、IBM Redbook 「IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios」にある小規模組織向けのシナリオを参照してください。

### 共用ネットワーク・ディレクトリーのクロール

Windows ファイル・システム・クローラーは、エンタープライズ・サーチ管理者に指定された読み取り許可に従って文書をクロールします。管理者は OmniFind Enterprise Edition サービスのアカウントです。

クロール対象のディレクトリーのユーザー ID とパスワードを指定できます。ただし、ユーザー ID とパスワードは、共用ネットワーク・ディレクトリーへの接続にのみ使用されます。クローラーは、ローカル・ドライブではなく、共用ネットワーク・ディレクトリーについてこのユーザーに設定された読み取り許可に従って、ファイルをクロールします。

ネットワーク・ディレクトリーへの接続は、OmniFind Enterprise Edition サービスを再始動するまで切断されません。接続の確立後は、誤りのユーザー ID とパスワードでディレクトリーにアクセスすることが可能になります。ただし、この接続は、エンタープライズ・サーチ・システムの制御下にある Windows ファイル・システムの検出およびクローラー・セッションにのみ許可されます。考えられるセキュリティ・リスクを回避するために、(OmniFind Enterprise Edition サービスが実行される際の) エンタープライズ・サーチ管理者のアカウントの許可が、適切に設定されるようにしてください。

将来、ネットワーク・ディレクトリーへの接続で問題が発生しないようにするために、同じネットワーク・ディレクトリーに対して同じユーザー ID とパスワードを指定してください。誤りのユーザー ID とパスワードを指定して OmniFind Enterprise Edition サービスを再始動すると、Windows ファイル・システム・クロー

ラーは、誤りのクレデンシャルでディレクトリーに接続を試行するために、クローラに失敗する可能性があります。正しいユーザー ID とパスワードを使用する別の Windows ファイル・システム・クローラーによってネットワーク接続が確立されていれば、クローラは正常に終了することが可能です。

## 構成の概要

Windows ファイル・システム・クローラーを作成または変更する場合は、エンタープライズ・サーチ管理コンソールにログインします。この場合、操作者はエンタープライズ・サーチ管理者役割の担当者か、またはクローラーを所有するコレクションのコレクション管理者でなければなりません。

クローラーを作成する場合、ウィザードに従って以下のタスクを実行します。

- クローラーの動作およびシステム・リソースの使用を制御するプロパティを指定します。クローラーのプロパティで、クローラーがクローラ・スペース内のすべてのサブディレクトリーをクローラする方法を制御します。
- ファイル・システムをクローラするスケジュールを設定します。
- クローラするサブディレクトリーを選択します。

クローラーでクローラするサブディレクトリーのレベル数も指定できます。リモート・ファイル・システムをクローラするには、クローラーがデータにアクセスできるようにするためのユーザー ID とパスワードも併せて指定します。

- サブディレクトリー内の文書を検索可能にするときのオプションを指定します。例えば、特定タイプの文書をクローラ・スペースから除外したり、クローラーが特定サブディレクトリーのファイルにアクセスできるようにユーザー ID とパスワードを指定したりすることができます。
- 文書レベルのセキュリティー・オプションを構成します。コレクションの作成時にセキュリティーが有効な場合、クローラーは、セキュリティー・データを索引内の文書に関連付けます。このデータによって、検索アプリケーションは、保管済みアクセス制御リストまたはセキュリティー・トークンに基づいたアクセス制御を実行できるようになります。

また、ユーザーが照会を実行するときにユーザーのクレデンシャルを検証するオプションを選択することもできます。この場合、システムは、ユーザーのクレデンシャルを索引付きセキュリティー・データではなく、元のデータ・ソースで維持されている現行のアクセス制御リストと比較します。

文書レベルのセキュリティーを施行するには、ユーザーおよびドメイン・アカウント情報がクローラー・サーバーに正しく構成されていることを確認する必要があります。

クローラーの作成中に「ヘルプ」をクリックすると、ウィザードのフィールドに関する説明や、データのクローラに必要な情報をクローラーに指定する方法が表示されます。

### 関連概念

318 ページの『Windows ファイル・システム文書に対する文書レベルのセキュリティーの施行』

321 ページの『Windows トラストド・ドメインのセキュア検索』

---

## データ・リスナー・アプリケーションのサポートの構成

外部クローラーを作成するデータ・リスナー API を使用することによって、エンタープライズ・サーチを拡張することができます。カスタム・データ・リスナー・アプリケーションは、コレクションへのデータの追加、コレクションからのデータの除去、または Web クローラーに対する URL へのアクセスおよび再アクセスの指示を行えます。

### 始める前に

データ・リスナー・アプリケーションを構成するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

**重要:** データ・リスナーは、将来のリリースではサポートされなくなります。エンタープライズ・サーチのクライアント・アプリケーションを開発するには、データ・リスナー API に代わって検索および索引 (SI-API) API を使用してください。以下の説明は、以前にデータ・リスナー用アプリケーションを作成したユーザー向けです。

### このタスクについて


クライアント・データ・リスナー・アプリケーションは、エンタープライズ・サーチのデフォルトのクローラーではクロールできないデータ・ソース・タイプのクロールを可能にします。データ・リスナー・アプリケーションを使用するためには、アプリケーションがコレクションのアクセスおよび更新を行えるようにするクレデンシャルを構成しておく必要があります。

クライアント・データ・リスナー・アプリケーションは、データ・リスナーに接続するときに、クライアント・アプリケーション ID とパスワード、更新されるコレクションの ID を渡す必要があります。この情報は、管理コンソールでアプリケーションについて構成した情報と同じでなければなりません。



データ・リスナーは、エンタープライズ・サーチ・システムが開始されるときに自動的に開始されます。管理コンソールでアプリケーションを構成した後でポート番号を変更した場合は、データ・リスナーを再始動する必要があります。

### 手順

データ・リスナー・アプリケーションを構成するには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「 編集」をクリックして、システム編集ビューに切り替えます。
3. 「データ・リスナー」ページで、「データ・リスナー・アプリケーションの構成」をクリックします。
4. 「データ・リスナー・アプリケーション」ページで、データ・リスナーがクライアント・アプリケーションからの要求の処理用に作成できるスレッドの数、およびデータ・リスナーが要求を listen するポート番号を指定します。また、パーサーが解析を開始するまで一時ストレージに保持できる、コレクション当たりの文書の最大数も指定します。
5. クライアント・アプリケーションについての情報を追加するため、「データ・リスナー・アプリケーションの追加」をクリックします。

6. 「データ・リスナー・アプリケーションの追加」ページで、クライアント・データ・リスナー・アプリケーションがエンタープライズ・サーチ・コレクションにアクセスできるようにする認証情報を指定します。データ・リスナーのクライアント ID は、エンタープライズ・サーチ・システム内で固有のものでなければなりません。
7. データ・リスナー・アプリケーションが更新できるコレクションを選択します。
  - すべてのコレクションをアプリケーションが更新できるようにするには、「すべてのコレクション」をクリックします。
  - 指定するコレクションのみをアプリケーションが更新できるようにするには、「特定のコレクション」をクリックします。

このオプションを選択すると、コレクション名のリストが表示されます。アプリケーションが更新できるコレクションごとに、その「**選択**」チェック・ボックスを選択します。
8. 「**OK**」をクリックします。
9. データ・リスナーのポート番号、または、一時ストレージに保持できる文書数を変更した場合は、データ・リスナーを再始動します。
  - a. 「 **モニター**」をクリックして、システム・モニター・ビューに移動します。
  - b. 「データ・リスナー」ページで、「 **再始動**」をクリックします。

#### 関連タスク

354 ページの『データ・リスナーのモニター』

---

## カスタム・クローラー・プラグイン

クローラーのプロパティを構成するときに、文書レベルのアクセス制御の実行に使用する Java クラスを指定できます。また、Java クラスを使用して、メタデータと文書コンテンツの追加、変更、または除去によって索引を更新することができます。また、プラグインを作成して、クローラーの機能をアーカイブ・ファイルをクロールするように拡張できます。

プラグインには、クローラーがクロールする各文書ごとに呼び出される Java クラスが含まれています。その Java クラスには、エンタープライズ・サーチ索引からの文書 ID (URI)、セキュリティー・トークン、メタデータ、および文書コンテンツが渡されます。クラスは、新規または変更されたセキュリティー・トークン、メタデータ、コンテンツを戻すことや、セキュリティー・トークン、メタデータ、コンテンツを除去することができます。

クロール・スペース内のすべての文書が 1 回クロールされた後は、プラグインは新規文書または修正された文書についてのみ呼び出されます。エンタープライズ・サーチ索引内にあるが、元のデータ・ソースでは更新されていない文書のセキュリティー・トークン、メタデータ、またはコンテンツを変更するには、クロール・スペース内のすべての文書のフル・クロールを開始してから、主索引を再作成します。

## プラグインによるセキュリティの施行

文書レベルのセキュリティは、クローラーがクロールする各文書に 1 つ以上のセキュリティ・トークン (コンマで区切られた文字列) を関連付けることによって施行されます。通常、セキュリティ・トークンとしてグループ ID が使用されます。

デフォルトで、各文書には、文書を全員に使用可能にするパブリック・トークンが割り当てられています。パブリック・トークンは、管理者が提供する値、またはクローラー対象文書中のフィールドから抽出された値に置き換えることができます。

プラグインを使用すると、独自のビジネス・ルールを適用して、クロールされる文書のセキュリティ・トークンの値を決定できます。各文書に関連付けられたセキュリティ・トークンは、索引に保管されます。セキュリティ・トークンを使用して、それらに一致する文書を選出するようフィルター操作し、ユーザーに表示を許可された文書だけが検索結果に戻されるようになります。

## プラグインによるメタデータの追加、変更、除去

文書の最終変更日時などの文書メタデータは、クロールされるすべての文書について作成されます。クローラー・プラグインを使用すると、独自のビジネス・ルールを適用して、索引付けされるメタデータの値を各文書について決定することができます。

メタデータは、名前と値のペアとして作成されます。ユーザーは、フリー・テキスト検索またはメタデータ・フィールド名を指定した照会で、メタデータを検索できます。

## プラグインによる文書コンテンツの追加、変更、除去

文書コンテンツには、検索可能コンテンツと、検索結果で動的な文書サマリーの一部分になる可能性があるコンテンツを含む文書構成部分が含まれます。クローラー・プラグインを使用すると、独自のビジネス・ルールを適用して、各文書の索引付けされるコンテンツを決定することができます。

## Web クローラー・プラグイン

Web クローラー用のアプリケーション・プログラミング・インターフェースで、文書がどのようにクロールされ、構文解析用に準備されるのかを制御できます。例えば、クローラーが文書を要求するときに使用されるフィールドを HTTP 要求ヘッダーに追加できます。文書がクロールされた後、構文解析およびトークン化される前に、コンテンツ、セキュリティ・トークン、およびメタデータを変更できます。また、文書がパーサーに送られるのを停止することもできます。

フォーム・ベースの認証の説明と、カスタム Web クローラー・プラグイン用に適応できるサンプル・プログラムについては、<http://www.ibm.com/developerworks/db2/library/techarticle/dm-0707nishitani> を参照してください。

## アーカイブ・ファイル・プラグイン

プラグインを作成して、クローラーを拡張し、ZIP と TAR 以外のアーカイブ・ファイル・フォーマットのクロールのサポートを使用可能にできます。例えば、LZH


フォーマットの文書のクローリングをサポートするプラグインを作成できます。

## unfenced モード

非 Web クローラーを構成する場合、プラグインを unfenced モードで実行するオプションを選択できます。このモードでは、プラグイン・プロセスはクローラー・プロセスの内部で実行します。こうすると、プラグインのパフォーマンスを向上できます。

**重要:** プラグインがこのモードで実行時に、回復不能な問題をプラグインが検出すると、クローラー・プロセスが終了することがあります。

### 関連概念

 クローラー・プラグイン

---

## アーカイブ・ファイルのクローリングのサポート

エンタープライズ・サーチ・クローラーは、アーカイブ内の個々のファイルを索引付けして検索できるように、アーカイブ・ファイル (ZIP や TAR ファイルなど) からファイルを抽出することができます。

### サポートされるアーカイブ・ファイルのフォーマット

次のアーカイブ・ファイルのフォーマットがサポートされます。

表 5. エンタープライズ・サーチ・クローラーによってサポートされるアーカイブ・ファイルのフォーマット

| ファイル拡張子       | MIME タイプ        | データ・タイプ | 注釈                                                                                                                                                                                                       |
|---------------|-----------------|---------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| .zip、.ZIP     | application/zip | zip     | <ul style="list-style-type: none"><li>• java.util.zip パッケージの機能に依存します</li><li>• デフレート (方式 8) 圧縮をサポートします<ul style="list-style-type: none"><li>- 暗号化ファイルのサポートなし</li><li>- zip64 のサポートなし</li></ul></li></ul> |
| .tar          | application/tar | tar     | サポートされる tar フォーマット: <ul style="list-style-type: none"><li>• GNU tar 1.13</li><li>• POSIX 1003.1-1998 (ustar)</li><li>• POSIX 1003.1-2001 (pax)</li></ul>                                                 |
| .tar、.gz、.tgz | 適用されない          | tgz     | java.util.zip パッケージの機能に依存します                                                                                                                                                                             |

### 制約事項とガイドライン

コード・ページ自動検出は、アーカイブ・ファイルから抽出されるファイルで使用できません。ファイルの抽出時に、クローラーは、プレーン・テキストおよび不明の文書タイプで使用するよう構成されているコード・ページ設定を使用します。エンタープライズ・サーチ管理コンソールを使用してクローラーの言語とコード・ページの設定を構成するときに、プレーン・テキスト文書用およびコード・ページを自動検出できない文書用にクローラーが使用するコード・ページを指定します。



アーカイブ・ファイル内のファイルの再クロールが必要となる時期を判別するために、クローラーは、各ファイルのアーカイブ・エントリー・ヘッダー・データの変更日付を使用します。クローラーをモニターすると、クロールされた文書について表示される統計値 (挿入、更新、削除された文書の統計値など) に、アーカイブ・ファイルから抽出されたファイルに関する情報が含まれます。

クローラーが LZH ファイルなど、他のアーカイブ・ファイル・フォーマットのアーカイブ・ファイルをクロールできるようにするには、クローラー・プラグインを作成し、そのプラグインをクローラーが使用するよう構成する必要があります。

## 移行

OmniFind Enterprise Edition バージョン 8.4 のインストールより前に存在したクローラーがアーカイブ・ファイルをクロールできるようにするには、クローラーのクローール・スペースを編集する必要があります。除外する MIME タイプと除外するファイル拡張子の設定に、.zip、.tar、.tgz、.gz ファイルへの参照が含まれないことを確認してください。

**重要:** Content Edition クローラーまたは DB2 Content Manager クローラーのこれらの設定を変更した場合には、変更を適用するためにすべての文書の再クロールが必要です。

---

## エンタープライズ・サーチ索引における URI フォーマット

エンタープライズ・サーチ索引内の各文書の Uniform Resource Identifier (URI) は、その文書をコレクションに追加したクローラーのタイプを示します。

コレクションのカテゴリ、有効範囲、クイック・リンクを構成する際、URI および URI パターンを指定できます。また、索引から文書を除去する必要がある場合、または、特定の URI に関する詳細状況情報を表示する必要があるときにも、URI を指定します。

文書の URI または URI パターンを判別するには、コレクションを検索してください。検索結果内の URI をクリックすれば、興味のある文書をリトリブできます。検索結果から URI をコピーして、その URI をエンタープライズ・サーチ管理コンソールで使用できます。例えば、URI パターンを指定して、その URI パターンに一致する文書を、エンタープライズ・サーチ・カテゴリに自動的に関連付けることができます。

## アーカイブ・ファイル

アーカイブ・ファイル (.zip ファイルや .tar ファイルなど) から抽出され、クロールされる文書の URI フォーマットは、次のとおりです。

```
Original_URI(?|&)ArchiveEntry=Entry_Name(&ArchiveEntry=Entry_Name)
```

パラメーター

*Original\_URI*

データ・ソース上のアーカイブ・ファイルの場所。

### *Entry\_Name*

アーカイブ・ファイル内のアーカイブ・エントリーの URL エンコード名。

### 例

```
file:///d:/Archive1.zip
file:///d:/Archive1.zip?ArchiveEntry=Folder1/PowerPoint.ppt
file:///d:/Archive1.zip?ArchiveEntry=Folder2/Text.txt
```

## Content Edition クローラー

サーバー・アクセス・モードの Content Edition クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
vbr:///Server_Name/Repository_System_ID/Repository_Persistent_ID
/Item_ID/Version_ID
/Item_Type/?[Page=Page_Number&] JNDI_properties
```

直接アクセス・モードの Content Edition クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
vbr:///Repository_System_ID/Repository_Persistent_ID
/Item_ID/Version_ID
/Item_Type/[?Page=Page_Number]
```

### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

#### *Server\_Name*

WebSphere Information Integrator Content Edition サーバーの名前。

#### *Repository\_System\_ID*

リポジトリのシステム ID。

#### *Repository\_Persistent\_ID*

リポジトリの永続 ID。

#### *Item\_ID*

項目の ID。

#### *Version\_ID*

バージョンの ID。バージョン ID がブランクの場合、この値は文書の最新バージョンを示します。

#### *Item\_Type*

項目のタイプ (CONTENT または FOLDER)。

#### *Page\_Number*

ページ番号。

#### *JNDI\_properties*

J2EE アプリケーション・クライアントの JNDI プロパティ。プロパティには、次の 2 つのタイプがあります。

#### **java.naming.factory.initial**

EJB ハンドルを作成するのに使用されるアプリケーション・サーバーのクラスの名前。

### java.naming.provider.url

EJB ハンドルを要求するのに使用されるアプリケーション・サーバーのネーミング・サービスへの URL。

例 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
```

## DB2 クローラー

DB2 クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
db2://Database_Name/Table_Name  
/Unique_Identifier_Column_Name1/Unique_Identifier_Value1  
[/Unique_Identifier_Column_Name2/Unique_Identifier_Value2/...  
/Unique_Identifier_Column_NameN/Unique_Identifier_ValueN]
```

パラメーター:

すべてのフィールドに URL エンコード方式が適用されます。

*Database\_Name*

データベースの内部名またはデータベースの別名。

*Table\_Name*

スキーマの名前を含む、ターゲット表の名前。

*Unique\_Identifier\_Column\_Name1*

表の先頭の固有 ID 列の名前。

*Unique\_Identifier\_Value1*

先頭の固有 ID 列の値。

*Unique\_Identifier\_Column\_NameN*

表の *n* 番目の固有 ID 列の名前。

*Unique\_Identifier\_ValueN*

*n* 番目の固有 ID 列の値。

例 ローカル、カタログ・データベース:

```
db2://LOCALDB/SCHEMA1.TABLE1/MODEL/ThinkPadA20
```

リモート、アンカタログ・データベース

```
db2://myserver.mycompany.com:50001/REMOTEDB/SCHEMA2.TABLE2/NAME/DAVID
```

## DB2 Content Manager クローラー

DB2 Content Manager クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
cm://Server_Name/Item_Type_Name/PID
```

## パラメーター

*PID* パラメーターには URL エンコードが適用されます。

*Server\_Name*

IBM DB2 Content Manager ライブラリー・サーバーの名前。

*Item\_Type\_Name*

ターゲット項目タイプの名前。

*PID* DB2 Content Manager 永続 ID

## 例

```
cm://cmsrvctg/ITEMTYPE1/92+3+ICM8+icmnlbdb12+ITEMTYPE159+26+A1001001A  
03F27B94411D1831718+A03F27B+94411D183171+14+1018
```

## Lotus Domino Document Manager クローラー

Lotus Domino Document Manager クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
dominodoc://Server_Name:Port_Number/Database_Replica_ID/Database_Path_and_Name  
/View_Universal_ID/Document_Universal_ID  
/?AttNo=Attachment_Number&AttName=Attachment_File_Name
```

## パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*Server\_Name*

Domino Document Manager サーバーの名前。

*Port\_Number*

オプション: Domino Document Manager サーバー用のポート番号。

*Database\_Replica\_ID*

データベース・レプリカの ID。

*Database\_Path\_and\_Name*

ターゲット Domino Document Manager サーバー上の文書 NSF データベースのパスとファイル名。

*View\_Universal\_ID*

Domino Document Manager 文書をクロールするのに使用される View Universal ID。

*Document\_Universal\_ID*

クロール対象文書で定義される Document Universal ID。

*Attachment\_Number*

オプション: ゼロから始まる、各添付ファイルの連続番号。

*Attachment\_File\_Name*

オプション: 添付ファイルのオリジナル名。

## 例

Domino Document Manager 文書:

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

文書添付ファイル:

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

## Exchange Server クローラー

Exchange Server クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

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

### パラメーター

*OWA\_Path*

プロトコルを含まない、Outlook Web Access (OWA) パス。

**useSSL=true**

オリジナルの OWA パスのプロトコルが HTTPS である場合に追加される。

例 文書本体:

```
exchange://exchangesvr.ibm.com/public/RootFolder1/Folder1/Document.EML
```

文書添付ファイル:

```
exchange://exchangesvr.ibm.com/public/RootFolder1/Folder1/Document.EML/AttachedFile.doc
```

SSL で使用可能:

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

## JDBC データベース・クローラー

JDBC データベース・クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
jdbc://DB_URL/Table_Name  
/Unique_Identifier_Column_Name1/Unique_Identifier_Value1  
/[Unique_Identifier_Column_Name2/Unique_Identifier_Value2  
/.../Unique_Identifier_Column_NameN/Unique_Identifier_ValueN]
```

### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*DB\_URL* データベースの URL。

*Table\_Name*

スキーマの名前を含む、ターゲット表の名前。

*Unique\_Identifier\_Column\_Name1*

表の先頭の固有 ID 列の名前。

*Unique\_Identifier\_Value1*

先頭の固有 ID 列の値。

*Unique\_Identifier\_Column\_NameN*

表の *n* 番目の固有 ID 列の名前。

*Unique\_Identifier\_ValueN*

*n* 番目の固有 ID 列の値。

**例:** DB2 データベース:  
jdbc:db2://host01.svl.ibm.com:50000/SAMPLE/DB2INST1.ORG/DEPTNUMB/51

Oracle データベース:  
jdbc:oracle:thin:@/host01.svl.ibm.com:1521:ora/SCOTT.EMP/EMPNO/7934

MS SQL Server 2000 データベース:  
jdbc:microsoft:sqlserver://host01.svl.ibm.com:1433;  
DatabaseName=Northwind/dbo.Region/RegionID/100

MS SQL Server 2005 データベース:  
jdbc:sqlserver://host01.svl.ibm.com:1433;  
DatabaseName=Northwind/dbo.Region/RegionID/100

## Lotus Notes クローラー

Lotus Notes クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
domino://Server_Name[:Port_Number]/Database_Replica_ID/Database_Path_and_Name  
/[View_Universal_ID]/Document_Universal_ID  
[?AttNo=Attachment_Number&AttName=Attachment_File_Name]
```

### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

#### *Server\_Name*

Lotus Notes サーバーの名前。

#### *Port\_Number*

Lotus Notes サーバーのポート番号。ポート番号はオプションです。

#### *Database\_Replica\_ID*

データベース・レプリカの ID。

#### *Database\_Path\_and\_Name*

ターゲット Lotus Notes サーバー上の NSF データベースのパスとファイル名。

#### *View\_Universal\_ID*

ターゲット・データベース上で定義された View Universal ID。この ID は、文書がビューまたはフォルダーから選択されたときのみ指定されます。ビューまたはフォルダーをクロール対象として指定しない場合 (例えば、データベース内の全文書をクロールするよう指定する場合)、View Universal ID は指定されません。

#### *Document\_Universal\_ID*

クローラーによってクロールされる文書内で定義された Document Universal ID。

#### *Attachment\_Number*

ゼロから始まる、各添付ファイルの連続番号。添付ファイル番号はオプションです。

#### *Attachment\_File\_Name*

添付ファイルのオリジナル名。添付ファイル名はオプションです。

例 ビューまたはフォルダーでクロールするために選択された文書:

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

ビューまたはフォルダーでクロールするために選択されていない文書:

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

文書添付ファイル:

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

## QuickPlace クローラー

QuickPlace クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
quickplace://Server_Name:Port_Number/Database_Replica_ID/Database_Path_and_Name  
/View_Universal_ID/Document_Universal_ID  
/?AttNo=Attachment_Number&AttName=Attachment_File_Name
```

### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*Server\_Name*

Lotus QuickPlace サーバーの名前。

*Port\_Number*

オプション: QuickPlace サーバー用のポート番号。

*Database\_Replica\_ID*

データベース・レプリカの ID。

*Database\_Path\_and\_Name*

ターゲット QuickPlace サーバー上の文書 NSF データベースのパスとファイル名。

*View\_Universal\_ID*

QuickPlace 文書をクロールするのに使用される View Universal ID。

*Document\_Universal\_ID*

クロール対象文書で定義される Document Universal ID。

*Attachment\_Number*

オプション: ゼロから始まる、各添付ファイルの連続番号。

*Attachment\_File\_Name*

オプション: 添付ファイルのオリジナル名。

例 文書:

```
quickplace://ltsvr.ibm.com/49257043000214B3/QuickPlace%5Csamplplace  
%5CPageLibrary4925704300021490.nsf  
/A7986FD2A9CD47090525670800167225  
/2B02B1DE3A82B2CE49257043001C2498
```

添付ページ:

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

## シード・リスト クローラー

シード・リスト クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
seedlist://Page_URL?pageID=Page_ID[&useSSL=true]
```

### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*Page\_URL*

文書の URL (各文書に固有)。

*Page\_ID*

文書のオブジェクト ID。

**useSSL** プロトコルが HTTPS の場合、&useSSL=true が URI に追加されます。それ以外の場合、useSSL は省略されます。

**例** HTTPS プロトコル:

```
seedlist://quickrserver.ibm.com:10035/lotus/mydoc?uri=dm:bec6090046f1cd5
2bc5cfcb06e9f4550&verb;=view&pageID;=N1FSZUR1MkJQNjZSMDZQMUMwM1FPNjZCQzY
2SUw2SUhPNk1RQ0M2Uk80Nk9PNjVCRUM2UUs2TDFDMA==&useSSL;=true
```

## UNIX ファイル・システム・クローラー

UNIX ファイル・システム・クローラーによってクロールされる文書の URI フォーマットは、次のとおりです。

```
file:///Directory_Name/File_Name
```

### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*Directory\_Name*

ディレクトリーの絶対パス名。

*File\_Name*

ファイルの名前。

**例**

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

## Web Content Management クローラー

Web Content Management クローラーによってクロールされる WebSphere Content Management 文書の URI フォーマットは、次のとおりです。

```
wcm://Page_URL?pageID=Page_ID[&useSSL=true]
```

### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*Page\_URL*

文書の URL (各文書に固有)。



*Page\_ID*

ページ ID。

**useSSL** プロトコルが HTTPS の場合、`useSSL=true` が URI に追加されます。それ以外の場合、`useSSL` は省略されます。

**例** HTTP プロトコル:

```
wcm://wp6server.ibm.com:9081/wps/wcm/myconnect/Web+Content
/Site01/SiteArea01/ContentTest01?pageID=
6QReDeJ9DI3R0663E03Q06L1E2MR47MHOC3Q862RD6J0863B0GJS86J9E0
```

HTTPS プロトコル:

```
wcm://wp6server.ibm.com:9444/wps/wcm/myconnect/Web+Content/Site01
/SiteArea01/ContentTest01?pageID=
6QReDeJ9DI3R0663E03Q06L1E2MR47MHOC3Q862RD6J0863B0GJS86J9E0&useSSL=true
```

## WebSphere Portal クローラー: WebSphere Portal バージョン 5

WebSphere Portal クローラーによってクロールされる WebSphere Portal バージョン 5 文書の URI フォーマットは、次のとおりです。

```
wps://Page_URL?portletDefID=Portlet_Def_ID&portletID=Portlet_ID
&pageID=Page_ID[&useSSL=true]
```

**パラメーター**

すべてのフィールドに URL エンコード方式が適用されます。

*Page\_URL*

文書の URL (各文書に固有)。

*Portlet\_Def\_ID*

ポートレット定義 ID。

*Portlet\_ID*

ポートレット ID。

*Page\_ID*

ページ ID。

**useSSL** プロトコルが HTTPS の場合、`useSSL=true` が URI に追加されます。それ以外の場合、`useSSL` は省略されます。

**例** 文書本体:

```
wps://wpserver.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+
LKnPy1vM0Y_QjzKCN4g3cQbJgQi0-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8gHRj/delta
/base64xml/L01DU1kvd0NrQUpORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!
?portletDefID=3_0_3S&pageID=6_0_6J
```

**例** SSL で使用可能:

```
wps://wpserver.ibm.com:9081/wps/myportal!/ut/p/kcxml/04_Sj9SPykssy0x+
LKnPy1vM0Y_QjzKCN4g3cQbJgQi0-pFQAW99X4_83FT9AP2C5IhyR0dFRQD8gHRj/delta
/base64xml/L01DU1kvd0NrQUpORUEvNFBVR0VoQSEvN18wXzZPLzZfMF80RA!!
?portletDefID=7_0_A4&pageID=6_0_6J&useSSL=true
```

## WebSphere Portal クローラー: WebSphere Portal バージョン 6

WebSphere Portal クローラーによってクロールされる WebSphere Portal バージョン 6 文書の URI フォーマットは、次のとおりです。

wp6://Page\_URL?portletURL=Portlet\_URL?portletDefID=Portlet\_Def\_ID  
&pageID=Page\_ID[&useSSL=true]

#### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*Page\_URL*

文書の URL (各文書に固有)。

*Portlet\_URL*

文書の固有な URL。

*Portlet\_Def\_ID*

ポータル定義 ID。

*Page\_ID*

ページ ID。

**useSSL** プロトコルが HTTPS の場合、useSSL=true が URI に追加されま  
す。それ以外の場合、useSSL は省略されます。

#### 例 HTTP プロトコル:

```
wp6://wp6server.ibm.com:9081/wps/myportal/!ut/p/c1/04_SB8K8xLLM9MSSzPy  
8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_  
Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGh  
gYWjobuhmaGDka0Bu5uHqFRpkaAAAEisaBQ!!  
?portletUrl=/wps/myportal/!ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59  
gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdF  
RUBTu-saA!!/d12/d0/Y2BiUZnBwMqsyBykYGBmS2tcDoDE40MebyRiU-woYGFOaG7oZu  
hgZGjggbuZh6uXY4ihAYOMGQ7ZOCBDawDcXPkM  
&portletDefID=6_24LS10811G1F102A0G6HEJUR10  
&pageID=3_24LS108110R6B02A0GGU94LN00
```

#### HTTPS プロトコル:

```
wp6://wp6server.ibm.com:9444/wps/myportal/!ut/p/c1/04_SB8K8xLLM9MSSzPy  
8xBz9CP0os3gjE59gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_  
Tzyc1P1C3IjDHQdFRUBTu-saA!!/d12/d0/Y2BkbGBgY1rDwMDEJ1XAwMggYxZvZOITbGh  
gYWjobuhmaGDka0Bu5uHq5RtqaAAA50L41Q!!  
?portletUrl=/wps/myportal/!ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz9CP0os3gjE59  
gQwMLQ0P_IDMnAyNHA3f3UEsTD1NjA6B8pFm8AQ7gaEBAAdjIPrz6_Tzyc1P1C3IjDHQdF  
RUBTu-saA!!/d12/d0/Y2BiUZnBwMqsyBykYGBmS2tcDoDE40MebyRiU-woYGFOaG7oZu  
hgZGjggbuZh6uXWZihAYOMGQ5Z31BDawAk73P2  
&portletDefID=6_24LS10811G1F102A0G6HEJMU10  
&pageID=3_24LS108110R6B02A0GGU94T410&useSSL=true
```

## Windows ファイル・システム・クローラー

Windows ファイル・システム・クローラーによってクロールされる文書の URI フ  
ォーマットは、次のとおりです。

file:///Directory\_Name/File\_Name

file:///Network\_Folder\_Name/Directory\_Name/File\_Name

#### パラメーター

すべてのフィールドに URL エンコード方式が適用されます。

*Directory\_Name*

ディレクトリーの絶対パス名。

*File\_Name*

ファイルの名前。

*Network\_Folder\_Name*

リモート・サーバー上のみにある文書の場合の、Windows ネットワーク上の共有フォルダーの名前。

**例** ローカル・ファイル・システム:

file:///d:/directory/test.doc

ネットワーク・ファイル・システム:

file:///filesvr.ibm.com/directory/file.doc



---

## パーサーの管理

文書の検索性を高めるために、文書およびメタデータが、エンタープライズ・サーチ索引に追加される前に、どのように構文解析、分析、および分類されるのかに関するオプションを指定できます。

文書コンテンツの構文解析および情報のリトリブ可能性の最適化のために指定できるオプションには、以下のものがあります。

### 中国語、日本語、韓国語の文書に対する構文解析オプションの構成

中国語、日本語、韓国語で書かれた文書を構文解析するために N-gram セグメンテーション使用のオプションを指定することができます。また、中国語と日本語の文書内の空白文字から改行文字を除去することもできます。

### ネイティブ XML 検索の使用可能化

ご使用のコレクションに XML 文書が含まれている場合、XPath および XML フラグメントなどの、ネイティブ XML 照会構文を使用した検索を可能にできます。ネイティブ XML 検索により、ユーザーは、さまざまな XML エlement間の関係に基づいた照会を指定できるようになります。

### カテゴリーの構成

類似 URI パターンを共有する文書、または特定の語句を含む文書をカテゴリーとしてまとめることができます。ユーザーは、コレクションの検索時に、検索結果を特定のカテゴリーに属する文書に限定することができます。

### 検索フィールドの構成

XML 文書内のエレメントを索引の検索フィールドにマップできます。また、HTML 文書のメタデータ・エレメントを検索フィールドにマップすることもできます。エンタープライズ・サーチ索引に検索フィールドを作成することで、ユーザーが XML および HTML 文書の特定部分を照会できるようになり、検索結果の精度を上げることができます。






### テキスト処理オプションの構成

カスタム・テキスト分析エンジンがエンタープライズ・サーチ・システムに追加されている場合、その 1 つを特定のコレクションで使用するよう選択できます。分析エンジンをコレクションに関連付けた後、コンテンツが言語的に分析され注釈が付けられるようにコンテンツをマップするオプションを指定できます。また、分析結果がどのようにエンタープライズ・サーチ索引または JDBC データベース表にマップされるのかを指定することもできます。

### ランキング調整クラスへのフィールドのマッピング

照会条件に一致するフィールドが含まれている文書が、照会条件に一致する他の文書よりも検索結果内で高くランク付けされるよう指定できます。フィールドをランキング調整クラスにマップするときには、どのコンテンツおよびメタデータ・フィールドがランキング調整されるのかを指定します。また、各ランキング調整クラスが文書のランク付けに使用するスコアを構成することもできます。

### 関連概念

-  セマンティック検索の言語サポート
  -  エンタープライズ・サーチに組み込まれているテキスト分析
  -  テキスト分析処理で使用される基本概念
  -  セマンティック検索アプリケーション
  -  セマンティック検索照会条件
- 187 ページの『言語とコード・ページのサポート』
- 171 ページの『文書フォーマットの検出』

---

## カテゴリーに関する作業

カテゴリーを使用すると、共通の特性を共有する文書をグループにまとめ、そのグループのメンバーに対する基準に合う文書だけを検索して、リトリートすることができます。

文書がカテゴリーに関連付けられていて、検索アプリケーションがこの機能をサポートする場合、ユーザーはカテゴリー名を指定してコレクションのサブセットを検索することができます。コレクション全体を検索する場合、ユーザーは検索結果を絞り込んで、同じカテゴリーにある文書だけを結果文書として参照できるようにすることができます。

カテゴリーを構成する際には、文書とカテゴリーの関連付けをクローラーに指示する規則を指定します。URI パターンを共有する文書をグループ化したり、または特定のコンテンツを含む文書 (特定の語句を含む文書、あるいは含まない文書など) をグループ化することができます。

カテゴリーを作成、管理するには、エンタープライズ・サーチ管理コンソールを使用します。

- コレクションの作成時に、カテゴリー化タイプを選択します。カテゴリーを使用しない、またはルール・ベースのカテゴリーを使用することを選択できます。
- コレクションの構文解析規則を構成する場合、必要に応じてカテゴリー化タイプを変更できます。文書のクロール、索引作成の終了後にカテゴリー化タイプを変更した場合、全文書を再クロールして主索引を再作成するまで、検索の品質が低下します。
- ルール・ベースのカテゴリーを使用する場合、管理コンソールを使用して、カテゴリー・ツリー、カテゴリー、およびカテゴリー規則を管理することができます。文書のクロール、索引作成の終了後にカテゴリーまたはカテゴリー規則を変更した場合、全文書を再クロールして主索引を再作成するまで、検索の品質が低下します。

## ルール・ベースのカテゴリー

エンタープライズ・サーチ・コレクションで文書をカテゴリーに関連付ける場合、規則を構成してどの文書を関連付けるか制御することができます。

エンタープライズ・サーチで作成したコレクション、および IBM WebSphere Portal コレクションから移行したルール・ベースのカテゴリに対してカテゴリ規則を作成できます。文書をカテゴリ化する際の規則を構成する場合、コレクション作成時、またはコレクションの解析オプションの指定時にルール・ベースのカテゴリを使用することを指定する必要があります。

パーサーは、指定した規則を使用して文書を 1 つ以上のカテゴリに関連付けます。

- 文書が特定のカテゴリ内の少なくとも 1 つの規則に該当した場合、パーサーはその文書をその特定カテゴリに関連付けます。
- 文書が複数のカテゴリ内の少なくとも 1 つの規則に該当した場合、パーサーはその文書をそれらすべてのカテゴリに関連付けます。
- 文書がカテゴリのどの規則にも該当しない場合は、パーサーはその文書をカテゴリに関連付けません。コレクションを検索すれば、この文書を検索し、それをリトリブできますが、特定のカテゴリを検索してもこの文書をリトリブできません。

コレクションのカテゴリ・ツリー (分類構造) を管理する場合、新規のカテゴリをカテゴリ階層のどこに追加するかを指定します。また、カテゴリ・ツリーを使用して、編集するカテゴリを選択し、文書をカテゴリ化する際の規則を追加、削除、および個別規則の内容を変更することもできます。

文書のカテゴリ化規則を構成する場合、エンタープライズ・サーチが、文書がカテゴリに属しているかどうかを判別する際に文書の URI を使用するか、文書内のコンテンツを使用するかを選択します。

## URI パターンによる規則

URI 規則を文書の URI に適用します。URI の一部 (パターン) を指定し、指定されたパターンを URI に持つ文書がその規則に該当することになります。

例えば、規則のテキストに「/hr/」と指定した場合、次に挙げる最初の URI は規則に該当しますが、2 番目の URI は該当しないことになります。

```
file:///corporate/hr/medicalform.doc  
http://company.com/human_resources/medicalform.htm
```

URI はすべてパターンとして扱うため、パターンの最初あるいは最後に指定したワイルドカード文字としてのアスタリスクは、システムによって無視されます。例えば、\*/hr/\* と /hr/ は、同一の URI とマッチングすることになります。

URI のパターン規則は、大/小文字の区別をしません。URI にスペースが含まれる場合、URI パターンは URI エンコード用のエンタープライズ・サーチ規則に従う必要があります。次の例は、Windows ファイル・システム・パスの URI を指定する場合の正しい方法と間違っている方法を示しています。

```
間違っている URI: file:///c:/program files/  
正しい URI: file:///c:/program+files/
```

## 文書コンテンツによる規則

文書コンテンツ規則の指定は、照会と同じフォーマットで行います。文書がその照会に対して有効な場合、その文書は規則に該当することになります。規則を構成する場合、文書に含まれる、または含まれてはならない語句を指定し、語のステミング規則を適用する言語を選択します。

例えば、次の規則では、語として「hr」を含むか、または句として「human resources」を含む文書が規則に該当することになります。

```
hr "human resources"
```

また別の例として、次の規則では、文書のタイトルに「health」という語が含まれるが、「employee benefits」という句が含まれない場合に、文書が規則に該当することを指定しています。

```
+title:health -title:"employee benefits"
```

コンテンツ規則では、検索および索引 API (SIAPI) 照会と同様の言語正規化が行われます。しかし、コンテンツ規則の構文でサポートされるのは、SIAPI 照会構文で使用可能な操作のサブセットです。以下の照会演算子のみをコンテンツ規則で使用できます。

- + 項の前に正符号を付けると、その項が文書に出現しなければならないことを示します。
- 項の前に負符号を付けると、その項が文書に出現してはならないことを示します。
- " " 複数の項を引用符で囲むと、その句が文書中に正確に出現する必要があることを示します。

フィールド名:

語句の前にフィールド名を付けると、その語句が、指定された文書フィールドで出現する必要がある (または出現してはならない) ことを指示します。

コレクションで検索可能フィールドとして構成されたすべてのコンテンツ・フィールドとメタデータ・フィールドがサポートされます。次の SIAPI フィールド・キーワードとフィールド・タイプは、サポートされません。

- site:
- url:
- link:
- docID:
- samegroupas:
- パラメトリック・フィールド
- セキュリティー・トークン
- 属性 (\$source、\$language、\$doctype など)



## カテゴリ・ツリー

カテゴリ・ツリーを使用すると、コレクション内のルール・ベースのすべてのカテゴリを表示できます。カテゴリ・ツリーを使って、カテゴリの作成、カテゴリの削除、およびカテゴリに文書を関連付けるときの規則を編集することもできます。

カテゴリ・ツリーは、分類構造とも呼ばれ、階層構造になっています。このツリーは、ルート・カテゴリを起点とし、他のカテゴリはすべてこのルート・カテゴリから分岐します。任意の数のカテゴリおよびサブカテゴリをネスト化して、文書の参照および検索のさまざまな選択肢をユーザーに提供することができます。

例えば、ある文書が複数のカテゴリで規則に該当する場合、その文書はそれらすべてのカテゴリに関連付けられます。ユーザーが特定のカテゴリを検索する、または、検索結果を参照するときに特定のカテゴリに属する文書を参照する場合、文書が複数のカテゴリに属していればユーザーがその文書を見つける可能性が高まります。

カテゴリ・ツリーを管理する場合、新規のカテゴリを既存のカテゴリ下にネストすることにより、どの文書がどのカテゴリ (1 つ以上) に属するかを調整することができます。カテゴリを作成するときに、そのカテゴリがルート・レベルで作成されるのか、別のカテゴリのサブカテゴリとして作成するのかを指定できます。また、カテゴリ・ツリーを使用して、コレクションからカテゴリを削除したり、文書をカテゴリに関連付けるときの規則を変更したりすることもできます。カテゴリの編集時には、カテゴリの名前変更、カテゴリ化規則の追加または削除、また個別規則の内容を変更することができます。

カテゴリ・ツリーを管理する場合、検索および参照動作のガイドラインとして以下に記載する説明を参考にしてください。

- ユーザーが上位レベルのカテゴリを検索すると、そのカテゴリ、およびそのカテゴリ以下のすべてのサブカテゴリが、検索基準に合致する文書の検索対象となります。ユーザーがサブカテゴリのないカテゴリを検索する場合、そのカテゴリだけが検索されます。
- ユーザーが検索結果を参照しているときに、特定のカテゴリに属する文書だけを参照するオプションを選択すると、そのカテゴリの文書だけが表示されます。サブカテゴリの名前も検索結果に表示されるので、ユーザーは異なるカテゴリに切り替え、文書のサブセットを一度に表示させることができます。

## カテゴリ化タイプの選択

カテゴリ化タイプを選択するときに、コレクションの中で文書をカテゴリに関連付けるために使用する方法を指定します。

### 始める前に

カテゴリ化タイプを変更するには、エンタープライズ・サーチ管理者の役割を持つメンバーであるか、変更するコレクションのコレクション管理者である必要があります。

### このタスクについて

カテゴリ化タイプは、コレクションが作成されるときに指定されます。必要であれば、コレクションの文書をカテゴリ化する方法を変更できます。コレクションに対して特に構成したルール・ベースのカテゴリを使用することも、カテゴリを使用しないこともできます。

**重要:** データをクロールし、コレクションの索引を作成したあとでカテゴリ化タイプを変更すると、索引の整合性がなくなります。検索結果の正確性を確保するために、コレクション内のすべての文書を再クロールしてから、主索引を再作成してください。

### 手順

カテゴリ化タイプを選択するには、次のようにします。

1. コレクションを編集して「構文解析」ページを選択し、「**カテゴリ化タイプの選択**」をクリックします。
2. 「カテゴリ化タイプの選択」ページで、次のオプションの 1 つを選択します。

**なし** このオプションは、このコレクションの中の文書をカテゴリ化しない場合に選択します。

#### ルール・ベース

このオプションは、このコレクションに対して特に構成した規則に従って文書をカテゴリ化する場合に選択します。

## カテゴリの構成

1 つのコレクションには、任意の数のカテゴリを作成でき、各カテゴリには任意の数の規則を含められます。カテゴリに自動的に関連付けられる文書は、その規則により決定されます。

### 始める前に

カテゴリを構成するには、エンタープライズ・サーチ管理者の役割を持つメンバーであるか、そのカテゴリが所属するコレクションのコレクション管理者である必要があります。

ルール・ベースのカテゴリを使用するオプションが、カテゴリ化タイプとして選択される必要があります。

文書をカテゴリに関連付ける規則をどのように指定するかについての例は、カテゴリを作成または編集しているときに、「ヘルプ」をクリックしてください。

### このタスクについて

検索アプリケーションでカテゴリのサポートが使用可能であれば、ユーザーは、カテゴリ名を指定することによって、コレクションのサブセットを検索できます。ユーザーはまた、検索結果の中でカテゴリを選択して、そのカテゴリに属する文書のみを参照することもできます。

**重要:** データをクロールし、コレクションの索引を作成したあとでカテゴリまたはカテゴリ規則を変更すると、索引の整合性がなくなります。検索結果の正確性を確保するために、コレクション内のすべての文書を再クロールしてから、主索引を再作成してください。

## 手順

カテゴリを構成するには、次のようにします。

1. コレクションを編集して「構文解析」ページを選択し、「**カテゴリ・ツリーの構成**」をクリックします。
2. 「カテゴリ・ツリー」ページで、カテゴリを追加するツリー内の場所を選択し、「**カテゴリの作成**」をクリックします。

ルートを選択すると、新規カテゴリがルート・レベルに作成されます。カテゴリ名を選択すると、新規カテゴリは、カテゴリ・ツリーの中の選択されたカテゴリの下にネストされます。

ウィザードが開き、文書を新規カテゴリに関連付ける規則の指定をガイドします。

- a. 「カテゴリの作成」ページで、カテゴリの記述名を入力し、「**次へ**」をクリックします。
- b. 「カテゴリ規則の作成」ページで、「**規則の追加**」をクリックします。
- c. 「カテゴリ規則の作成」ページで、「**規則名**」フィールドに規則の固有の名前を入力します。この名前はコレクションの中の全カテゴリで固有でなければなりません。
- d. 文書をこのカテゴリに関連付けるために使用する規則を指定して、「**OK**」をクリックします。
  - 文書がカテゴリに属するかどうかを判断するために文書の **URI** を使用する場合は、「**URI パターン**」をクリックして、URI パターンを指定します。

指定したテキストが **URI** に存在すれば、文書はそのカテゴリに関連付けられます。

例: `file:///c:/program+files/finance`

- 検索可能コンテンツの照会によって、文書がカテゴリに属するかどうかを判断する場合は、「**文書コンテンツ**」をクリックし、文書の言語を選択して、文書コンテンツに出現しなければならない語句、または出現してはならない語句を指定します。

規則の指定は、照会と同じ形式で行いますが、包含 (+)、除外 (-)、句 (" ")、フィールド名 (フィールド名:) の照会演算子のみを使用できます。N-gram セグメンテーションは、コンテンツ規則ではサポートされません。

指定した語の包含または除外に適合する文書がそのカテゴリに関連付けられます。

例: `+finance -accounting +title:"fiscal year"`

- e. 「**完了**」をクリックします。

新規カテゴリーが、このコレクションに所属する他のカテゴリーと一緒に、「カテゴリー・ツリー」ページにリストされます。

#### 関連資料

135 ページの『エンタープライズ・サーチ索引における URI フォーマット』

---

## XML 検索フィールドの作業

ユーザーが XML 文書の特定部分を検索できるようにしたい場合、XML エlementを検索フィールドにマップします。

XML エlementを検索フィールドにマップするには、エンタープライズ・サーチ管理コンソールを使用します。

通常、XML 文書内のテキスト・データはすべて索引付けされます。XML エlementを検索フィールドにマップすると、XML 文書の構造情報を使用して、より明確な照会をサポートすることができます。例えば、XML エlement名をフィールド名にマップすることによって、XML エlement内のデータをフィールド名で検索して検索結果に戻せるようにすることができます。XML エlementが検索フィールドになるための基準として XML エlementの属性 (およびその属性値) を指定することによって、マッピングの精度を高めることが可能です。

## XML 検索フィールド

XML 検索フィールドは、ユーザーによる XML 文書の特定部分の照会を可能にします。

XML 文書は、半構造化テキストと構造化されていないテキストの両方を含むため、ますます一般的になってきています。XML の構造はカプセル化されていて、テキストを囲む XML エlementによって明示的に定義されたコンテキストを使用します。例えば、作成者の名前が次のように表されます。

```
<author>John Smith</author>
```

このコンテキストでは、テキスト John Smith によって、XML 文書の作成者が示されます。

エンタープライズ・サーチでは、XML エlementを検索フィールド名に関連付ける、つまりマップすることができます。コレクションの構文解析オプションを構成するとき、どの XML エlementがどの検索フィールド名にマップされるのかを指定します。XML エlementを検索フィールドにマップすると、ユーザーは照会においてフィールド名を指定することで、それらのElementの値を検索できるようになります。指定されたフィールドを検索する照会を使用すると、全文書コンテンツを検索するフリー・テキスト照会よりも精度の高い検索結果が得られます。

例えば、コレクション中に XML 文書が含まれていて、<title> Elementと <author> Elementを索引内で検索フィールドとしてマークされるように指定すると、ユーザーがこれらのElementを直接照会できるようになります。

author:Smith の検索では、author というフィールドにマップされたElementが Smith である XML 文書を検出します。

別の例として、<summary> という XML エlement に、検索結果に表示すると便利な情報が含まれるとします。 <summary> Element を検索フィールドにマップし、この Element の値が検索結果に表示されるように指定すると、この Element の内容が結果文書に入ります。

XML Element をフィールド名にマップすると、その Element 内のテキストが、指定したフィールド名で検索可能になります。ただし、XML Element に属性が含まれる場合、属性値は索引付けされず、直接検索できません。この属性値を照会するには、エンタープライズ・サーチ管理コンソールで構文解析オプションを構成して、ネイティブ XML 検索を使用可能にする必要があります。ネイティブ XML 検索を使用可能にすると、XML 文書の構造が索引で利用可能になり、XPath 照会制約を指定することによって文書を照会できます。次に例を示します。

```
@xmlxp:.'/'.$elementName[@attributeName="attributeValue"]
```

#### 関連概念



セマンティック検索照会条件

#### 関連タスク

169 ページの『ネイティブ XML 検索のサポートの使用可能化』

## XML Element を検索フィールドにマップ

XML Element を検索フィールドにマップするときには、ユーザーが、照会においてフィールド名を指定することで検索可能な XML Element を指定します。

### 始める前に

XML Element を検索フィールドにマップするには、エンタープライズ・サーチ管理者の役割を持つメンバーであるか、その XML 文書が所属するコレクションのコレクション管理者である必要があります。

### 制約事項

コレクションに対して XML フィールド・マッピングは複数存在可能ですが、XML ルート・Element・マッピングは 1 つのみです。XML 文書のルート・Element は、それに応じてそのマッピングを適用します。

### このタスクについて

XML フィールド・マッピングの作成や、既存の XML フィールド・マッピングでフィールドの追加、変更、削除を行った場合、それらの変更内容が有効になるのは、パーサーの再始動後です。新規マッピングおよび変更されたマッピングは、パーサーの再始動後に構文解析される新しいデータに適用され、既に索引内にあるデータには影響しません。既に索引内にある文書を更新するには、文書を再度クローリングして索引付けを行う必要があります。

このタスクでは、以下の XML 文書をサンプルとして使用することで、従業員 (personnel) レコードをマップし、ユーザーが特定の Element を直接照会できるようにする方法を示します。

```
<?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>

```

## 手順

この例の XML エlementを検索フィールドにマップするには、次のようにします。

1. コレクションを編集、「構文解析」ページを選択、そして「フィールドへの **XML エlementのマッピング**」をクリックします。
2. 「XML フィールドのマッピング」ページで、「**XML マッピングの作成**」をクリックします。「XML フィールド・マッピングの作成」ページが開きます。
3. 「**XML ルート・Element名**」フィールドに、ルート・Element名 `personnel` を入力します。

ここで指定する名前が、検索したい XML 文書内のルート・Elementと完全に一致するようにしてください。エンタープライズ・サーチは、XML 文書の構文解析および索引付けを行う際、ルート・Element名に従って、どのマッピングを使用するのを選択します。

4. 「**XML マッピング名**」フィールドに、この XML フィールド・マッピング規則セットの名前を入力します。

XML マッピング規則セットを作成した後、この名前が「XML フィールド・マッピング」ページに表示されるので、マッピング規則の追加、削除、または変更を行う対象としてこの名前を選択できます。

5. XML Element `jobrole` を、`jobrole` という名前の検索フィールドにマップします。
  - a. 「**フィールド名**」フィールドに `jobrole` と入力します。

**ヒント:** フィールド名でサポートされていない ASCII 文字とメタキャラクターについては、「ヘルプ」をクリックしてください。

- b. 「**XML Element名**」フィールドに `jobrole` と入力します。
- c. `jobrole` フィールドを照会できるようにするには、「**フィールド検索**」チェック・ボックスを選択します。
- d. 検索語が `jobrole` フィールドの値全体に一致する (つまり、他の語がフィールドに存在しない) 場合にのみ一致となるようにするには、「**完全一致**」チェック・ボックスを選択します。
- e. `jobrole` フィールドの値で検索結果をソートできるようにするには、「**ソート可能**」チェック・ボックスを選択します。

フィールドは、検索アプリケーションがこの機能をサポートする場合にのみソート可能なフィールドとして表示されます。エンタープライズ・サーチの

サンプル検索アプリケーションは、検索結果をソートするために選択できるフィールドのリストの中に、XML エlementからマップされるフィールドを含みません。

- f. 検索結果に `jobrole` フィールドの値を表示できるようにするには、「**検索結果**」チェック・ボックスを選択します。
6. XML Element `jobroles` を、同じ検索フィールドにマップします。
  - a. 「**フィールドの追加**」をクリックして、フィールド・マッピング規則のリストに空白行を追加します。
  - b. 「**フィールド名**」フィールドに `jobrole` と入力します。
  - c. 「**XML Element名**」フィールドに `jobroles` と入力します。

**ヒント:** XML Element名は、検索フィールド名と一致する必要はなく、複数の XML Elementを同じ検索フィールドにマップできます。

- d. `jobrole` フィールドの照会、検索語がフィールド値に完全一致する必要があることの指定、フィールドを使用した検索結果のソート、検索結果でのフィールドの表示を可能にするには、該当のチェック・ボックスを選択します。
7. XML Element `section` (その属性 `ID` が値 `expertise` を含む) を検索フィールド `expertise` にマップします。
  - a. 「**フィールドの追加**」をクリックして、フィールド・マッピング規則のリストに空白行を追加します。
  - b. 「**フィールド名**」フィールドに `expertise` と入力します。
  - c. 「**XML Element名**」フィールドに `section` と入力します。
  - d. 「**XML 属性名**」フィールドに `id` と入力します。
  - e. 「**XML 属性値**」フィールドに `expertise` と入力します。
  - f. `expertise` フィールドを照会できるようにするには、「**フィールド検索**」チェック・ボックスを選択します。
  - g. 検索語が `expertise` フィールドの値全体に一致する (つまり、他の語がフィールドに存在しない) 場合にのみ一致となるようにするには、「**完全一致**」チェック・ボックスを選択します。
  - h. `expertise` フィールドの値で検索結果をソートできるようにするには、検索アプリケーションがこの機能をサポートしている場合は、「**ソート可能**」チェック・ボックスを選択します。
  - i. 検索結果に `expertise` フィールドの値を表示できるようにするには、「**検索結果**」チェック・ボックスを選択します。

### 例:

検索 (search) 製品を業務とする部門の全員を検出するには、次のような照会を指定します。

```
jobrole:search
```

言語 (linguistics) を専門にする部門の全員を検出するには、次のような照会を指定します。

```
expertise:linguistics
```

---

## HTML 検索フィールドの作業

ユーザーが HTML 文書の特定メタデータ部分を検索できるようにする場合、HTML メタデータ・エレメントを検索フィールドにマップします。

HTML メタデータ・エレメントを検索フィールドにマップするには、エンタープライズ・サーチ管理コンソールを使用します。

HTML メタデータ・エレメントを検索フィールドにマップすると、ユーザーが、よりの確な照会で HTML 文書を検索できるようになります。

## HTML 検索フィールド

HTML 検索フィールドにより、ユーザーは HTML 文書の属性を照会できます。

HTML 文書のメタデータ・エレメントは、文書に関する情報を提供するという点、そのフォーマット方法、および Web 上のアクセス許可の方法の点において、文書の属性と類似しています。次に例を示します。

```
<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="This topic describes an IBM product." />
<meta name="format" content="XHTML" />
```

エンタープライズ・サーチは、HTML メタデータ・エレメントの名前を検索フィールド名に関連付ける、すなわちマップすることができます。コレクションの構文解析オプションを設定するときに、どの HTML メタデータ・エレメントをどの検索フィールド名にマップするのかを指定します。HTML メタデータ・エレメントを検索フィールドにマップすると、ユーザーは照会に検索フィールド名を指定することで、それらのエレメントを含む文書を検索できます。特定フィールドを検索する照会を使用すると、全文書コンテンツを検索するフリー・テキスト照会よりも精度の高い検索結果が得られます。

例えば、コレクションに HTML 文書が含まれていて、copyright メタデータ・エレメントと abstract メタデータ・エレメントを検索フィールドとして索引付けすると指定すると、ユーザーはこれらの特定エレメントを照会できるようになります。

「copyright:IBM」で検索すると、copyright メタデータに IBM という語のある HTML 文書を検出できます。

HTML メタデータ・エレメントを検索フィールドにマップする際には、すべての HTML メタデータ・エレメントをマップするか、メタデータ・エレメントの Dublin Core セットに属するエレメントのみをマップするか、指定した HTML メタデータ・エレメントのみをマップするかを指定します。Dublin Core メタデータ・エレメント・セット内のエレメントの説明については、次の Dublin Core Metadata Initiative の Web サイトを参照してください。

<http://dublincore.org/documents/dcmi-terms/#H2>

すべての HTML メタデータ・エレメント、またはすべての Dublin Core メタデータ・エレメントのマッピング作成を選択した場合、デフォルトの検索フィールド名はメタデータ・エレメント名に一致します。デフォルトの検索フィールド名をオー



オーバーライドして個々のエレメントに別の検索オプションを指定するには、個別に構成するエレメントのリストにエレメントを追加します。

## 検索フィールドへの HTML メタデータ・エレメントのマッピング

HTML メタデータ・エレメントを検索フィールドにマップするときには、ユーザーが照会にフィールド名を指定することで検索可能な HTML エレメントを指定します。

### 始める前に

HTML メタデータ・エレメントを検索フィールドにマップするには、エンタープライズ・サーチ管理者役割のメンバーであるか、その HTML 文書が所属するコレクションのコレクション管理者である必要があります。

### このタスクについて

HTML フィールド・マッピングの作成や、既存の HTML フィールド・マッピングでフィールドの追加、変更、削除を行った場合、それらの変更内容が有効になるのは、パーサーの再始動後です。新規マッピングおよび変更されたマッピングは、パーサーの再始動後に構文解析される新しいデータに適用され、既に索引内にあるデータには影響しません。既に索引内にある文書を更新するには、文書を再度クロールして索引付けを行う必要があります。

### 手順

検索フィールドに HTML メタデータ・エレメントをマップするには、次のようにします。

1. コレクションを編集して「構文解析」ページを選択し、「**HTML メタデータをフィールドにマップ**」をクリックします。
2. 「HTML フィールド・マッピング」ページで、検索フィールドにマップする HTML メタデータ・エレメントを指定します。
  - すべての HTML エレメントまたは Dublin Core エレメント・セットに属するすべてのエレメントをマップする場合、フィールド名でフィールドを検索できるようにするか、フィールドを検索結果に表示するかを指定します。また、ユーザーの検索語がメタデータ・エレメント値全体に完全に一致する（つまり、ユーザーの検索語に一致する語以外にフィールド内に語がない）必要があるか、このフィールドで検索結果をソートできるかも指定します。

システムは、フィールド名としてメタデータ・エレメント名を自動的に使用します。特定エレメントのデフォルト・フィールド名をオーバーライドする、または特定エレメントで異なる検索オプションを使用する場合は、エレメント名を固有のフィールド名にマップし、そのエレメントで使用する検索オプションを指定します。

- 特定のメタデータ・エレメントを検索フィールドにマップするには、次のようにします。
  - a. 「**フィールドの追加**」をクリックして、フィールド・マッピング規則のリストに空白行を追加します。

- b. マップする HTML メタデータ・エレメントと関連付ける名前を入力します。ユーザーは、このコレクション内の HTML 文書を照会するときに、このフィールド名を指定できます。

**ヒント:** フィールド名でサポートされていない ASCII 文字とメタキャラクターについては、「ヘルプ」をクリックしてください。

- c. マップするメタデータ・エレメントの名前を入力します。
- d. このフィールドを照会できるようにするには、「**フィールド検索**」チェック・ボックスを選択します。

**ヒント:** このチェック・ボックスが選択されていないと、そのフィールドはフィールド照会やフリー・テキスト照会で検索できません。この方法は、例えば、すべての HTML メタデータ・エレメントまたは Dublin Core メタデータ・エレメントを含めるオプションを選択したが、特定のフィールドは検索されないようにする場合に便利です。

- e. このフィールドを照会できるようにし、照会条件がフィールド値全体に一致する（つまり、他の語がフィールドに存在しない）場合にのみ文書が一致するように指定するには、「**完全一致**」チェック・ボックスを選択します。
- f. このフィールドで検索結果を英字順にソート（ストリング・ソート）できるようにするには、「**ソート可能**」チェック・ボックスを選択します。

フィールドは、検索アプリケーションがこの機能をサポートする場合にのみソート可能なフィールドとして表示されます。エンタープライズ・サーチのサンプル検索アプリケーションは、検索結果をソートするために選択できるフィールドのリストの中に、HTML メタデータ・エレメントからマップされるフィールドを含みません。

- g. このフィールドのデータ・タイプが、DECIMAL、DOUBLE、INTEGER、SHORT、TIME、または TIMESTAMP で、かつ、このフィールドの検索時にユーザーがパラメトリック照会を指定できるようにするか、またはこのフィールドの値に従って数値順に結果をソートできるようにするには、「**パラメトリック検索**」チェック・ボックスを選択します。
- h. 検索結果にこのフィールドを表示できるようにするには、「**検索結果**」チェック・ボックスを選択します。

### 例:

ユーザーは、マップされたフィールド名を照会して、特定のメタデータを持つ HTML 文書を検索できます。例えば、description という名前の HTML メタデータ・エレメントを、abstract という名前の検索フィールドにマップしてある場合、以下のような照会を入力して、Thinkpad コンピューターに関する HTML 文書を検索できます。

```
abstract:thinkpad
```

---

## カスタム・テキスト処理

カスタム・テキスト処理アルゴリズムをエンタープライズ・サーチ・コレクションと統合することにより、検索結果の品質と精度を高めることができます。

OmniFind Enterprise Edition は、IBM Unstructured Information Management Architecture (UIMA) をサポートします。これは、テキスト分析機能の作成、検出、合成、デプロイを行うためのフレームワークです。アプリケーション開発者は、コンテンツが検索される分析アルゴリズムを作成し、テストしてから、エンタープライズ・サーチにアーカイブを使用するのに必要なすべてのリソースが含まれている処理エンジン・アーカイブ (.pear ファイル) を作成します。カスタム・テキスト分析アルゴリズムを使ってコレクションを検索できるようにするために、アーカイブ (分析エンジンが含まれています) をエンタープライズ・サーチ・システムに追加する必要があります。

テキスト分析エンジン内の分析論理コンポーネントをアノテーター といいます。各アノテーターは、特定の言語分析タスクを実行します。テキスト処理エンジンは、任意の数のアノテーターが含まれているものであっても、また、それぞれに固有のカスタム・アノテーターが含まれている幾つかの分析エンジンの複合であっても構いません。

アノテーターによって作成された情報を、分析結果 といいます。分析結果は、検索したい情報に相当するもので、共通分析構造 と呼ばれるデータ構造に書き込まれます。

コレクションについてテキスト処理オプションを構成する際には、以下の操作を行ってください。

- コレクション内の文書に注釈を付けるのに使用するテキスト分析エンジンを選択します。
- コレクションに分かりやすいマークアップを持つ XML 文書が含まれており、このマークアップをカスタム・テキスト分析で使用する場合は、マッピング・ファイルをコレクションと関連付けて、XML マッピングの出力を共通分析構造にマップすることができます。

例えば、共通分析構造内で、<addressee> エlementおよび <customer> エlementの内容を Person 注釈にマップできます。こうすると、これらの注釈にカスタム・アノテーターからアクセスできるようになり、その他の情報を検出できます (例えば、Person の性別を検出できます)。Person 注釈をエンタープライズ・サーチ索引にマップすることもできます。こうすると、ユーザーは、元の XML エlementを知らなくても、Person を検索することができます。

ユーザーが照会で元の XML エlementを指定できるようにしたい場合、XML マッピングを定義する必要はありません。その代わりに、構文解析オプションを構成し、コレクションについてネイティブ XML 検索を使用可能に設定できます。







- 共通分析構造をエンタープライズ・サーチ索引にマップします。これにより、注釈が付けられた文書をセマンティック検索で検索できるようになります。

例えば、アノテーターによって検出されるエンティティと関係に応じて、ユーザーは、同じ文で発生する概念 (例えば、特定の人物と競争者名) やキーワードと概念 (例えば、名前 Alex と電話番号) を検索できます。

- 共通分析構造をリレーショナル・データベースにマップします。データを IBM DB2 表または Oracle 表にマップできます。このタイプのマッピングによって、分析の結果をデータ・マイニングのようなデータベース・アプリケーションで使

用することができるようになります。また、SQL 照会を使用して、エンタープライズ・サーチの外部にあるデータを検索することができるようになります。

#### 関連概念

-  [カスタム・テキスト分析の組み込み](#)
-  [テキスト分析処理で使用される基本概念](#)
-  [カスタム分析の組み込みのワークフロー](#)
-  [テキスト分析アルゴリズム](#)
-  [セマンティック検索アプリケーション](#)
-  [セマンティック検索照会条件](#)

## システムへのテキスト分析エンジンの追加

カスタム・テキスト分析エンジンを作成する場合は、システムにそれを追加してからでないと、エンタープライズ・サーチ用に使用できません。コレクションは、エンジンを使用して、文書の分析と注釈付けを行い、検索結果の精度を高めることができます。

### 始める前に

システムにテキスト分析エンジンを追加するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### このタスクについて


アプリケーション開発者は、テキスト分析用の UIMA フレームワークに従う処理エンジン・アーカイブ (.pear) を作成できます。アーカイブには、エンタープライズ・サーチ・コレクションを検索するのに必要なリソースのすべてが含まれます。カスタム・テキスト分析アルゴリズムを使ってコレクションを検索できるようにするために、アーカイブ (分析エンジンが含まれています) をエンタープライズ・サーチ・システムに追加する必要があります。

テキスト分析エンジンをシステムに追加後、その表示名を変更して、XML ソースを表示するオプションを選択できます。XML ソースは、何の情報がこのエンジンで作成されたのかを示します。

テキスト分析エンジンが既にこのコレクションと関連付けられている場合、テキスト分析エンジンをシステムから除去することはできません。





### 手順

カスタム・テキスト分析エンジンをエンタープライズ・サーチ・システムに追加するには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「 編集」をクリックして、システム編集ビューに切り替えます。
3. 「構文解析」ページで、「**テキスト分析エンジンの構成**」をクリックします。

4. 「テキスト分析エンジン」 ページで、「**テキスト分析エンジンの追加**」をクリックします。
5. 「テキスト分析エンジンの追加」 ページで、新規エンジンの記述名を入力します。 システムは、この表示名を使用して、管理コンソール全体でテキスト分析エンジンを識別します。
6. .pear ファイルのロケーションを指定します。 ファイルが 8 MB より小さい場合、そのファイルはローカル・コンピューター上に置くことが可能で、参照してファイルを見つけることができます。ファイルが 8 MB より大きい場合、そのファイルは索引サーバー上に存在しなければならず、そのファイルの完全修飾パスの入力が必要です。
7. 「**OK**」をクリックします。 テキスト分析エンジンが「テキスト分析エンジン」ページにリストされます。

#### 関連概念

-  カスタム分析の組み込みのワークフロー
-  カスタム・テキスト分析の組み込み
-  テキスト分析処理で使用される基本概念
-  分析および検索における XML マークアップ

#### 関連タスク

-  XML エlement から共通分析構造へのマッピング・ファイルの作成

## テキスト分析エンジンとコレクションとの関連付け

カスタム・テキスト分析エンジンがエンタープライズ・サーチ・システムと関連付けられている場合、そのうちの 1 つを、このコレクションで使用するよう選択できます。そうすると、ユーザーは、コレクションを検索するときにセマンティック照会を指定し、検索結果の品質と精度を高めることができます。

#### 始める前に

テキスト分析エンジンをコレクションと関連付けるには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

#### このタスクについて

テキスト分析エンジンが既にこのコレクションと関連付けられている場合、別のエンジンを関連付けるときに、以下のアクションが発生します。

- 「**カスタム分析なし**」を選択した場合、コレクションについて以前に定義したすべてのテキスト分析マッピングはリセットされます。コレクションはシステム・デフォルト値を使用し始めます。
- 別のカスタム・テキスト分析エンジンを選択した場合、コレクションについて以前に定義したすべてのテキスト分析マッピングは保存されます。例えば、engine\_1 から engine\_2 に変更した場合、engine\_2 は、engine\_1 について構成された XML マッピング・ファイルを継承します。

## 手順

テキスト分析エンジンをコレクションと関連付けるには、次のようにします。

1. コレクションを編集、「構文解析」ページを選択、そして「**テキスト処理オプションの構成**」をクリックします。
2. 「**テキスト分析エンジンの選択**」をクリックします。 エンタープライズ・サーチ・システムにカスタム・テキスト分析エンジンが追加されていない場合、またはコレクションがデフォルトの分析アルゴリズムを使用している場合、エンジン名は **Default** です。
3. 「コレクションのテキスト分析エンジンの選択」ページで、このコレクションで使用したいエンジンの名前を選択します。 選択可能なテキスト分析エンジンがない場合、または「**カスタム分析なし**」を選択した場合、パーサーは、文書に注釈を付け、索引の準備をするときに、デフォルトのテキスト分析規則を適用します。

### 関連概念



カスタム分析の組み込みのワークフロー



カスタム・テキスト分析の組み込み



テキスト分析処理で使用される基本概念

## XML エLEMENTの共通分析構造へのマッピング

コレクションに分かりやすいマークアップを持つ XML 文書が含まれており、このマークアップを使用して、ユーザーがセマンティック検索でエンタープライズ・サーチ索引またはリレーショナル・データベースを検索できるようにする場合、XML エLEMENTを共通分析構造にマップすることができます。

### 始める前に

XML エLEMENTを共通分析構造にマップするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 制約事項

マッピング・ファイルの最大サイズは 8 MB です。

### このタスクについて

カスタム・テキスト分析プロセスが XML 文書内の特定のELEMENTにアクセスできるようにしたり、セマンティック検索で使用するために共通のタイプにいくつかの XML ELEMENTをマップしたりする場合、カスタム・マッピング・ファイルを作成できます。マッピング・ファイルは、テキスト分析のための UIMA フレームワークに従うものでなければなりません。

カスタム・テキスト分析エンジンを使用するコレクションにマッピング・ファイルを追加すると、ソース文書内の XML ELEMENTを共通分析構造内の注釈にマップすることができます。こうすると、これらの注釈をカスタム・テキスト分析エンジ

ンで使用することができます。共通分析構造を索引にマップすることができ、これにより、セマンティック検索でコレクションを検索するときに注釈を照会できるようになります。

例えば、共通分析構造内で、`addressee` エlementおよび `customer` エlementの内容を `Person` 注釈にマップできます。こうすると、これらの注釈にカスタム・アノテーターからアクセスできるようになり、その他の情報を検出できます (例えば、`Person` の性別を検出できます)。 `Person` 注釈をエンタープライズ・サーチ索引にマップすることもできます。こうすると、ユーザーは、元の XML エlementを知らなくても、`Person` を検索することができます。





ユーザーが照会で元の XML エlementを指定できるようにする場合、マッピング・ファイルの構成は不要です。その代わりに、構文解析オプションを構成し、コレクションについてネイティブ XML 検索を使用可能に設定できます。

### 手順

XML エlementを共通分析構造にマップするには、次のようにします。

1. コレクションを編集、「構文解析」ページを選択、そして「テキスト処理オプションの構成」をクリックします。
2. 「XML エlementを共通分析構造にマップ」領域で、「マッピングの追加」をクリックします。
3. 「XML エlementを共通分析構造にマップ」ページでマッピング・ファイルの記述表示名を入力します。
4. ファイルのロケーションを指定します。マッピング・ファイルがローカル・システムにある場合、参照してパスを見つけることができます。マッピング・ファイルが索引サーバーにある場合は、完全修飾パスを入力してください。
5. 「OK」をクリックします。新規マッピング・ファイルが、「テキスト処理オプション」ページに追加されます。

### 関連概念

-  [カスタム分析の組み込みのワークフロー](#)
-  [カスタム・テキスト分析の組み込み](#)
-  [テキスト分析処理で使用される基本概念](#)
-  [分析および検索における XML マークアップ](#)

### 関連タスク

-  [XML エlementから共通分析構造へのマッピング・ファイルの作成](#)

## 共通分析構造の索引へのマッピング

共通分析構造からのどのテキスト分析結果が、索引にマップされ、セマンティック検索を使用してコレクションを照会するユーザーから使用できるかを指定できます。

### 始める前に

共通分析構造を索引にマップするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 制約事項

マッピング・ファイルの最大サイズは 8 MB です。

### このタスクについて

共通分析構造をエンタープライズ・サーチ索引にマップすることにより、ユーザーが意味的に的確な照会を指定して、検索結果の品質を高められるようにします。





例えば、アノテーターによって検出されるエンティティと関係に応じて、ユーザーは、同じ文で発生する概念 (例えば、特定の人物と競争者名)、あるいはキーワードと概念(例えば、名前 Alex と電話番号) を検索できます。

### 手順


共通分析構造を索引にマップするには、次のようにします。

1. コレクションを編集、「構文解析」ページを選択、そして「テキスト処理オプションの構成」をクリックします。
2. 「共通分析構造を索引にマップ」領域で「マッピング・ファイルの選択」をクリックします。
3. 「このコレクションのマッピング・ファイルの選択」ページでエンタープライズ・サーチ索引で使用するマッピング・ファイルを選択します。
  - エンタープライズ・サーチ索引でデフォルトのマッピング規則を使用する場合、**Default** を選択します。
  - カスタム共通分析構造を索引にマップするには、マッピング・ファイルのロケーションを指定します。ファイルがローカル・システムにある場合、参照してファイルを見つけることができます。ファイルが索引サーバーにある場合は、完全修飾パスを入力してください。
4. 「OK」をクリックします。指定したマッピング・ファイルが、「テキスト処理オプション」ページに表示されます。

### 関連概念

-  カスタム分析の組み込みのワークフロー
-  カスタム・テキスト分析の組み込み
-  テキスト分析処理で使用される基本概念
-  カスタム分析結果の索引マッピング

### 関連タスク

-  共通分析構造から索引へのマッピング・ファイルの作成



## 共通分析構造のリレーショナル・データベースへのマッピング

共通分析構造からのどのテキスト分析結果が、データベース・アプリケーションで使用するためにリレーショナル・データベースにマップされるかを指定できます。

### 始める前に

共通分析構造をリレーショナル・データベースにマップするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 制約事項

マッピング・ファイルの最大サイズは 8 MB です。

### このタスクについて

共通分析構造をリレーショナル・データベース表にマップすることによって、データベース・アプリケーションがデータを使用できるようになります。例えば、ユーザーは、エンタープライズ・サーチの外部の SQL 照会を指定して、テキスト分析によって追加された注釈を検索できます。また、データ・マイニング・アプリケーションの情報を使用するなど、一層のテキスト処理の情報を使用できます。

一度に 1 つのマッピング・ファイルを使用して、共通分析構造をリレーショナル・データベースにマップできます。新しいマッピング・ファイルをアップロードすると、現行のマッピング・ファイルが上書きされます。

### 手順

共通分析構造をリレーショナル・データベースにマップするには、次のようにします。

1. コレクションを編集、「構文解析」ページを選択、そして「テキスト処理オプションの構成」をクリックします。
2. 「共通分析構造をリレーショナル・データベースにマップ」領域で「マッピングの追加」をクリックします。
3. 「共通分析構造をリレーショナル・データベースにマップ」ページで、リレーショナル・データベースに情報をマップするのに使用するマッピング・ファイルの記述表示名を入力します。
4. マッピング・ファイルのロケーションを指定します。ファイルがローカル・システムにある場合、参照してファイルを見つけることができます。ファイルが索引サーバーにある場合は、完全修飾パスを入力してください。
5. 「OK」をクリックします。マッピング・ファイルの表示名が、「テキスト処理オプション」ページに表示されます。

### 関連概念


 [カスタム分析の組み込みのワークフロー](#)

 [カスタム・テキスト分析の組み込み](#)

 [テキスト分析処理で使用される基本概念](#)

 選択した分析結果のデータベース・マッピング

### 関連タスク

 共通分析構造からデータベースへのマッピング・ファイルの作成

---

## パーサー・サービス用スレッドの構成

十分なメモリー・リソースがある場合、パーサーが文書の構文解析に使用できるスレッドの数を増やすことができます。

### 始める前に

コレクションが多数ある場合などに、パーサー・スレッド数を増やすことができます。ただし、追加スレッドをサポートするのに十分なメモリーがシステムにあることを確認してください。スレッドが 1 つのパーサーは 200 MB のメモリーを必要とします。スレッドを 1 つ増やすごとに 50 MB のメモリーが追加が必要になります。

パーサー用に開始されるスレッドの数を構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順

パーサー・スレッドの数を構成するには、次のようにします。

1. コレクションを編集、「構文解析」ページを選択、そして「**構文解析オプションの構成**」をクリックします。
2. パーサー開始時に開始されるパーサー・スレッドの最大数を指定し、「**OK**」をクリックします。
3. パーサーを再始動します。

---

## 複合語の拡張分析の使用可能化

パーサーが複合語に対して拡張分析を使用できるようにすることにより、検索の品質を高めることができます。拡張分析を使用すると、複合語は分解され、各パーツを 1 つの語であるように扱うことができるようになります。

### 始める前に

複合語の構文解析に関するオプションを指定するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

言語によっては、複数の語からなる単語を空白を入れずに 1 つの語にまとめるものがあります (複合語)。複合語の拡張分析と分解は、ドイツ語のような言語の検索で有用であり、韓国語のような言語の検索では特に重要です。

複合語の拡張分析を使用可能にしておくと、ユーザーは、ワイルドカード文字を使用して複数の照会条件の組み合わせを表す必要なく、語句の検索を行うことができます。例えば、Organ (臓器) を検索すると、Organspender (臓器ドナー) が含まれている文書が戻されますが、Organisation (組織) が含まれている文書は戻されません。ワイルドカード照会 Organ\* では Organ の後にどのようなストリングが続いていてもすべて戻されますが、この検索はそれとは異なり、複合語を構成する言語的に可能なサブワード全体を突き合わせます。

別の例では、複合語 Mustermann は、索引に別個に格納される 2 つのトークン (muster と mann) に分割されます。ワイルドカード照会 Musterma\* が入力された場合、検索処理では Musterma を分解された語の接頭部として識別できません。この結果、用語 Mustermann を使用する文書は検出されません。複合語のワイルドカード照会をユーザーが入力できるようにする場合は、複合語の拡張分析を使用可能にしないでください。

同義語やランキング調整ワードなどのユーザー定義の語も、照会で個々の語として使用される複合パーツに当てはまります。

### 手順

複合語の拡張分析を使用可能にするには、次のようにします。

1. コレクションを編集して「構文解析」を選択し、「**構文解析オプションの構成**」をクリックします。
2. 「**複合語の拡張分析を使用可能にする**」チェック・ボックスを選択し、「**OK**」をクリックします。

### 関連概念

205 ページの『照会内のワイルドカード文字』



セマンティック検索の言語サポート



エンタープライズ・サーチに組み込まれているテキスト分析

---

## ネイティブ XML 検索のサポートの使用可能化

コレクションに XML 文書が含まれている場合、ユーザーがコレクションのネイティブ XML 検索を使用可能にすることにより文書を検索するときに XML マークアップを使用できるようにすることができます。

### 始める前に

XML 文書をネイティブ XML 検索で検索するためのサポートを使用可能にするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

XPath または XML フラグメントなど、ネイティブ XML 検索は、照会内で文書の XML マークアップを活用して検索結果の精度を高めることができます。ユーザーは、特定の XML エlement または属性内で照会条件が発生するよう指定できます。

例えば、XML 形式のコンピューター小売業者からの送り状に、<company> エレメントと <computertype> エレメントが組み込まれている <order> 項目が含まれていることがあります。IBM ノートブックの注文が含まれている送り状をリトリートするために、IBM とノートブックのキーワード検索により、Dell ノートブック・コンピューターおよび IBM デスクトップ・モデルが含まれている文書がリトリートされます。XML 検索を使用すると、<company> エレメントに IBM が出現すること、<computertype> エレメントに notebook が出現すること、さらに、両方のエレメントが同じ <order> エレメントの下にあることを指定できます。この方法で、IBM ノートブックを明記している送り状をリトリートします。

## 手順

ユーザーがネイティブ XML 検索でコレクションを検索できるようにするには、次のようにします。

1. コレクションを編集、「構文解析」ページを選択、そして「**構文解析オプションの構成**」をクリックします。
2. 「**XML 文書の検索でネイティブ XML 検索を可能にする**」を選択します。
3. 「**OK**」をクリックします。

## 関連概念



セマンティック検索の言語サポート



エンタープライズ・サーチに組み込まれているテキスト分析



セマンティック検索アプリケーション



セマンティック検索照会条件

154 ページの『XML 検索フィールド』

---

## 文書フォーマットの検出

エンタープライズ・サーチでは、URL 拡張子や MIME タイプのデフォルト・マッピングを使用して、文書タイプや各文書で使用するパーサー・タイプを判別します。

parserTypes.cfg 構成ファイルを編集することによって、URL 拡張子や MIME タイプからパーサー・タイプへのデフォルト・マッピングをオーバーライドし、拡張することができます。parserTypes.cfg ファイルでは、ファイル拡張子や MIME タイプをパーサー・タイプにマップする規則を定義します。例えば、.content といったファイル拡張子をマップして、そのタイプの文書が HTML パーサーで解析されるように指定することができます。

文書フォーマットによって、その内部表記が異なります。エンタープライズ・サーチ・システムでは、文書の解析に内部フィルターやサード・パーティーのフィルターを使用します。多くの文書が、個々のフォーマットに専用のパーサー・サービスで解析されます。

文書フォーマットの検出とパーサーの割り当ては次のように行われます。

1. 文書フォーマットを検出するアルゴリズムでは、処理される文書の URL の拡張子を検査します。
2. システムが文書の MIME タイプを検査します。これは、クローラーによって設定されるメタデータの一部です。
3. システムが、各文書に適切なパーサー・タイプを割り当てようとします。HTML、テキスト (TXT)、XML 文書の場合、システムは各文書フォーマットに特定のパーサー・タイプを割り当てます。

他のすべての文書フォーマットに対して、システムは Stellent パーサーを使用します。Stellent 文書フィルタリング・テクノロジーは現在 Oracle が所有しています。本書で Stellent と言及することは、Oracle Outside In Content Access テクノロジーを指すのと同じことです。

Stellent パーサーは、数百の文書フォーマットをサポートしますが、エンタープライズ・サーチでは文書フィルターの一部のみが使用可能になっています。ただし、構成ファイルを編集して、他の文書タイプを Stellent パーサーで解析できるようにすることが可能です。

**重要:** デフォルトのシステム構成でエンタープライズ・サーチに使用可能に設定されている文書フィルターのサブセットに属していない、追加する文書フィルターは、テストされておらずサポートされません。

4. parserTypes.cfg ファイルが有効でない場合、デフォルトのマッピングを使用して、文書タイプと使用するパーサーを判別します。システムは以下の順番で、文書タイプの判別を行います。
  - a. URL 拡張子を parserTypes.cfg ファイルのユーザー定義拡張子規則と比較します。

- b. MIME タイプを parserTypes.cfg ファイルのユーザー定義 MIME タイプ規則と比較します。
  - c. URL 拡張子をエンタープライズ・サーチのデフォルト規則と比較します。
  - d. MIME タイプをエンタープライズ・サーチのデフォルト MIME タイプ規則と比較します。
5. システムが文書の文書フォーマットを識別できない場合、その文書は拒否されます。文書タイプがサポートされないことを示すエラー・メッセージが表示されることもあります。

パーサー・タイプとして Stellent が割り当てられている場合、Stellent が文書フォーマットを認識できないとエラー・メッセージが表示されることがあります。エラーは、以下の場合に発生する可能性があります。

- 文書が壊れている。
- 文書のフォーマットが Stellent でサポートされるものでない。この問題を解決するには、拒否されたファイル・フォーマットを stellentTypes.cfg ファイルに追加する必要があります。また、parserTypes.cfg ファイルを更新して、拒否された文書フォーマットの MIME タイプまたは拡張子が Stellent パーサーに関連付けられることを指定する必要もあります。

---

## デフォルトでサポートされる文書タイプ

文書フォーマットを検出する際、特定の文書タイプのみが判定されます。

次の文書フォーマットは、組み込みのコレクション・パーサー・サービスによって自動的に検出および解析されるネイティブ・タイプです。

HTML  
プレーン・テキスト  
XML

デフォルトで、次の文書フォーマットは、Stellent パーサーによって解析されます。

Adobe PDF (Portable Document Format)  
Lotus 1-2-3®  
Lotus Freelance Graphics®  
Lotus Word Pro®  
Just System Ichitaro  
Microsoft Excel (2007 までのバージョン)  
Microsoft PowerPoint (2007 までのバージョン)  
Microsoft Visio  
Microsoft Word (2007 までのバージョン)  
RTF (リッチ・テキスト・フォーマット)  
StarOffice/OpenOffice Calc  
StarOffice/OpenOffice Impress  
StarOffice/OpenOffice Draw  
StarOffice/OpenOffice Writer

Microsoft Office の Office Open XML ファイル形式および OpenOffice の OpenDocument 形式は、構成ファイルに変更を加える必要なしに処理されます。

他のタイプの文書を構文解析するには、構成ファイル (parserTypes.cfg および stellenttypes.cfg) を更新して、コレクション・パーサー・サービスまたは Stellent フィルターに特定の文書タイプをマッピングする規則を指定する必要があります。

**制約事項:** テキストの論理読み取り順を一致させるために PDF ファイル内の双方向テキストを処理することは、Stellent ビューアー・テクノロジーの範囲外です。Stellent パーサーは、PDF ファイルから抽出するテキストの順序は保証しません。PDF ファイル内の双方向テキストでは、テキストが構文解析される順序は、テキストの論理読み取り順と一致しない可能性があります。この制約のために、ヘブライ語やアラビア語などの中東の言語で作成された PDF ファイルの処理で問題が発生します。これらの言語では、ほとんどが右から左に記述されます (双方向)。

---

## コレクション・パーサーと Stellent パーサーに関連付けられた文書タイプ

クロール・スペース内の文書が正確にかつ効率的に構文解析されるように、どのタイプの文書をコレクション・パーサーで構文解析し、どれを Stellent 文書フィルターで構文解析するかを指定する構成ファイルを作成できます。

エンタープライズ・サーチ・コレクションでは、ほとんどの文書フォーマットが標準装備の HTML パーサーまたは XML パーサーで処理されます。一定タイプの文書 (ポストスクリプト文書など) が、通常、構文解析されず、他に Stellent 解析機能で処理される文書タイプ (Microsoft Word、Microsoft Excel、Microsoft PowerPoint、Lotus Freelance、Lotus 123、PDF、RT、および一太郎文書タイプなど) があります。

メタデータが誤りを招く可能性があり、プレーン・テキストや HTML 文書が Stellent パーサーに間違えて送信されて、それから標準装備のパーサーの 1 つに戻されるので、パフォーマンスに影響を与えかねない状態になります。また、他の文書で、文書タイプを検出できず、その文書がスキップされることがあります。こうした状況を避けるために、構成ファイルを作成して種々のタイプの文書をどこでどのように構文解析するかを制御することができます。

文書タイプとコレクション・パーサーや Stellent パーサーとの関連付けには、以下の作業が含まれます。

1. コレクション・パーサー用の文書タイプを構成する。このステップでは、文書タイプをコレクションが使用するパーサーにマップする構成ファイルを作成します。コレクションごとに、こうした構成ファイルを 1 つ作成することができます。
2. Stellent パーサー用の文書タイプを構成する。このステップでは、文書タイプをコレクションが使用する Stellent 文書フィルターにマップする構成ファイルを作成します。コレクションごとに、こうした構成ファイルを 1 つ作成することができます。
3. パーサーの停止と再始動。変更を有効にするには、エンタープライズ・サーチ管理コンソールを使用して、文書タイプを構成したコレクションをモニターしてから、パーサーを停止して再始動します。

---

## 文書タイプとコレクション・パーサーとの関連付け

特定の文書タイプをコレクション・パーサーと関連付けるには、`parserTypes.cfg` 構成ファイルを作成します。エンタープライズ・サーチ管理コンソールでは、このタスクのサポートはありません。

### 始める前に

この操作を完了するには、エンタープライズ・サーチ管理者としてログインする必要があります。

### このタスクについて

構成ファイルが存在しない場合、コレクション・パーサーはデフォルトのパーサー・サービス規則を使用します。構成ファイルが存在する場合、ファイル内の規則は次のものを指定します。

- URL 拡張子や MIME タイプのパーサー・タイプへのマップ。
- メタデータが不完全であるためにタイプが不明な文書の構文解析方法。

`parserTypes.cfg` ファイルのフォーマットは、各行が以下のいずれかの規則となる一連の行です。

#### **EXTENSION** *extension parser*

指定の拡張子で終わる URL をもつすべての文書が、指定のパーサーによって処理されます。拡張子にピリオドを含めないでください。比較では大/小文字を区別しません。

#### **CONTENTTYPE** *type/subtype parser*

指定のタイプ/サブタイプに一致するコンテンツ・タイプをもつすべての文書が、指定のパーサーによって処理されます。文書のコンテンツ・タイプ `t/s` を指定した場合、一致は、`t` が `type` に等しく、かつ、`s` が `subtype` に等しいか、またはサブタイプがワイルドカード文字 (アスタリスク `*`) である場合に発生します。

#### **UNKNOWN** *parser*

不明な拡張子とコンテンツ・タイプ (すなわち、クローラーにより使用可能にされていない) をもつすべての文書が、指定のパーサーによって処理されます。

#### **DEFAULT** *parser*

他のどの規則の対象にもならないすべての文書が、指定のパーサーによって処理されます。

どの場合でも、`parser` には、`text`、`html`、`xml`、`stellent`、または `none` を指定する必要があります。ここで、`none` は、そのタイプの文書が構文解析されないことを意味します。

1 つの文書に複数の規則が一致する場合は、規則の出現順は関係なく、より明確な特定の規則が優先されます。

- **EXTENSION** 規則は、**CONTENTTYPE** 規則よりも明確です。



- サブタイプが含まれている CONTENTTYPE 規則は、ワイルドカード文字をもつ規則よりも明確です。例えば、コンテンツ・タイプ `application/postscript` の規則は、`application/*` の規則よりも優先されます。
- 同じ拡張子またはコンテンツ・タイプに規則が 2 つあってはなりません。規則が 2 つある場合、どちらの規則が優先されるかはインプリメンテーション次第です。

## 手順

文書タイプをコレクション・パーサーと関連付けるには、次のようにします。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. 以下のように、構成ファイルを作成します。ここで、`collection_ID` は、構成したいコレクションを示します。  
`ES_NODE_ROOT/master_config/collection_ID.parserdriver/parserTypes.cfg`
3. テキスト・エディターを使用して構成ファイルを編集し、パーサー・サービス規則を指定したら、そのファイルを保存して終了します。
4. 変更を有効にするために、エンタープライズ・サーチ管理コンソールを使用してコレクションのパーサーをモニターし、パーサーを停止して、再始動します。

## 例

この例では、組み込み HTML パーサーは、拡張子 `txt`、`htm`、または `html` をもつか、`text/` で始まるコンテンツ・タイプをもつ、あるいは不明な拡張子とコンテンツ・タイプをもつ、すべての文書进行处理します。組み込み XML パーサーは、拡張子 `xml` をもつか、またはコンテンツ・タイプ `text/xml` をもつ、すべての文書进行处理します。`application/` で始まるコンテンツ・タイプをもつ文書を含め、他の文書はすべて、Stellent パーサーに送信されます。

```
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
```

### 関連タスク

177 ページの『一部の HTML タグの置換規則の変更』

179 ページの『文書タイプと Stellent パーサーとの関連付け』

### 関連資料

176 ページの『デフォルトのコレクション・パーサー・サービス規則』

## デフォルトのコレクション・パーサー・サービス規則

ファイル・タイプとコンテンツ・タイプをコレクションのパーサーにマッピングする構成ファイルを作成しない場合、文書の構文解析にはデフォルトの規則が使用されます。

コレクション・パーサーで使用されるデフォルトの規則は以下のようになります。

```
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 jttdc stellent
EXTENSION jtddc stellent
EXTENSION jtddx stellent
EXTENSION ps none
EXTENSION xml xml
EXTENSION txt text
EXTENSION htm html
EXTENSION html html
EXTENSION shtml html
EXTENSION xhtml html
EXTENSION asp html

CONTENTTYPE application/postscript none
CONTENTTYPE application/* stellent
CONTENTTYPE text/rtf stellent
CONTENTTYPE text/richtext stellent
CONTENTTYPE text/xml xml
CONTENTTYPE text/html html
CONTENTTYPE text/plain text

UNKNOWN none
DEFAULT none
```

### 関連タスク

174 ページの『文書タイプとコレクション・パーサーとの関連付け』

## 不明な文書タイプの構文解析

文書タイプが不明な場合に (例えば、文書にファイル拡張子がない、または文書に MIME タイプが関連付けされていない)、パーサーが文書を廃棄しないようにする規則を構成できます。

このタスクについて

パーサーがファイル・フォーマットを認識できない場合、パーサーはデフォルトの HTML パーサーで文書を構文解析しようとしています。内容が HTML フォーマットでない場合、パーサーはその文書を拒否します。すべての拒否された文書のレコードが、索引サーバー上の次のロケーションに書き込まれます。

ES\_NODE\_ROOT/data/collection\_ID/dropped\_doc\_logs/dropped\_docs\_pd\_date.log

### 手順

この状態を回避するには、不明な文書タイプに ASCII パーサーを使用するようにパーサーを構成します。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. 次のパーサー構成ファイルを作成または編集します。ここで、*collection\_ID* は構成するコレクションを示します。

ES\_NODE\_ROOT/master\_config/collection\_ID.parserdriver/parserTypes.cfg

3. 次の規則をファイルに追加します。

UNKNOWN text

4. 変更を有効にするために、エンタープライズ・サーチ管理コンソールを使用してコレクションのパーサーをモニターし、パーサーを停止して、再始動します。

---

## 一部の HTML タグの置換規則の変更

パーサーが HTML 文書を構文解析するときに使用する HTML タグ置換規則を変更できます。エンタープライズ・サーチ管理コンソールでは、このタスクのサポートはありません。

### 始める前に

この操作を完了するには、エンタープライズ・サーチ管理者としてログインする必要があります。

### このタスクについて

エンタープライズ・サーチ索引のテキスト情報を提供するために、HTML パーサーは、Web ブラウザーがタグを処理するのと同じ方法でタグの意味をモデル化する他の文字で、マークアップ情報 (HTML タグ) を置換します。例えば、パラグラフ・タグ (<p>) は、共通分析構造 (CAS) および索引ではパラグラフ区切り文字になります。置換規則は、同じパラグラフにテキスト情報のどの部分が表示されるかに影響します。

ある HTML タグの意味をさらに HTML 標準に近づけるために、コレクションのパーサー構成ファイルを更新して、置換規則の一部を変更することができます。

**ヒント:** HTML 文書内の HTML タグを処理する方法を制御できる他の方法については、<http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011251> を参照してください。

### 手順

HTML タグの置換規則を変更するには、次の手順で行います。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. ES\_NODE\_ROOT/master\_config/collection\_ID/parserdriver/collection.properties ファイルを編集します。ここで、collection\_ID は構成するコレクションを示します。
3. 次のパラメーターを true に設定します。  
trevis.tokenizer.newHtmlTagReplacement=true

次の HTML タグの置換規則が、下に示す値に変更されます。他のすべての HTML タグの置換規則は、デフォルトの規則を継続して使用します。

```
dfn    EMPTY
div    NEW LINE
q      BLANK
span   EMPTY
```

4. 変更を有効にするために、エンタープライズ・サーチ管理コンソールを使用してコレクションのパーサーをモニターし、パーサーを停止して、再始動します。

#### 関連タスク

174 ページの『文書タイプとコレクション・パーサーとの関連付け』

#### 関連資料

『デフォルトの HTML 置換規則』

## デフォルトの HTML 置換規則

HTML パーサーは、HTML 文書内のマークアップ情報を解釈および置換して、エンタープライズ・サーチ索引用のテキスト情報を提供します。

エンタープライズ・サーチ用の HTML パーサーは、HTML タグの解析時に以下の置換規則を使用します。第 1 列は HTML タグ名を示します。第 2 列は置換値を示します。

```
comment  BLANK
a         EMPTY
applet    EMPTY
area     EMPTY
b        EMPTY
base     EMPTY
big      EMPTY
body     EMPTY
br       NEW LINE
caption  EMPTY
center   EMPTY
del      EMPTY
dfn     PARAGRAPH
dir     PARAGRAPH
div     BLANK
dl     PARAGRAPH
em     EMPTY
form   EMPTY
frame  PARAGRAPH
h1     PARAGRAPH
h2     PARAGRAPH
h3     PARAGRAPH
h4     PARAGRAPH
h5     PARAGRAPH
```

h6	PARAGRAPH
hr	PARAGRAPH
iframe	EMPTY
img	NEW LINE
li	BLANK
meta	EMPTY
object	EMPTY
ol	BLANK
option	EMPTY
p	PARAGRAPH
q	PARAGRAPH
samp	PARAGRAPH
script	EMPTY
select	EMPTY
spacer	BLANK
span	NEW LINE
strike	EMPTY
strong	EMPTY
style	EMPTY
table	PARAGRAPH
td	NEW LINE
title	PARAGRAPH
tr	NEW LINE
ul	BLANK
xmp	PARAGRAPH

#### 関連タスク

177 ページの『一部の HTML タグの置換規則の変更』

---

## 文書タイプと Stellent パーサーとの関連付け

Stellent 文書フィルターで構文解析する文書のタイプを指定するには、`stellenttypes.cfg` 構成ファイルを作成します。エンタープライズ・サーチ管理コンソールでは、このタスクのサポートはありません。

#### Oracle Outside In テクノロジー:

Stellent 文書フィルタリング・テクノロジーは現在 Oracle が所有しています。本書で Stellent と言及することは、Oracle Outside In Content Access テクノロジーを指すのと同じことです。

#### 制約事項

OmniFind Enterprise Edition では、次の文書タイプに Stellent フィルターをサポートします。

- Adobe PDF (Portable Document Format)
- Lotus 1-2-3
- Lotus Freelance Graphics
- Lotus Word Pro
- 一太郎
- Microsoft Excel
- Microsoft PowerPoint
- Microsoft Visio
- Microsoft Word
- RTF (リッチ・テキスト・フォーマット)

エンタープライズ・サーチ索引に追加の文書タイプを組み込む際に、文書フォーマットが Stellent フィルターでサポートされる場合は、stellenttypes.cfg 構成ファイルに構文解析規則を構成できます。

**重要:** 前のリスト内の文書タイプに一致しない、追加する文書フォーマットは、エンタープライズ・サーチではテストされておらず、サポートされません。

Stellent 文書フォーマットの詳細リストについては、[http://www.oracle.com/technology/products/content-management/oit/ds\\_oitFiles.pdf](http://www.oracle.com/technology/products/content-management/oit/ds_oitFiles.pdf) の「*Outside In Technology: Supported File Formats*」を参照してください。

## 始める前に

この操作を完了するには、エンタープライズ・サーチ管理者としてログインする必要があります。

## このタスクについて

stellenttypes.cfg 構成ファイルは、次のものを指定します。

- 受け入れる規則。Stellent パーサーによって構文解析されるファイル・タイプ。ファイル・タイプは、Stellent ライブラリーが認識するファイル・タイプの 1 つに一致します。
- ネイティブ規則。コレクション・パーサーに送り返して、組み込みパーサーの 1 つで処理するファイル・タイプ。このアクションが必要になるのは、メタデータが紛らわしいために、コレクション・パーサーが誤って文書を Stellent パーサーに送信する場合があるからです。
- 拒否する規則。エンタープライズ・サーチでサポートされていないためにリジェクトされるファイル・タイプ。

構成ファイルが指定されたが、存在しない場合、パーサーは開始できません。

stellent.properties ファイル内の OutsideInSupportedTypes プロパティーに構成ファイルが指定されていない場合、Stellent パーサーのデフォルトの構文解析規則が使用されます。

構成ファイルには、文書タイプと、それらの取り扱い方法がリストされています。ファイルのフォーマットは一連の行で、各行が以下のフォーマットの 1 つと一致する規則です。

```
accept DEFAULT
accept ALL doc_type
accept stellent_type doc_type
native DEFAULT
native stellent_type doctype
reject stellent_type
```

ここで、それぞれ次のことを表します。

*doc\_type*

*doctype* 照会トークンに使用される値です。文書を文書タイプで検索できます。例えば、ユーザーは、PDF 文書を検索するために `$doctype::pdf` を指定できます。

### *stellent\_type*

Stellent ライブラリー内のフィルター・タイプ値の 1 つです (例えば、FI\_123R1)。

#### **DEFAULT**

規則のタイプに応じて、受け入れるタイプまたはネイティブ・タイプのリストにすべてのデフォルトの規則が含まれることを意味します。このオプションを選択すると、デフォルトの構成を置き換える代わりに、展開することができます。

**A11** 明示的にリストされないすべてのタイプが指定の `doctype` トークンを使用して受け入れられることを意味します。

構成ファイル内の規則は、以下のように処理されます。

1. *stellent\_type* について `reject` 規則がある場合、その文書は受け入れられません。
2. *stellent\_type* の `native` 規則 (`native DEFAULT` が指定されている場合は、デフォルトの構文解析規則を含む) がある場合、文書は、この規則で指定された *doc\_type* トークンの値に加えて組み込みパーサーに戻されます。 *doc\_type* の値は、`txt`、`htm`、または `xml` のいずれかでなければなりません。それらは、それぞれ、プレーン・テキスト、HTML、または XML を示します。
3. *stellent\_type* について `accept` 規則がある場合 (`accept DEFAULT` が指定されている場合は、デフォルト・リストを含む)、文書は受け入れられます。
4. さらに、`accept ALL` が指定されている場合、その文書は受け入れられます。
5. 上記以外の場合、文書はリジェクトされ、構文解析されません。

文書タイプが受け入れられる場合、適用された規則に指定されている *doc\_type* 値が使用されます。この値は、構文解析されたコンテンツに加えて、コレクション・パーサーに戻されます。

#### **手順**

文書タイプを `Stellent` パーサーと関連付けるには、次のようにします。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. `ES_NODE_ROOT/master_config/collection_ID.stellent/stellent.properties` ファイルを編集します。ここで、*collection\_ID* は、構成したいコレクションを示します。
3. `OutsideInSupportedTypes` プロパティに、作成している構成ファイルの絶対パスを指定します。

例えば、1 つのコレクションについて以下の構成ファイルを作成し、それを他のコレクション固有ファイルと一緒に保管できます。

```
ES_NODE_ROOT/master_config/collection_ID.stellent/stellenttypes.cfg
```

また、以下の構成ファイルを作成すると、同じ設定をすべてのコレクションで使用し、このファイルを他のシステム・レベルのファイルと一緒に保管することもできます。この方法の場合、ステップ 2 に指定されているとおり、各コレクションの `stellent.properties` ファイルに、必ずこのパスを指定してください。

ES\_INSTALL\_ROOT/default\_config/stellent/stellenttypes.cfg

4. テキスト・エディターを使用して構成ファイルを作成し、Stellent 構文解析規則を指定したら、そのファイルを保存して終了します。
5. 変更を有効にするために、エンタープライズ・サーチ管理コンソールを使用してコレクションのパーサーをモニターし、パーサーを停止して、再始動します。

## 例

以下の構成ファイルでは、Stellent セッションは、サポートされる文書タイプのデフォルトのリストのほかに、Microsoft Visio フォーマットの文書を受け入れます。

```
accept DEFAULT
accept FI_VISI03 visio
accept FI_VISI04 visio
accept FI_VISI05 visio
accept FI_VISI06 visio
```

以下の構成ファイルでは、Postscript 文書がサポートされるため、ps という文書タイプで検索が可能です。。X ピクスマップ・フォーマット (XPM) の文書は組み込みテキスト・パーサーに戻され、PNG イメージ・フォーマットの文書は拒否されます。他のファイル・タイプはすべて受け入れられ、other という文書タイプで検索可能です。

```
accept DEFAULT
accept FI_POSTSCRIPT ps
native FI_XPIXMAP txt
reject FI_PNG
accept ALL other
```

## 関連タスク

174 ページの『文書タイプとコレクション・パーサーとの関連付け』

## 関連資料

『Stellent パーサーのデフォルトの解析規則』

## Stellent パーサーのデフォルトの解析規則

ファイル・タイプを Stellant 文書フィルターにマップする構成ファイルを作成しないと、パーサーはデフォルトの規則を使用して文書を解析します。

Stellent パーサーは、以下の文書タイプを受け入れて解析します。

ACCEPT FI_123R1	123
ACCEPT FI_123R2	123
ACCEPT FI_123R3	123
ACCEPT FI_123R4	123
ACCEPT FI_123R6	123
ACCEPT FI_123R9	123
ACCEPT FI_EXCEL	xls
ACCEPT FI_EXCEL2000	xls
ACCEPT FI_EXCEL2002	xls
ACCEPT FI_EXCEL2003	xls
ACCEPT FI_EXCEL2007	xlsx
ACCEPT FI_EXCEL3	xls
ACCEPT FI_EXCEL4	xls
ACCEPT FI_EXCEL5	xls
ACCEPT FI_EXCEL97	xls
ACCEPT FI_EXTPOWERPOINT4	ppt
ACCEPT FI_EXTPOWERPOINTMAC4	ppt
ACCEPT FI_FREELANCE	prz
ACCEPT FI_FREELANCE3	prz



ACCEPT FI_ICHITAR03	jxw
ACCEPT FI_ICHITAR04	jsw
ACCEPT FI_ICHITAR08	jtd
ACCEPT FI_PDF	pdf
ACCEPT FI_PDFMACBIN	pdf
ACCEPT FI_POWERPOINT2	ppt
ACCEPT FI_POWERPOINT2000	ppt
ACCEPT FI_POWERPOINT2007	pptx
ACCEPT FI_POWERPOINT3	ppt
ACCEPT FI_POWERPOINT4	ppt
ACCEPT FI_POWERPOINT7	ppt
ACCEPT FI_POWERPOINT9597	ppt
ACCEPT FI_POWERPOINT97	ppt
ACCEPT FI_POWERPOINTMAC3	ppt
ACCEPT FI_POWERPOINTMAC4	ppt
ACCEPT FI_POWERPOINTMACB3	ppt
ACCEPT FI_POWERPOINTMACB4	ppt
ACCEPT FI_RTF	rtf
ACCEPT FI_RTFJ	rtf
ACCEPT FI_STAROFFICEWRITER8	odt
ACCEPT FI_STAROFFICEDRAW8	odg
ACCEPT FI_STAROFFICEIMPRESS8	odp
ACCEPT FI_STAROFFICECALC8	ods
ACCEPT FI_STAROFFICECALC6	sxc
ACCEPT FI_STAROFFICEDRAW6	sxd
ACCEPT FI_STAROFFICEIMPRESS6	sxi
ACCEPT FI_STAROFFICEWRITER6	sxw
ACCEPT FI_STAROFFICECALC52	sdc
ACCEPT FI_STAROFFICEIMPRESS52	sdd
ACCEPT FI_STAROFFICEWRITER52	sdw
ACCEPT FI_VISIO3	vsd
ACCEPT FI_VISIO4	vsd
ACCEPT FI_VISIO5	vsd
ACCEPT FI_VISIO6	vsd
ACCEPT FI_VISIO2003	vsd
ACCEPT FI_WINWORD1	doc
ACCEPT FI_WINWORD1COMPLEX	doc
ACCEPT FI_WINWORD1J	doc
ACCEPT FI_WINWORD2	doc
ACCEPT FI_WINWORD2000	doc
ACCEPT FI_WINWORD2002	doc
ACCEPT FI_WINWORD2003	doc
ACCEPT FI_WINWORD2007	docx
ACCEPT FI_WINWORD5J	doc
ACCEPT FI_WINWORD6	doc
ACCEPT FI_WINWORD7	doc
ACCEPT FI_WINWORD97	doc
ACCEPT FI_WORD4	doc
ACCEPT FI_WORD5	doc
ACCEPT FI_WORD6	doc
ACCEPT FI_WORDPRO	lwp
ACCEPT FI_WORDPRO97	lwp

Stellent パーサーは以下の文書タイプをコレクション・パーサーに返します。返された文書タイプはいずれかの組み込みパーサーで処理されます。

NATIVE FI_7BITTEXT	txt
NATIVE FI_ANSI	txt
NATIVE FI_ANSI8	txt
NATIVE FI_ARABIC_710	txt
NATIVE FI_ARABIC_720	txt
NATIVE FI_ARABIC_WINDOWS	txt
NATIVE FI_ASCII	txt
NATIVE FI_ASCII8	txt
NATIVE FI_CENTRALEU_1250	txt
NATIVE FI_CHINESEBIG5	txt
NATIVE FI_CHINESEGB	txt

NATIVE FI_CYRILLIC1251	txt
NATIVE FI_CYRILLICKOI8	txt
NATIVE FI_EBCDIC_1026	txt
NATIVE FI_EBCDIC_273	txt
NATIVE FI_EBCDIC_277	txt
NATIVE FI_EBCDIC_278	txt
NATIVE FI_EBCDIC_280	txt
NATIVE FI_EBCDIC_284	txt
NATIVE FI_EBCDIC_285	txt
NATIVE FI_EBCDIC_297	txt
NATIVE FI_EBCDIC_37	txt
NATIVE FI_EBCDIC_500	txt
NATIVE FI_EBCDIC_870	txt
NATIVE FI_EBCDIC_871	txt
NATIVE FI_HANGEUL	txt
NATIVE FI_HEBREW_E0	txt
NATIVE FI_HEBREW_OLDCODE	txt
NATIVE FI_HEBREW_PC8	txt
NATIVE FI_HEBREW_WINDOWS	txt
NATIVE FI_HTML	htm
NATIVE FI_HTML_ARABIC_ASM0708	htm
NATIVE FI_HTML_ARABIC_DOS	htm
NATIVE FI_HTML_ARABIC_ISO	htm
NATIVE FI_HTML_ARABIC_MAC	htm
NATIVE FI_HTML_ARABIC_WINDOWS	htm
NATIVE FI_HTML_BALTIC_ISO	htm
NATIVE FI_HTML_BALTIC_WINDOWS	htm
NATIVE FI_HTML_CENTRALEUROPEAN_DOS	htm
NATIVE FI_HTML_CENTRALEUROPEAN_ISO	htm
NATIVE FI_HTML_CENTRALEUROPEAN_MAC	htm
NATIVE FI_HTML_CENTRALEUROPEAN_WINDOWS	htm
NATIVE FI_HTML_CHINESEBIG5	htm
NATIVE FI_HTML_CHINESEEUC	htm
NATIVE FI_HTML_CHINESEGB	htm
NATIVE FI_HTML_CHINESESIMPLIFIED_EUC	htm
NATIVE FI_HTML_CHINESESIMPLIFIED_WINDOWS	htm
NATIVE FI_HTML_CHINESETRADITIONAL_WINDOWS	htm
NATIVE FI_HTML_CYRILLIC_DOS	htm
NATIVE FI_HTML_CYRILLIC_ISO	htm
NATIVE FI_HTML_CYRILLIC_KOI8R	htm
NATIVE FI_HTML_CYRILLIC_MAC	htm
NATIVE FI_HTML_CYRILLIC_WINDOWS	htm
NATIVE FI_HTML_CYRILLIC1251	htm
NATIVE FI_HTML_CYRILLICKOI8	htm
NATIVE FI_HTML_EBCDIC_1026	htm
NATIVE FI_HTML_EBCDIC_273	htm
NATIVE FI_HTML_EBCDIC_277	htm
NATIVE FI_HTML_EBCDIC_278	htm
NATIVE FI_HTML_EBCDIC_280	htm
NATIVE FI_HTML_EBCDIC_284	htm
NATIVE FI_HTML_EBCDIC_285	htm
NATIVE FI_HTML_EBCDIC_297	htm
NATIVE FI_HTML_EBCDIC_37	htm
NATIVE FI_HTML_EBCDIC_500	htm
NATIVE FI_HTML_EBCDIC_870	htm
NATIVE FI_HTML_EBCDIC_871	htm
NATIVE FI_HTML_GREEK_ISO	htm
NATIVE FI_HTML_GREEK_MAC	htm
NATIVE FI_HTML_GREEK_WINDOWS	htm
NATIVE FI_HTML_HEBREW_DOS	htm
NATIVE FI_HTML_HEBREW_ISO_VISUAL	htm
NATIVE FI_HTML_HEBREW_WINDOWS	htm
NATIVE FI_HTML_JAPANESE_MAC	htm
NATIVE FI_HTML_JAPANESE_SHIFTJIS	htm
NATIVE FI_HTML_JAPANESE_EUC	htm
NATIVE FI_HTML_JAPANESE_JIS	htm
NATIVE FI_HTML_JAPANESE_SJIS	htm

NATIVE FI_HTML_KOREAN_JOHAB	htm
NATIVE FI_HTML_KOREAN_WINDOWS	htm
NATIVE FI_HTML_KOREANHANGUL	htm
NATIVE FI_HTML_LATIN2	htm
NATIVE FI_HTML_RUSSIAN_DOS	htm
NATIVE FI_HTML_THAI_WINDOWS	htm
NATIVE FI_HTML_TURKISH_DOS	htm
NATIVE FI_HTML_TURKISH_ISO	htm
NATIVE FI_HTML_TURKISH_MAC	htm
NATIVE FI_HTML_TURKISH_WINDOWS	htm
NATIVE FI_HTML_VIETNAMESE_WINDOWS	htm
NATIVE FI_HTML_WESTERNEUROPEAN_ISO	htm
NATIVE FI_HTML_WESTERNEUROPEAN_MAC	htm
NATIVE FI_HTML_WESTERNEUROPEAN_WINDOWS	htm
NATIVE FI_HTMLUNICODE	htm
NATIVE FI_JAPANESE_EUC	txt
NATIVE FI_JAPANESE_JIS	txt
NATIVE FI_LATIN2	txt
NATIVE FI_MAC	txt
NATIVE FI_MAC8	txt
NATIVE FI_PP2KHTML	htm
NATIVE FI_SHIFTJIS	txt
NATIVE FI_UNICODE	txt
NATIVE FI_UTF8	txt
NATIVE FI_W2KHTML	htm
NATIVE FI_WML	xml
NATIVE FI_WML_CHINESEBIG5	xml
NATIVE FI_WML_CHINESEEUC	xml
NATIVE FI_WML_CHINESEGB	xml
NATIVE FI_WML_CYRILLIC1251	xml
NATIVE FI_WML_CYRILLICKO18	xml
NATIVE FI_WML_JAPANESEEUC	xml
NATIVE FI_WML_JAPANESEJIS	xml
NATIVE FI_WML_JAPANESESJIS	xml
NATIVE FI_WML_KOREANHANGUL	xml
NATIVE FI_WML_LATIN2	xml
NATIVE FI_XHTML	htm
NATIVE FI_XL2KHTML	htm
NATIVE FI_XML	xml
NATIVE FI_XML_DOCTYPE_HTML	htm

### 関連タスク

179 ページの『文書タイプと Stellent パーサーとの関連付け』



---

## 言語とコード・ページのサポート

エンタープライズ・サーチの言語処理は、パーサーと検索サーバーによって処理が異なります。

言語処理のために、パーサーは言語とロケールを区別しません。ただし、ユーザーが複数言語の文書を含むコレクションを検索する場合、検索サーバーは、検索結果を特定の言語またはロケールに限定できるようにします。

例えば、英語の文書のメタデータで、文書ロケールに `en_US` が指定された場合、この文書は英語の文書 (`en`) として、また英語の米国ロケールを使用する文書 (`en_US`) として索引付けされます。こうした索引付けにより、数字、日付、時刻などのロケール固有情報を正しく表すことができます。このコレクションを検索すると、`en` の文書を検索しても、`en_US` の文書を検索しても、この文書を検出できます。

文書が `en` などの言語コードのみによって索引付けされる場合、文書は、ロケールではなく言語コードのみによって索引付けされます。例えば、このコレクションで `en_US` の文書を検索した場合、文書は検出されません。

エンタープライズ・サーチ・システムは、ISO 639 規格に文書化された以下の言語と 2 文字言語コードをサポートします。

### 単純テキスト言語:

- en= 英語
- sq= アルバニア語
- az= アゼルバイジャン語 - ラテン文字
- bg= ブルガリア語
- be= ベラルーシ語
- ca= カタロニア語
- hr= クロアチア語
- cs= チェコ語
- da= デンマーク語
- nl= オランダ語
- et= エストニア語
- fi= フィンランド語
- fr= フランス語
- de= ドイツ語
- el= ギリシャ語
- hu= ハンガリー語
- is= アイスランド語
- id= インドネシア語
- in= インドネシア語
- it= イタリア語
- kk= カザフ語
- lv= ラトビア語
- lt= リトアニア語
- lo= ラオス語
- mk= マケドニア語
- ms= マレー語
- mt= マルタ語
- no= ノルウェー語
- nb= ノルウェー語 (ブークモール)
- pl= ポーランド語
- pt= ポルトガル語
- ro= ルーマニア語
- ru= ロシア語

sr= セルビア語 (キリル文字)  
sh= セルビア語 (ラテン文字)  
sk= スロバキア語  
sl= スロベニア語  
es= スペイン語  
sv= スウェーデン語  
tr= トルコ語  
uk= ウクライナ語  
cy= ウェールズ語

#### 表意文字言語:

中国語 (簡体字) と中国語 (繁体字) に対して、2 文字言語コードの代わりに、拡張された言語コードが使用されます。

zh-CN= 中国語 (簡体字)  
zh-TW= 中国語 (繁体字)  
ja= 日本語  
ko= 韓国語

#### 複合テキスト言語:

ar= アラビア語  
as= アッサム語  
bn= ベンガル語  
gu= グジャラート語  
iw= ヘブライ語  
he= ヘブライ語  
hi= ヒンディ語  
kn= カンナダ語  
ml= マラヤーラム語  
mr= マラーティー語  
or= オリヤー語  
pa= パンジャブ語  
ta= タミール語  
te= テルグ語  
th= タイ語  
ur= ウルドゥー語  
vi= ベトナム語

エンタープライズ・サーチ・システムは、これらの言語の多くを自動的に検出でき、プレーン・テキスト文書で使用されるコード・ページを自動的に検出できます。クローラーの構成時に、使用する明示的な言語またはコード・ページを指定する場合は、言語およびコード・ページの自動検出を使用不可にすることができます。

---

## 言語自動検出

エンタープライズ・サーチ・システムは、事実上、任意の言語の文書を処理できます。

文書の言語が以下のいずれかである場合、システムは、その言語を自動的に検出できます。文書の言語が分かっている場合は、システムに言語を自動的に検出させる代わりに、使用する言語をクローラーの構成時に指定することができます。

アラビア語  
ブルガリア語  
チェコ語  
中国語 (簡体字)  
中国語 (繁体字)  
デンマーク語  
オランダ語  
英語  
フィンランド語

フランス語 (カナダ)  
フランス語 (フランス)  
ドイツ語 (ドイツ)  
ドイツ語 (スイス)  
ギリシャ語  
ヘブライ語  
ハンガリー語  
アイスランド語  
イタリア語  
日本語  
韓国語  
ノルウェー語 (ブークモール)  
ポーランド語  
ポルトガル語 (ブラジル)  
ポルトガル語 (ポルトガル)  
ルーマニア語  
ロシア語  
スペイン語  
スウェーデン語  
タイ語  
トルコ語

---

## コード・ページ自動検出

エンタープライズ・サーチ・システムは、さまざまなコード・ページの文書をサポートします。

テキスト・ファイルの場合、システムは以下のコード・ページを自動的に検出します。他の文書フォーマットの場合、システムは、HTML メタデータ・エレメントなど、文書内のメタデータを使用して、コード・ページを検出します。文書のコード・ページが分かっている場合は、システムにコード・ページを自動的に検出させる代わりに、使用するコード・ページをクローラーの構成時に指定することができます。

### Unicode エンコード形式:

UTF-8  
UTF-16BE  
UTF-16LE

### マルチバイト・エンコード形式:

Shift-JIS  
ISO-2022-CN  
ISO-2022-JP  
ISO-2022-KR  
GB18030  
EUC-JP  
EUC-KR

### 1 バイト・エンコード形式:

ISO-8859-1: デンマーク語、オランダ語、ドイツ語、英語、フランス語、イタリア語、ノルウェー語、ポルトガル語、スペイン語、スウェーデン語  
ISO-8859-2: チェコ語、ハンガリー語、ポーランド語、ルーマニア語  
ISO-8859-5: ロシア語  
ISO-8859-6: アラビア語  
ISO-8859-7: ギリシャ語  
ISO-8859-8: ヘブライ語、視覚的順序のヘブライ語  
ISO-8859-9: トルコ語  
Windows-1250: チェコ語、ハンガリー語、ポーランド語、ルーマニア語  
Windows-1251: ロシア語  
Windows-1252: デンマーク語、オランダ語、ドイツ語、英語、フランス語、イタリア語、ノルウェー語、ポルトガル語、スペイン語、

Windows-1253: スウェーデン語  
Windows-1253: ギリシャ語  
Windows-1254: トルコ語  
Windows-1255: ヘブライ語  
Windows-1256: アラビア語  
KOI8-R: ロシア語

文字セット検出は、不正確な処理です。コード・ページ検出処理では、バイト・データの特徴に最も適合する文字セット (charset) を識別しようとはしますが、この処理は本来ある程度統計的なもので、結果が正しいと保証することはできません。

精度を最大にするために、入力データは主として 1 つの言語のものにしてください。その言語の最低で数百バイトのプレーン・テキストも必要です。

検出されたエンコードとサポートされるエンコードが一致しない場合、システムは、コレクションのデフォルトのコード・ページを使用します。

---

## 中国語、日本語、韓国語の文書の言語分析

中国語、日本語、韓国語で書かれた文書の検索性を高めるには、言語分析オプションを指定します。

中国語、日本語、韓国語の文書の場合、パーサーで字句解析に N-gram セグメンテーション方式を使用することを指定できます。また、中国語および日本語の文書では、パーサーが空白から改行文字を除去するように設定できます。

### N-gram セグメンテーション

コレクションの作成時に、ワードの区切りに空白文字を使用しない言語で書かれた文書の構文解析に使用する字句解析のタイプを選択します。

unicode ベースの空白文字のセグメンテーションでは、ワードとワードの間の区切りとして空白文字を使用します。N-gram セグメンテーションは、任意の数の文字が並ぶ連続を単一ワードとみなします。ワードの区切りに空白文字を使用しない中国語、日本語、韓国語のような言語では、N-gram セグメンテーションを使用すると、unicode ベースの空白文字のセグメンテーションよりも質のよい検索結果を戻すことができます。

コレクションの作成時に文書の構文解析に使用するセグメンテーション方式を選択します。コレクション作成後は、構文解析オプションを表示して設定を確認することはできますが、変更することはできません。

エンタープライズ・サーチ・コレクションの完全な N-gram 構文解析およびトークン化のサポートを構成する方法の詳細、および完全な N-gram サポートを構成されたコレクションで文字が処理される仕組みを理解するには、<http://www.ibm.com/support/docview.wss?rs=63&uid=swg27011088> を参照してください。

### テキストからの空白文字の除去

テキストから空白文字を除去するようにパーサーを構成することができます。

始める前に



この操作を完了するには、エンタープライズ・サーチ管理者としてログインする必要があります。

### このタスクについて

コレクションに対してこのオプションを使用可能にすると、パーサーは、2つの文字を分け隔てる連続した空白文字を除去します。例えば、文書が、中国語や日本語など、語境界の区切りに空白文字を使用しない言語のものである場合に、空白文字を除去したいことがあります。

パーサーが空白文字を除去するように構成する場合、2バイト文字セット (DBCS) 文字の間に存在する空白文字のみを除去するか、文字の前後状況に関係なくすべての空白文字を除去するかを指定できます。後者のオプションは、例えば、日本語文書に英語のテキストが含まれていて、英語のテキストからも空白文字を除去する場合に使用します。

パーサーは、以下の文字を除去します。

- タブ (0x09)
- LF または改行 (0x0A)
- CR または復帰 (0x0D)

変更を有効にするために、パーサーを停止し、再開します。既に索引に保管されている文書に変更を適用するには、目的の文書をもう一度クロールしてから、主索引を再作成します。

### 手順

テキストから空白文字を除去するには、次のようにします。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. テキスト・エディターを使用して以下のファイルを編集します。ここで、*collection\_ID* は、コレクションの作成時にそのコレクションについて指定された ID です (または、システムによって割り当てられたもの)。

`ES_NODE_ROOT/master_config/collection_ID.parserdriver/collection.properties`

3. どのように空白文字を除去するかを指定します。
  - DBCS 文字間の空白文字を除去する場合は、`removeCjNewlineChars` プロパティの値を `true` に設定します。

```
removeCjNewlineChars=true
```

- 文書内の場所を問わず空白文字を除去する場合は、`removeCjNewlineChars` プロパティの値を `true` に設定し、`removeCjNewlineCharsMode` プロパティの値を `all` に設定します。

```
removeCjNewlineChars=true  
removeCjNewlineCharsMode=all
```



---

## 索引の管理

ユーザーが常に最新の情報にアクセスできるようにするため、エンタープライズ・サーチでは、各コレクションの索引を作成し、定期的にその内容を更新して索引を保守します。

クローラーが収集したデータを検索可能にするためには、索引を作成する必要があります。コレクションを最初に作成すると、エンタープライズ・サーチによって、当初クロールされたすべてのデータの索引が作成されます。クローラーが新規または変更されたデータ・ソースをクロールすると、新規の内容に対応する更新 (差分索引 と呼ぶ) が作成されます。その更新は最終的に、ベースとなる索引にマージされる必要があります。このマージ・プロセスを主索引 の作成と呼びます。差分索引または主索引が作成されると、その新規の内容が検索サーバーにコピーされ、検索可能になります。

クローラーは、データ収集を連続的に、または定期的なスケジュールに基づいて行います。索引を頻繁に更新すれば、ユーザーが最新のデータを検索できるようになります。継続的に更新されている索引は、最終的に再作成する必要があります。索引は、大きくなるにつれて、システム・リソースの消費が増大します。最適なパフォーマンスを維持するために、主索引を定期的に作成してください。

主索引を作成する頻度は、以下のような条件によって異なります。

- システム・リソース (ファイル・システム・スペース、プロセッサの速度、およびメモリー)
- クロールおよび再クロールの必要がある文書数
- クロール対象のデータ・タイプ
- カテゴリー規則の変更頻度 (変更は、主索引作成が行われるまで有効になりません)
- クローラーをスケジュールされていた時間に実行するのではなく強制始動した回数

コレクションの文書が数百万あり、ほとんどが Web 文書で作成されている場合、主索引の作成を 1 日に 1 回、索引の更新を 1、2 時間ごとに行うのがその目安となります。

現行の、検索可能な索引を保守するには、次の作業を行います。

- 索引作成のスケジュールを指定する
- 索引スケジュールを変更する
- 索引スケジュールを使用可能/使用不可に設定する
- 索引の並行作成を構成する

索引のユーザーのビューに影響するオプションを指定するには、次の作業を行います。

- 照会でのワイルドカード文字のサポートを構成する

- ユーザーが検索できる文書の範囲を制限する有効範囲を構成する
- 検索結果で同じソースからの文書を縮小表示する
- 索引から URI を除去する

#### 関連タスク

350 ページの『コレクションの索引アクティビティのモニター』

351 ページの『エンタープライズ・サーチの索引キューのモニター』

---

## 索引作成のスケジューリング

主索引の作成と、索引の新しいコンテンツでの更新を行うスケジュールを指定できます。

### 始める前に

索引作成をスケジュールするには、エンタープライズ・サーチ管理者役割のメンバーであるか、コレクションのコレクション管理者である必要があります。

### このタスクについて

ユーザーが検索するソースで常に最新情報にアクセスできるようにするため、索引の作成が定期的に行われるようにスケジュールしてください。主索引の作成では、索引全体が再作成されます。索引作成プロセスは、クローラーが収集し、パーサーが分析したすべてのデータを読み込みます。差分索引の作成では、最後に行われた主索引の作成の後にクロールされた情報が、検索可能になります。

索引作成をスケジュールするオプションは、デフォルトで選択されます。このオプションが選択されていると、スケジューラー・プロセスは、エンタープライズ・サーチ・システムの開始時に主索引と差分索引の作成タスクをスケジュールします。必要であればいつでも「システム始動時に有効化」チェック・ボックスをクリアして、スケジュール済みの索引作成が実行されないようにできます。例えば、問題をトラブルシューティングするときなどに、スケジュールを無効にしたいことがあります。

システム・リソースを節約し、パフォーマンスを向上するために、索引に適用する必要のある変更が行われたかどうかを確認する検査がシステムによって自動的に行われます。索引に適用する必要のある変更が検出されない場合、スケジュールされた作成要求は廃棄されます。

### 手順

索引作成をスケジュールするには、次のようにします。

1. コレクションを編集して「索引」ページを選択し、「索引作成のスケジュール」をクリックします。
2. 索引が新しいコンテンツで更新される頻度を指定するには、「差分索引の作成スケジュールの指定」領域の「索引作成のスケジュール」ページで以下のオプションを指定します。
  - a. 「開始日時」領域の「年」、「月」、「日」、「時」「分」フィールドに、最初の差分索引が作成される日時を指定します。

- b. 「更新間隔」領域の「日」、「時間」、および「分」フィールドに、差分索引が作成される頻度を指定します。

通常、差分索引は、1 時間または 2 時間ごとのように、頻繁に作成します。ソース・コンテンツの変更頻度に応じて、間隔の長短を調整してください。例えば、毎時間 (0 日と 1 時間)、あるいは 12 時間ごと (0 日と 12 時間) などと指定します。

- 3. 索引が完全に再作成される頻度を指定するには、「主索引の作成スケジュールの指定」領域の以下のオプションを指定します。
  - a. 「開始日時」領域の「年」、「月」、「日」、「時」「分」フィールドに、主索引が最初に作成される日時を指定します。
  - b. 「更新間隔」領域の「日」、「時間」、および「分」フィールドに、主索引が作成される頻度を指定します。

通常、索引全体は、24 時間ごとのように、定期的に作成します。ソース・コンテンツの変更頻度に応じて、間隔の長短を調整してください。例えば、12 時間ごと (0 日と 12 時間)、あるいは 2.5 日間ごと (2 日と 12 時間) などと指定します。

- 4. 「OK」をクリックします。

## 索引スケジュールの変更

索引作成のスケジュールを変更することができます。

### 始める前に

索引スケジュールを変更するには、操作者がエンタープライズ・サーチ管理者の役割のメンバーか、またはそのコレクションのコレクション管理者である必要があります。

### 手順

索引スケジュールを変更するには、次のようにします。

- 1. コレクションを編集して「索引」ページを選択し、「月」、「日」、「年」、および「時」フィールドで該当する値を変更します。索引の更新頻度と、主索引の作成頻度を指定します。
- 2. 「適用」をクリックします。

## 索引スケジュールの使用可能および使用不可設定

索引作成のスケジュールを使用可能または使用不可に設定できます。

### 始める前に

索引スケジュールを使用可能または使用不可にするには、操作者がエンタープライズ・サーチ管理者の役割のメンバーか、またはそのコレクションのコレクション管理者である必要があります。

### このタスクについて

スケジュール済みの索引作成を実行しないようにする必要がある場合、索引スケジュールを使用不可にすることができます。例えば、問題が発生した場合に、トラブルシューティングを行うために、スケジュールされた日時に索引作成が行われないようにそのスケジュールを使用不可にすることができます。

スケジュールの使用可能および使用不可設定は、コレクションの編集時、およびコレクションのモニター時に行うことができます。

### 手順

1. コレクションの編集で索引スケジュールを使用可能または使用不可にするには、次のステップに従います。
  - a. 変更するコレクションを編集します。
  - b. 「索引」 ページで、「システム始動時に有効化」 チェック・ボックスを選択または選択解除すると、索引更新のスケジュールが使用可能または使用不可になります。
  - c. 「システム始動時に有効化」 チェック・ボックスを選択または選択解除すると、主索引作成のスケジュールが使用可能または使用不可になります。
  - d. 「適用」 をクリックします。
2. コレクションのモニターで索引スケジュールを使用可能または使用不可にするには、次のステップに従います。
  - a. 変更するコレクションをモニターします。
  - b. 「索引」 ページで、索引がスケジュール済みの場合にスケジュールされた日時に索引が作成されないようにするには、「 スケジュールを使用不可にする」 をクリックします。これで、索引は、スケジュールを使用可能にするか、「 開始」 をクリックして索引作成プロセスを開始するまで作成されなくなります。
  - c. 索引がスケジュール済みであるが、索引作成のスケジュールが使用不可にされている場合、「 スケジュールを使用可能にする」 をクリックします。

索引は、索引スケジュールで指定した日時に作成されるようにキューに入れます。

## 索引の並行作成の構成

索引作成要求の同時処理が可能なコレクションの数を指定することによって、索引作成リソースの使用を制御します。十分なシステム・リソースがある場合、索引の更新と主索引の作成が同時に行われるように設定することにより、検索の質を高めることができます。

### 始める前に

システムの索引作成オプションを指定するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### このタスクについて

エンタープライズ・サーチは、複数のコレクションでリソースを共有することによって複数の索引を同時に作成でき、それによって、複数のコレクションの索引作成

要求が並行して処理されることが可能になります。プロセスを共用すると、大規模な索引の作成によって、キューで作成を待っている他の索引の可用性が阻害されることがありません。

索引作成が要求またはスケジュールされると、索引キューに入れられて処理される順番を待ちます。各コレクションには独自の索引があるので、さまざまなコレクションからの複数の索引作成要求が同時に索引キューにある場合があります。システムの索引オプションを構成するとき、いくつかのコレクションが索引作成用リソースを共用して要求を同時に処理させることができるのかを指定します。


また、索引を更新する要求が、そのコレクションの主索引が作成されるのと同時に処理されるように指定することもできます。このオプションを使用可能にすると、実行に長く時間のかかる主索引作成が処理されている間に、検索サーバーは (差分索引により) 最新の文書でリフレッシュされます。ただし、索引作成は、リソース集約的なプロセスです。索引作成中には、大量のシステム・メモリーとディスク・スペースが消費されます。ディスク・スペースまたはメモリーが不十分な場合にこのオプションを使用可能に設定すると、システムの全体的パフォーマンスが低下することがあります。

索引の並行作成数を増やした場合、キューに既に入っている索引作成要求は自動的に開始されません。変更内容は、この値の変更後にキューに入れられた新しい索引作成に影響します。

索引の並行作成数を減らした場合、現行の索引作成は自動的に停止されません。変更内容は、現行の索引作成が停止した後 (これにより、キューに入れられた索引作成が開始できるようになります) に有効になります。

## 手順

システムの索引作成オプションを指定するには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2.  「編集」をクリックして、システム編集ビューに切り替えます。
3. 「索引」ページで、「索引作成オプションの構成」をクリックします。
4. 「システム・レベルの索引付けオプション」ページで、システム・リソースを共用し、同時に索引作成要求を処理することが可能なコレクションの数を入力します。

索引作成用リソースを共用するコレクションの数は、エンタープライズ・サーチ・システム内のコレクション数より大きくすることはできません。例えば、コレクションが 5 個ある場合、5 以下の数を入力する必要があります。

5. 個々のコレクションの索引作成を同時に複数サポートできる十分なシステム・リソースがある場合、差分索引の作成と主索引の作成が同時に実行されるようにするオプションを選択できます。

**制約事項:** いずれかのコレクションの索引作成が進行中にこのオプションを選択すると、索引作成が失敗することがあります。差分および主索引の同時作成を使用可能にする前に、スケジュールされた索引作成をすべて使用不可にして、現行

の索引作成が停止するのを待ちます (または、可能な場合は完了する前に終了させます)。このオプションを使用可能にした後、使用不可にしたスケジュールされた索引作成を使用可能にします。

---

## 変更を検出した場合にのみ索引を作成

スケジュールされた索引作成の場合、索引に適用する必要のある変更が検出された場合にのみ作成が行われます。索引作成を開始して、システムで変更を検査するかどうかを選択するには、**startIndexBuild** コマンドを使用します。

### 始める前に

変更が検出された場合にのみ索引作成を行うことを手動で指定するには、エンタープライズ・サーチ管理者としてログインする必要があります。

### 制約事項

主索引が少なくとも 2 回作成されるまで、索引に影響することのある変更の検査は行われません。この実装は、検索結果の縮小についての有効範囲定義や規則の変更などの、主索引が少なくとも 2 回作成された後でないと検出できない変更に対処します。

### このタスクについて

索引作成では、CPU、メモリー、および I/O 帯域幅などのシステム・リソースを消費します。システム・リソースを節約し、パフォーマンスを向上するために、エンタープライズ・サーチ・システムは、索引に適用する必要のある変更が行われたかどうかを検出して、変更を適用する必要がある場合にのみ索引を作成します。次の索引変更イベントが、システムによって検出されます。

#### 主索引の作成のみ

- パターンによる URI の除去。索引から文書を除去します。
- システムからのクローラーの除去。索引から文書を除去します。

#### 主および差分索引の作成

- 検索結果の縮小についての有効範囲の定義または規則の変更。
- クローラー設定値またはプッシュ API による文書の直接の追加または除去。
- 索引内の文書の静的ランキング・スコアに影響する、ランキング調整要因の変更。
- 照会でワイルドカード文字を使用できる規則の変更。

スケジュールされた索引作成要求はすべて、変更検出の対象になります。索引作成は、索引に適用する必要のある保留中の変更がある場合にのみ行われます。変更が検出されなかった場合、索引作成要求は廃棄されます。

変更の発生に関係なく索引作成を強制するには、エンタープライズ・サーチ管理コンソールを使用して索引作成を開始します。また、**startIndexBuild** コマンドを使用し、手動で索引作成を開始して、要求の処理に進む前にシステムで変更を検査するかどうかを指定できます。



## 手順

コマンド行から索引作成を開始し、変更が検出された場合にのみ作成要求の処理に進むには、次のようにします。

1. エンタープライズ・サーチ管理者としてログインします。マルチサーバー構成では、索引サーバーにログインします。
2. 次のコマンドを入力します。

```
esadmin controller startIndexBuild -オプション
```

オプション:

**-cid** *collection\_ID*

索引作成を開始する対象のコレクションのコレクション ID。

**-buildType** *build\_type*

作成する索引のタイプを指定します。有効な値は、main および delta です。

**-detectChanges**

オプション。索引作成要求で、変更判別テストを行うように要求します。索引に適用する必要がある変更が検出された場合にのみ、索引作成の処理に進みます。

次の例では、col\_1 コレクションの主索引作成の開始要求は、索引に適用する必要がある変更が検出された場合にのみ処理されます。

```
esadmin controller startIndexBuild -cid col_1 -buildType main  
-detectChanges
```

---

## 索引作成の停止

エンタープライズ・サーチ管理コンソールの代わりに **stopIndex** コマンドを使用して、主索引作成および差分索引作成を停止することができます。

### 手順

管理コンソールの代わりにコマンドを使用して索引作成を停止するには、次のようにします。

1. エンタープライズ・サーチ管理者としてログインします。マルチサーバー構成では、索引サーバーにログインします。
2. 次のコマンドを入力します。

```
esadmin stopIndex -cid collection_id -buildType typeここで、それぞれ次のことを表します。
```

**-cid** コレクション ID

索引を所有するコレクションのコレクション ID を指定します。

*type*

停止する索引作成のタイプを指定します。許可される値は、main または delta です。

例

```
esadmin stopIndex -cid coll -buildType delta
```

---

## 索引の検索可能ビューに影響するオプション

文書の索引作成が終わった後、ユーザーが文書を検索したり検索結果内の文書を表示する方法を制御するオプションを指定できます。

索引に関するユーザーのビューを制御するオプションを指定するために、以下の作業を実行できます。

- 照会条件内でのワイルドカード文字のサポートを構成する。ワイルドカード照会のサポートを索引に構築するか、照会処理の間に照会条件を展開するオプションを指定できます。
- ユーザーが検索できる文書の範囲を制限する有効範囲を構成する。ユーザーがコレクションを検索するとき、ユーザーは索引全体ではなく、有効範囲にある文書のみを検索します。
- 検索結果で同じソースからの文書を縮小表示する。索引の URI あるいは URI パターンと一致する文書をグループ化して、検索結果に先頭の結果文書のみを表示することができます (ユーザーは、結果文書を縮小表示するオプションを指定できます)。
- 索引から URI を除去する。一時的に、ユーザーが索引の特定文書を検索できないようにする必要がある場合があります。

一部のタイプのクローラー、およびセキュリティーが使用不可のコレクションの場合、検索結果で同じかほぼ同じ文書を複数表示しないように、重複文書検出が使用されます。

## 文書の検索に対する索引付きオプション

クロールされたデータを検索するオプションを構成するとき、または、XML および HTML メタデータ・エレメントを検索フィールドにマップするときに、文書がどのように検索され、検索結果に表示されるのかを指定します。

指定する検索オプションは、文書と共に索引に保管されます。それらのオプションを使用して、検索結果を照会できるユーザーおよび検索結果を表示できるユーザーを制限することができます。

### クローラー・オプション:

フィールドを含んでいるデータ・ソースをクロールするクローラーを構成するとき、フィールドが検索可能かどうか、どのように検索できるのか、検索結果に戻されるかどうかを、以下のオプションを指定することによって制御できます。

- フリー・テキスト検索
- フィールド検索
- 完全一致
- ソート可能
- パラメトリック検索
- 検索結果

- 文書コンテンツ

#### XML および HTML フィールド・マッピング・オプション:

パーサーを構成し、XML エlementおよび HTML メタデータ・Elementを索引内の検索可能フィールドにマップすることを指定する場合、以下のオプションを指定します。

- フィールド検索
- 完全一致
- ソート可能
- 検索結果

すべてのElementや Dublin Core メタデータ・Element・セット内のElementではなく、特定の HTML メタデータ・Elementに対して検索オプションを構成する場合は、数値を含むフィールドをパラメトリック照会で検索するように指定することもできます。

### フリー・テキスト検索

エンタープライズ・サーチ索引は、さまざまなデータ・ソースからのコンテンツで構成されるフルテキスト索引です。単純な照会を自然言語で指定することによって、コンテンツを検索できます。検索プロセスは、フィールドおよび文書コンテンツを検索し、照会に関連する文書を検出します。

フリー・テキスト照会でフィールドを検索できるようにするには、クローラーの構成時に「フリー・テキスト検索」チェック・ボックスを選択します。タイトル、キーワード、および説明フィールドを検索するには、このチェック・ボックスと「フィールド検索」チェック・ボックスを選択します。

#### 例 1:

フリー・テキスト検索は、次の例に示すように単純な照会です。

bicycle chain

文書中に含まれるべき語、含まれてはならない語を示したい場合、特殊な表記を入れればそのような指定が可能です。例えば、正符号 (+) を単語の前に置くと、文書にその語と一致する語がなければならないことを指定します。また、負符号 (-) を単語の前に置くと、その語を含む文書を検索結果から除外します。複数の語を引用符 (") で囲むと、その句全体に正確に一致するものを対象として検索が行われます。

#### 例 2:

次のフリー・テキスト照会では、検索一致となるのは文書に「science fiction」に完全一致する句が含まれ、「robot」の語を含まない場合に限られます。

+ "science fiction" -robot

## フィールド検索

フィールド検索を使用すると、照会対象を文書内の特定のデータ・フィールドおよびメタデータ・フィールドに制限することができます。例えば、文書のタイトルに特定の語が含まなければならないといった指定をすることができます。

フィールド名でフィールドを検索できるようにするには、クローラーの構成時、または XML および HTML エLEMENTのフィールド・マッピング・オプションの構成時に、「**フィールド検索**」チェック・ボックスを選択します。タイトル、キーワード、および説明フィールドを検索するには、このチェック・ボックスと「**フリー・テキスト検索**」チェック・ボックスを選択します。

### 例:

エンタープライズ・サーチでフィールド検索を指定するには、照会にフィールド名とそのフィールドになければならない語句とを指定します。

次の照会では、タイトル・フィールドに「ibm」という語と「enterprise search」という句を含む文書を検索します。

```
title:ibm title:"enterprise search"
```

## 完全一致

完全一致検索では、正確な照会を指定可能にすることで、検索結果の品質を高めることができます。完全一致検索を使用すると、フィールドと XML エLEMENTを照会してフィールド値全体または XML エLEMENT値全体が照会条件に一致する文書のみをリトリブすることができます。フィールドの内容が検索条件より少なかったり多かったりした場合は、一致にはなりません。

**ヒント:** 完全一致が存在するかどうかを判別する際、システムは照会条件を小文字に変換し、照会ストリング内の余分なスペースを除去して、ワイルドカード文字パターン・マッチングを実行します。ただし、レンマタイゼーションや同義語検索は行われず、ストップワードは除去されません。XML エLEMENTの完全一致では、ネストされたELEMENTを含まないELEMENT名と、ELEMENTの値全体が、照会条件に正確に一致することが必要です。

完全一致のフィールドは、エンタープライズ・サーチ (SIAPI) 照会構文または XMLFrag2 照会構文 (XPath 照会をサポートされません) を使用して検索できます。照会条件の前に等号 (=) が付くと、完全一致検索が行われることを示します。

完全一致でフィールドを検索できるようにするには、クローラーの構成時、または XML および HTML エLEMENTのフィールド・マッピング・オプションの構成時に、「**完全一致**」チェック・ボックスを選択します。コレクションの解析オプションの構成時に、ネイティブ XML 検索による XML 文書検索を使用可能にした場合は、すべての XML エLEMENTを検索語の完全一致で検索できます。

### 例 1:

color という名前のフィールドに、値 dark blue が含まれます。

- 次の完全一致照会は、照会に他の語が含まれないため、一致となります。

```
color:"dark blue"
```

- 次の完全一致照会は、dark blue の他に skirt という語が照会に含まれるため、一致となりません。

```
color:"dark blue skirt"
```

- 次の完全一致照会は、color フィールドに dark という語も含まれるため、一致となりません。

```
color:=blue
```

## 例 2:

完全一致でない次の XMLFrag2 照会では、<diagnosis>intraductal carcinoma comedo type</diagnosis> と示された文書や、照会条件に正確に一致しない他の語を示された文書も戻される可能性があります。

```
@xmlf2::'<diagnosis>intraductal carcinoma</diagnosis>'
```

完全一致である次の XMLFrag2 照会では、XML エlement 値の内容全体が照会条件に一致する文書のみが戻されるようになります。

```
@xmlf2::'=<diagnosis>intraductal carcinoma</diagnosis>'
```

## ソート可能

データ・ソースがフィールドを含む場合や、ユーザーが XML または HTML 文書を検索する場合に、特定フィールドの値で結果をソートできるようにしたいことがあります。エンタープライズ・サーチのサンプル検索アプリケーションでは、ソート可能フィールドとして構成されたすべてのフィールドの名前がリストされます。ユーザーは、結果に関連性や文書日付でソートするのではなく、リストされたフィールドの 1 つを選択して、結果を英字順に (ストリング・ソートに従う) ソートするように選択できます。また、文書のソートを昇順にするか、降順にするかも選択できます。

ソート・フィールドを含まない結果文書は、検索結果の最後にリストされます。ソート・フィールドを含むが、そのフィールドがソート可能と構成される前に索引付けされた結果文書も、検索結果の最後にリストされます。

ユーザーが検索結果をフィールドの値で英字順にソートできるようにするには、クローラーの構成時、または XML および HTML エlement のフィールド・マッピング・オプションの構成時に、「ソート可能」チェック・ボックスを選択します。フィールドに数値を含む場合は、「パラメトリック検索」チェック・ボックスを選択し、検索結果の数値でのソートにそのフィールドの値を使用できることを指定します。

## パラメトリック検索

パラメトリック検索は、フィールド検索の一種で、数値フィールド、日付フィールド、およびメタデータに対して比較、評価を行う照会を指定できます。例えば、一定サイズの文書、または一定の日付以降に書かれた文書などを検索することができます。また、属性値が指定した値より大きい、小さい、または等しいなどの条件に当てはまる文書を検索することもできます。

結果をフィールドの値に従って数値でソートするには、そのフィールドをパラメトリック検索に使用可能に設定する必要があります。

パラメトリック照会でフィールドを検索するか、または結果を数値でソートできるようにするには、クローラーの構成時、または特定の HTML メタデータ・エレメントのフィールド・マッピング・オプションの構成時に、「**パラメトリック検索**」チェック・ボックスを選択します。

#### 例 1:

次の照会では、価格がちょうど 50 ドル (ないし price フィールドに指標付けられている任意の通貨単位) の品目を検索します。

```
#price::=50
```

#### 例 2:

次の照会では、ファイル・サイズが 1024 より大きく 2048 以下の文書を検索します。

```
#filesize::>1024<=2048
```

## 検索結果

検索対象にはするが検索結果には表示させたくないフィールドや、照会を行わないが検索結果には表示させたいフィールドがある場合があります。例として、財務データを照会して意味のあるレポートを取得する必要があるが、結果には従業員の給与と名前を表示する必要がある場合などが考えられます。

フィールドを検索結果に表示できるようにするには、クローラーの構成時、または XML および HTML エレメントのフィールド・マッピング・オプションの構成時に、「**検索結果**」チェック・ボックスを選択します。

## 文書コンテンツ

Web 文書など、一定の文書タイプでは、文書全体がコンテンツと見なされます。フィールドを含む文書などの他の文書タイプでは、メタデータではなく、有用なコンテンツをどのフィールドが含むかを指定できます。

フィールドが文書コンテンツを構成することを指定するには、クローラーの構成時に「**文書コンテンツ**」チェック・ボックスを選択します。「**文書コンテンツ**」と「**フリー・テキスト検索**」の両方のチェック・ボックスが選択された場合、フィールドの値は、重複文書の検出に使用され、検索結果の動的文書サマリー領域の一部になります。

### 関連概念

 [照会構文](#)

## 重複文書の検出

重複文書の検出は、同じ内容やほぼ同じ内容を持つ複数の文書が検索結果に含まれないようにする技法です。

同じ (またはほぼ同じ) 文書のコピーが検索結果に複数リストされると、検索の質が低下することがあります。重複文書の分析は、次の条件が両方とも満たされた場合にのみ行われます。

- コレクションがリンク・ベースのランキング・モデルを使用している。このモデルは、Web クローラーや WebSphere Portal クローラーなど、Web サイトをクロールするクローラーに適用されます。
- コレクション・セキュリティが無効になっている。

グローバル分析では、索引付けプロセスで、各文書の文書コンテンツをスキャンして、重複を検出します。2つの文書が同じ文書コンテンツであると、これらの文書は重複文書と見なされます。

重複検出の分析時に文書メタデータも考慮したい場合は、コレクション用にクローラーを構成するときに「**文書コンテンツ**」チェック・ボックスを選択し、メタデータをクロールするオプションを指定する必要があります。この場合、クローラーはメタデータ・フィールドを文書コンテンツとしてクロールし、重複文書の内容を分析するときにメタデータを含めます。同様の分析は、HTML および XML 文書を解析するオプションを構成し、「**文書コンテンツ**」チェック・ボックスを選択したときにも行われます。

フィールドまたはメタデータ・フィールドが文書コンテンツになるように指定したときは、それらのフィールドの内容が検索結果で文書の動的サマリーに追加されます。このことは、文書を検索結果に表示するかどうかに影響を及ぼすことがあります。検索アプリケーションで近似重複の検出が有効になっている (`setProperty` メソッドの `NearDuplicateDetection` プロパティが `Yes` に設定されている) 場合は、ユーザーが検索結果を表示すると、類似したタイトルやサマリーを持つ文書は抑制されます。ユーザーは、抑制された近似重複文書を表示するリンクをクリックすることができます。

重複文書のグループでは、1つの文書がマスターで、他は重複となります。重複文書のグループに含まれる文書は、すべて同じ正規表現の内容を持っています。索引付けでは、マスター文書の内容 (トークン) が索引付けされます。重複文書の場合は、メタデータ・トークンのみが索引付けされます。マスター文書が索引から削除されると、次の重複文書がマスター文書になります。ユーザーがコレクションを検索したときは、マスター文書のみが返されます。

#### 関連概念

290 ページの『重複文書の分析とコレクション・セキュリティ』

## 照会内のワイルドカード文字

ユーザーが照会条件にワイルドカード文字を含め、指定のパターンと一致するワードを検索できるようにすることができます。

ワイルドカード照会条件は、アスタリスク (\*) が含まれている用語です。ワイルドカード文字が含まれている照会をユーザーが実行する場合、検索結果には、その照会条件に一致する索引内のすべての文書のほか、ワイルドカード文字で表されるパターンに一致する索引内のすべての文書が含まれます。例えば、照会条件 `sea*` の末尾ワイルドカードは、`search`、`season`、および `seals` と一致が可能です。

索引に対してワイルドカード文字オプションを構成する場合、ユーザーが照会でワイルドカード文字を指定できるようにするかどうか、また、その場合、どのようにこのサポートが提供されるかを選択します。

- 文書のすべての部分に対して、ワイルドカード文字パターンを突き合わせた、ワードでの検索を使用可能にできたり、あるいはフィールドに対するパターン・マッチングを制限できたりします。
- すべてのフィールドがワイルドカード文字を含む照会をサポートすることができるようにしたり、あるいは指定したフィールドに対するパターン・マッチングを制限できたりします。
- 照会条件の最終文字にワイルドカード文字を制限できたり (末尾ワイルドカード文字)、あるいはワイルドカード文字を照会条件のいずれの場所にも出現させることができたりします。(ワイルドカード文字は、フィールド名に出現させることはできません。)
- ワイルドカード文字が出現することを、どこに許可するかによって、どのように照会条件を展開するかを選択できます (ワイルドカード文字を含む照会条件は、マッチングする索引の用語のすべてに対して展開されます)。索引は、すべての可能な用語の展開を保管できたり、あるいは検索プロセスは、照会処理中に用語を展開できたりします。

ワイルドカード文字の設定に対する変更は、次回の主索引の作成後に有効になります。

## 索引展開

索引内の用語の展開を含めるには、出現するマッチングに対する照会条件のワイルドカード文字パターンに、ワード内のいくつかの先行文字がマッチングするかを指定します。少なくともこの数の文字をもつ照会条件のみ (\* を除く)、結果を返します。例えば、4 を指定すると、照会条件は、マッチングとして出現するために、最小で 4 文字を指定する必要があります。

4 を指定すると、ワード `technology` は、照会条件の `tech*` および照会条件 `techno*` とマッチングしますが、照会条件 `te*` とはマッチングしません。

差分索引または主索引の作成では、元の用語に加えて、文書内の各用語の考えられるすべての展開が索引付けされます。この方法は、照会処理中に用語を展開するのに追加の時間は必要でないという利点があります。ただし、この方法の場合、索引のサイズが大きくなります。つまり、大きい索引に対応できるような十分なシステム・リソースが使用可能でなければなりません。

これは、コレクションが相対的に小さい場合、または索引を作成するためのスペースと時間よりも照会応答時間を重視する場合に、とても役に立つ方法です。例えば、カタログまたは従業員ディレクトリーを検索するのに、この方法を選択できます。

このアプローチは、末尾ワイルドカード文字のサポートを使用可能にした場合にのみ有効です。照会条件にいずれの場所にも出現するワイルドカード文字のサポートを使用可能にした場合、索引内に用語の展開を含めるオプションを選択することはできません。



## 照会展開

ワイルドカード文字が含まれている照会をユーザーが実行するときに照会を展開し、パターン・マッチング規則を適用するには、一致を構成する照会条件のバリエーションの数を指定します。例えば、50 を指定した場合、照会条件の最大 50 のバリエーションを、その照会条件の一致とみなすことができます。

この例を示すと、照会条件 `tech*` はワード `technical`、`technique`、`technology` と、さらに、文字 `tech` で始まる最大 50 個の異なるワードに一致します。

照会展開により、索引のサイズはあまり影響を受けませんが、照会パフォーマンスは低下する可能性があります。検索プロセスは、ワイルドカード文字設定に指定した限度まで、ワイルドカード照会条件の考えられるすべての展開に対して繰り返す必要があります。

この方法は、コレクションが相対的に大きく、しかも、索引を作成するためのスペースと時間が最小でなければならない場合に、とても役立ちます。例えば、E メール・リポジトリ用にこの方法を選択します。その場合、索引は、急速に変化する文書に応じて増加する必要がありますが、照会応答時間はあまり重要ではありません。

このアプローチは、末尾ワイルドカードのサポートを使用可能にしているか、または照会条件のいずれの場所にも出現するワイルドカード文字のサポートを使用可能にしているかどうかにかかわらず有効です。

## 照会でのワイルドカード文字のサポート

ワイルドカード照会条件の展開のセットには、ワイルドカード文字を任意の文字のシーケンスと置換することによって取得できる索引内のすべての用語が含まれます。そのセットは、次のように判断されます。

- コレクションが、照会条件のいずれの場所でも出現できるワイルドカード文字をサポートする場合、アスタリスクを含むいずれの照会条件も、ワイルドカード用語と解釈されます。
- そのセットには、最大限、エンタープライズ・サーチ管理者が構成する展開の最大数が含まれます。索引に、この展開数を超過が含まれている場合、無視されます。(検索結果は、ワイルドカード展開が無視されたかどうかを示します。)
- ワイルドカード文字サポートが、あるセットのフィールドに制限される場合、そのセットには、指定されたフィールドの 1 つに表示される用語のみが含まれます。用語は、少なくとも、索引内の 1 つの文書の 1 つのフィールドにのみ表示される必要があります。
- 照会条件が「フィールド用語」である場合、ワイルドカード文字は、フィールド指定子の後に表示される必要があります (例えば、`fieldname:*sphere`)。フィールド名にコロン (:) を含めることはできません。
- ワイルドカード文字サポートが、あるセットのフィールドに制限される場合、ワイルドカード照会条件のフィールド名は、エンタープライズ・サーチ管理者コンソールで指定されたフィールドの 1 つである必要があります。そうでないと、用語に対する展開を 1 つも検出しません。

- ワイルドカード文字は、プレーン・テキスト用語でのみサポートされ、XML エlement名、属性名、または属性値ではサポートされません。ワイルドカード文字でのみ構成される用語はサポートされません。

## ワイルドカード文字の索引への影響

索引展開に基づくワイルドカード文字をサポートすると、索引のサイズや索引作成の時間が増大します。

索引展開により、用語のすべての接頭部と、その用語自体が索引付けされます。例えば、用語 `support` に対して、以下の用語が索引付けされます。

```
s su sup supp suppo suppor support
```

索引に保管される用語の数は、語の平均長の因子により増大します。索引圧縮では、索引のサイズが削減されますが、著しくは減りません。索引作成に必要な時間は、語の平均長に従って増えます。

英語の単語の平均長は 5 文字から 6 文字であるため、英語の文書の索引は、約 4 という係数で増大します。N-gram トークン化による索引は、各 N-gram が 2 文字を含むため、約 2 倍に増大します。

索引展開は、以下の状況でワイルドカード文字サポートに推奨されます。

- コレクションが十分に小さく、展開された索引で使用されるスペースと時間でパフォーマンス上の問題が生じない。
- ユーザー (またはエンタープライズ) の要件を満たすために、ワイルドカード文字のすべての可能な展開を検索結果に含める。

ワイルドカード展開における接頭部の最小長を指定することで、索引への影響を緩和することができます。例えば、接頭部の最小長が 3 に設定されると、`support` という語に対して、`s` と `su` という接頭部は索引付けされません。そのため、英語の索引は、係数 4 ではなく、係数 3 で増大することになります。

ワイルドカード文字サポートの照会展開の方式では、索引に接頭部は書き込まれません。用語は照会の実行時に展開され、索引は、その展開のサポートに必要な小さなデータ構造の分だけ増大します。通常、照会展開による索引は、ワイルドカード文字サポートなしの索引より、10 % から 20 % 大きくなり、索引作成に追加でかかる時間は 10 % 未満です。展開の最大数の構成は、索引のサイズや索引作成時間に影響しません。

## 照会でのワイルドカード文字のオプションの構成

エンタープライズ・サーチ・コレクションの索引作成オプションを構成するとき、ユーザーが照会条件にワイルドカード文字を含められるようにしたいかどうかを指定できます。

### 始める前に

ワイルドカード文字のオプションを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、その索引が所属するコレクションのコレクション管理者である必要があります。

### このタスクについて

ワイルドカード文字のオプションを指定した場合、変更内容は、次回、主索引作成が行われると有効になります。

## 手順

照会でのワイルドカード文字のサポートを構成するには、次のようにします。

1. コレクションを編集、「索引」ページを選択、そして「ワイルドカード文字のオプションの構成」をクリックします。
2. 「ワイルドカード文字のオプションの構成」ページで、「照会でのワイルドカード文字のサポート」チェック・ボックスを選択します。
3. オプション: フリー・テキストを検索する照会でワイルドカード文字をサポートするように指定することができます。例えば、名前付きフィールドを検索しないフリー・テキスト照会 `tech*` は、このチェック・ボックスが選択されている場合のみ、展開された結果（「`technology`」または「`technique`」など）を戻します。
4. ワイルドカード文字をサポートするフィールドの指定:
  - フィールドを検索する照会でワイルドカード文字を処理しないことを指定するには、「フィールドなし」を選択します。
  - ワイルドカード文字を含む照会をサポートする文書のフィールドすべてを使用可能にするには、「すべてのフィールド」を選択します。
  - ワイルドカード文字のサポートを幾つかのフィールドに制限するには、「特定のフィールド」を選択して、フィールド名を入力します。展開された結果は、指定したフィールドに対してのみ戻されます。例えば、照会 `author:john*` は、「`author`」フィールドがワイルドカード文字をサポートするように指定した場合にのみ、展開された結果を戻します。
5. ワイルドカード文字が照会条件の最終位置に出現しなければならないかどうか（末尾ワイルドカード）、またはワイルドカード文字が非限定であり、照会条件のいずれの場所にも出現できるかどうかを指定します。

ワイルドカードの位置とタイプを選択する場合、どのようにワイルドカード文字のサポートを使用可能にしたいかをもまた指定する必要があります。詳細については、管理コンソールで「ヘルプ」をクリックしてください。

## 有効範囲

ユーザーにコレクションの限定ビューを提供するには、有効範囲を構成します。

有効範囲とは、索引内の関連した URI のグループのことです。有効範囲を構成することで、コレクションの中のユーザーが見ることができる文書を制限します。ユーザーがコレクションを検索するとき、ユーザーは索引全体ではなく、有効範囲内の文書のみを検索します。この機能を使用するには、検索アプリケーションで有効範囲の検索をサポートしている必要があります。

有効範囲を作成するとき、ユーザーが検索できる索引中の URI の範囲を指定します。ユーザーが検索できる文書を制限することで、検索結果内の文書をユーザーが探している情報に特定できることとなります。

例えば、技術サポート部門の URI を含んだ有効範囲と、人事管理部門の URI を含む別の有効範囲を作成することができます。検索アプリケーションが有効範囲をサ

ポートする場合、技術サポート部門のユーザーは、技術サポート部門有効範囲から文書をリトリートし、人事管理部門のユーザーは、人事管理部門有効範囲から文書をリトリートします。

有効範囲は必要なだけ作成できますが、多すぎるとパフォーマンスに影響を及ぼす場合があります。ほとんどの検索要求が、1 つまたは 2 つの有効範囲のみをフィルタリングする必要があるように有効範囲を構成してください。有効範囲には URI 全体または URI パターンを含めることができるので、同じ文書が複数の有効範囲に属することがあります。

有効範囲の構成時、変更を有効にするために、主索引の作成を 2 回行わなければならない場合があります。コレクションの最初の主索引が作成される前に有効範囲を構成すると、ユーザーはコレクションを検索できても、検索結果に有効範囲データは表示されません。検索結果が URI の有効範囲を確実に反映するようにするため、主索引をもう一度作成してください。

主索引が作成された後に有効範囲を構成した場合、次の主索引作成が行われたときに、その変更内容が有効になります。

## 有効範囲の構成

エンタープライズ・サーチ・コレクションの有効範囲を構成するときに、ユーザーが検索を許可される索引内の文書の範囲に合うように URI または URI パターンを指定します。

### 始める前に

有効範囲を構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、その有効範囲が所属するコレクションのコレクション管理者である必要があります。

### このタスクについて

検索アプリケーションで有効範囲のサポートが使用可能に設定されている場合、ユーザーはコレクションを検索する際に、有効範囲の境界を定義する URI に一致する文書のみを検索できます。

有効範囲の構成時、変更を有効にするために、主索引の作成を 2 回行わなければならない場合があります。最初の索引作成が行われる前に有効範囲を構成すると、ユーザーはコレクションを検索できても、検索結果に有効範囲データは表示されません。検索結果が URI の有効範囲を確実に反映するようにするため、主索引をもう一度作成してください。

主索引が作成された後に有効範囲を構成した場合、次の主索引作成が行われたときに、その変更内容が有効になります。

### 手順

有効範囲を構成するには、次のようにします。

1. コレクションを編集、「索引」ページを選択、そして「有効範囲の構成」をクリックします。

2. 「有効範囲」ページで、「有効範囲の作成」をクリックします。
3. 有効範囲の名前と、有効範囲の境界を定義する URI および URI パターンを指定します。有効範囲から除外したい URI および URI パターンを指定することもできます。
4. 「OK」をクリックします。

新規有効範囲が、このコレクションに属している他の有効範囲と共に「有効範囲」ページにリストされます。

#### 関連資料

135 ページの『エンタープライズ・サーチ索引における URI フォーマット』

## 縮小表示した URI

エンタープライズ・サーチは、ソースからの文書のうち URI 接頭部が同じものが検索結果に縮小表示されるように、検索結果を編成できます。

結果が縮小表示される場合、先頭の結果は、通常、左揃えで表示されます。それより下のランキングの結果は、グループ化され、先頭の結果より字下げされて表示されます。

URI 接頭部が異なる結果文書を単一のグループとして縮小表示するには、作成する 1 つのグループ名に、それらの接頭部を関連付けます。例えば、財務データを管理するサーバーが 3 つある場合、3 つのサーバーすべてからの文書を検索結果においてグループ化し、先頭の結果文書の下に低いランキングの結果を縮小表示できます。

検索アプリケーションは、URI 接頭部あるいはグループ名を使用して、検索結果で文書を縮小表示できます。エンタープライズ・サーチのサンプル検索アプリケーションでは、先頭の 2 つの検索結果文書が表示されます。同じ URI 接頭部を持つ (あるいは、同じ URI グループに属す) 3 つ以上の結果文書が返された場合、結果を縮小表示させるオプションを選択できます。

ユーザーは、エンタープライズ・サーチ照会構文 (`samegroupas:URI prefix`) を使用して、照会で指定された URI 接頭部と同じグループ内のすべての文書を検索できます。

### URI 接頭部とグループ名の編成方法

管理コンソールを使用して検索結果の縮小表示規則を構成するときに、縮小表示する文書の URI 接頭部を指定します。また、オプションで、それらの URI 接頭部を 1 つのグループ名に関連付けることができます。

設定する URI 接頭部の順序は重要です。索引サーバーは、コレクションの中の各 URI の値を計算するときに、URI 接頭部の順序を使用します。それぞれの URI ごとに以下ようになります。

1. 索引サーバーは、検索結果の縮小表示規則にある URI 接頭部を順番にスキャンします。
2. 索引サーバーは、索引内の文書と接頭部が一致する最初の URI 接頭部を検出すると、グループ名 (あるいは、規則でグループ名を指定していない場合は、URI 接頭部) をその文書の追加の検索語として関連付けます。

Web 文書が URI 接頭部と一致しない場合、索引サーバーは、URL のホスト名を URI 接頭部として使用します。 NNTP 文書が URI 接頭部と一致しない場合、索引サーバーは、参照ヘッダーの値にある最初のメッセージ ID を URI 接頭部として使用します。

検索結果で縮小表示する URI 接頭部のリストに URI 接頭部を追加した後、索引サーバーがその URI 接頭部をスキャンして、追加の検索語として索引内の文書に関連付ける際の望ましい順序にそれを配置してください。

- URI 接頭部を追加しても、それをグループ名に関連付けないときは、個別の URI 接頭部を選択して、リスト内でそれを上下に移動させます。
- URI 接頭部を追加してそれをグループ名に関連付ける場合に、リスト内で URI 接頭部を上下に移動させるときはいつでも、同じグループに属している URI 接頭部のグループ全体を移動させます。グループ内の URI 接頭部の順序は問題ではありません。個々の URI 接頭部を選択すれば、自動的にグループ全体が選択されます。

## 検索結果内の URI の縮小表示

同じ URI 接頭部を持つソースからの検索結果文書をグループ化して縮小表示するオプションを指定することができます。異なる URI 接頭部をもつ結果文書をまとめて縮小表示できるようにするグループ名も作成できます。

### 始める前に

検索結果を縮小表示するオプションを指定するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

検索結果の縮小表示について行った変更は、次に主索引作成が行われるまで、有効になりません。



### 手順

検索結果を縮小表示するためのオプションを指定するには、次のようにします。

1. コレクションを編集し、「索引」ページを選択して「**検索結果の縮小**」をクリックします。
2. 「検索結果の縮小」ページで、「**URI 接頭部の追加**」をクリックします。
3. 「結果を縮小するための URI 接頭部の追加」ページで、検索結果で縮小表示したい文書の URI 接頭部を入力します。次に例を示します。

```
http://finance/ROI/  
http://server1.com/finance/  
db2://LOCALDB/SCHEMA1.TABLE1/  
exchange://exchangesvr.ibm.com/public/TeamRoom/Folder1/
```

4. この URI 接頭部と関連付けたい記述グループ名を入力できます。いくつかのソースからの結果文書を単一グループとして縮小表示するために、各 URI 接頭部を追加するときに同じグループ名を入力してください。
5. 「**OK**」をクリックします。

6. 「検索結果の縮小」ページで、新しい規則を、索引サーバーにスキャンさせる順序で並べます。
  - URI 接頭部を追加し、それをグループ名と関連付けていない場合、新しい URI 接頭部は、リストの一番下に示されます。矢印キーを使用して、新しい接頭部を正しい位置に移動します。
  - 新しい URI 接頭部をグループ名と関連付けた場合、その URI 接頭部は、同じグループに所属する URI 接頭部のセットの一番下に示されます。矢印キーを使用して、URI 接頭部のグループ全体を正しい位置に移動します。
7. URI 接頭部またはグループ名を変更するには、目的の URI 接頭部を選択して、「 編集」をクリックします。
8. リストから URI 接頭部を除去するには、目的の URI 接頭部を選択して、「 除去」をクリックします。

## 索引から URI を除去

コレクション内の文書をユーザーが検索できないようにするため、それらの文書の URI を索引から除去することができます。

### 始める前に

索引から URI を除去するには、エンタープライズ・サーチ管理者役割のメンバーであるか、コレクションのコレクション管理者である必要があります。

### このタスクについて

完全修飾 URI を指定すると、その URI はユーザーの検索結果に現れなくなります。ただし、ユーザーが同じ照会を実行し、その照会の結果文書が検索キャッシュに残っている場合には、除去した URI のキャッシュされた結果ページが検索結果に戻され続けます。次回、主索引または差分索引の作成が行われるまで、検索キャッシュは更新されず、URI は索引から除去されません。

複数の URI を除去するための URI パターンを指定した場合、次回、主索引の作成が行われるまで、指定されたパターンに一致する URI がユーザーの検索結果に戻され続けます。

索引から URI を除去しても、その URI はクロール・スペースから除去されるわけではありません。クローラーが次回にその文書をクロールすると、その URI が索引内に作成され、再び検索に使用可能になります。URI をクロール・スペースから除去するには、該当する文書を除外するようにクロール規則を更新し、クローラーをいったん停止して、再始動する必要があります。

### 手順

特定の文書の URI を索引から除去するには、次のようにします。

1. コレクションを編集、「索引」ページを選択、そして「索引から URI を除去」をクリックします。
2. 「索引から URI を除去」ページで、索引から除去したい URI (または URI パターン) を入力します。

例:

http://domain.org/hr/\*  
db2://knowledgeManagement/ROI\*  
cm://enterprise/finance\*

#### 関連資料

135 ページの『エンタープライズ・サーチ索引における URI フォーマット』



---

## 検索サーバーの管理

検索サーバーに指定できるオプションとして、検索結果を戻すキャッシュ・スペースを使用する、検索結果に表示する文書サマリーの最大長を制御する、検索品質向上のためにカスタム辞書を関連付ける、照会に特定の用語が使用されたときに事前定義しておいた URI を検索結果に戻す、などがあります。

ユーザーが照会を実行すると、検索サーバーは索引を使用して、該当する文書を素早く探し出します。検索サーバーは、構文解析され、トークン化されたデータが保管されているエンタープライズ・サーチ・データ・ストアを使用して、該当する文書のメタデータをリトリブします。メタデータには、文書の URI、タイトル、説明、日付、データ・タイプなど (さらに多くのデータを格納することもできます) を入れることができます。

コレクションに対して検索サーバーを設定するときに、照会の処理方法に影響するオプションを指定します。これには、照会のパフォーマンスを左右する可能性のあるオプションも含まれます。

### 検索キャッシュの構成

照会のパフォーマンスを最適化するために、検索結果 (照会の応答) をキャッシュに保管するように指定できます。また、キャッシュに入れる検索結果に割り当てるスペース量を設定できます。

### 文書サマリーの表示最大長の設定

結果文書のほとんどで、ユーザーにとってその文書がリトリブしたいものかどうかを見極められるように文書コンテンツのサマリーを表示します。この要約情報を表示するために、検索結果上で使用するスペースを指定することができます。

### 異なるデフォルト言語の指定

コレクションの文書を検索するためのデフォルト言語は、コレクション作成時に指定されていますが、必要に応じて別の言語を指定することができます。

### カスタム辞書の関連付け

アプリケーション開発者によって、同義語、ストップワード、ランキング調整ワードに関するカスタム辞書が作成されている場合、コレクションの検索時にその辞書を指定することができます。

### クイック・リンクの構成

一定のキーワードおよび句に対して戻す URI を事前設定できます。ユーザーが照会でキーワードまたは句を指定すると、事前定義された URI が検索結果と共に戻されます。クイック・リンク URI は、検索サーバーが索引検索によって戻す URI に追加される形で戻されます。

### 関連概念

229 ページの『文書ランキング』

233 ページの『カスタム・ランキング調整ワード辞書』

---

## 検索キャッシュ

検索サーバーの負荷が比較的高い場合、検索結果をキャッシュに入れることでパフォーマンスが向上します。

検索サーバーが検索要求を処理する場合、まず、同じ照会の結果がキャッシュに既に存在していないかチェックされます。検索サーバーが該当する照会応答を検出できれば、検索結果をそのまますぐにユーザーに戻すことができます。該当の照会応答が見つからないと、検索サーバーは索引を検索します。

検索キャッシュが満杯になると、最も古い検索結果および頻度の低い照会の結果が、新たな検索結果用の場所を空けるためにキャッシュから出されます。

エンタープライズ・サーチ管理コンソールから、検索キャッシュを使用可能に設定し、また、キャッシュの容量 (同時にキャッシュに入れることができる照会応答の数) を指定できます。

検索キャッシュの設定を変更した場合は、検索サーバーを再始動して、その変更を有効にする必要があります。

---

## 検索キャッシュの構成

コレクションの検索キャッシュを使用可能または使用不可にできます。また、検索キャッシュのサイズを制御するオプションも指定できます。

### 始める前に

コレクションの検索キャッシュを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順

検索キャッシュを構成するには、次のようにします。

1. コレクションを編集して「検索」ページを選択し、「**検索サーバー・オプションの構成**」をクリックします。
2. 「検索サーバー・オプション」ページで、「**検索キャッシュを使用**」チェックボックスを選択します。
3. 「**キャッシュ項目の最大数**」フィールドに、検索キャッシュが保持できる照会応答の最大数を入力します。
4. 「**OK**」をクリックします。
5. 変更を有効にするために、検索サーバーをモニターし、サーバー・プロセスを再始動します。

---

## カスタム同義語辞書

検索結果の質を高めるために、ユーザーがコレクションを検索するときに、照会条件の同義語を検索するようにできます。

同義語辞書を作成し、それをエンタープライズ・サーチ・システムに追加して、コレクションに関連付けると、ユーザーは、そのコレクションを検索するときに、照会条件の同義語を含む文書を検索できます。このように照会を拡張することで、ユーザーは、正確に照会条件と一致する文書だけでなく、興味のある文書をすべて検出できる可能性が高くなります。同義語辞書を作成するときに、どのワードとどのワードが互いに同義語であるかを定義するので、ユーザーがあらゆるバリエーションを考慮して照会条件を指定しなくても、関係のある文書を見つけられるようになります。

例えば、ユーザーの組織で、部門や設備などを頭字語や省略語を使用して参照している場合もありますし、ユーザーのコレクション内の文書に、業界固有の用語が含まれている場合もあります。同義語辞書を作成すれば、頭字語 (例えば ACL など) を含む照会で、その頭字語を拡張したもの (例えば、ACLs、アクセス制御リスト (access control lists)、アクセス制御 (access controls) など) を扱う文書を返すことができます。

エンタープライズ・サーチ照会言語では、照会条件の前にチルド演算子を付加して同義語を指定することをサポートしています。例えば、`~WAS` という照会では、WebSphere Application Server を扱う文書を返すことができます。また、アプリケーション開発者は、照会のプロパティによって同義語サポートを使用可能にでき、これには特別な構文は必要ありません。

同義語辞書には、ワードのさまざまな変形が含まれ、以下のような特性があります。

- ワードは、言語に固有なものではなく、異なる言語で使用されることがある。コレクションごとに 1 つだけ同義語辞書があります。
- ワードは、語尾変化をしない。あらゆるすべての語尾変化は、同義語リストに追加する必要があります。例えば、ワードの単数形と複数形の語尾変化の場合もあります (ACL と ACLs など)。

同義語辞書に追加する用語のほとんどは、厳密な意味で同等です。つまり、用語 A が用語 B の同義語である場合、すなわち B は A の同義語です。A が照会で使用されるたびに、B が使用され、その反対もまた真です。

ただし、ある用語の異なる用途に対応する用語を追加することもでき、これにはその用語の汎用的変形、あるいはより特定のな変形も含まれます。例えば、1 つの同義語グループは `building` と `house` を含み、もう 1 つのグループは `bank`、`shore`、と `credit union` を含みます。

用語と用語の関係が厳密でないほど、検索結果は大きくなり、検索結果によっては照会に関係ないものになる場合があります。検索および索引 API では、ユーザーが検索要求を実行したときに、適切な同義語を選択できるようにする方法や、ユーザーに照会条件がどのような同義語に拡張されたかを示す方法もあります。

同義語辞書を作成するには、そのコレクションの主題についてのエキスパートが XML フォーマットで同義語リストを作成するか、あるいは、アプリケーション開発者と一緒に作業して XML ファイルを作成する必要があります。エンタープライズ・サーチ・ツール `essydictbuilder` を使用して、XML ファイルをバイナリー (.dic) ファイルに変換する必要があります。

エンタープライズ・サーチ管理者は、このバイナリー・ファイルをシステムにアップロードし、表示名を割り当てます。コレクション管理者は、コレクションに検索サーバー・オプションを設定するときに、コレクション内の文書の検索に使用する同義語辞書を選択できます。

**制約事項:** カスタム同義語辞書をシステムに追加後は、それを編集することができません。コレクションに使用可能な同義語を修正するには、以下のようにします。

1. ソース XML ファイルを更新します。
2. XML ソースを新規ディクショナリー・ファイルに変換します。
3. 古い同義語辞書を使用するコレクションからその辞書を削除します。
4. 古い同義語辞書をシステムから削除します。
5. 新しい同義語辞書をシステムに追加します。
6. 新しい同義語辞書を使用するコレクションにその辞書を関連付けます。

これらのステップを含むスクリプトを作成し、そのスクリプトを使用してエンタープライズ・サーチ・システムに辞書を再デプロイできます。

#### 関連概念



検索アプリケーションの同義語サポート

#### 関連タスク



同義語に使用できる XML ファイルの作成



同義語辞書の作成

222 ページの『カスタム辞書の再デプロイ』

## システムへの同義語辞書の追加

コレクション内で文書を検索するためにカスタム同義語辞書を作成する場合、辞書をエンタープライズ・サーチ・システムと関連付ける必要があります。コレクションの検索に使用したい同義語辞書を、後で選択できます。

#### 始める前に


エンタープライズ・サーチ照会で使用するためにカスタム同義語辞書を追加するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

#### 制約事項

同義語辞書の最大サイズは 8 MB です。

#### 手順

同義語をエンタープライズ・サーチ・システムと関連付けるには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「編集」をクリックして、システム編集ビューに切り替えます。
3. 「検索」ページで、「同義語辞書の構成」をクリックします。

4. 「同義語辞書の構成」ページで、「**同義語辞書の追加**」をクリックします。
5. 「同義語辞書の追加」ページで、同義語辞書の固有の表示名を入力し、オプションで、説明を入力します。
6. .dic ファイルのロケーションを指定します。ファイルがローカル・システムにある場合、参照してファイルを見つけることができます。ファイルが索引サーバーにある場合は、完全修飾パスを入力してください。
7. 「**OK**」をクリックします。カスタム同義語辞書が、エンタープライズ・サーチ・システムに追加され、コレクションの検索に使用できるようになります。

## 同義語辞書とコレクションとの関連付け

同義語辞書がエンタープライズ・サーチ・システムと関連付けられている場合、そのうちの 1 つを選択して、コレクションの検索時に使用できます。照会条件が辞書内の用語に一致した場合、その用語の同義語が含まれている結果文書も、検索結果に入れて戻されます。

### 始める前に

コレクションの同義語辞書を選択するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順

同義語辞書をコレクションと関連付けるには、次のようにします。

1. コレクションを編集して「検索」ページを選択し、「**検索サーバー・オプションの構成**」をクリックします。
2. 「検索サーバー・オプション」ページの「**同義語辞書名**」フィールドで、ユーザーがこのコレクションを照会するときを使用したい同義語辞書を選択します。

使用できる同義語辞書のリストに、エンタープライズ・サーチ・システムに追加されたすべての同義語辞書が含まれます。

3. 「**OK**」をクリックします。

---

## カスタムのストップワード辞書

検索結果の質を高めるために、照会処理中に一定の語が自動的に照会条件から除去されるよう指定できます。

ストップワード辞書には、頻繁に使用されるために照会条件としては実用的でない、エンタープライズ特有の語句が含まれます。これらの語を照会から除外することによって、わずかに関連するだけのおびただしい数の結果文書をユーザーが受け取るのを防止できます (照会内の他の用語に一致する文書のみが戻されます)。照会処理中に、検索サーバーが照会からストップワードを除去します。除去される語には、カスタム辞書中のストップワードと、エンタープライズ・サーチに事前定義されたストップワード (一般的な前置詞や冠詞など) があります。

エンタープライズ・サーチでは、言語固有のストップワード認識がデフォルトで実行されます。このプロセスによって、頻繁に使用される一般的な語 (例えば、a や

the) が照会から除去されます。カスタムのストップワード辞書を定義する必要があるのは、エンタープライズ特有またはドメイン特有のストップワードのみについてです。

照会が処理される際、ストップワードはスペル提案が行われる前に除去されます。照会内のすべての語がストップワードである場合は、照会処理の間にストップワードは除去されません。検索結果が戻されるようにするために、照会条件のすべてがストップワードのとき、ストップワードの除去は無効になります。例えば、car という語がストップワードで、car を検索した場合、検索結果には、car という語に一致する文書が含まれます。car volvo を検索した場合は、volvo という語に一致する文書のみが検索結果に含まれます。

ストップワード辞書を作成するには、そのコレクションの主題についてのエキスパートが XML フォーマットでストップワード・リストを作成するか、あるいは、アプリケーション開発者と一緒に作業して XML ファイルを作成する必要があります。エンタープライズ・サーチ・ツール **esstopworddictbuilder** を使用して、XML ファイルをバイナリー (.dic) ファイルに変換する必要があります。

エンタープライズ・サーチ管理者は、このバイナリー・ファイルをシステムにアップロードし、表示名を割り当てます。コレクション管理者は、コレクションに検索サーバー・オプションを設定するときに、コレクション内の文書の検索に使用するストップワード辞書を選択できます。

**制約事項:** カスタムのストップワード辞書をシステムに追加した後は、それを編集することができません。照会処理に使用可能なストップワードを訂正するには、以下の手順を実行する必要があります。

1. ソース XML ファイルを更新します。
2. XML ソースを新規ディクショナリー・ファイルに変換します。
3. 古いストップワード辞書を使用するコレクションからその辞書を削除します。
4. 古いストップワード辞書をシステムから削除します。
5. 新しいストップワード辞書をシステムに追加します。
6. 新しいストップワード辞書を使用するコレクションにその辞書を関連付けます。


これらのステップを含むスクリプトを作成し、そのスクリプトを使用してエンタープライズ・サーチ・システムに辞書を再デプロイできます。

#### 関連概念

 [カスタムのストップワード辞書](#)

#### 関連タスク

 [ストップワードに使用できる XML ファイルの作成](#)

 [ストップワード辞書の作成](#)

[222 ページの『カスタム辞書の再デプロイ』](#)

## システムへのストップワード辞書の追加

照会から語を除去するためのカスタムのストップワード辞書を作成した場合、それらの辞書をエンタープライズ・サーチ・システムに追加する必要があります。どのストップワード辞書をコレクションの検索に使用するのかは、後で選択できます。

### 始める前に


カスタムのストップワード辞書をシステムに追加するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### 制約事項

ストップワード辞書の最大サイズは 8 MB です。

### 手順

カスタムのストップワード辞書をエンタープライズ・サーチ・システムと関連付けるには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「編集」をクリックして、システム編集ビューに切り替えます。
3. 「検索」ページで、「ストップワード辞書の構成」をクリックします。
4. 「ストップワード辞書の構成」ページで、「ストップワード辞書の追加」をクリックします。
5. 「ストップワード辞書の追加」ページで、辞書の固有の表示名を入力します。
6. .dic ファイルのロケーションを指定します。ファイルがローカル・システムにある場合、参照してファイルを見つけることができます。ファイルが索引サーバーにある場合は、完全修飾パスを入力してください。
7. 「OK」をクリックします。カスタムのストップワード辞書が、エンタープライズ・サーチ・システムに追加され、コレクションの検索に使用できるようになります。

## ストップワード辞書とコレクションとの関連付け

ストップワード辞書がエンタープライズ・サーチ・システムと関連付けられている場合、そのうちの 1 つをコレクションの検索時に使用するよう選択できます。照会条件に一致する語が辞書内にある場合、その語は処理される前に照会から除去されます。

### 始める前に

コレクションに対してストップワード辞書を選択するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順

ストップワード辞書をコレクションと関連付けるには、次のようにします。

1. コレクションを編集して「検索」ページを選択し、「検索サーバー・オプションの構成」をクリックします。

2. 「検索サーバー・オプション」ページの「ストップワード辞書名」フィールドに、ユーザーがこのコレクションを照会するときを使用するようにしたいストップワード辞書の名前を入力します。

使用できるストップワード辞書のリストに、エンタープライズ・サーチ・システムに追加されたすべてのストップワード辞書が含まれます。

3. 「OK」をクリックします。

---

## カスタム辞書の再デプロイ

同義語、ストップワード、またはランキング調整ワード辞書の変更に、エンタープライズ・サーチ管理コンソールを使用することはできません。しかし、スクリプトにステップを組み込み、そのスクリプトを使用して辞書を再デプロイすることはできます。

### 始める前に

カスタム辞書を再デプロイするには、エンタープライズ・サーチ管理者としてログインする必要があります。

### 手順

エンタープライズ・サーチ用のカスタム辞書を再デプロイするには、次の手順で行います。

1. 索引サーバーで、次のいずれかのファイルを開いて、再デプロイする辞書の名前を見つけてみます。
  - 同義語辞書の名前を見つけるには、ES\_NODE\_ROOT/master\_config/SynonymConfiguration.xml を開きます。
  - ストップワード辞書の名前を見つけるには、ES\_NODE\_ROOT/master\_config/StopWordDictionaryConfiguration.xml を開きます。
  - ランキング調整ワード辞書の名前を見つけるには、ES\_NODE\_ROOT/master\_config/BoostingWordDictionaryConfiguration.xml を開きます。

次に、hello と goodbye という名前でアップロードされた 2 つの同義語辞書が指定された SynonymConfiguration.xml ファイルの例を示します。

```
% cat $ES_NODE_ROOT/master_config/SynonymConfiguration.xml
```

```
<SynonymConfiguration>
<Synonyms>
<Synonym Name="hello" ID="SynonymId_1">
<Filename>synonym_hello1.dic</Filename>
<Timestamp>1169766691776</Timestamp>
</Synonym>
<Synonym Name="goodbye" ID="SynonymId_2">
<Filename>synonym_goodbye2.dic</Filename>
<Timestamp>1169767224839</Timestamp>
</Synonym>
</Synonyms>
</SynonymConfiguration>
```

辞書がアップロードされると、固有なファイル ID が割り当てられます。上の例では、この ID は synonym\_hello1.dic と synonym\_goodbye2.dic です。索引サーバー上のこのファイルの絶対パスは、次のとおりです。



- 同義語辞書のパスは、ES\_NODE\_ROOT/data/custom\_dictionary/synonym\_\*.dic です。
- ストップワード辞書のパスは、ES\_NODE\_ROOT/data/custom\_dictionary/stopwordDictionary\_\*.dic です。
- ランキング調整ワード辞書のパスは、ES\_NODE\_ROOT/data/custom\_dictionary/boostingwordDictionary\_\*.dic です。

AIX、Linux、および Solaris では、**ls** コマンドを使用して使用可能な辞書をリストできます。次に例を示します。

```
% ls -l $ES_NODE_ROOT/data/custom_dictionary/synonym_*.dic
-rw-rw-r-- 1 esuser users 9 Jan 25 15:11 /home/esuser/node/data/custom_dictionary/synonym_hello1.dic
-rw-rw-r-- 1 esuser users 9 Jan 25 15:11 /home/esuser/node/data/custom_dictionary/synonym_goodbye2.dic
```

2. 更新する辞書を特定したら、エンタープライズ・サーチ・システムを停止します。
3. 更新する辞書を新しい辞書ファイルで上書きします。例えば、`synonym_hello1.dic` または `synonym_goodbye2.dic` を上書きします。
4. エンタープライズ・サーチ・システムが 4 つのサーバーのうちの 2 つで稼働する場合は、検索サーバー上の辞書ファイルも手動で上書きします。辞書ファイルは索引サーバー上と同じパスに存在します (ES\_NODE\_ROOT/data/custom\_dictionary/)。
5. エンタープライズ・サーチ・システムを再始動します。

#### 関連概念

216 ページの『カスタム同義語辞書』

219 ページの『カスタムのストップワード辞書』

233 ページの『カスタム・ランキング調整ワード辞書』

---

## 動的要約

動的要約は、結果文書中のどの語句がユーザーの検索対象概念を最もよく表しているかを判別する技法です。

エンタープライズ・サーチで、動的要約は、文書中のさまざまな検索語を多く含む文を捕捉しようとする機能です。検索結果には、いくつかの文、または各文の一部が選択されて表示されます。対象となる検索語は、検索結果の中で、HTML レンダリングによって強調表示されます。

コレクションの検索サーバー・オプションを構成するときに、検索結果内での文書サマリーの最大表示長を指定できます。サマリーには強調表示文字が含まれるので、検索アプリケーションに戻されるバッファは指定の最大値より大きくなる場合があります。ただし、表示長が指定の最大値を超えることはありませんが、(ソース文書から抽出されるサマリー・データに従って) サマリーはより短くなる可能性があります。

## 管理コンソールの文書サマリーのカスタマイズ

エンタープライズ・サーチ管理コンソールで検索サーバーのオプションを指定することによって、文書サマリーに表示される情報量をカスタマイズすることができます。

### 始める前に

コレクションのサマリー表示長を制御するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

文書サマリーの最大表示長に指定する値は、各サマリーに入れることのできるセンテンスの数に指定する値と連動して機能します。文書サマリーが最も短くなる値が優先されます。

例えば、センテンス数の制限を 4 に指定した場合、表示長では 4 つのセンテンスに含まれる文字総数より多い文字が許可されていても、文書サマリーに含まれるセンテンスの数は 4 になります。他にも、例えばセンテンス数が 10 に制限され、表示長が 500 文字に制限されているという組み合わせを想定すると、結果の文書サマリーのセンテンス数が 10 より少なくなることもあります。

### 手順

文書サマリーの表示長を構成するには、次のようにします。

1. コレクションを編集して「検索」ページを選択し、「**検索サーバー・オプションの構成**」をクリックします。
2. 「検索サーバー・オプション」ページで、文書サマリーの最大表示長を指定します。ユーザーが検索結果を表示するときの文書サマリーは、ここで指定する値を超えなくなります。
3. 各文書サマリーが含むことのできるセンテンスの数を指定します (サマリーは 10 個までのセンテンスを含むことができます)。
4. 「**OK**」をクリックします。
5. 変更を有効にするために、検索サーバーをモニターし、サーバー・プロセスを再始動します。

## プロパティの編集による文書サマリーのカスタマイズ

エンタープライズ・サーチ照会の各結果文書にサマリーが含まれています。プロパティ・ファイルの編集によって、各サマリーに含まれる情報をカスタマイズすることができます。

### このタスクについて

`ES_NODE_ROOT/master_config/collection_ID.runtime.node1/runtime-generic.properties` ファイル内の以下のプロパティの値を変更することにより、検索結果の説明をカスタマイズできます。

### **MinWordsPerSentence**

サマリー内の各センテンスの最小ワード数。MinWordsPerSentence 値よりも多くの数のワードを含むことができるセンテンスがない場合、サマリーにはこれより短いセンテンスが組み込まれます。デフォルト値は 4 です。

### **MaxWordsPerSentence**

サマリーに組み込まれる各センテンスの最大ワード数。センテンスにこの制限を超える数のワードがある場合、サマリーにはセンテンスの一部 (MaxWordsPerSentence 値までの数の照会条件を含む部分) のみが組み込まれます。そのセンテンスの残りは除外されます。デフォルト値は 20 です。

センテンスは、文書サマリーのためにプロプラエタリーに従って選択されます。プロプラエタリーは、検索語を含むすべてのセンテンスの関連性を決定する内部アルゴリズムです。関連性による選択は、センテンスがセンテンスの長さでフィルターに掛けられる前に行われます。

### **NumberOfReturnedSentences**

文書の説明を構成する文の数。デフォルト値は 5 です。

### **MaxSentencesPerDocument**

説明を作成するプロセスで候補と見なされる文書内の文の最大数。デフォルト値は 1000 です。

## **手順**

検索結果内の文書サマリーをカスタマイズするには、次のようにします。

1. 検索サーバーに、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. テキスト・エディターを使用して以下のファイルを編集します。ここで、`coll_ID` は、コレクションの作成時にそのコレクションについて指定された ID です (または、システムによって割り当てられたもの)。

```
ES_NODE_ROOT/master_config/coll_ID.runtime.nodel/runtime-generic.properties
```

**ヒント:** コレクション名とその ID との間でのマッピングを確認するためには、`ES_NODE_ROOT/master_config/collections.ini` ファイルを参照してください。

3. カスタマイズしたいプロパティを変更してから、ファイルを保存して終了します。
4. 検索サーバーを一度停止してから再始動して、変更を適用します。

---

## **クイック・リンクに関する作業**

クイック・リンクは、ユーザーが特定の語句を含む照会を実行したときに必ず検索結果に戻される文書を指しています。

コレクションのクイック・リンクを構成するには、エンタープライズ・サーチ管理コンソールを使用します。

## クイック・リンク

クイック・リンクによって、照会条件に関連すると事前指定された文書へのリンクをユーザーに提供できます。

クイック・リンクは、照会に特定の語句が含まれる場合に、エンタープライズ・サーチが自動的に検索結果に入れるようにする URI です。通常、クイック・リンク URI は結果リストの先頭に表示され、照会に適合するとあらかじめ設定した文書をユーザーが確実に参照できるようになっています。

クイック・リンクは、他の検索結果に追加する形で戻されます。検索プロセスでは、照会条件に一致する文書がないか索引を検索し、それらの文書の URI をクイック・リンクの URI に追加して戻します。

クイック・リンクを構成する場合、その URI を示すタイトルと要約を指定して、ユーザーがその文書の内容を把握し、それがそのユーザーがリトリブする文書かどうかをすぐに判断できるようにすることができます。

例えば、URI 「<http://www.ibm.com/education/us/>」ならば、「米国 IBM 教育関係サイト」というタイトルをつけ、要約として「米国内の教育関係者、専門家、および学生向けソリューション、製品、およびリソースの紹介」などと設定します。

エンタープライズ・サーチ・コレクションでクイック・リンクを使用するには、クイック・リンクを表示するオプションが検索アプリケーションで使用可能になっていなければなりません。検索アプリケーションによっては、ユーザーがコレクションの検索時にクイック・リンクを戻す機能を使用可能/使用不可に設定できる場合があります。

## クイック・リンクの構成

エンタープライズ・サーチ・コレクションのクイック・リンクを作成するため、文書の URI を、検索結果に含まれたものをトリガーするキーワードと関連付けます。

### 始める前に

クイック・リンクを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのクイック・リンクが所属するコレクションのコレクション管理者である必要があります。

### このタスクについて

クイック・リンクのキーワードおよび URI の指定方法の例を参照するには、クイック・リンクを作成または編集しているときに「ヘルプ」をクリックしてください。

変更を有効にするのに、検索サーバーを再始動する必要はありません。

### 手順

クイック・リンクを構成するには、次のようにします。

1. コレクションを編集、「検索」ページを選択、そして「クイック・リンクの構成」をクリックします。

2. 「クイック・リンク」ページで、「クイック・リンクの作成」をクリックします。
3. このクイック・リンクが検索結果に戻されるようにするためのキーワードと句、この照会に関連すると事前判定した文書の URI、このクイック・リンクのその他のオプションを指定します。

1 行につき、1 つ以上のキーワードを指定するか、1 つの句 (2 つ以上の語を引用符で囲んだもの) を指定します。キーワードはスペースで区切ります (コンマを使用してキーワードを区切ることはできません)。Enter キーを押すと改行されます。

4. 「OK」をクリックします。

「クイック・リンク」ページに、このコレクションに属している他のクイック・リンクと共に新しいクイック・リンクがリストされます。

#### 関連資料

135 ページの『エンタープライズ・サーチ索引における URI フォーマット』



---

## 文書ランキング

ユーザーがコレクションを検索するときに、検索プロセスは、照会に指定された使用条件に照合して、最も関連性の高い結果を戻します。

検索サーバーは、最も関連性の高い検索結果を生成するために、豊富な照会構文をサポートし、テキスト・ベースのスコアリングや静的ランキングなどの幾つかの技法を使用します。次のように、検索結果での文書の重要度を操作するオプションを構成することによって、ランキングのデフォルト動作を拡張することができます。

- カスタムのランキング調整ワード辞書を作成して、指定されたランキング調整ワードが含まれている文書が検索結果内でどのようにランク付けされるのかを操作できます。
- 指定された URI パターンに一致する文書のスコアを操作できます。
- ランキング調整クラスにマップされたフィールドが含まれている文書のスコアを操作できます。

### 関連概念

237 ページの『ランキング調整クラスに基づいた文書ランキング』

233 ページの『カスタム・ランキング調整ワード辞書』

---

## テキスト・ベースのスコアリング

エンタープライズ・サーチでは、照会中の用語に一致した各文書について、動的にスコアを計算します。

文書のテキスト・スコアで、文書内の照会条件の重要度を表します。照会に一致する各文書のテキストのスコアを計算する場合、エンタープライズ・サーチでは次のようなさまざまな要素を計算に入れます。

- 用語で文書を区別するかどうか。例えば、照会条件がある文書には出現するが、他の文書には出現しない場合、この用語はその文書には重要であり、他の文書と区別します。ほとんどの文書に出現する照会条件は、選択した文書セットに出現する照会条件に比べて、文書のスコアに対する貢献度合いは少なくなります。
- 文書内の照会条件の出現回数。用語のスコアは、文書内のその用語の各出現を基に計算されます。文書内の照会条件の出現回数が増えるほど、その文書のスコアは高くなります。例えば、thinkpad を検索する場合、この用語が何回も使用されている文書は、出現の少ない他の文書よりも検索結果では高いランクになります。
- HTML 文書の場合、照会条件の属性 (つまり、用語の内容) が考慮されます。用語のスコアは、その用語の属性 (ロケーション、太字、イタリック、アンカーなど) を基に計算されます。一般的に、文書タイトルに出現する用語は、パラグラフで出現する用語よりもスコアが高くなります。強調された用語 (太字テキストなど) は、プレーン・テキストの用語よりもスコアが高くなります。パーサー設定で、属性の重要度を構成できます。
- 照会条件の重み。文書内の用語の重要度をカスタマイズする場合、用語のランキング調整値を構成できます。この場合、ランキング調整ワード辞書 (用語とその

ランキング調整値を含む) をコレクションと関連付けします。この辞書は検索時に使用され、辞書内の用語のランキング調整値が文書のスコアに寄与します。ランキング調整値が高くなると、文書スコアに対するその用語の寄与が高くなります。

- 文書内の照会条件の接近性。文書内で照会条件がお互いに接近して出現する場合、その字句親和性がテキスト・スコアの計算に使用されます。例えば、2 つの文書があると想定します。1 つは、市内の駐車場について説明する文書です (車と駐車場は接近している)。もう 1 つは、駐車場の近くの車のショールームについて説明する文書です (車と駐車場は接近していない)。駐車場 (car park) を検索する場合、最初の文書の利用の接近性により、最初の文書は 2 番目の文書よりもランクが高くなります。
- 各文書の長さ、および語彙の豊富さ (固有なワード数など) も、文書のスコアを決定する要因です。

---

## 静的ランキング

文書のタイプによっては、検索結果における文書の重要度を高くする静的ランキング要素を関連付けることができます。

コレクションを作成する場合は、「**文書の重要度**」オプションを指定します。選択した文書重要度のタイプで、静的ランキング要因がコレクション内の文書と関連付けられているかどうかを判別します。静的ランキングを使用するコレクションをユーザーを検索する場合、静的ランキング要因が、検索結果に戻される文書のランク付けに影響します。

Web コンテンツの場合、静的ランキング要因はリンクに基づきます。他の文書からある文書へのリンク数、およびそれらのリンク元によって、検索結果におけるその文書の適合度を高くすることができます。

日付フィールドまたは日付メタデータを含む文書の場合、ランキング要因は、文書の日付に基づきます。クローラーによって指定される文書日付フィールドは、クローラー構成のセットアップ方法により、文書の最終変更日付または文書の最終クローリング日付になります。

文書の日付で、適合度を高くすることができます。例えば、NNTP ニュースグループではより最近の記事を古い記事よりも適合度を高くする設定ができます。データ・ソースに複数の日付値がある場合は、クローラーを構成するときに、文書の適合度を判定する上で、どの日付を最も重視するかを選択することができます。

静的ランキングをコレクションに使用する場合、異なるランキング・タイプを使用するデータ・ソースを同一のコレクションに混在させないように注意してください。例えば、ある文書に対して静的ランキング要素としてリンク数を使用したい場合、コレクション内の文書が Web 文書だけになっていることを確認してください。ランキング・モデルの異なるソースが同一コレクションに結合されている場合、文書ランキングの正確さが低くなり、検索結果の順序が予想したものでないことがあります。

また、コレクション内の文書に、静的ランキングを適用できるフィールドと値が含まれていることも確認してください。例えば、文書の日付に基づいて静的ランキン



グを使用するように構成されたコレクションがあり、コレクション内のクローラーが特定のフィールドを文書日付として使用するように構成されているとします。文書にそのフィールドが含まれていない場合、文書の重要度は適切にランク付けされず、検索結果の順序は予期したものでないことがあります。

## リンク・ベースのランキングについての考慮事項

静的ランキングは、URI パターンのランキング調整のためのスコアの割り当てなどの要因と共に、文書の静的スコアに寄与し、文書の重要度に影響を与えます。リンク・ベースのランキング・モデルは通常、Web コレクションに適用されます。それは、このモデルでは、文書へのリンク数に基づいて文書の静的ランクが計算されるからです。他の多くの数の文書からリンクされる文書は、適合度が高いとランクされます。

この理由から、非 Web コレクションまたは混在コレクション (Web 文書および Web 以外の文書を含むコレクション) 用にこのモデルを構成する場合、Web 以外の文書にはリンクの概念がないので、検索品質が低下することがあります。

リンク・ベースのランキングが使用可能な場合は、重複文書検出も使用可能です。重複文書の静的ランクはマスター文書と同じです。重複グループ内の文書に URI パターン・ランキング調整要因が構成されていない場合、重複文書の静的スコアはすべて同じです。

---

## 静的文書ランキングのデフォルト値の復元

コレクションの作成時に静的文書ランキング・オプションを構成した場合、コレクションの `runtime.properties` ファイルを編集することによって、プロパティをデフォルト値に戻すことができます。

### 始める前に

デフォルトの文書ランキング値をコレクションに復元するには、エンタープライズ・サーチ管理者でなければなりません。

### このタスクについて

コレクションのデフォルト文書ランキング値を復元するには、そのコレクションとエンタープライズ・サーチ・システム内のすべての検索サーバーの `runtime.properties` ファイルを更新する必要があります。マルチサーバー構成において、`runtime.properties` ファイルは、索引サーバー上の `ES_NODE_ROOT/master_config/collection_ID.runtime.node_ID` ディレクトリにあります。ここで、`collection_ID` はコレクションの ID で、`node_ID` は検索サーバーの ID です。

例えば、マルチサーバー・エンタープライズ・サーチ・システムで `col1` コレクションを更新する場合、そのコレクションと 2 つの検索サーバー (`node3` と `node4`) の `runtime.properties` を更新します。

```
ES_NODE_ROOT/master_config/col1.runtime.node3/runtime.properties
ES_NODE_ROOT/master_config/col1.runtime.node4/runtime.properties
```

### 手順

コレクションについて文書ランキング値をデフォルト値に復元するには、次のようにします。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。
2. デフォルト・ランキング値を復元するコレクションのコレクション ID を確認します。コレクション ID は、ES\_NODE\_ROOT/master\_config/collections.ini ファイルにあります。見やすくするためにこのファイルをソートします。次の例で、coll がコレクション ID です。

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

3. 復元するコレクションの runtime.properties ファイルを編集し、以下のように変更します。

- a. 以下のプロパティを削除します。

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- b. runtime.properties ファイルで trevi.sourcetype=1 と指定されている (文書がリンクでランキング調整されることを示す) 場合、ES\_INSTALL\_ROOT/default\_config/runtime.1/runtime.properties ファイルを編集して、次のデフォルト・プロパティを runtime.properties ファイルにコピー・アンド・ペーストします。

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- c. runtime.properties ファイルで trevi.sourcetype=2 と指定されている (文書が日付でランキング調整されることを示す) 場合、ES\_INSTALL\_ROOT/default\_config/runtime.2/runtime.properties ファイルを編集して、次のデフォルト・プロパティを runtime.properties ファイルにコピー・アンド・ペーストします。

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- d. runtime.properties ファイルで trevi.sourcetype=3 と指定されている (コレクションの文書のランキング調整に静的ランキング要素が使用されないことを示す) 場合、ES\_INSTALL\_ROOT/default\_config/runtime.0/runtime.properties ファイルを編集して、次のデフォルト・プロパティを runtime.properties ファイルにコピー・アンド・ペーストします。

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- マルチサーバー構成で、ステップ 3 (232 ページ) を繰り返し、同じコレクションおよび 2 番目の検索サーバーの `runtime.properties` ファイルを更新します。
- 管理コンソールで「検索」ページをモニターし、このコレクションの検索プロセスを再始動します。

必要に応じて、デフォルトの文書ランキング値を復元したい各コレクションについて、このステップを繰り返します。

---

## カスタム・ランキング調整ワード辞書

検索結果の質を向上させるために、カスタム・ランキング調整ワード辞書を作成して、検索結果における文書のランクを調整することができます。

照会でランキング調整ワード辞書にある用語が指定されると、その用語が含まれている文書の重要度が、辞書の実語に構成されているランキング調整要因に応じて高くなったり、低くなったりします。

ランキング調整ワード辞書を使用して、ユーザーが一定の照会条件を指定したとき一定の文書が戻されるようにすることができます。例えば、自動車を話題にした多数の文書を含むコレクションがあるとします。このような文書の場合は、車種やメーカーの名前など自動車に関連する語が重要であると思われます。検索結果のランキングを調整するため、ランキング調整ワード辞書でキー・ワード (車種やメーカーなど) にランキング調整値を割り当てることができます。ユーザーがコレクションを検索し、これらのキー・ワードのいずれかを含む照会を指定すると、検索結果で、自動車に関する文書が他の文書よりも高くランクされます。

ランキング調整要因の範囲は -10 から 10 までです。照会処理時に、検索サーバーは、正の数のランキング調整要因の実語が含まれている文書の重要度を高くし、負の数のランキング調整要因の実語が含まれている文書の重要度を低くします。

例えば、ランキング調整要因が高い照会条件が含まれている文書は、ランキング調整要因が適用されない場合に比べて、高くランクされます。(ランキング調整要因のみが、文書のスコアに影響する係数です。)

辞書を作成する場合、同じランキング調整要因を複数の用語に割り当てることができます。辞書には、単一の実語または複数の用語を含めることができます。複数の用語は、1 つの句として突き合わせが行われます。

ランキング調整値によって重み付けされた用語が OR 演算子を使用する照会 (例えば `this | that`) で指定された場合、その照会条件に対して重み付けされた平均が計算されます。計算結果の集約スコアが、その OR 照会オペランドのすべての出現に使用されます。個々の OR 照会オペランドについて、個々のスコアが計算されることはありません。

ランキング調整ワード辞書に基づくランキング調整は、フィールド照会条件ではサポートされません。照会条件の解析時に、フィールド名ではなく照会テキストのみ

が、文書スコアの計算に使用されます。フィールド内の照会条件にランキング調整要因を適用するためには、フィールド名をランキング調整クラスにマップすることができます。

ランキング調整ワード辞書を作成するには、そのコレクションの主題についてのエキスパートが XML フォーマットでランキング調整ワード・リストを作成するか、アプリケーション開発者と一緒に作業して XML ファイルを作成する必要があります。エンタープライズ・サーチ・ツール **esboosttermdictbuilder** を使用して、XML ファイルをバイナリー (.dic) ファイルに変換する必要があります。

エンタープライズ・サーチ管理者は、このバイナリー・ファイルをシステムにアップロードし、表示名を割り当てます。コレクション管理者は、コレクションの検索サーバー・オプションを構成するときに、コレクション内の文書の検索に使用するランキング調整ワード辞書を選択します。

**制約事項:** カスタム・ランキング調整ワード辞書は、システムに追加した後で編集することはできません。照会処理に使用可能なランキング調整ワードを変更するには、以下を行う必要があります。

1. ソース XML ファイルを更新します。
2. XML ソースを新規ディクショナリー・ファイルに変換します。
3. ランキング調整ワード辞書を使用するコレクションから古い辞書を削除します。
4. システムからランキング調整ワード辞書を削除します。
5. 新規のランキング調整ワード辞書をシステムに追加します。
6. 新規ランキング調整ワード辞書を、その辞書を使用するコレクションに関連付けます。

これらのステップを含むスクリプトを作成してから、そのスクリプトを使用して辞書をエンタープライズ・サーチ・システムに再デプロイすることができます。

#### 関連概念

229 ページの『文書ランキング』



カスタム・ランキング調整ワード辞書

#### 関連タスク



ランキング調整ワードに使用できる XML ファイルの作成



ランキング調整ワード辞書の作成

222 ページの『カスタム辞書の再デプロイ』

## ランキング調整ワード辞書のシステムへの追加

カスタム・ランキング調整ワード辞書を作成する場合、辞書をエンタープライズ・サーチ・システムに関連付ける必要があります。そうすることにより、コレクションの検索に使用したいランキング調整ワード辞書を後で選択することができます。

始める前に


カスタム・ランキング調整ワード辞書をシステムに追加するには、エンタープライズ・サーチ管理者役割の担当者である必要があります。

### 制約事項

ランキング調整ワード辞書の最大サイズは 8 MB です。

### 手順

カスタム・ランキング調整ワードをエンタープライズ・サーチ・システムに関連付けるには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「 編集」をクリックして、システム編集ビューに移動します。
3. 「検索」ページで、「ランキング調整ワード辞書の構成」をクリックします。
4. 「ランキング調整ワード辞書の構成」ページで、「ランキング調整ワード辞書の追加」をクリックします。
5. 「ランキング調整ワード辞書の追加」ページで、辞書の固有の表示名を入力し、オプションで、説明を入力します。
6. .dic ファイルのロケーションを指定します。 ファイルがローカル・システムにある場合、参照してファイルを見つけることができます。ファイルが索引サーバーにある場合は、完全修飾パスを入力してください。
7. 「OK」をクリックします。 カスタム・ランキング調整ワード辞書が、エンタープライズ・サーチ・システムに追加され、コレクションの検索に使用できるようになります。

## ランキング調整ワード辞書とコレクションとの関連付け

ランキング調整ワード辞書がエンタープライズ・サーチ・システムと関連付けられている場合、そのうちの 1 つを選択して、コレクションの検索時に使用できます。照会条件が辞書内の用語に一致した場合、その用語が含まれている文書の重要度が、辞書用語に割り当てられているランキング調整要因に応じて高くなったり、低くなったりします。

### 始める前に

コレクションのランキング調整ワード辞書を選択するには、エンタープライズ・サーチ管理者役割の担当者であるか、そのコレクションのコレクション管理者である必要があります。

### 手順

ランキング調整ワード辞書をコレクションと関連付けるには、次のようにします。

1. コレクションを編集して「検索」ページを選択し、「検索サーバー・オプションの構成」をクリックします。
2. 「検索サーバー・オプション」ページの「ランキング調整ワード辞書名」フィールドで、ユーザーがこのコレクションを照会するときに使用するランキング調整ワード辞書の名前を入力します。

使用できる辞書のリストに、エンタープライズ・サーチ・システムに追加されたすべてのランキング調整ワード辞書が含まれます。

3. 「OK」をクリックします。

---

## URI パターンに基づいた文書ランキング

URI パターンにランキング調整要因を割り当てることによって、文書の重要度を増減できます。

すべての文書には、索引に追加されるときにデフォルトの静的ランキング・スコアが割り当てられます。デフォルトのスコアは、コレクションに対して静的ランキングが使用可能かどうかによって変わり、使用可能な場合には静的ランキング・タイプ (文書日付、または、Web 文書の場合はリンクする他の文書の数) によって変わります。

URI パターンにランキング調整要因を割り当てることによって、文書の相対的重要度を操作できます。ランキング調整要因は、デフォルトの静的ランキング・スコアおよび他の要因と共に、文書の最終的な静的スコアを決定するのに使用されます。

構成する URI パターンの順序は重要です。索引サーバーは、コレクション内の各文書の値を計算する際、URI パターンをリストされている順に評価します。それぞれの URI ごとに以下ようになります。

1. 索引サーバーは、URI パターンを順番にスキャンします。
2. 索引サーバーは、索引内に一致する文書がある最初の URI を検出すると、その URI パターンに構成されたランキング調整要因をその文書に適用します。
3. どの URI パターンにも一致しない文書には、デフォルトの静的ランキング・スコアが使用されます。

URI パターンにランキング調整要因を構成する場合は、その後、索引サーバーにスキャンさせたい順序でその URI パターンを配置する必要があります。

## URI パターンに一致する文書のスコアの操作

デフォルトの静的ランキング・スコアにランキング調整要因を適用することによって、URI パターンに一致する文書の重要度を増減することができます。

### 始める前に

URI パターンに一致する文書の重要度を操作するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

構成するランキング調整要因は、デフォルトの静的ランキング・スコアと共に使用されて、指定された URI パターンに一致するすべての文書について新しい静的スコアが計算されます。

ランキング調整要因は、静的スコアのみを調整するものであり、文書の最終的なリンクを決定する計算に影響する 1 つの要素にすぎません。例えば、多数からリンク

される文書 (この場合は初期スコアが高くなります) がある場合、リンクがまったく  
ない文書は常にそれよりも低くランク付けされます。

## 手順

URI パターンに一致する文書のスコアを操作するには、次のようにします。



1. コレクションを編集して「索引」ページを選択し、「URI パターン・マッチングによるスコアの操作」をクリックします。
2. 「URI パターン・マッチングによるスコアの操作」ページで、「URI パターンの追加」をクリックします。
3. 検索結果内での重要度を増減したい文書の URI パターンを指定します。次に例を示します。

```
http://domain.org/hr/*  
db2://*ROI*  
*/afs/*
```

4. ランキング調整要因に -10 と 10 の間の値を入力します。このランキング調整要因に基づいて、指定した URI パターンに一致するすべての文書の最終的な静的スコアが計算されます。
5. 「OK」をクリックします。
6. 「URI パターン・マッチングによるスコアの操作」ページで、新しい URI パターンを、索引サーバーにスキャンさせたい順序で配置します。

索引サーバーは、URI がリストされている順序で、静的ランキング・スコアを計算します。最良の結果を得るには、より特定度の高い URI を先にリストします。以下の例で、/forms サブディレクトリーは http://www.ibm.com/hr/\* URI パターンに一致します。/forms サブディレクトリー内の文書のスコアが確実に正しく計算されるようにするには、/forms サブディレクトリーの URI パターンを先にリストします。

```
http://www.ibm.com/hr/forms/* 8  
http://www.ibm.com/hr/* -2
```

7. URI パターンまたはランキング調整要因を変更するには、目的の URI パターンを選択して、「 編集」をクリックします。
8. リストから URI パターンを除去するには、目的の URI パターンを選択して、「 除去」をクリックします。
9. 以前に索引作成された文書にランキング調整要因を適用するには、主索引を再作成します。

---

## ランキング調整クラスに基づいた文書ランキング

フィールドをランキング調整クラスにマッピングすることにより、検索結果における文書のランクを調整することができます。

文書の構文解析時に、パーサーは、トークンが属しているフィールドに応じて ランキング調整クラス を文書トークンに割り当てます。これらのランキング調整クラスは、索引に組み込まれ、照会の評価時に、結果文書のランクに影響するスコアの計算に使用されます。

スコアの計算方法に影響を与えるには、ランキング調整クラスのランキング調整要因を構成します。照会条件が、ランキング調整クラスにマップされているフィールドのトークンに一致すると、トークンのこのオカレンスに課せられている係数が文書の合計スコアに影響します。スコアは、ランキング調整クラスに構成されているランキング調整要因を適用して計算されます。

例えば、タイトル・フィールドのスコアのランキングを上げると、照会条件がタイトルにある場合、オカレンスによって文書スコアが上昇するため、検索結果における文書のランクが上昇します。

文書ランキングを調整するには、エンタープライズ・サーチ管理コンソールを使用して、ランキング調整クラスのランキング調整要因を指定し、フィールドをランキング調整クラスにマップします。エンタープライズ・サーチ用に 16 のランキング調整クラスが事前構成されています。そのうちの 8 つのランキング調整クラスはコンテンツ・フィールドで使用するように設計されており、それ以外の 8 つのランキング調整クラスはメタデータ・フィールド用に設計されています。デフォルトのランキング調整クラスに関連付けられているスコアを編集したり、ランキング調整クラスに別のフィールドや追加フィールドを関連付けたりすることができます。

フィールドのマッピングを変更する場合には、以前に索引付けされていた文書に変更が適用されるように、文書を再度クロールおよび構文解析する必要があります。ランキング調整クラスに指定されている係数を変更する場合には、検索サーバーをモニターし、変更が有効になるように、検索サーバー処理を停止して再始動してください。

## 重複文書の検出と文書サマリー

フィールドをランキング調整クラスにマップする際には、そのフィールドが重複文書の検出に使用されるか、また、そのフィールドの内容を検索結果の文書サマリーに含めることができるかを指定する必要があります。

- フィールドが重複文書の検出に使用される場合、フィールドはコンテンツ・フィールドであるとみなされ、コンテンツ・フィールド用に設計されているランキング調整クラスのみが選択対象となります。こうしたタイプのフィールドの内容は、検索結果の動的文書サマリーで使用できます。
- フィールドが重複文書の検出に使用されない場合、フィールドはメタデータ・フィールドであるとみなされ、メタデータ・フィールド用に設計されているランキング調整クラスのみが選択対象となります。この場合、指定されているフィールド以外すべての点において同じである 2 つの文書は、互いに重複しているとみなされ、フィールドは動的文書サマリーで使用されません。

## 高い再呼び出し値と低い再呼び出し値

照会を評価する際に、検索プロセスは、戻される結果文書数を見積もります。しきい値によって、照会が持つとみなされる値が低い再呼び出しの値か、高い再呼び出しの値か、その間の値かが決まります。

### 低い再呼び出し値

結果文書の見積もり数が下限しきい値を下回る場合、照会は低い再呼び出し照会であるとみなされます。



### 高い再呼び出し値

結果文書の見積もり数が上限しきい値を上回る場合、照会は高い再呼び出し照会であるとみなされます。

### 混合再呼び出し値

文書の見積もり数が 2 つのしきい値の間である場合、照会の再呼び出し値は 2 つのしきい値の混合になります。

各ランキング調整クラスは、照会処理時に、低い再呼び出し照会および高い再呼び出し照会に関連付けられます。低いランキング調整要因は低い再呼び出し照会の相対重要度に影響し、高いランキング調整要因は高い再呼び出し照会の相対重要度に影響します。2 つのランキング調整要因の混合は、混合の再呼び出し値を持つ照会の相対重要度に影響します。

ランキング調整要因の値は、文書内の照会条件の各オカレンスの相対重要度を制御します。文書内の照会条件の各オカレンスは、対応するランキング調整要因に応じてカウントされます。

コレクションにランキング調整クラスを構成する場合は、デフォルトのランキング調整要因を編集できます。例えば、タイトル・フィールドにある照会条件が、通常のテキストにある照会条件に比べて 5 倍にカウントされるようにランキング調整要因を指定することができます。

#### 関連概念

229 ページの『文書ランキング』

## ランキング調整クラスへのフィールドのマッピング

フィールド名をランキング調整クラスにマッピングすることにより、フィールドの相対重要度に影響を与えることができます。

### 始める前に

フィールドをランキング調整クラスにマップするには、エンタープライズ・サーチ管理者役割の担当者であるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

システムはランキング調整要因を使用して、ランキング調整クラスにマップされているフィールド内に照会条件が含まれている文書のランキングを調整します。

エンタープライズ・サーチでは、他の定義特性を持たない内部フィールドや通常のテキスト用にいくつかのマッピングが予約されています。予約フィールド以外のフィールドを予約フィールドが使用するランキング調整クラスにマップすることはできませんが、予約フィールドを編集したり削除することはできません。

### 手順

フィールドをランキング調整クラスにマップするには、次のようにします。

1. コレクションを編集し、「構文解析」ページを選択して、「フィールドをランキング調整クラスにマップ」をクリックします。

2. 「フィールドをランキング調整クラスにマップ」ページで、「フィールドの追加」をクリックします。
3. 「フィールドをランキング調整クラスに追加」ページで、ランキング調整クラスにマップしたいフィールドの名前を入力します。



クローलされたソースや外部ソースにあるフィールド名、XML エlementによりマップされるフィールド名、HTML メタデータ・Elementによりマップされるフィールド名、事前定義されているフィールド名を指定できます。

4. フィールドが重複文書の検出に使用されるかどうかを指定します。チェック・ボックスを選択すると、選択可能なランキング調整クラスのリストに、コンテンツ・フィールドに適用されるクラスが表示されます。このフィールドを含む文書が検索結果に戻されると、フィールドの内容が文書サマリー域に表示されます。

チェック・ボックスをクリアすると、選択可能なランキング調整クラスのリストに、メタデータ・フィールドに適用されるクラスが表示されます。フィールドの内容は、検索結果の文書サマリー域に表示されません。

5. ランキング調整クラスを選択して、「OK」をクリックします。

追加したフィールドが「フィールドをランキング調整クラスにマップ」ページに表示されます。オプションを選択して、ランキング調整クラスを編集し、このフィールドが含まれている文書のスコアの決定に使用されるさまざまなランキング調整要因を構成します。

6. フィールドが重複文書の検出に使用されるかどうかを変更する場合、またはフィールドを別のランキング調整クラスにマップする場合は、「 編集」をクリックします。(エンタープライズ・サーチによって予約されているフィールドは編集できません。)
7. ランキング調整クラスからフィールドを除去する場合は、「 除去」をクリックします。(エンタープライズ・サーチによって予約されているフィールドは除去できません。)
8. 以前に索引付けされている文書に変更を適用するには、文書を再度クロールして索引付けします。

#### 関連概念

229 ページの『文書ランキング』

## ランキング調整クラスのランキング調整要因の構成

ランキング調整クラスに対して構成するランキング調整要因は、結果文書内の特定のフィールドの存在が照会にどのように関連するかについてのユーザーの見積もりを表します。ランキング調整クラスに最も高いランキング調整要因が指定されている場合、そのランキング調整クラスにマップされているフィールドが含まれている結果文書の重要度が高くなります。

#### 始める前に

ランキング調整クラスのランキング調整要因を構成するには、エンタープライズ・サーチ管理者役割の担当者であるか、そのコレクションのコレクション管理者である必要があります。

#### このタスクについて

システムは、ランキング調整クラスに構成されているランキング調整要因、デフォルトの静的ランキング・スコア、およびその他の係数を使用して、ランキング調整クラスにマップされているフィールドが含まれている結果文書の新規スコアを計算します。

## 手順

ランキング調整クラスのランキング調整要因を構成するには、次のようにします。

1. コレクションを編集し、「構文解析」ページを選択して、「**フィールドをランキング調整クラスにマップ**」をクリックします。
2. 「**フィールドをランキング調整クラスにマップ**」ページで、「**ランキング調整クラスの編集**」をクリックします。
3. 「**ランキング調整クラス**」ページで、変更したいランキング調整クラスを指定し、「 **編集**」をクリックします。
4. 「**ランキング調整クラスの編集**」ページで、ランキング調整要因の最大値と最小値に新規の値を指定します。両方の係数に同じ値を指定することもできます。
5. 「**OK**」をクリックします。
6. 変更を有効にするために、検索サーバーをモニターして、検索処理の停止および再始動のアイコンを選択します。ユーザーが照会を実行すると、このランキング調整クラスにマップされているフィールドが含まれている結果文書の相対重要度が、新規ランキング調整要因によって判別されます。

## 関連概念

229 ページの『文書ランキング』

## ランキング調整クラスのデフォルト値

エンタープライズ・サーチでは、検索結果における文書のランク付けに影響を与える 16 のランキング調整クラスが用意されています。

エンタープライズ・サーチでは、他の定義特性を持たないフィールドおよびテキストのスコアを計算するために、以下のフィールドが予約されています。

```
es_special_field.regular_text  
es_special_field.default_field  
es_special_field.default_metadata_field
```

予約フィールド以外のフィールドを予約フィールドが使用するランキング調整クラスにマップすることはできますが、予約フィールドを編集したり削除することはできません。

予約フィールド以外のすべてのフィールドについては、システムが文書のランク計算に使用するランキング調整要因を編集することができます。任意の数のフィールドを任意のランキング調整クラス (予約フィールドによって使用されるランキング調整クラスも含む) にマップすることができます。

以下の表に、ランキング調整クラス名、低い再呼び出し値を持つ照会用のデフォルト・ランキング調整要因、高い再呼び出し値を持つ照会用のデフォルト・ランキング調整要因、デフォルト構成でランキング調整クラスにマップされている事前定義フィールド名を示します。

デフォルトのランキング調整要因は、コレクション作成時にコレクションに対して選択された静的ランキング・メソッドによって異なります。オプションには、静的でないランキング、文書に対するリンク数によって決まるランク (Web ソースの場合)、文書日付によって決まるランクがあります。

表 6. ランキング調整クラスのデフォルト値

ランキング調整クラス名	デフォルト時の低いランキング調整要因および高いランキング調整要因			事前定義フィールド・マッピング
	静的でないランキング	文書リンク	文書日付	
コンテンツ・クラス A	低: 4 高: 2	低: 6 高: 1	低: 4 高: 2	es_special_field.regular_text
コンテンツ・クラス B	低: 5 高: 4	低: 7 高: 3	低: 5 高: 4	es_special_field.html_emphasized_text  次の HTML エレメントを含む: b、big、caption、dfn、em、h4、h5、h6、strong
コンテンツ・クラス C	低: 7 高: 4	低: 9 高: 3	低: 7 高: 4	es_special_field.html_headers  次の HTML エレメントを含む : h1、h2、h3
コンテンツ・クラス D	低: 2 高: 5	低: 1 高: 5	低: 2 高: 5	タイトル
コンテンツ・クラス E	低: 1 高: 1	低: 5 高: 10	低: 1 高: 1	es_special_field.anchor
コンテンツ・クラス F	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	es_special_field.anchor_same_dir
コンテンツ・クラス G	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	es_special_field.anchor_same_host
コンテンツ・クラス H	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	es_special_field.default_field
メタデータ・クラス A	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	es_special_field.default_metadata_field
メタデータ・クラス B	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	
メタデータ・クラス C	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	
メタデータ・クラス D	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	
メタデータ・クラス E	低: 1 高: 1	低: 5 高: 1	低: 1 高: 1	キーワード
メタデータ・クラス F	低: 1 高: 1	低: 3 高: 1	低: 1 高: 1	es_special_field.urlhost
メタデータ・クラス G	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	es_special_field.urlpath
メタデータ・クラス H	低: 1 高: 1	低: 1 高: 1	低: 1 高: 1	説明

## 関連概念

229 ページの『文書ランキング』

---

## エンタープライズ・サーチの検索アプリケーション

検索アプリケーションは、エンタープライズ・サーチ・システム内のコレクションおよび外部ソースの検索を可能にします。検索アプリケーションはいくつでも作成できます。また、1つの検索アプリケーションで、任意の数のコレクションおよび外部ソースを検索できます。

### サンプル検索アプリケーション

サンプル検索アプリケーションによって、エンタープライズ・サーチで使用可能な多くの検索、取り出し機能を実際に試すことができます。サンプル・アプリケーションは、企業の目標を反映する、対話式のカスタム検索アプリケーションを作成するために、IBM 検索および索引 API (SI-API) の使用方法を示す実施例でもあります。

デフォルトの構成ファイルのプロパティを変更しないかぎり、サンプル検索アプリケーションを使用して、エンタープライズ・サーチ・システム内のすべてのアクティブなコレクションおよび外部ソースを検索することができます。新しいコレクションおよび外部ソースをユーザーに使用可能にする前に、サンプル検索アプリケーションを使用してそれらをテストできます。

サンプル検索アプリケーションは、自動的にすべてのコレクションおよび外部ソースと関連付けられます。実稼働環境では、エンタープライズ・サーチ管理者が、さまざまなコレクションをどの検索アプリケーションで検索できるかを制御します。

### カスタム検索アプリケーション

作成する検索アプリケーションは、IBM WebSphere Application Server 環境で独立型 Web アプリケーションとして実行することも、IBM WebSphere Portal 環境でポートレットとして起動することもできます。検索および索引 API を使用することによって、サンプル検索アプリケーションのように、両方の環境でシームレスに機能する検索アプリケーションを設計できます。

検索アプリケーションのカスタマイズの支援として、検索アプリケーション・カスタマイザーを使用することができます。このアプリケーションでは、グラフィカル・インターフェースで選択を行い、それを行うことにより起こった変更の影響を表示することができます。変更を保存すると、検索アプリケーションの構成ファイルを更新します。

#### ヒント:

検索アプリケーション・カスタマイザーの使用法、および検索アプリケーションを WebSphere Portal のポートレットとしてインストールする方法の詳細な例については、IBM Redbook、IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios を参照してください。


#### 関連概念

200 ページの『文書の検索に対する索引付きオプション』

291 ページの『検索アプリケーション ID』  
292 ページの『文書レベルのセキュリティ』

 [検索および索引 API の概要](#)

 [照会構文](#)

 [照会プロパティの設定](#)

---

## コレクションとの検索アプリケーションの関連付け

新規の検索アプリケーションを使用するには、その検索アプリケーションを、それが検索できるコレクションに関連付けておく必要があります。

### 始める前に

検索アプリケーションを、それが検索できるコレクションに関連付けるには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### 手順

検索アプリケーションを 1 つ以上のコレクションに関連付けるには、次のようにします。

1. 管理コンソールのツールバーで「**セキュリティ**」をクリックします。
2. 「検索アプリケーション」ページで、「**検索アプリケーションの構成**」をクリックします。
3. 「検索アプリケーションの構成」ページで、「**検索アプリケーションの追加**」をクリックします。
4. 検索アプリケーションの名前を入力します。
5. アプリケーションが検索できるコレクションを選択します。
  - システムに追加するすべてのコレクションを検索アプリケーションがアクセスできるようにするには、「**すべてのコレクションと外部ソース**」をクリックします。
  - 指定するコレクションのみを検索アプリケーションがアクセスできるようにするには、「**特定のコレクションと外部ソース**」をクリックします。

このオプションを選択すると、コレクション名および外部ソース名のリストが表示されます。アプリケーションが検索できるコレクションごとに、その「**選択**」チェック・ボックスを選択します。

6. 「**OK**」をクリックします。

---

## サンプル検索アプリケーションの機能

エンタープライズ・サーチのサンプル検索アプリケーションを使用して、カスタム検索アプリケーションに構築できる検索機能のほとんどを実際に試してみることができます。

サンプル検索アプリケーションを使用して、一度にすべてのコレクションおよび外部ソースの検索を行うことができます。アプリケーション・プロパティのデフォルトが変更されていない限り、このアプリケーションを使用して、エンタープライズ・サーチ・システム内のすべてのコレクションおよび外部ソースを検索できます。

## 照会機能

これらの機能を使用して、次のことができます。

- 単純なフリー・テキストの照会を指定する。
- 検索結果の精度を上げるために、より複雑な照会を指定する。例えば、特定のフィールドまたは XML 文書を検索したり、照会構文を使用して特定の語句が含まれる、または除外された文書を検索したりできます。
- 検索する対象のコレクションおよび外部ソースを指定する。
- 特定のソース・タイプまたはすべてのソース・タイプを検索する。
- 特定のタイプの文書を検索する。例えば、Microsoft Word 文書のみを検索したり、PDF 形式の文書のみを検索できます。
- 使用する照会条件の言語を指定する。検索対象にする文書の言語も指定できます。
- コレクションの特定のサブセットを検索する。例えば、検索アプリケーションは、事前に定義された文書の範囲 (有効範囲) にビューを制限できます。あるいは、指定されたカテゴリーに属する文書のみを検索する照会を実行できます。
- 照会を拡張して、照会条件の同義語を含める。同義語辞書がコレクションに関連付けられている場合、照会条件の同義語を含む文書が検索結果に戻されます。

## 検索結果機能

これらの機能を使用して、次のことができます。

- 照会条件に一致する検索結果を表示する。
- 各ページに表示する結果文書数を制御し、結果セットの間で前方または後方を参照する。
- 結果文書についての詳細を表示、または非表示にする。例えば、文書の要旨を表示したり、あるいは、各結果文書内のフィールドの名前などの詳細を表示できます。
- 同じソースからの文書を縮小表示する。例えば、1 つのソースが 100 件の文書を返す場合、結果セットの中では、もっとも関連する 2 文書がグループ化されて表示されます。残りの 98 文書は、同じソースからもっと多くの文書を表示するオプションを選択すると表示されます。
- 関連性、文書日付、特定フィールドの値で文書をソートする。日付またはフィールドでソートする場合、結果の表示を昇順にするか、降順にするかを指定できます。
- 照会ストリングの中にミススペルの可能性のある語があれば、スペルの修正候補のプロンプトを出す。
- 結果の文書が属するカテゴリーについての情報を表示し (コレクションがカテゴリーを使用する場合)、特定のカテゴリーに属する文書のみを参照する。
- 照会条件を追加で指定し、検索結果内でさらに検索を行う。

## 文書検索機能

これらの機能を使用して、次のことができます。

- 文書 URI をクリックして文書を Web ブラウザーで開き、文書をリトリートする。 Lotus Notes クローラーまたは Lotus Domino Document Manager クローラーが DIIOP プロトコルを使用するように構成されている場合、これらのクローラーでクロールされた文書は、 Web ブラウザーの代わりに、 Lotus Notes クライアント・ビューアー・アプリケーションで表示できます。

文書レベルのセキュリティがクローラーに対して構成されている場合、セキュア・コンテンツへのアクセスを認可されているユーザーのみが文書をリトリートできます。

- クイック・リンクをクリックして、文書をリトリートする。クイック・リンクは、キーワードを URI に関連付けます。指定されたキーワードが照会の中に含まれていれば、関連した URI (これは、これらのキーワードに高い関連があると事前に決められている) が検索結果の先頭に表示されます。

---

## 検索アプリケーションのプロパティ

検索アプリケーションの構成ファイルを編集することによって、環境に合わせたオプションを指定したり、アプリケーションの外観を変更したり、ユーザーが検索アプリケーションを開始した後に使用できるオプションを制御したりすることができます。

プロパティの編集には、検索アプリケーション・カスタマイザーも使用できます。カスタマイザーで選択を行った場合には、変更の影響を即時に確認することができます。コレクションの検索や検索結果の表示のために指定したオプションで問題がなければ、そのオプションを保存して検索アプリケーションの構成ファイルを更新します。

**重要:** 検索アプリケーションを WebSphere Portal 内のポートレットとして実行する場合は、検索アプリケーション・カスタマイザーを使用して対話式で検索アプリケーションを変更することはできません。プロパティを編集し、ポートレット・インスタンスを WebSphere Portal 管理インターフェースで構成する必要があります。

エンタープライズ・サーチのサンプル検索アプリケーションの構成ファイルは、`config.properties` ファイルです。ここでは、このファイル内のプロパティについて説明し、デフォルトのプロパティを示します。カスタム検索アプリケーションの構成ファイルを作成する場合は、そうしたファイルのプロパティや、そのプロパティに指定される値が異なることがあります。

### 環境パラメーター

検索アプリケーションの操作を制御するオプションを指定できます。

#### **applicationName**

有効な検索アプリケーションの名前を指定します。デフォルト値は `Default` です。

別のアプリケーションをデフォルト・アプリケーションとして使用する場合、このデフォルト値を変更してください。



**ヒント:** アプリケーション名が Default の場合、このサンプル検索アプリケーションを使用して、すべてのコレクションおよび外部ソースを 1 回の照会で検索できます。

#### **timeout**

検索サーバーからの応答をどれだけ待機すれば検索要求がタイムアウトになるのかを秒数で指定します。この数値は整数でなければなりません (60.5 や sixty ではなく、60)。タイムアウト値を指定しない場合、デフォルト値は 30 秒です。

#### **hostname**

ご使用の WebSphere Application Server インスタンスをサポートするように構成された Web サーバーの、完全修飾ホスト名を指定します。デフォルト値は localhost です。

検索アプリケーションが正しく機能するようにするため、デフォルト値を、WebSphere Application Server が使用すると構成されている完全修飾ホスト名に変更してください。例えば、ローカル・コンピューターのホスト名が MyMachine で、Web サーバーのホスト名が www.ibm.com の場合、www.ibm.com を指定します。

#### **protocol**

Web サーバーと通信するためのプロトコルの http または https を指定します。ブランクの場合、デフォルト値は http です。

#### **port**

ご使用の WebSphere Application Server インスタンスをサポートするように構成された Web サーバーのポート番号を指定します。デフォルト値は、プロトコルが HTTP の場合の標準である 80 です。HTTPS プロトコルに使用される標準ポートは 443 です。

#### **trustStore**

HTTPS プロトコルを使用する場合、鍵ストア・ファイル (公開鍵を含むデータベース・ファイル) の完全修飾パスを指定します。この情報は、トラストストア と呼ばれ、トラステッド通信のために Secure Sockets Layer (SSL) プロトコルを使用できるようにします。Windows のパスを指定する場合、円記号は 2 つ目の円記号を付けてエスケープします。例:

x:¥¥Application Server¥¥webserver.key

#### **trustPassword**

HTTPS プロトコルを使用する場合、指定した鍵ストア・ファイルのパスワードを指定します。

#### **username**

この値は、検索アプリケーションによって、検索アプリケーションへのログイン時にユーザーが指定したユーザー名に自動的に設定されます。このユーザー名は、ユーザー認証のデフォルト動作をオーバーライドする場合にのみ指定します。このフィールドは、WebSphere Application Server でグローバル・セキュリティーを使用可能に設定した場合にのみ使用されます。

#### **password**

この値は、検索アプリケーションによって、検索アプリケーションへのログイン時にユーザーが指定したパスワードに自動的に設定されます。このパスワードは、ユーザー名を指定した場合にのみ指定します。このフィールド

は、WebSphere Application Server でグローバル・セキュリティーを使用可能に設定した場合にのみ使用されます。

**ssoCookieName**

シングル・サインオン (SSO) トークン・ストリングを含む Cookie の名前を指定します。デフォルト値は LtpaToken です。

**proxyHost**

検索サーバーへのアクセスにプロキシー・サーバーが必要な場合に、プロキシー・サーバーの完全修飾ホスト名を指定します。

**proxyPort**

指定したプロキシー・サーバー・ホストのポートを指定します。

**proxyUser**

プロキシー・サーバーが基本認証を必要とする場合に、プロキシー・サーバーへのログインに使用するユーザー名を指定します。

**proxyPassword**

指定したプロキシー・サーバー・ユーザー名のパスワードを指定します。

**filter** 検索結果内にリストされた文書をリトリートするのに使用するクラスを指定します。デフォルトのクラスは、`com.ibm.es.api.filters.SetDocumentURIFilterFetch` です。代わりにカスタムのクラスを使用して文書をリトリートする場合のみ、この値を変更してください。

**logging.level**

ログに出力する詳細の量を以下のように指定します。

**OFF** どのメッセージもログに記録されません。

**SEVERE**

重大な失敗を示すメッセージがログに記録されます。これはデフォルト値です。

**INFO** 通知メッセージがログに記録されます。

**FINE** 少量の詳細を含むトレース・メッセージをログに記録します。(このオプションは、Java `java.util.logging.Level` クラスの FINE ログ・レベルに対応します。)

**ALL** すべてのメッセージがログに記録されます。

## ソース・タイプ用アイコン

結果文書が属しているデータ・ソースのタイプを表すイメージをカスタマイズできます。ソース・タイプ用の以下のアイコンが `config.properties` ファイルに事前定義されています。これらは、OmniFind Enterprise Edition のインストール時にサポートされるクローラーと外部ソースを示します。

**documentSource.vbr.icon**

文書が Content Edition クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/sourceVBR.gif` です。



#### **documentSource.db2.icon**

文書が DB2 クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceDB2.gif です。



#### **documentSource.cm.icon**

文書が DB2 Content Manager クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceCM.gif です。



#### **documentSource.dominodoc.icon**

文書が Lotus Domino Document Manager クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceDominoDoc.gif です。



#### **documentSource.exchange.icon**

文書が Exchange Server クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceExchange.gif です。



#### **documentSource.database.icon**

文書が JDBC データベース・クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceJDBC.gif です。



#### **documentSource.nntp.icon**

文書が NNTP クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceNNTP.gif です。



#### **documentSource.notes.icon**

文書が Lotus Notes クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceNotes.gif です。



#### **documentSource.quickplace.icon**

文書が QuickPlace クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceWorkplace.gif です。



#### **documentSource.seedlist.icon**

文書がシード・リスト・クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceSeedlist.gif です。



#### **documentSource.unixfs.icon**

文書が UNIX ファイル・システム・クローラーによってクロールされたこ

とを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceUnixFS.gif です。



#### **documentSource.web.icon**

文書が Web クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceWeb.gif です。



#### **documentSource.wcm.icon**

文書が Web Content Management クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceWorkplace.gif です。



#### **documentSource.wps.icon**

文書が WebSphere Portal クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceWPS.gif です。



#### **documentSource.winfs.icon**

文書が Windows ファイル・システム・クローラーによってクロールされたことを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceWindowsFS.gif です。



#### **documentSource.ldap.icon**

文書が LDAP サーバー用に作成された外部ソースに属していることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceLDAP.gif です。



#### **documentSource.jdbc.icon**

文書が Java Database Connectivity (JDBC) データベース表用に作成された外部ソースに属していることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは /images/sourceJDBC.gif です。

## クライアント・ビューアー用アイコン

結果文書は Web ブラウザーで表示できます。DIIOOP プロトコルの使用を構成された Lotus Notes クローラーまたは Lotus Domino Document Manager クローラーによってクロールされた文書は、Lotus Notes クライアント・ビューアー・アプリケーションでも表示できます。

クライアント・ビューアー・アプリケーションで文書を表示できるようにするには、必ず、次のプロパティーが true に設定されるようにしてください。

```
clientViewer.show=true
```

クライアント・ビューアー・アプリケーションを表すイメージは、カスタマイズできます。次の例で、Lotus Notes アイコンは、文書をビューアー・アプリケーションで表示できることを示しています。

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

検索結果には、アイコンとクライアント・ビューアー・アプリケーションへのリンクが、以下のように表示されます。



## クライアント・ビューアー

### 文書フィールド

フィールドがあるデータ・ソースのタイプに対して、どのフィールドが結果文書に表示されるのかを制御できます。

**fields.URI prefix=**スペースで区切ったフィールド名リスト

URI 接頭部にあるコロン文字 (:) の前に円記号 (¥) を付けて、コロン文字をエスケープする必要があります。フィールド名リストを次の行に継続する場合、前の行が円記号 (¥) で終わるようにしてください。例えば、次のようにします。

```
fields.db2¥://=databasename tablename
fields.domino¥://=databasetitle filename creator
fields.dominodoc¥://=librarydbtitle documentdbtitle filename author
fields.exchange¥://=from creator
fields.file¥://=directory filename
fields.https¥://=documentID
fields.http¥://=documentID
fields.jdbc¥://=databasename tablename
fields.news¥://=group from
fields.quickplace¥://=placetitle roomtitle creator
fields.seedlist¥://=author
fields.vbr¥://=itemname repositorytype revisionuser
fields.wcm¥://=author owner modifier
fields.web¥://=
fields.wp6¥://=
fields.wps¥://=
```

### フィールド用アイコン

フィールドがあるデータ・ソースのタイプおよび文書に対して、フィールドを表すイメージをカスタマイズできます。文書サマリーより上のすべてのフィールドに、識別イメージが含まれます。フィールド用の以下のアイコンが `config.properties` ファイルに事前定義されています。



#### **field.icon.databasetitle**

文書タイトルが入っているフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/notesdb.gif` です。



#### **field.icon.databasesname**

文書が属しているデータベースの名前が入っているフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/db2.gif` です。



#### **field.icon.tablename**

文書が属している表の名前が入っているフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/table.gif` です。

**field.icon.directory**

文書が属しているディレクトリーの名前が入っているフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/closedFolder.gif` です。

**field.icon.filename**

文書のファイル名が入っているフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/document.gif` です。

**field.icon.documentID**

文書 ID が入っているフィールドであることを示すイメージ・ファイルのパスと名前を指定します。例えば、URL のイメージを指定するが、ユーザーにイメージを表示しない場合に、Web 文書でこのブランク・イメージを使用します。デフォルトのアイコンは `/images/dot.gif` です。

**field.icon.group**

文書 ID が入っているフィールドであることを示すイメージ・ファイルのパスと名前を指定します。例えば、URL のイメージを指定するが、ユーザーにイメージを表示しない場合に、Web 文書でこのブランク・イメージを使用します。デフォルトのアイコンは `/images/document.gif` です。

**field.icon.from**

文書を送信した人を識別するフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/author.gif` です。

**field.icon.creator**

文書の作成者を識別するフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/author.gif` です。

**field.icon.author**

文書の執筆者を識別するフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/author.gif` です。

**field.icon.revisionuser**

文書を改訂した人を識別するフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/author.gif` です。

**field.icon.owner**

文書の所有者を識別するフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/author.gif` です。



### **field.icon.modifier**

文書を変更した人を識別するフィールドであることを示すイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/author.gif` です。

## デフォルトのフィールド用アイコン

検索結果に表示されるフィールド用にアイコンが何も構成されていない場合に使用するイメージを指定できます。以下のデフォルトのフィールド用アイコンが `config.properties` ファイルに事前定義されています。



### **field.defaultIcon**

検索結果内のフィールド用のデフォルトのアイコンであるイメージ・ファイルのパスと名前を指定します。デフォルトのアイコンは `/images/database.gif` です。

## 日付フィールド

どのフィールドが日付フィールドであるかを指定できます。ここで指定するフィールド名は、検索結果内で日付データのようにフォーマットされます。日付のフォーマットは、Web ブラウザーでのロケール設定に一致します。

**date.fields=**スペースで区切ったフィールド名リスト

フィールド名リストを次の行に継続する場合、前の行が円記号 (¥) で終わるようにしてください。

例:

```
date.fields=modifieddate createddate
```

## 文書タイトル

文書タイトルのテキストをより分かりやすいデータで置換することにより、文書の代替タイトルを指定できます (つまり、タイトルを簡潔化 (*clean*) できます)。例えば、「Slide 1」のような内容を表示していないラベルが含まれた一連の文書タイトルがある場合、タイトルをそのまま表示する代わりに、検索結果で「Slide 1」を抑制するよう指定できます。結果文書を識別できるように、より意味のあるフィールド (ファイル名など) を代わりに使用することが考えられます。

また、文書タイトルから無意味な語を除去することにより、文書の代替タイトルを指定できます (つまり、タイトルを切り捨て (*truncate*) できます)。例えば、多くの結果文書が「Microsoft Word -」で始まる場合、反復される先頭のテキストを抑制することによって、検索結果を読みやすくすることができます。

**titles.clean=**コンマで区切ったタイトル・リスト

**titles.truncatePrefix=**コンマで区切った接頭部リスト

コンマで区切られたこれらのリストには、コンマを除き、スペースおよびその他の文字を入れることができます。リストを次の行に継続する場合、前の行が円記号 (¥) で終わるようにしてください。

例:

titles.clean=Slide 1, Layout 1, untitled, ¥  
Untitled Document, PowerPoint Presentation, ¥  
(no title for this page)

titles.truncatePrefix=Microsoft Word -, Microsoft Powerpoint -

## ユーザー・プリファレンスのデフォルト値

検索アプリケーションの「プリファレンス」ページのデフォルト値を指定できます。ユーザーがプリファレンスを変更した場合、そのユーザーの現行セッションのみ新しい値が有効になります。以下のプリファレンスが `config.properties` ファイルに事前定義されています。

### **preferences.resultsRange=10**

検索結果の各ページに 10 個の結果文書をリストできることを指定します。

### **preferences.siteCollapsing=Yes**

同じソースの URI が検索結果に縮小表示されることを指定します。サイト縮小表示は、結果が関連性でソートされる場合のみ有効です。Web および NNTP データ・ソースの場合、ルート・サイト URI (`www.ibm.com` など) に一致する URI が自動的に縮小表示されます。その他のデータ・ソース・タイプの場合やパス・レベルがこれより深い Web サイト (`www.ibm.com/hr` など) の場合は、エンタープライズ・サーチ管理コンソールでサイト縮小表示規則が構成されている必要があります。

### **preferences.spellCorrections=Yes**

ミスペルの可能性がある語を含んでいる照会をユーザーが実行した場合にスペル修正候補が表示されることを指定します。スペル提案が算出される前に、ストップワードが常に除去されることに注意してください。

### **preferences.extendedHighlighting=No**

文書サマリーのフィールドに加えて、追加のフィールド (文書タイトルなど) で、照会条件が強調表示されないことを指定します。

## デフォルトのコレクションおよび外部ソース

「プリファレンス」ページと「拡張検索」ページで事前選択されるコレクションおよび外部ソースを指定できます。ユーザーはこのデフォルトのセットを編集して、デフォルトで使用可能にされているより少ないコレクションおよび外部ソースを検索対象にできます。コレクションおよび外部ソースのセットをここで制限すると、ユーザーはプリファレンスまたは拡張検索オプションを変更する際に、検索アプリケーションで使用可能になっているコレクションまたは外部ソースのうち任意のものを選択できます。

### **preferences.defaultCollections=\***

#### **preferences.defaultCollections=**スペースで区切ったコレクション ID リスト

すべてのコレクションおよび外部ソースを検索対象として使用可能にするには、アスタリスク (\*) を指定します (コレクションおよび外部ソースは、エンタープライズ・サーチ管理コンソールで、検索アプリケーションと関連付けられていなければなりません)。これは、`config.properties` ファイルでのデフォルト設定です。



プリファレンスまたは拡張検索オプションを変更していないユーザーが何を検索対象とするのかを制限するには、デフォルトでユーザーが検索するようにしたいコレクションおよび外部ソースのコレクション ID を指定します。

例:

```
preferences.defaultCollections=*  
preferences.defaultCollections=coll_id1 coll_id2
```

## 検索結果の追加情報

検索結果で提供される情報量をカスタマイズすることができ、また、ユーザーが検索結果をフィルターできるかどうかを制御できます。config.properties ファイル中のデフォルト設定は、次のとおりです。

### refreshButton.show=false

基本検索ページに「リフレッシュ」ボタンが表示されるかどうかを制御します。「リフレッシュ」ボタンは、拡張検索では常に使用できます。このオプションを true に設定すると、ユーザーは検索に使用可能なコレクションおよび外部ソースのリストをリフレッシュできます。

検索アプリケーション・カスタマイザーを使用する場合、「リフレッシュ」ボタンは不要です。

検索アプリケーション・カスタマイザーを使用しない場合、構成ファイルに加えた変更をテストするときに、「リフレッシュ」ボタンを表示するとよいでしょう。変更を保存した後、「リフレッシュ」をクリックすると、その変更が検索アプリケーションにどのように影響するのかが確認できます。「リフレッシュ」ボタンがない場合、WebSphere Application Server で ESSearchServer アプリケーションを再始動しないと、変更は有効になりません。

検索に使用可能なコレクションまたは外部ソースがない場合 (例えば、誤ったホスト名が指定されている場合や、検索サーバーが開始されていない場合、ESSearchServer アプリケーションが WebSphere Application Server で開始されなかった場合)、問題のトラブルシューティングに役立つよう、「リフレッシュ」ボタンが自動的に表示されます。

### builtQueryString.show=false

結果文書リストの前にある領域での、完全に展開された照会構文の表示を制御します。処理された実際の照会が表示されるようにするには、このオプションを true に設定してください。

### extraQueryData.show=false

照会に関する追加情報の表示を制御します。ACL 制約、検索されるコレクションや外部ソースの名前、照会言語に関する情報を表示する場合に、このオプションを true に設定します。

### refineResults.show=true

ユーザーが追加の照会条件を指定して検索結果を絞り込むことができるかどうかを制御します。このオプションを true に設定すると、「結果内の検索」というラベルが付いた照会ボックスが、検索結果ページの下部に表示されます。

### sorting.show=true

検索結果のソートに関するオプションを表示するかどうかを制御します。こ

のオプションを `false` に設定すると、検索結果をソートするための「ソート基準」および「ソート順序」オプションが抑制されます。

#### **sourceTypeFilter.show=true**

ソース・タイプによる結果のフィルター処理に関するオプションを検索結果に表示するかどうかを制御します。ユーザーがソース・タイプで結果をフィルター処理できないようにする場合、このオプションを `false` に設定します。

ユーザーが文書タイプで結果をフィルター処理できないようにするには、構成ファイル内の選択した文書タイプ・エントリー、またはすべての文書タイプ・エントリー (`documentType.label=document_types`) を削除します。

#### **filter.showOnTwoLines=true**

ソース・タイプによる結果のフィルター処理とファイル・タイプによる結果のフィルター処理に関するオプションを、検索結果に 1 行で表示するか、2 行で表示するかを制御します。ユーザーは、検索結果の表示中に、ソース・タイプの選択とファイル・タイプの選択を行って、選択したフィルターに一致する結果文書のみを表示できます。

検索結果の表示に使用できるスペース量を最大化するには、このプロパティを `false` に設定します。使用可能なフィルターが 1 行を超える場合は特に、このプロパティを `true` に設定して各フィルターが別の行に表示されるようにすると、フィルターの読みやすさが増します。

#### **clientViewer.show=true**

結果文書の表示に Lotus Notes クライアント・ビューアー・アプリケーションを使用するかどうかを制御します。Lotus Domino 文書の表示にビューアー・アプリケーションを使用しない場合、このオプションを `false` に設定します。

#### **showDetails.show=true**

検索結果での「詳細表示」および「詳細を非表示」リンクの表示を制御します。ユーザーが結果文書に関する追加の詳細を表示できるようにしない場合は、このオプションを `false` に設定します。

#### **showDetailsImage.show=true**

ウィンドウでの結果文書に関する詳細の表示を制御します。ユーザーが、文書 URI にカーソルを合わせることによって結果文書に関する追加の詳細を表示できるようにしない場合は、このオプションを `false` に設定します。

#### **numberSearchResultsReturned.show=true**

検索結果の総数を表示するかどうかを制御します。検索結果に戻された文書の数をユーザーに表示しない場合は、このオプションを `false` に設定します。

#### **showMessage.error=true**

エラー・メッセージの表示を制御します。検索アプリケーションの上部にエラー・メッセージを表示しない場合、このオプションを `false` に設定します。

#### **showMessage.warning=true**

警告メッセージの表示を制御します。検索アプリケーションの上部に警告メッセージを表示しない場合、このオプションを `false` に設定します。

### **showMessage.info=true**

通知メッセージの表示を制御します。検索アプリケーションの上部に通知メッセージを表示しない場合、このオプションを `false` に設定します。

### **showMessage.success=true**

成功メッセージの表示を制御します。検索アプリケーションの上部に、アクションの正常終了を示すメッセージを表示しない場合、このオプションを `false` に設定します。

## **カスタム・バナーおよびロゴ**

検索アプリケーションの上部にあるバナー領域に表示されるイメージをカスタマイズできます。例えば、OmniFind Enterprise Edition のデフォルトのイメージを、ご使用のエンタープライズ・ブランドを反映するイメージで置き換えることができます。バナーが表示されないようにするには、以下の一方または両方の行をコメント行にしてください。 `banner.icon` プロパティは、バナー領域の左側に表示されるグラフィックを示します。 `banner2.icon` プロパティは、バナー領域の右側に表示されるグラフィックを示します。

```
banner.icon=/images/WS_II_OFEdition.gif  
banner2.icon=/images/WS_II_mosaic.gif
```

## **カスタム背景イメージ**

検索アプリケーションのページ背景に表示されるイメージをカスタマイズできます。例えば、エンタープライズ・サーチのデフォルトのイメージを、ご使用のエンタープライズ・ブランドを反映するイメージで置き換えたい場合があります。背景イメージがページに表示されないようにするには、以下の行の 1 つ以上をコメント行にしてください。

```
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
```

## **リンク**

`config.properties` ファイルの `Links` 領域にあるプロパティは、検索アプリケーション・ページの名前が、ツールバーおよびタブ付きページに表示される代わりに、各ページにリンクとして表示されるようにします。リンクの表示は、検索アプリケーションをポートレットとして実行し、ポータル・ページで検索アプリケーションの表示に使用されるスペース量を最小化する場合に役立ちます。

ツールバーおよびタブ付きページでオプションを選択することによって検索アプリケーションをナビゲートするほうが望ましい場合は、これらの行をコメント化してください。

## **検索タブ**

`config.properties` ファイルの `Search tabs` 領域にあるプロパティは、検索アプリケーションの「検索」ビュー内のタブ付きページ（「基本検索」、「拡張検

索」、および「カテゴリー・ツリー」) に使用される Java Server Pages (JSP) の名前を指定します。Java プログラミングおよび JSP に精通している場合以外は、これらのページを編集しないでください。

この領域のカスタマイズ例には、以下のものがあります。

- 検索アプリケーションを、タブ付きページに別の外観を提供するカスタム JSP に誘導する。
- カテゴリー・ツリーのエントリーをコメント化する。例えば、コレクションに対してカテゴリーを構成しない場合、検索アプリケーションで「カテゴリー・ツリー」ページを表示する必要はありません。
- タブ付きページのエントリーを `config.properties` ファイルの Toolbars 領域にコピーし、これらの行をコメント化する。例えば、ツールバーのみが表示されればよく、タブ付きページはまったく表示する必要がない場合があります。

## ツールバー

`config.properties` ファイルの Toolbars 領域にあるプロパティは、検索アプリケーションのツールバーに使用される Java Server Pages (JSP) の名前を指定します。Java プログラミングおよび JSP に精通している場合以外は、これらのページを編集しないでください。

この領域のカスタマイズ例には、以下のものがあります。

- 検索アプリケーションを、ツールバーに別の外観を提供するカスタム JSP に誘導する。
- 表示されないようにする項目について、ツールバーのエントリーをコメント化する。例えば、「製品情報」ページへのリンクをツールバーに組み込む必要がない場合があります。
- 「拡張検索」ページを表示する機能を `config.properties` ファイルのタブ領域から移動し、このオプションをツールバーでのみ使用可能にする。

## 分かりやすい文書タイプのラベル

実際の文書タイプ名を、より正確で分かりやすい語にマップすることによって、文書タイプ・フィルターの読みやすさを向上させることができます。検索アプリケーションで使用可能な文書タイプは、検索および索引 API (SI-API) の `AvailableDocumentTypes` クラスで定義されています。利便性のため、使用可能な文書タイプは、`config.properties` ファイルの最後にもリストされています。

**documentType.ラベル=スペース**で区切った文書タイプのリスト

検索結果で文書タイプ・フィルター行に表示される名前と、ユーザーがそのフィルターを選択したときに表示される実際の文書タイプのリストを指定します。

例えば、ラベル `html` を指定し、さまざまな Web 文書のファイル拡張子および MIME タイプをその名前にマップします。ユーザーが検索結果をフィルターに掛けるために `html` をクリックすると、指定された拡張子および MIME タイプの文書のみが表示されます。

文書タイプの以下のマッピングが `config.properties` ファイルに事前定義されています。

```
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
```

文書タイプ・ラベルに指定した値が `application.properties` ファイルの中のプロパティの名前と一致すると、ここで指定した値ではなく、`application.properties` ファイルの中のプロパティの値が表示されます。例えば、ファイル・タイプ・フィルターのラベルとして `documentType.unixfs` を指定すると、`application.properties` ファイルの中の `unixfs` プロパティの値 (**UNIX ファイル・システム**) が、クリックできるファイル・タイプ・フィルター名として表示されます。

## カスタム・フィルター

検索結果の表示をフィルター操作するためのカスタム照会を指定できます。

### `filterCustom`.ラベル =照会条件

検索結果でカスタム・フィルター行に表示される名前と、ユーザーがそのフィルターを選択したときに検索結果を絞り込むための照会を指定します (ユーザーは、検索結果の表示中に、カスタム・フィルターを選択して、事前定義済み照会に一致する結果文書のみを表示できます)。

以下の例では、人的資源 (`hr`) データベースに属している文書のみが表示されるよう、検索結果がフィルターに掛けられます。

```
filterCustom.HR_database_only=databasename::hr
```

ユーザーが検索結果をフィルターに掛けるために **HR\_database\_only** をクリックすると、照会 `databasename::hr` が処理されます。検索結果が表示されるとき、`hr` データベースからの文書のみがリストされます。

カスタム・フィルター・ラベルに指定した値が `application.properties` ファイルの中のプロパティの名前と一致すると、ここで指定した値ではなく、`application.properties` ファイルの中のプロパティの値が表示されます。例えば、カスタム・フィルター・ラベルとして

```
filterCustom.hostData=
```

を指定すると、`application.properties` ファイルの中の `hostData` プロパティの値 (**サーバー設定**) が、クリックできるカスタム・フィルター名として表示されます。

`config.properties` ファイルには、例として幾つかのカスタム・フィルターがコメント化されて含まれています。

## 重複検出

文書がエンタープライズ・サーチ索引に追加されると、ユーザーが検索結果に何度も出てくる同じ文書を参照しないように、重複を除去するための分析が行われます。検索結果をさらにフィルタリングするために、お互いにほとんど同じである文書を抑制し、それらの文書が検索結果に表示されないようにするオプションを指定することができます。

### **preferences.nearDuplicateDetection=No**

ほぼ同一の文書が照会処理にフィルタリングされないように指定します。

Yes を指定すると、類似したタイトルやサマリーを持つ文書は、ユーザーが検索結果を表示する際に抑制されます。一部の文書が結果セット内の他の文書と類似しているために省略されたことを、メッセージで通知します。ユーザーは、抑制を無効にするリンクをクリックして、結果セット内のすべての文書を表示することができます。

ほぼ同一の文書を抑制するには、検索アプリケーション用の Search and Index API (SI-API) Query オブジェクトで、NearDuplicateDetection ストリングを Yes に設定した setProperty メソッド (例えば、`query.setProperty("NearDuplicateDetection", "Yes");`) を指定する必要があります。

## **上位結果の分析 (メタデータ・フィールドのバー・チャート)**

上位結果の分析を表すバー・チャートを表示するオプションを指定することができます。デフォルトでは、上位 500 件の結果が表示されます。各チャートは 1 つのメタデータ・フィールドに対応し、チャート内の各バーはフィールド値に対応します。バーの長さは、そのフィールド値の発生頻度を示します。バーが長いほど、そのフィールド値が多く出現しています。

クローラーのクロール・スペースの構成時に、メタデータ・フィールドのオプションを指定することができます。メタデータ・フィールドの上位結果チャートを構成するには、そのフィールドが検索可能フィールドであること、および検索結果でそのフィールドを返すことができることを、指定する必要があります。

バー・チャート用に構成するプロパティの形式は次のとおりです。ここで、**数** は検索アプリケーションでそのバー・チャートを一意的に識別する番号、**オプション** はバー・チャート・オプション、**値** はそのオプションの値です。

`topResultsCharts数.オプション=値`

**topResultsCharts数.titleKey=アプリケーション・キー**

チャートのタイトルを指定します。ここで、**数** は検索アプリケーションでそのチャートを一意的に識別する番号、**アプリケーション・キー** はチャート・タイトルのラベルです。このラベルには、検索アプリケーションの `application.properties` ファイルにあるキーまたはここで指定する値を使用できます。次の例では、チャート・タイトルのラベルは、`application.properties` ファイル

`topResultsCharts3.titleKey=topResults.mostRecentDocuments` にある `topResults.mostRecentDocuments` キーの値によって指定されます。

もう 1 つの例として、チャートのタイトルは、ここで指定したとおりの値になります。`topResultsCharts3.titleKey=Organizations`

**topResultsCharts数.enable=true**

ユーザーが検索結果を表示するときに、このバー・チャートを表示するかどうかを指定します。 `false` を指定すると、チャートは表示されません。

**topResultsCharts数.fieldName=フィールド名**

このチャートについて値を分析するメタデータ・フィールドの名前を指定し

ます。例えば、`databaseTitle` などです。追加するチャートごとに異なるフィールド名を指定する必要があります (同じフィールドを複数のチャートで使用することはできません)。

**topResultsCharts数.maxValues.collapsed=数**

このチャートに表示する縮小表示された項目の数を指定します。例えば、上位 5 個の最頻出値をこのフィールドに表示するには 5 を指定します。

**topResultsCharts数.maxValues.expanded=数**

チャートの表示が完全に拡張されたときに、このチャートに表示する項目の数を指定します。例えば、チャートの拡張時に 10 個以内の異なるフィールド値を表示するには 10 を指定します。

**topResultsCharts数.fieldValueSeparator=character**

分析するフィールド内の値を区切る文字を指定します。例えば、セミコロンで区切られた複数の値 (`agent;seller;broker` など) がフィールドに含まれている場合は、このプロパティを使用してセミコロン (;) をフィールド値の区切り文字として識別できます。その結果、各値は独立した項目としてバー・チャートに追加されます。このオプションを指定しないと、フィールド値全体が 1 つの項目としてチャートに追加されます。

**topResultsCharts数.canUserChangeFieldName=true**

検索結果を表示するときに、ユーザーが別のフィールドを選択して、そのフィールドの上位結果を表示できるかどうかを指定します。 `false` を指定すると、ユーザーは検索結果を表示するときに、分析される別のフィールドを選択できません。

`true` を指定すると、最初の上位 500 件の結果にあったすべてのフィールドのリストが現行フィールドの結果と一緒に表示されます。ユーザーがこのリストからフィールドを選択すると、チャート・ラベルが「動的フィールド・チャート」に変わり、選択したフィールドのバー・チャート結果は、ユーザーが別のフィールドを選択するかブラウザを閉じるまで表示されます。次回にユーザーが検索アプリケーションを実行すると、元のフィールドの結果が表示されます。

**topResultsCharts数.width=数**

バー・チャートの表示幅をピクセル数で指定します。例えば、300 などです。


**topResultsCharts数.barheight=数**

バー・チャート内の各バーの高さをピクセル数で指定します。例えば、10 などです。

**topResultsCharts数.color=#カラー・コード**

バー・チャート内のバーの基本カラーの 16 進コードを指定します。デフォルト値は青色 (#0309C0) です。

**topResultsCharts数.color.gradient=#カラー・コード**

バー・チャート内のバーのグラデーション・カラーを指定します。デフォルト値は青緑色 (#00FFFF) です。特定のフィールド値の結果の件数が高頻度から低頻度へと変化するにつれて、バーの色調は `color` オプションに指定された値から `color.gradient` オプションに指定された値へと変わっていきます。例: 

### **topResultsCharts数.sortKey=frequency**

バー・チャート内の項目のソート方法を指定します。サポートされる値:

**none** バー・チャート内の項目はソートされません。

**label** バー・チャート内の項目はフィールド値に従ってソートされます。カスタム・ラベルを `topResultsCharts.数listOfLabels.prefixKey=フィールド名` プロパティーで構成すると、バー・チャート内の項目はラベル名によってもソートされます。

#### **frequency**

バー・チャート内の項目は、フィールド値ごとの返された結果の件数に従ってソートされます。

### **topResultsCharts数.sortOrder=descending**

**label** または **frequency** によってソートされる項目の場合は、ソート順を指定します。サポートされる値:

#### **ascending**

**label** によってソートされる項目は、a から z のアルファベット順にリストされます。**frequency** によってソートされる項目の場合は、発生頻度の高いフィールド値が発生頻度の低いフィールド値よりリスト内で下位に表示されます。

#### **descending**

**label** によってソートされる項目は、z から a の逆アルファベット順にリストされます。**frequency** によってソートされる項目の場合は、発生頻度の高いフィールド値が発生頻度の低いフィールド値よりリスト内で上位に表示されます。

### **topResultsCharts数.listOfLabels.prefixKey=フィールド名**

オプション。一定のフィールド値について情報が常に表示されるように指定できます。ここで **フィールド名** は、このチャートについて値が分析されるフィールドを示します。このプロパティーに指定したラベルは、そのフィールド値が検索結果に現れない場合でも、常にバー・チャートに表示されません。

フィールドのラベルはいくつでも構成できます。各ラベルごとに、次の形式を持つ 2 つのプロパティーを指定します。

フィールド名数.**value=値**

フィールド名数.**displayValue=display\_value**

各要素の意味:

**フィールド名**

ラベルを構成するフィールドの名前です。

**number**

ラベルを一意的に識別する番号です。

**value** 分析結果に表示する値です。

**display\_value**

バー・チャートに表示するラベルです。このラベルには、検索アプリケーションの `application.properties` ファイルにあるキーまたはここで指定する値を使用できます。



例:

```
topResultsCharts1.listOfLabels.prefixKey=databasetitle

databasetitle1.value=JK Enterprises Articles & Papers
databasetitle1.displayValue=Articles & papers

databasetitle2.value=JK Enterprises Blank Forms
databasetitle2.displayValue=Blank forms

databasetitle3.value=JK Enterprises Bulletins & Guidelines
databasetitle3.displayValue=Bulletins & Guidelines
```

## 上位結果の分析 (カスタム HTML)

カスタム Java クラスで

`com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` API を拡張することによって上位結果チャートを表示するオプションを指定できます。この方法によって結果を表示する場合は、HTML を使用して、メタデータ・フィールドを検索する照会だけでなく、検索結果の戻りもフォーマットできます。エンタープライズ・サーチ・コレクション内の文書の結果や、外部リポジトリおよび Web サイトの検索の結果を返すことができます。

検索アプリケーションのチャートはいくつでも構成できます。次のプロパティを使用して各チャートにタイトルを割り当てます。

**topResultsCharts数.titleKey=アプリケーション・キー**

チャートのタイトルを指定します。ここで、**数** は検索アプリケーションでそのチャートを一意的に識別する番号、**アプリケーション・キー** はチャート・タイトルのラベルです。このラベルには、検索アプリケーションの `application.properties` ファイルにあるキーまたはここで指定する値を使用できます。次の例では、チャート・タイトルのラベルは、`application.properties` ファイル `topResultsCharts3.titleKey=topResults.mostRecentDocuments` にある `topResults.mostRecentDocuments` キーの値によって指定されます。

チャート内の各項目は 1 つの検索結果値に対応し、次の形式を持つ一組のプロパティから構成されます。ここで、**数** はそのチャートを一意的に識別する番号、**オプション** はチャート・オプション、**値** はそのオプションの値です。

`topResultsCharts数.オプション=値`

**topResultsCharts数.enable=true**

ユーザーが検索結果を表示するときに、このチャートを表示するかどうかを指定します。 `false` を指定すると、チャートは表示されません。

**topResultsCharts数.maxValues.collapsed=数**

このチャートに表示する縮小表示された項目の数を指定します。例えば、基準と一致する上位 5 個の最頻出結果を表示するには 5 を指定します。

**topResultsCharts数.width=数**

チャートの表示幅をピクセル数で指定します。例えば、400 などです。

**topResultsCharts数.dynamicChartClass=custom\_class**

`com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` API を拡張するカスタム Java クラスの名前を指定し、チャートでの出力の表示方法を定義します。次に例を示します。

```
topResultsCharts.3.dynamicChartClass=com.ibm.es.searchui.charts.servlet.  
DynamicMostRecentDocuments
```

以下のプロパティは、例としてサンプル検索アプリケーションのデフォルトの `config.properties` ファイルに含まれているもので、サンプルの `DogearSearchResults` Java クラスによって使用されます。カスタム検索アプリケーションに組み込むことができる照会パラメーターについては、Dogear API の資料を参照してください。

以下のプロパティは検索アプリケーション・カスタマイザーでは設定できません。

**topResultsCharts数.xml.fileName=style\_file**

チャート内の上位結果の表示をフォーマットするために使用する XSL スタイル・シートのパスと名前を指定します。例えば、`/styles/dogear.xml` です。

**topResultsCharts数.url=url**

検索する URL を指定します。

**topResultsCharts数.url.parameters=ps=数**

ページ・サイズを指定します。例えば、`ps=3` はページ・サイズを 3 に制限します。

これは、単に Dogear REST API で定義されたとおりのパラメーター・ストリングです。

## 上位結果の分析 (結果の最大数)

**topResult.resultSize=数**

上位結果分析から返す結果の件数のデフォルト値および最大値は 500 です。この値は減らすことはできますが、増やすことはできません。例えば、上位結果分析で要求タイムアウトの問題が発生した場合は、より低い値を指定することができます。このプロパティは検索アプリケーション・カスタマイザーでは設定できません。

### 関連概念



照会プロパティの設定



上位結果の表示用の Java クラス

## サンプル検索アプリケーション・プロパティの編集

エンタープライズ・サーチのサンプル検索アプリケーションは、システムにあるすべてのアクティブなコレクションおよび外部ソースを検索できます。構成ファイルを編集して、ご使用の Web サーバー環境に合うようにオプションを指定したり、デフォルト・アプリケーションとして別の検索アプリケーションを使用したり、検索アプリケーションの開始時に表示されるオプションを制御したりできます。

### このタスクについて

インストール・プログラムは、エンタープライズ・サーチ用の検索サーバー上の IBM WebSphere Application Server に、エンタープライズ・サーチのサンプル検索

アプリケーションをデプロイします。この検索アプリケーションを構成するには、そのアプリケーションと一緒にデプロイされる構成ファイル `config.properties` を編集します。

変更を有効にするには、WebSphere Application Server で ESSearchServer アプリケーションを停止し、再始動する必要があります。

## 手順

サンプル検索アプリケーションのプロパティを編集するには、次のようにします。

1. エンタープライズ・サーチ管理者として検索サーバーにログインします。
2. 標準のテキスト・エディターで、`config.properties` ファイルを編集します。

`config.properties` ファイルは、以下の場所にインストールされています。ここで、`ES_INSTALL_ROOT` は、検索サーバー上の OmniFind Enterprise Edition インストール・ディレクトリーです。

```
ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/  
    ESSearchApplication.war/WEB-INF/config.properties
```

3. ご使用の Web サーバー環境に関する情報および検索プリファレンスを指定するプロパティを編集し、ファイルを保存してクローズします。ファイル中のポンド記号文字 (#) は、コメント行を示します。
4. ESSearchServer アプリケーションを停止し、再始動します。

### AIX、Linux、または Solaris

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```


### Windows

```
stopServer ESSearchServer  
startServer ESSearchServer
```

これらのスクリプトは、`WAS_INSTALL_ROOT/AppServer/bin` ディレクトリーにあります。

- WebSphere Application Server バージョン 5 の場合、デフォルトのインストール・パスは、AIX システムでは `/usr/WebSphere`、Linux または Solaris システムでは `/opt/WebSphere`、Windows システムでは `C:\Program Files\WebSphere` です。
- WebSphere Application Server バージョン 6 の場合、デフォルトのインストール・パスは、AIX システムでは `/usr/IBM/WebSphere`、Linux または Solaris システムでは `/opt/IBM/WebSphere`、Windows システムでは `C:\Program Files\IBM\WebSphere` です。

## 関連概念

 照会プロパティの設定

 上位結果の表示用の Java クラス

## 関連タスク

274 ページの『検索サーバーがセキュア (SSL) 要求のみを受け入れるようにする構成』

## 検索アプリケーションのカスタマイズ

検索アプリケーション・カスタマイザーは、エンタープライズ・サーチの検索アプリケーションやカスタム検索アプリケーションのカスタマイズに使用できるグラフィカル・インターフェースです。

### 制約事項

検索アプリケーション・カスタマイザーは、スタンドアロン・アプリケーションとして使用できます。WebSphere Portal 内の検索アプリケーション・カスタマイザーを起動して、ポートレットとして実行する検索アプリケーションをカスタマイズすることはできません。ポートレットとして実行する検索アプリケーションをカスタマイズするには、WebSphere Portal 管理インターフェースを使用してプロパティを編集し、ポートレット・インスタンスを構成する必要があります。

### このタスクについて

検索アプリケーション・カスタマイザーを使用すると、行う変更を視覚化することや、構成ファイルを編集せずに検索アプリケーションを変更することが可能になります。例えば、バナー・イメージや背景イメージの変更、検索インターフェースのレイアウトの変更、検索結果の処理オプションの指定を行えます。

検索アプリケーション・カスタマイザーで選択を行うと、選択の影響が表示されます。変更内容を保存すると、検索アプリケーションの構成ファイルが更新されます。

変更を有効にするには、WebSphere Application Server で ESSearchServer アプリケーションを停止し、再始動する必要があります。

### ヒント:

検索アプリケーション・カスタマイザーの使用法の詳細な例については、IBM Redbook、IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios を参照してください。

### 手順

検索アプリケーションをカスタマイズするには、次のようにします。

1. エンタープライズ・サーチのサンプル検索アプリケーションをカスタマイズするには、Web ブラウザーに検索アプリケーション・カスタマイザーの URL を入力します。次に例を示します。

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

`SearchServer.com` は、検索サーバーのホスト名です。

Web サーバーがポート 80 を使用するように構成されていない場合は、正しいポート番号を指定することも必要です。次に例を示します。

<http://SearchServer.com:9080/ESSearchApplication/palette.do>

**ヒント:** エンタープライズ・サーチ管理者であれば、エンタープライズ・サーチ管理コンソールで「**検索カスタマイザー**」オプションを選択して、検索アプリケーション・カスタマイザーを開くこともできます。

2. カスタム検索アプリケーションをカスタマイズするには、検索アプリケーション・カスタマイザーの URL を入力し、検索アプリケーションの構成ファイルの名前を付加します。次に例を示します。

<http://SearchServer.com/ESSearchApplication/palette.do?configFile=/WEB-INF/myConfig.properties>

指定したファイルが存在しない場合、サンプル検索アプリケーションの `config.properties` ファイルの値が表示されます。

**ヒント:** 検索アプリケーションで使用する構成ファイルを指定するために、検索アプリケーション・カスタマイザーの開始後に「**ロード**」をクリックしてファイルの名前を指定することもできます。

3. WebSphere Application Server でグローバル・セキュリティーが使用可能になっている場合は、有効なユーザー ID とパスワードを使ってログインします。
4. 検索サーバーに関する情報、表示するメッセージのタイプ、照会および検索結果オプション、検索結果におけるさまざまな文書タイプの識別に使用するイメージの名前など、カスタマイズするオプションを選択します。

オプション指定のヘルプについては、「**カスタマイザーのヘルプ**」をクリックします。検索結果がどう表示されるかなど、変更による影響を確認するには、照会を入力して「**検索**」をクリックします。

5. 行った選択でよければ、「**保存**」をクリックして構成ファイルを更新します。「**リセット**」をクリックすると、検索アプリケーション・カスタマイザーで表示されるオプションが、最後に保存された構成ファイルでの値に戻されます。
6. 検索サーバーで、エンタープライズ・サーチ管理者としてログインし、ESSearchServer アプリケーションを一度停止してから再始動します。

#### **AIX、Linux、または Solaris**

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

#### **Windows**

```
stopServer ESSearchServer  
startServer ESSearchServer
```

これらのスクリプトは、`WAS_INSTALL_ROOT/AppServer/bin` ディレクトリーにあります。

- WebSphere Application Server バージョン 5 の場合、デフォルトのインストール・パスは、AIX システムでは `/usr/WebSphere`、Linux または Solaris システムでは `/opt/WebSphere`、Windows システムでは `C:\Program Files\WebSphere` です。

- WebSphere Application Server バージョン 6 の場合、デフォルトのインストール・パスは、AIX システムでは /usr/IBM/WebSphere、Linux または Solaris システムでは /opt/IBM/WebSphere、Windows システムでは C:\Program Files\IBM\WebSphere です。

#### 関連タスク

274 ページの『検索サーバーがセキュア (SSL) 要求のみを受け入れるようにする構成』

## サンプル検索アプリケーションの複製

独自の検索アプリケーションを作成するためのモデルとしてサンプル検索アプリケーションを使用するには、`config.properties` ファイルをコピーするか、検索アプリケーション・カスタマイザーを使用できます。

#### このタスクについて

より簡単にカスタム検索アプリケーションを作成するには、サンプル検索アプリケーション用に指定した構成オプションを複製してから、変更するオプションをカスタマイズすることができます。

サンプル検索アプリケーションを複製した後、構成ファイルの名前を指定して、新規の検索アプリケーションを開始します。また、検索アプリケーション・カスタマイザーを使用して構成ファイルの名前を指定し、新規の検索アプリケーションをカスタマイズすることもできます。

サンプル検索アプリケーションを複製することによって、特定の目的または対象者用の検索アプリケーションを迅速に作成することができます。例えば、人事部門の従業員向けの検索アプリケーションと、営業担当者向けの検索アプリケーションを作成する場合があります。

変更を有効にするには、WebSphere Application Server で ESSearchServer アプリケーションを一度停止してから、再始動する必要があります。

#### 手順

サンプル検索アプリケーションを複製するには、次のようにします。

1. 構成ファイルを編集する場合:
  - a. サンプル検索アプリケーションの `config.properties` ファイルをコピーし、ファイル名を変更します。

`config.properties` ファイルは、以下の場所にインストールされています。ここで、`ES_INSTALL_ROOT` は、検索サーバー上の OmniFind Enterprise Edition インストール・ディレクトリーです。

```
ES_INSTALL_ROOT/installedApps/ESSearchApplication.ear/  
ESSearchApplication.war/WEB-INF/config.properties
```

作成したファイルは、`WEB-INF` サブディレクトリーに置かなければなりません。

- b. カスタム検索アプリケーションで使用するプロパティを編集し、ファイルを保存します。少なくとも、検索アプリケーションの名前を指定する `applicationName` プロパティは変更してください。
2. 検索アプリケーション・カスタマイザーを使用してサンプル検索アプリケーションを複製する場合:

- a. 作成する構成ファイルの名前を付加して検索アプリケーション・カスタマイザーを開始します。以下の例では、`myNewFile.properties` というファイルが作成されます。

```
http://ESServer.com/ESSearchApplication/palette.do?configFile=/WEB-INF/myNewFile.properties
```

ファイルがまだ存在していないため、サンプル検索アプリケーションの構成ファイル `config.properties` で指定された値が使用されます。

**ヒント:** また、検索アプリケーションの構成ファイルは、検索アプリケーション・カスタマイザーの開始後に「ロード」をクリックしてファイルの名前を指定することによっても作成できます。ファイルは、カスタマイズ・オプションを保存するために「保存」をクリックしたときに作成されます。

- b. WebSphere Application Server でグローバル・セキュリティーが使用可能になっている場合は、有効なユーザー ID とパスワードを使ってログインします。
- c. 検索アプリケーションの名前を指定し、コレクションの検索に使用するオプションを指定して、「保存」をクリックします。指定した変更内容が、WEB-INF サブディレクトリー内の新しい構成ファイルに保存されます。

オプション指定のヘルプについては、「**カスタマイザーのヘルプ**」をクリックします。検索結果がどう表示されるかなど、変更による影響を確認するには、照会を入力して「**検索**」をクリックします。

3. 検索サーバーで、エンタープライズ・サーチ管理者としてログインし、ESSearchServer アプリケーションを一度停止してから再始動します。

#### **AIX、Linux、または Solaris**

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

#### **Windows**

```
stopServer ESSearchServer  
startServer ESSearchServer
```

これらのスクリプトは、`WAS_INSTALL_ROOT/AppServer/bin` ディレクトリーにあります。

- WebSphere Application Server バージョン 5 の場合、デフォルトのインストール・パスは、AIX システムでは `/usr/WebSphere`、Linux または Solaris システムでは `/opt/WebSphere`、Windows システムでは `C:%Program Files%WebSphere` です。

- WebSphere Application Server バージョン 6 の場合、デフォルトのインストール・パスは、AIX システムでは /usr/IBM/WebSphere、Linux または Solaris システムでは /opt/IBM/WebSphere、Windows システムでは C:\Program Files\IBM\WebSphere です。

## 上位結果の分析

ユーザーが一連の検索結果を絞り込めるようにするには、上位結果を分析するオプションを構成します。

上位結果の分析は本質的には、分析値が発生した頻度に従って、現行の一連の検索結果を分類することです。検索結果をフィルタリングするには、分析される値の 1 つを選択します。これにより、選択された値が新規キーワードとして検索条件に自動的に追加されます。上位結果分析の値とは、ユーザーが複雑な検索構文や拡張検索構文を使用しなくても、検索結果を微調整できるようにすることです。

エンタープライズ・サーチ・アプリケーションでは、バー・チャートを使用して、検索結果に最も頻繁に出現するメタデータ・フィールド値をグラフィックで表示できます。また、カスタム Java クラスを作成して、エンタープライズ・サーチ・ソース以外の結果を含む上位結果を任意の HTML フォーマットで表示することもできます。

### 制約事項

使用する言語に該当するフォントが、WebSphere Application Server および検索アプリケーションを実行するコンピューターにインストールされていることを確認します。このステップは、バー・チャートを生成するときに、そのフォントが、チャート・ラベルの文字をサポートするフォントに設定されていることの確認のために必要です。このことは、アジア言語の場合は特に重要です。WebSphere Application Server をインストール後にフォントをインストールした場合は、変更を有効にするために WebSphere Application Server を再始動する必要があります。

検索アプリケーションをスタンドアロン・アプリケーションとして実行する場合は、検索アプリケーション・カスタマイザーを使用するか、または検索アプリケーションの構成ファイルを編集して、上位結果のチャートを構成できます。検索アプリケーションを WebSphere Portal 内のポータルレットとして実行する場合は、WebSphere Portal 管理インターフェースを使用してプロパティを編集し、ポータルレット・インスタンスを構成する必要があります。上位結果の分析のオプションの構成には、検索アプリケーション・カスタマイザーは使用できません。

### このタスクについて

上位結果をグラフィックで表現するには、最も頻繁に発生するメタデータ・フィールド値をバー・チャートで示します。また、Java クラスを使用して、検索アプリケーションを拡張し、HTML を使用して上位結果を順不同のリストに示すなどの、別の形式で上位結果を表示することもできます。

#### メタデータ・フィールド用のバー・チャート

メタデータ・フィールドを分析し、この分析結果をバー・チャートで示すためのオプションを指定できます。このチャートは、ユーザー照会の結果の他



に表示されます。チャートはそれぞれ、単一メタデータ・フィールド (文書サイズ、作成者、日付など) に対応し、チャート内のバーはそれぞれ特定のフィールド値に対応します。

バーの長さは、特定のフィールドに異なる値を含む他の文書を基準に、そのフィールドの値を含む文書の数を示します。バーが長いほど、そのフィールド値が多く出現しています。ユーザーが、バー・チャートのバーをクリックして結果を微調整することができます。選択したバーで表現されているフィールド値は、追加照会条件として追加され、追加検索条件によって新規検索結果が絞り込まれます。

クローラーのクローラール・スペースを構成すると、メタデータ・フィールドの検索オプションを指定するオプションを選択できます。例えば、メタデータ・フィールドをフリー・テキストとして検索できるか、フィールド名で検索できるか、検索結果に表示できるか、パラメトリック・データとして検索できるかどうかなどを指定できます。検索結果にメタデータ・フィールドのチャートを表示するには、クローラーのメタデータ・フィールド・オプションを構成する必要があります。少なくとも、フィールドはフィールド検索可能であり、検索結果に表示可能であることを指定する必要があります。

メタデータ・フィールドのオプションの構成時に、「完全一致」オプションを選択すると、ユーザーがバー・チャートからバーを選択したときに完全一致照会が実行されます。完全一致検索とは、フィールド値全体と照会条件が一致した場合にのみ、結果が戻されることを指定します。フィールドの内容が検索条件より少なかったり多かったりした場合は、一致にはなりません。

メタデータ・フィールドのオプションの構成時に、「完全一致」オプションを選択しない場合は、ユーザーがバー・チャートからバーを選択したときにフィールド照会が実行されます。この場合には、フィールドに追加の照会条件が出現する必要がありますが、フィールドの値全体と一致する必要はありません。

## HTML でフォーマット設定された結果

`com.ibm.es.searchui.charts.servlet.AbstractDynamicChart` API をカスタム Java クラスで拡張することができます。結果の表示にこの方式を使用する場合は、メタデータ・フィールドを検索する照会だけでなく、検索結果の表示のフォーマット設定に HTML を使用できます。結果は、エンタープライズ・サーチ・コレクション内の文書に対して、または外部リポジトリおよび Web サイトの検索から戻すことができます。

チャートはそれぞれ単一の Java クラスに対応し、このクラスで、上位結果をチャートでどのように表現するかを指定します。例えば、エンタープライズ・サーチ・サンプル・コードに提供されているサンプル・クラス

`com.ibm.es.searchui.charts.servlet.DynamicMostRecentDocuments` は、上位結果を順不同リストとして提示します。文書は日付別にソートされ、文書のタイトルおよび日付のみが表示されます。

エンタープライズ・サーチで提供される別のサンプル・クラス

`com.ibm.es.searchui.charts.servlet.DogearSearchResults` は、ユーザーの照会に関連した Lotus Connections Dogear からのブックマークのリストをユーザーにどのように提供するかを示します。

ユーザーがカスタム・フォーマット設定された出力内の検索結果をクリックすると、文書が新しいブラウザ・ウィンドウに表示されます。

## 手順

次の手順では、検索アプリケーション・カスタマイザーを使用して、メタデータ・フィールドを分析することを指定する方法を示します。最も頻繁に出現するメタデータ・フィールド値をバー・チャートにグラフィックで示します。

1. 検索アプリケーション・カスタマイザーを開きます。まだ表示されていない場合は、カスタマイズする検索アプリケーションの構成ファイルをロードして、「適用」をクリックします。
2. 「上位結果チャート」までスクロールダウンして、「チャートの追加」をクリックします。
3. チャート・オプションのリストが表示されたら、「カスタム・チャート」チェック・ボックスはオフのままにしておき（このオプションは、上位結果を分析して戻すカスタム Java クラスを作成する場合にのみ選択します）、「チャートを有効にする」チェック・ボックスを選択して、ユーザーが検索結果を表示するときにバー・チャートを表示するようにします。
4. 「チャート・タイトル」フィールドに、チャートの説明ラベルを指定します。ここで指定した値で、「新しいチャートの行」プレースホルダー・テキストが置換されます。
5. 「動的フィールド選択を有効にする」チェック・ボックスを選択するかどうかを決定します。このオプションを使用可能にすると、ユーザーが検索結果を表示するときに別のフィールドを選択して、そのフィールドの上位結果分析を表示できます。

このオプションを使用して、ユーザーが検索結果セットを微調整できるようにする必要がある場合があります。例えば、女性従業員に関する情報を検索するためにデータベースを検索することがあります。従業員の性別を指定するフィールドを検索する照会（例えば、sex:F）を入力すると、上位結果の値がバー・チャートにバーで表示されます。次に、その初期結果セットに存在するフィールドのリストから職種を選択します。バー・チャートには、女性が行う先頭の職種の上位結果が表示されます。最後に、使用可能なフィールドのリストからデザイナーを選択します。結果には、デザイナーとして働いている女性従業員に関する情報が示されます。

6. 「メタデータ・フィールド名」フィールドに、このチャートの、値を分析する対象のメタデータ・フィールドの名前を入力します。
7. フィールドに区切り文字で区切られた複数の値を含む場合は、「フィールド値の区切り文字」フィールドにその文字を指定します。例えばフィールドに、セミコロンで区切られた顧客の名と姓などの、2つの値が含まれる場合は、ここにセミコロンを指定して、バー・チャートに各値を別個に追加できます。区切り文字を指定しないと、フィールド全体の値が単一項目として分析されます。
8. バーのサイズおよび色などのバー・チャートの表示オプションと、チャートが縮小または展開されたときに表示するバーの数を指定します。
9. 分析の結果をソートするためのオプションを指定します。例えば、フィールド値または値の出現頻度でソートできます。

- 特定のフィールド値についての情報を、そのフィールドの値が検索結果に出現していなくても、常に表示するようにする場合は、「カスタム・ラベル」 域の「**行の追加**」をクリックします。検索結果にフィールドの値が戻されていない場合でも、バー・チャートに表示されるフィールド値を指定し、バー・チャート内のそのバーのラベルを指定します。

例えば、競争相手 (JK Enterprises など) が、その頻度が上位結果として表示するに値しない場合でも、検索結果に出現しているかどうかを常に確認する必要がある場合は、JK Enterprises を分析対象のソース値として指定してから、バー・チャート内のこのバーを識別する説明ラベルを指定します。

- チャート・タイトルの横の「**適用**」をクリックして、このチャートに指定したオプションを適用します。
- 検索結果が検索アプリケーション域に既に表示されている場合、上位結果分析が現行検索に適用され、新しいチャートが表示されます。表示されていない場合は、バー・チャートの表示をテストする照会を入力します。
- チャートに問題がなければ、「**保存**」をクリックして、検索アプリケーションの構成ファイルを更新します。

#### 関連概念

 [照会プロパティの設定](#)

 [上位結果の表示用の Java クラス](#)

---

## 検索アプリケーションのアクセス

検索アプリケーションにアクセスするには、Web ブラウザーで URL を指定します。

### 始める前に

検索アプリケーションをご使用の Web サーバー環境に対して構成する必要があります。

### このタスクについて

サンプル検索アプリケーションは、エンタープライズ・サーチ用の検索サーバーにインストールされています。このアプリケーションをそのまま使用してコレクションおよび外部ソースをテストし、テストが終わった後で、それらをユーザーに対して使用可能にすることができます。また、このアプリケーションをモデルとして使用して、独自の検索アプリケーションを作成することもできます。

### 手順

検索アプリケーションを開始するには、次のようにします。

- Web ブラウザーで検索アプリケーションの URL を入力します。次に例を示します。

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

*SearchServer.com* は、検索サーバーのホスト名です。

Web サーバーがポート 80 を使用するように構成されていない場合は、正しいポート番号を指定することも必要です。次に例を示します。

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

2. カスタム検索アプリケーションを開始するには、サンプル検索アプリケーションの URL を入力し、検索アプリケーションの構成ファイルの名前を付加します。次に例を示します。

```
http://SearchServer.com/ESSearchApplication/search.do?configFile=/WEB-INF/myConfig.properties
```

指定したファイルが存在しない場合、エンタープライズ・サーチのサンプル検索アプリケーションが表示されます。

3. WebSphere Application Server でグローバル・セキュリティーが使用可能になっている場合は、有効なユーザー ID とパスワードを使ってアプリケーションにログインします。

検索アプリケーションで有効なコレクションのいずれかでセキュリティーが有効に設定されている場合や、照会処理中にユーザー・クレデンシャルを検証するよう構成されたクローラーがセキュア・コレクションに含まれている場合は、ユーザー・プロファイルを構成できます。「マイ・プロファイル」ページで、セキュア・ドメインにアクセスするためのクレデンシャルを指定してください。これにより、ログインせずに、それらのドメインを検索できるようになります。

クローラーがシングル・サインオン (SSO) セキュリティーをサポートする場合、ユーザー・プロファイルを作成せずにセキュア・ドメインを検索できます。

4. 「検索」ページで照会を実行します。「設定」ページで検索に選択されたすべてのコレクションと外部ソースが検索されます。

---

## 検索サーバーがセキュア (SSL) 要求のみを受け入れるようにする構成

検索サーバーの HTTP インターフェースを無効にし、SSL およびセキュアな HTTPS インターフェースを通してのみ検索要求を受け入れるようにサーバーを構成することができます。

### このタスクについて

検索要求を処理するときに Secure Sockets Layer (SSL) プロトコルのみを使用するように検索サーバーを構成する場合、HTTP インターフェースを無効にする必要があります。また、検索サーバーと、エンタープライズ・サーチの検索ポートレットがインストールされた WebSphere Portal サーバーといったすべてのクライアント・コンピューターに、同じ鍵ストア・ファイルが保管されていることを確認する必要があります。鍵ストア・ファイルはトラストストアとも呼ばれ、トラステッド通信に使用される SSL を使用可能にする公開鍵が含まれています。

### 手順

検索サーバーをセキュアな要求のみを受け入れるように構成するには、次のようにします。

1. エンタープライズ・サーチ管理者としてログインします。 マルチサーバー・インストールの場合、索引サーバーで次のステップを行ってください。
  - a. エンタープライズ・サーチ・システムを停止します。
 

```
esadmin system stopall
```
  - b. `ES_NODE_ROOT/master_config/nodes.ini` ファイルを編集します。
  - c. `node_ID.searchserverport` 値を、HTTP ポート (通常 80) から HTTPS ポート (通常 443) に変更し、ファイルを保存します。 マルチサーバー・インストールの場合、両方の `node_ID.searchserverport` 値を (それぞれの検索サーバー用の値に) 更新します。
2. 次のステップを行って、検索サーバーを更新します。 マルチサーバー・インストールの場合、両方の検索サーバーで次のステップを行ってください。
  - a. マルチサーバー・インストールの場合、検索サーバーにエンタープライズ・サーチ管理者としてログインします。
  - b. `ES_NODE_ROOT/nodeinfo/es.cfg` ファイルを編集します。
  - c. TrustStore プロパティを更新して、SSL 鍵ストア・ファイルの完全修飾パスを指定します。
  - d. HTTPProtocol プロパティを更新して HTTPS を指定し、ファイルを保存します。
  - e. 次のコマンドを入力します。ここで、`trustStore_password` は、鍵ストア・ファイルのパスワードです。このコマンドは、パスワード値を暗号化し、`es.cfg` ファイル内の `TrustStorePassword` 値を更新します。

#### AIX、Linux、または Solaris

```
eschangetrustpw.sh trustStore_password
```

#### Windows

```
eschangetrustpw trustStore_password
```

- f. 検索アプリケーションの `config.properties` ファイル内の `trustStore` および `trustPassword` プロパティが、鍵ストア・ファイルの正しい完全修飾パスとパスワードを示していることを確認してください。 この情報は、`config.properties` ファイルを編集するか検索アプリケーション・カスタマイザーを使用して、確認または変更できます。
  - g. エンタープライズ・サーチの検索ポートレットを使用する場合は、`trustStore` および `trustPassword` ポートレット・パラメーターが、鍵ストア・ファイルの正しい絶対パスおよびパスワードを指定していることを確認してください。 WebSphere Portal 管理インターフェースでポートレット管理オプションを使用して、この情報を確認または変更します。
3. エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```

#### 関連タスク

264 ページの『サンプル検索アプリケーション・プロパティの編集』

266 ページの『検索アプリケーションのカスタマイズ』

384 ページの『WebSphere Portal バージョン 5.1 でのエンタープライズ・サーチのセットアップ』

## 検索サーバーがプロキシ・サーバーを介して要求を受け入れるようにする構成

検索サーバーがプロキシ・サーバーを介して要求を受け入れるように構成することができます。

### 手順

プロキシ・サーバーを介して検索サーバーに要求を実行依頼できるようにするには、次のようにします。

1. エンタープライズ・サーチ管理者でログインし、エンタープライズ・サーチ・システムを停止します。 マルチサーバー構成では、索引サーバーにログインします。

```
esadmin system stopall
```

2. `ES_NODE_ROOT/master_config/nodes.ini` ファイルを編集します。
  - a. `ProxyServer` プロパティを変更して、プロキシ・サーバーの完全修飾ホスト名を指定します。
  - b. `ProxyServerPort` プロパティを変更して、プロキシ・サーバーのポートを指定します。
  - c. オプション: プロキシ・サーバーがすべての要求の認証を必要とする場合、`ProxyServerUserName` プロパティを更新して、プロキシ・サーバーの有効なユーザー名を指定します。
  - d. オプション: プロキシ・サーバーがすべての要求の認証を必要とする場合、次のコマンドを入力します。ここで、`proxyServer_password` は、指定されたプロキシ・サーバーのユーザー名のパスワードです。このコマンドは、パスワード値を暗号化し、`es.cfg` ファイル内の `ProxyServerUserPassword` 値を更新します。

#### AIX、Linux、または Solaris

```
eschangeproxypw.sh proxyServer_password
```

#### Windows

```
eschangeproxypw proxyServer_password
```

3. 検索アプリケーションの構成ファイル (`.properties` ファイル) 内の `proxyHost` および `proxyPort` プロパティが、プロキシ・サーバーの正しい完全修飾ホスト名とポート番号を指定していることを確認してください。

プロキシ・サーバーが認証を必要とする場合、`proxyUser` および `proxyPassword` プロパティが、プロキシ・サーバーの有効なユーザー名とパスワードを指定していることを確認してください。この情報は、構成ファイルを編集するか検索アプリケーション・カスタマイザーを使用して、確認または変更できます。

4. エンタープライズ・サーチの検索ポートレットを使用する場合は、`proxyHost` および `proxyPort` ポートレット・パラメーターが、プロキシ・サーバーの正しい完全修飾ホスト名およびポート番号を指定していることを確認してください。

プロキシ・サーバーが認証を必要とする場合、proxyUser および proxyPassword プロパティが、プロキシ・サーバーの有効なユーザー名とパスワードを指定していることを確認してください。WebSphere Portal 管理インターフェイスでポートレット管理オプションを使用して、この情報を確認または変更します。

5. エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```





---

## 外部ソースのサポート

外部ソースとは、そのデータ・ソース内の文書のクロール、構文解析、索引作成を必要とせずにエンタープライズ・サーチ・アプリケーションでの検索が可能なデータ・ソースです。

以下のタイプのデータ・ソースを外部ソースとして検索できます。

- Java database connectivity (JDBC) プロトコルをサポートするデータベース。IBM DB2、Oracle、Microsoft SQL Server 2000、および Microsoft SQL Server 2005 データベースのみがサポートされます。検索可能に設定されている各表ごとに、個別の外部ソースが作成されます。

**制約事項:** SQL Server 2005 データベースに対するサポートは、Variant データ・タイプを含まない表に制限されています。SQL Server 2005 の JDBC ドライバーは AIX システムでサポートされません。

- Lightweight Directory Access Protocol (LDAP) サーバー。各 LDAP サーバーごとに、1 つの外部ソースが作成されます。

外部ソースに関する情報を構成した後、その外部ソースを少なくとも 1 つの検索アプリケーションと関連付ける必要があります。そうすると、ユーザーは、データのクロール、構文解析、および索引作成によってエンタープライズ・サーチ用に作成されたコレクションを照会すると同時に、その外部ソースを検索できるようになります。

### 関連概念



検索および索引 API フェデレーター

---

## システムへの外部ソースの追加

外部ソースをエンタープライズ・サーチ・システムに追加するときには、追加するソースのタイプを指定します。ウィザードのガイドに従って、データ・ソースに関する情報、および、データ・ソースがどのように検索できるのかについての情報を指定します。

### 始める前に

システムに外部ソースを追加するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### 制約事項

Oracle データベースを外部ソースとして検索するには、エンタープライズ・サーチ用の検索サーバーに Oracle クライアント・プログラムがインストールされていなければなりません。

Microsoft SQL Server 2005 の JDBC ドライバーは AIX システムでサポートされません。

## このタスクについて

外部ソースに関する情報をシステムに追加するときに、ユーザーがエンタープライズ・サーチ・アプリケーションでそのソースを照会できるようにします。

Lightweight Directory Access Protocol (LDAP) サーバーおよび Java Database Connectivity (JDBC) データベース表を、検索対象として使用可能にできます。

LDAP サーバーを構成するときには、ウィザードのガイドに従って、そのサーバーにシステムが接続するための情報を指定し、そのサーバーが検索される方法について、オプションを指定します。

JDBC データベースを構成するときには、ウィザードのガイドに従って、そのサーバーにシステムが接続するための情報を指定し、検索に使用できる表を選択し、それらの表に含まれるデータが検索される方法について、オプションを指定します。システムに追加する各表ごとに、個別に検索可能な外部ソースが作成されます。

SQL Server 2000 ドライバーについては、<http://www.microsoft.com/downloads/details.aspx?familyid=07287B11-0502-461A-B138-2AA54BFDC03A&displaylang=en> を参照してください。SQL Server 2005 ドライバーについては、<http://www.microsoft.com/downloads/details.aspx?familyid=e22bc83b-32ff-4474-a44a-22b6ae2c4e17&displaylang=en> を参照してください。

## 手順

外部ソースをシステムに追加するには、次のようにします。

1. JDBC データベースをエンタープライズ・サーチ・システムに組み込むには、外部ソースを追加する前に、次のステップを行います。システムが適切な JDBC ドライバーを一度に見つけられるようになるためには、このステップの実行が必要です。
  - a. クローラー・サーバーで、エンタープライズ・サーチ管理者としてログインします。
  - b. `ES_INSTALL_ROOT/configurations/interfaces/discovery_interface.ini` ファイルを編集し、JDBC ドライバーへのクラスパスとライブラリー・ファイルへのパスを含むように、`CLASSPATH` および `LD_LIBPATH` 環境変数を指定します。
  - c. `ES_INSTALL_ROOT/configurations/interfaces/customcommunication_interface.ini` ファイルを編集し、JDBC ドライバーへのクラスパスとライブラリー・ファイルへのパスを含むように、`CLASSPATH` および `LD_LIBPATH` 環境変数を指定します。
  - d. オプション: ローカルまたはカタログ・データベースに Oracle JDBC ドライバーを使用するには、Oracle ライブラリー・パスを `LD_LIBPATH` 環境変数に追加し (例: `LD_LIBPATH=.../home/oracle/OraHome1/lib32`)、ライブラリー・パスを指定して Oracle インストール・ディレクトリー変数をエクスポートするように、`escrset.sh` ファイルを編集します。次に例を示します。

```
ORACLE_HOME=/home/oracle/OraHome1
export ORACLE_HOME
```
  - e. 共通通信層 (CCL) を含め、エンタープライズ・サーチ・システムを再始動します。

### AIX、Linux、または Solaris

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

### Windows コマンド・プロンプト

```
esadmin stop
stopccl
startccl
esadmin start
```

### Windows サービス管理ツール

バックグラウンドで CCL を開始するには、次のようにします。

- 1) esadmin stop と入力します。
  - 2) 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」 により、Windows サービスを起動します。
  - 3) 「**IBM OmniFind Enterprise Edition**」を右クリックし、「停止」をクリックします。サービスが停止したら、「開始」をクリックします。
  - 4) esadmin start と入力します。
2. 「外部ソース」をクリックして「外部ソース」ビューを開きます。
  3. 「外部ソースの追加」をクリックします。
  4. 追加する外部ソースのタイプとして、LDAP サーバーまたは JDBC データベースのいずれかを選択します。
  5. 外部ソースの構成を始めるため、「次へ」をクリックします。

作成するソースのタイプに対応するウィザードが開きます。ウィザードのプロンプトに従って、外部ソースを構成してください。ウィザードのどのページでも、「ヘルプ」をクリックすると、指定可能なオプションの説明が表示されます。

DB2 および Oracle データベースの接続情報を構成する際には、以下に示すデフォルトの JDBC ドライバーと場所を参考にしてください。

#### DB2: レガシー JDBC ドライバー

ドライバー名: COM.ibm.db2.jdbc.app.DB2Driver  
サンプルの場所: *db2\_install\_root*/java/db2java.zip

#### DB2: ユニバーサル JDBC ドライバー

ドライバー名: com.ibm.db2.jcc.DB2Driver  
サンプルの場所:  
*db2\_install\_root*/java/db2jcc.jar  
*db2\_install\_root*/java/db2jcc\_license\_cu.jar

#### Oracle

ドライバー名: oracle.jdbc.driver.OracleDriver  
サンプルの場所: *oracle\_home*/jdbc/lib/ojdbc14.jar

#### Microsoft SQL Server 2000

ドライバー名: com.microsoft.jdbc.sqlserver.SQLServerDriver  
サンプルの場所:

```
mssql_jdbc_home/lib/mssqlserver.jar
mssql_jdbc_home/lib/msbase.jar
mssql_jdbc_home/lib/msutil.jar
```

#### Microsoft SQL Server 2005

ドライバー名: com.microsoft.sqlserver.jdbc.SQLServerDriver  
サンプルの場所: *install\_dir*/sqljdbc\_1.0/locale/sqljdbc.jar  
例: *install\_dir*/sqljdbc\_1.0/enu/sqljdbc.jar

- 外部ソース検索に関するオプションの指定が終わったら、「完了」をクリックします。

新しい外部ソースが、システムに以前に追加された他の外部ソースと共に、「外部ソース」ビューにリストされます。

#### 関連概念



検索および索引 API フェデレーター

---

## 外部ソースと検索アプリケーションの関連付け

外部ソースは、少なくとも 1 つの検索アプリケーションを関連付けないと検索できません。

### 始める前に

検索アプリケーションを、検索できる外部ソースと関連付けるには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### 手順

検索アプリケーションを 1 つ以上の外部ソースに関連付けるには、次のようにします。

- 管理コンソールのツールバーで「**セキュリティ**」をクリックします。
- 「検索アプリケーション」ページで、「**検索アプリケーションの構成**」をクリックします。
- 「検索アプリケーションの構成」ページで、「**検索アプリケーションの追加**」をクリックします。
- 検索アプリケーションの名前を入力します。
- そのアプリケーションが検索できる外部ソースを次のように選択します。
  - システムに追加するすべての外部ソースを検索アプリケーションがアクセスできるようにするには、「**すべてのコレクションと外部ソース**」をクリックします。
  - 指定する外部ソースのみを検索アプリケーションがアクセスできるようにするには、「**特定のコレクションと外部ソース**」をクリックします。

このオプションを選択すると、コレクション名および外部ソース名のリストが表示されます。アプリケーションが検索できる外部ソースの「**選択**」チェック・ボックスを選択します。

- 「**OK**」をクリックします。

## 関連概念

 [検索および索引 API フェデレーター](#)



---

## エンタープライズ・サーチのセキュリティー

エンタープライズ・サーチのセキュリティー機構によって、無許可でソースが検索されないようにし、管理機能の使用を特定のユーザーに制限することができます。

エンタープライズ・サーチによって、ユーザーは広い範囲のデータ・ソースを検索できます。コンテンツに対するアクセス許可を持つユーザーのみがアクセスすることを確認し、また許可されたユーザーのみが管理コンソールにアクセスできることを確実にするために、エンタープライズ・サーチは複数のレベルにおいてセキュリティーの調整および施行を行います。

### ヒント:

エンタープライズ・サーチのセキュリティーを構成する方法の詳細な例については、IBM Redbook、IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios を参照してください。シナリオで、LDAP リポジトリを備えた WebSphere Application Server でグローバル・セキュリティーを使用可能に設定して、ID 管理コンポーネントをセットアップし、さまざまなクローラーを構成して文書レベルのセキュリティーを確実に施行する方法を示します。

### Web サーバー

セキュリティーの第 1 レベルは Web サーバーです。WebSphere Application Server でグローバル・セキュリティーを使用可能にすると、ユーザーに管理役割を割り当てたり、システムを管理するユーザーを認証することができます。ユーザーが管理コンソールにログインすると、そのユーザーが管理することを許可された機能とコレクションのみが使用可能になります。

また、検索アプリケーションは WebSphere Application Server のセキュリティー・メカニズムを使用して、コレクションを検索するユーザーを認証できます。

### コレクション・レベルのセキュリティー

コレクションを作成するときに、コレクション・レベルのセキュリティーを有効にできます。コレクションの作成後にこの設定を変更することはできません。コレクション・レベルのセキュリティーを有効にしている場合は、あとで文書レベルのセキュリティー制御を指定することはできません。

コレクション・レベルのセキュリティーが有効である場合:

- エンタープライズ・サーチのグローバル分析処理により、重複文書を索引付けする別の規則が適用されます。
- クロール時に文書にセキュリティー・トークンを関連付ける、照会処理中に現行のクレデンシャルの検証を必要とする、Web 文書内のアンカー・テキストを索引付けするかどうかを指定するなど、文書レベルのセキュリティーを施行するオプションを構成できます。
- (個別のユーザーではなく) 検索アプリケーションを検索可能なコレクションと外部ソースにマッピングすることによって、セキュリティーを施行

できます。そして、標準のアクセス制御メカニズムを使用して、ユーザーに対し、検索アプリケーションへのアクセスの許可や拒否を行います。

コレクション・セキュリティを有効にすることと、検索の質とでトレードオフが生じます。コレクション・セキュリティを有効にすると、文書ごとに索引付けされる情報量が減ります。この副次作用は、照会によっては結果の数が少なくなることです。

### 文書レベルのセキュリティ

コレクションのクローラーを構成するときに、文書レベルのセキュリティを使用可能にできます。例えば、データがクローラーによって収集される際に、データにセキュリティ・トークンを関連付けるオプションを指定できます。検索アプリケーションは、文書と共に索引に保管されているこのトークンを使用して、アクセス制御を実行し、適切なクレデンシャルを持つユーザーのみがデータを照会して検索結果を表示できるようにします。

一部のタイプのデータ・ソースでは、照会処理中にユーザーのログイン・クレデンシャルを現行のアクセス制御で検証するオプションを構成できます。こうしてさらに重ねられたセキュリティにより、ユーザーの特権がネイティブのデータ・ソースでリアルタイムに検証されるようになります。この機能では、文書とそのセキュリティ・トークンが索引付けされた後にユーザーのクレデンシャルが変わった場合に対して保護することができます。

グローバル分析のアンカー・テキスト処理フェーズは、通常、ある文書 (ソース文書) に現れるテキストを、そのテキストが必ずしも現れるとは限らない別の文書 (ターゲット文書) に関連付けます。Web クローラーを構成する際、Web クローラーがクロールを許可されていない文書にテキストがリンクしている場合に、索引からアンカー・テキストを除外するかどうかを指定できます。

コレクションに対するセキュリティは、索引付けされたコンテンツを保護するためにエンタープライズ・サーチが使用できる認証とアクセス制御機構を越えて拡張されます。転送中に、悪意のある、あるいは無許可のユーザーにデータをアクセスさせないための安全機能も存在します。例えば、検索サーバーは、Secure Sockets Layer (SSL)、Secure Shell (SSH)、および Secure Hypertext Transfer Protocol (HTTPS) といったプロトコルを使用して、索引サーバーおよび検索アプリケーションと通信します。

暗号化によってさらにセキュリティが強化されます。例えば、エンタープライズ・サーチ管理者のパスワードが製品のインストール時に指定されますが、これは、暗号化されたフォーマットで保管されます。ユーザーがユーザー・プロファイルで指定するパスワードも、暗号化されたフォーマットで保管されます。

セキュリティを向上させるには、サーバー・ハードウェアが適切に隔離されており、無許可の侵入から保護されている必要があります。ファイアウォールをインストールすると、使用しているネットワークの別の部分からの侵入から、エンタープライズ・サーチ・サーバーを保護できます。また、エンタープライズ・サーチ・サーバーに、オープン・ポートがないことも確認してください。システムが、エンタープライズ・サーチのアクティビティとアプリケーション用に明示的に割り当てられたポートでのみ要求を listen するように、システムを構成してください。



---

## インストールのセキュリティー

OmniFind Enterprise Edition のインストール・プログラムでは、ユーザーがエンタープライズ・サーチ・コレクションを管理または検索するときにセキュリティーを施行するための環境を設定します。

### エンタープライズ・サーチ管理者 ID

インストール処理中に、インストーラーは、エンタープライズ・サーチ管理者に使用するユーザー ID とパスワードを要求します。システムは、指定されたクレデンシャルを使用して、以下を行います。

- 管理タスクの実行時にエンタープライズ・サーチ管理者を認証します。
- エンタープライズ・サーチの内部データベースを作成します。
- エンタープライズ・サーチのすべてのセッションまたはプロセスを開始します。

インストール中に指定されるユーザー ID は、システム管理特権を持つ有効なオペレーティング・システム・ユーザー ID でなければなりません。インストール・プログラムは、適切に暗号化したクレデンシャルを各エンタープライズ・サーチ・サーバー上のプロパティ・ファイルに保管します。

### 暗号化

機密データを保護するために、暗号化を使用して、エンタープライズ・サーチ・システムで伝送される全メッセージの認証データ部分がエンコードされます。この処理では、認証 ID とパスワードのみが暗号化されるため、オーバーヘッドは少ししか発生しません。システムによって (構成ファイル、エンタープライズ・サーチ・データベースなどに) 保管されるすべてのパスワードも、暗号化されます。

### WebSphere Application Server のセキュリティー

WebSphere Application Server が検索サーバーに前もってインストールされていなかった場合、インストール・プログラムは、グローバル・セキュリティーを使用不可にして、この製品をサイレント・インストールします。後でグローバル・セキュリティーを WebSphere Application Server で使用可能にした場合には、WebSphere Application Server が、エンタープライズ・サーチ管理者を認証する責任を負いません。

グローバル・セキュリティーを使用可能にした場合、インストール中に指定されたエンタープライズ・サーチ管理者 ID とパスワードを、WebSphere Application Server ユーザー・レジストリー (Lightweight Directory Access Protocol (LDAP) ディレクトリーなど) に追加してください。

OmniFind Enterprise Edition のインストール後にグローバル・セキュリティーを使用可能にした場合は、構成値を更新し、コマンド **eschangewaspw** を実行して WebSphere Application Server クレデンシャルを暗号化してエンタープライズ・サーチのプロパティ・ファイルに保管してください。

---

## 認証とアクセス制御

無許可のユーザーからコンテンツを保護し、管理機能へのアクセスを制御するために、エンタープライズ・サーチは、ユーザー認証と許可（アクセス制御）をサポートします。

### 認証

認証とは、システムへのアクセスを要求するユーザーの身元を、システムが検査する処理です。アクセス制御は一般にリソースへのアクセスを要求するユーザーの ID を基にしているため、認証は、セキュリティーを実効的にするために不可欠です。

エンタープライズ・サーチ・ユーザーの認証は、クレデンシャルによって実装されます。クレデンシャルは、最小でユーザー ID とパスワードから構成されます。

管理コンソールにアクセスするユーザーを認証するために、エンタープライズ・サーチは、WebSphere Application Server で提供される認証サポートを利用します。

### 許可 (アクセス制御)

許可とは、データ・アクセスまたはアクション実行の権限をシステムが付与または取り消すメカニズムです。多くの場合、ユーザーは、何らかの形の認証を使用してシステムにログインする必要があります。アクセス制御のメカニズムでは、ユーザーの ID をアクセス制御リスト (ACL) と比較し、ユーザーが実行できる、または実行できない操作を判別します。アクセス制御には、以下が含まれます。

- ファイル許可。ファイルの作成、読み取り、編集、削除の権限など。
- プログラム許可。プログラム実行の権限など。
- データ許可。データベースにおける情報のリトリブまたは更新の権限など。

---

## 管理役割

エンタープライズ・サーチは管理コンソールのさまざまな機能へのアクセスを制御するために、役割の概念を使用します。

OmniFind Enterprise Edition (OmniFind Enterprise Edition) のインストール時に、インストーラーは、エンタープライズ・サーチ管理者のユーザー ID とパスワードを構成します。管理コンソールに最初にアクセスするときは、このユーザーとしてログインしなければなりません。WebSphere Application Server でグローバル・セキュリティーを使用可能に設定していない場合、このユーザー ID がエンタープライズ・サーチ管理コンソールのアクセスに使用できる唯一のユーザー ID となります。

WebSphere Application Server でグローバル・セキュリティーを使用可能にしていれば、追加のユーザーをエンタープライズ・サーチ管理ユーザーとして登録できます。ユーザーを役割に割り当てることによって、アクセスを特定のコレクションに制限し、それぞれの管理ユーザーが実行できる機能を制御できます。エンタープライズ・サーチ内の管理役割に割り当てるユーザー ID は、WebSphere Application Server ユーザー登録の中になければなりません。

管理ユーザーがログインするときに、エンタープライズ・サーチはユーザー ID を認証します。ユーザーが管理できるコレクションと機能のみが、コンソールの中で選択可能になっています。

次の管理役割にユーザーを登録できます。

#### **エンタープライズ・サーチ管理者**

これらのユーザーは、コレクションを作成し、エンタープライズ・サーチ・システムのすべての側面を管理する権限を持ちます。OmniFind Enterprise Edition をインストールする場合は、最初のエンタープライズ・サーチ管理ユーザーのユーザー ID とパスワードを指定します。最初にログインした後に、このユーザーは、他のユーザーをエンタープライズ・サーチ管理者の役割に割り当てることができます。

#### **コレクション管理者**

これらのユーザーは、管理権限があるコレクションの編集、モニター、操作制御を行えます。これらのユーザーは、コレクションを作成することはできません。エンタープライズ・サーチ管理者がコレクション管理者を認可した場合のみ、コレクション管理者は、システム・レベルのアクティビティをモニターおよび操作することができます。

#### **オペレーター**

これらのユーザーは、管理権限があるコレクションのモニターと操作制御を行えます。これらのユーザーは、コレクション・アクティビティの開始と停止はできますが、コレクションの作成や編集はできません。エンタープライズ・サーチ管理者がオペレーターを認可した場合のみ、オペレーターは、システム・レベルのアクティビティをモニターおよび操作することができます。

#### **モニター担当者**

これらのユーザーは、管理権限があるコレクションをモニターできます。操作の制御 (サーバーの開始や停止など)、コレクションの作成、またはコレクションの編集はできません。エンタープライズ・サーチ管理者がモニター担当者を認可した場合のみ、モニター担当者は、システム・レベルのアクティビティを監視することができます (しかし、操作はできません)。

## **管理ユーザーの構成**

管理役割を構成することで、コレクションへのアクセスを制限でき、各管理ユーザーが実行できる機能を制御できます。

#### **始める前に**

管理役割にユーザーを割り当てる前に、WebSphere Application Server でセキュリティが使用可能にされていることを確認してください。また、WebSphere Application Server ユーザー・レジストリー内にユーザー ID が存在することを確認してください。

管理ユーザーを構成するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

#### **手順**

ユーザーを管理役割に割り当てるには、次のようにします。

1. 「**セキュリティ**」をクリックして「セキュリティ」ビューを開きます。
2. 「管理役割」ページで、「**ユーザーの追加**」をクリックします。
3. 登録したいユーザーのユーザー ID を入力し、適切な管理役割を選択します。
4. このユーザーをエンタープライズ・サーチ管理者として登録していない場合は、このユーザーが「**システム**」ツールバーからページにアクセスできるかどうかを指定します。

例えば、一部のオペレーターまたはコレクション管理者に、システム・レベルのログ・ファイルのモニターを許可したい場合があります。

5. このユーザーをエンタープライズ・サーチ管理者として登録していない場合は、このユーザーが管理できるコレクションおよび外部ソースを選択します。

個々のコレクションおよび外部ソースのチェック・ボックスを選択するか、あるいはこのユーザーがすべてのコレクションおよび外部ソースを管理できるようにします。

---

## コレクション・レベルのセキュリティ

コレクション・レベルのセキュリティを提供するために、コンテンツの索引付けのオプションと、検索アプリケーションが特定のコレクションを検索できるようにするオプションを構成します。

コレクションを作成するときに、コレクション・セキュリティを有効にするオプションを選択できます。このオプションを選択すれば、あとで文書レベルのセキュリティ制御を構成することができます。コレクション・セキュリティが有効である場合、エンタープライズ・サーチのグローバル分析処理により、重複文書を索引付けする別の規則も適用されます。

検索アプリケーションを作成した後、検索アプリケーション ID によって、検索アプリケーションが検索できるコレクションおよび外部ソースと、検索アプリケーションにアクセスできるユーザーを指定できます。

## 重複文書の分析とコレクション・セキュリティ

コレクション・セキュリティを有効にすると、グローバル分析プロセスは、コレクションの中の重複文書の識別は行いません。

グローバル分析の間、索引付けプロセスは、互いに重複している (または、ほとんど重複している) 文書を識別します。そして、これらの文書をすべて、コンテンツの 1 つの正規表現に関連付けます。重複文書を識別することで、検索結果に、同じ (あるいはほとんど同じ) コンテンツの複数の文書が入ることがなくなります。

コレクションを作成するときにコレクション・セキュリティを有効にすると、重複文書は識別されず、それらの文書は共通の正規表現に関連付けられません。代わりに、各文書は、別個に索引付けされます。これによって、ユーザーがそのクレデンシャルに適合するセキュリティ・トークンの文書のみを検索するように、各文

書のセキュリティー制御が評価されるようになります。2つの文書のコンテンツがほとんど同じである場合がありますが、異なるアクセス制御リストを使用して、セキュリティーを施行します。

例えば、document\_A と document\_B の2つの重複文書ある場合に、ユーザーには document\_B へのアクセス権限しかないと想定します。document\_B が重複検出で除去されると、このユーザーは、document\_A に設定されたアクセス制約のために、検索結果に文書を表示できません。

重複文書分析を使用不可にすればコレクションの中の文書のセキュリティーは強化されますが、検索結果の中に同じ文書の複数のコピーが返されると、検索の質は低下するかもしれません。

## 検索アプリケーション ID

さまざまなコレクションに対する検索性は、検索アプリケーションを検索可能なコレクションや外部ソースにマッピングすることによって制御されます。Default というアプリケーションでは、サンプル検索アプリケーションをそのまま使用して、すべてのコレクションと外部ソースを検索することができます。

すべての検索アプリケーションは、有効なアプリケーション名 (APPID) をエンタープライズ・サーチ・アプリケーション・プログラミング・インターフェース (API) に渡す必要があります。この APPID に関連付けられたコレクションと外部ソースのみが、検索アプリケーションによって検索可能になります。

検索アプリケーションがコレクションまたは外部ソースにアクセスして検索を開始する前に、エンタープライズ・サーチ管理者は、検索アプリケーションを、それが検索できる特定のコレクションおよびソースに関連付ける必要があります。検索アプリケーションには、エンタープライズ・サーチ・システム内のすべてのコレクションおよび外部ソースを検索させることも、あるいは指定したコレクションおよび外部ソースのみを検索させることもできます。

サンプル検索アプリケーション (ESSearchApplication) のプロパティ・ファイルに、使用するアプリケーション名が指定されています。このプロパティ・ファイルのデフォルトの場所は、

```
ES_INSTALL_ROOT¥installedApps¥ESSearchApplication.ear¥
ESSearchApplication.war¥WEB-INF¥config.properties です。
```

アプリケーション名の初期値は Default です。この値を変更した場合は、ESSearchApplication アプリケーションが検索できるコレクションと外部ソースのリストを変更します。

どのユーザーがどのコレクションを検索できるかを制御するには、WebSphere Application Server の標準のアクセス制御機能を使用し、この機能で URL へのアクセスを制限する場合と同じようにして、ユーザー (またはユーザー・グループ) をクライアント・アプリケーションに関連付けてください。例えば、ユーザーの検索アプリケーションを起動する URL へのアクセスを制限することができます。

検索アプリケーション ID の詳細と、ユーザーのカスタム検索アプリケーションにセキュリティー制御を組み込む方法については、エンタープライズ・サーチの Search and Index API (検索および索引 API) の説明を参照してください。

## 関連概念

 [検索および索引 API の概要](#)

 [検索および索引 API のセキュリティ](#)

---

## 文書レベルのセキュリティ

コレクションが作成されたときに、そのコレクションに対するセキュリティを有効にすると、文書レベルのセキュリティ制御を構成することができます。文書レベルのアクセス制御により、検索要求を実行したユーザーが表示を許可されている文書のみが、検索結果に含まれるようになります。

エンタープライズ・サーチ・システムでは、文書レベルのセキュリティ制御を構成するために、多くの方法がサポートされています。

- 文書が索引に追加される前に、文書を事前にフィルターに掛け、セキュリティ・トークンと関連付けることができます。
- 一部のデータ・タイプでは、ユーザーのログイン・クレデンシャルを現行のアクセス制御データに照らして検証するために、検索結果に後でフィルターを掛けることができます。エンタープライズ・サーチ ID 管理コンポーネントでは、異なるリポジトリへのアクセスにユーザーが必要とするさまざまなクレデンシャルを暗号化し、暗号化されたクレデンシャルをプロファイルに保管することができます。検索されるソースが、シングル・サインオン (SSO) セキュリティを提供する製品によって保護されている場合、ID 管理コンポーネントは、ユーザーによるプロファイルの作成を必要とせず、文書へのアクセスを制御できます。
- ほとんどのクローラー・タイプの場合、カスタム Java クラス (プラグイン) を使用して、セキュリティ・トークンを索引内の文書と関連付けられます。
- Web クローラーでクロールされる文書の場合、禁止文書へのリンクを含む文書内のアンカー・テキストは、索引から除外できます。

## 関連概念

 [アプリケーションのセキュリティ](#)

## 検索結果の事前フィルターと事後フィルター

検索要求を実行したユーザーが表示を許可されている文書のみが検索結果に含まれるように、文書をフィルターに掛けるには、2つの異なる方法があります。

- 1つ目の方法は、文書のネイティブのアクセス制御リスト (ACL) をクロール時に索引に複製し、検索エンジンにユーザー・クレデンシャルと複製した文書 ACL を比較させるものです。文書を事前にフィルターに掛け、索引に追加する文書を制御すると、最善のパフォーマンスになります。ただし、索引にある多様なバックエンド・ソースのすべてのセキュリティ・ポリシーをモデル化して、均一の方法で比較ロジックを実装することは困難です。また、この方法では、ソースの ACL で行われる変更に対してそれほどすぐに対応できません。
- 2番目の方法は、バックエンド・ソースで現行のセキュリティ・データを調べ、結果セットの文書に後でフィルターを掛けるものです。この方法では、当該のバックエンド・ソースにより、ユーザーに戻される文書を最終的に決定することができ、結果セットが現行のアクセス制御を反映したものになります。ただ

し、この方法では、すべてのバックエンド・ソースとの接続が存在する必要があるため、検索のパフォーマンスが低下します。ソースがアクセス不能の場合、文書へのリンクが、ユーザーが表示を許可されていない文書と共に、フィルターにより結果セットから除外されなければなりません。

**重要:** マルチサーバー構成では、一部のソース・タイプについて事後フィルター処理がクローラー・サーバーで行われます。クローラー・サーバーが保守のためにサービス停止になると、ユーザーは、エンタープライズ・サーチ・コレクションを照会したときに、何も結果を得られません。さらに、アクセス制御に必要なバックエンド・サーバーがアクセス不能である場合、結果が戻されません。

エンタープライズ・サーチでは、アクセス制御の実行をサポートするために、この 2 つの方法を組み合わせて使用しています。この設計により、元の文書リポジトリの正確なセキュリティー・ポリシーを維持しながら、最適なパフォーマンスを実現します。索引にハイレベルのアクセス制御データを保管することによって、システムが一時的な (可能性としては、小さくなった) 結果セットを提供し、その後、この結果セットに事後フィルター処理を行って、現行のアクセス制御を検証します。ここでは、ユーザーが、文書を所有するリポジトリへのアクセス権を持つ場合、その文書へのアクセス権も持つ可能性が高いということを前提としています。

索引に保管されるアクセス制御データは、クローラー・タイプによって異なります。例えば、Lotus Notes クローラーは、データベース・レベルとサーバー・レベルのアクセス制御を保管でき、QuickPlace クローラーは、サーバー、プレース、ルールのアクセス制御を保管できます。

エンタープライズ・サーチ・システムのすべてのデータ・ソース・タイプで、クロール時にネイティブのアクセス制御リストを索引付けする機能がサポートされています。一部のデータ・ソース・タイプでは、結果セットを事後フィルター処理して、ユーザーの現行のクレデンシャルを検証する機能もサポートされます (このタイプのサポートは、ネイティブのセキュリティー・メカニズム、またはエンタープライズ・サーチ ID 管理コンポーネントを通して提供されます)。

こうした 2 本の柱からなるセキュリティー設計には、以下のタスクが含まれます。

- クロール中のネイティブ ACL 情報の抽出。
- サーバーおよびデータベース ACL 情報の索引への保管。
- ログイン時、またはセッション初期化時の、ユーザーのセキュリティー・コンテキストの作成。このタスクは、1 人のユーザーが多様なバックエンド・ソースのアクセスに使用する異なった ID に対処する必要があります。
- ユーザーのセキュリティー・コンテキストを使用した検索の処理と、リポジトリ・レベルでユーザーがアクセス権を持つ文書のみを含む一時結果セットの生成。
- 結果セットに文書を提供したバックエンド・ソースで現行のネイティブ ACL 情報を調べることによる、一時結果セットの事後フィルター処理。

## 保管されたセキュリティー・トークンによる検証

コレクションが作成されたときに、そのコレクションに対するセキュリティーを有効にした場合、索引にセキュリティー・データを保管することで、文書レベルのセキュリティー制御を構成することができます。

デフォルトで、各文書には、文書を全員に使用可能にするパブリック・トークンが割り当てられています。コレクションに対してセキュリティが有効になっている場合、パブリック・トークンは、管理者が提供する値、またはクロール対象文書中のフィールドから抽出された値に置き換えることができます。クローラーを構成するときに、そのクローラーがクロールした文書にアクセスできるユーザーを制限するために、セキュリティ・トークンを使用することを指定します。

コレクション管理者は、クローラーの構成時に、個別の表やファイル・システムなどに対してセキュリティ・オプションを指定できます (つまり、クロール・スペース内のデータ・ソースごとに異なるセキュリティ規則を構成可能です)。管理者は、以下を行うことができます。

- 文書がパブリックである (すべてのユーザーがその文書を検索できる) ことを指定する
- ユーザー定義のセキュリティ・トークンを各文書に割り当てる
- クロールされたデータのフィールドからセキュリティ・トークンを抽出し、抽出したトークンを各文書に割り当てる

セキュリティ・トークン (デフォルトのパブリック・トークンを除く) は完全にユーザー定義です。セキュリティ・トークンは、ユーザー ID、グループ ID、ユーザー役割、またはデータ・ソースに有効であると管理者が判断したその他の値を表す場合があります。

例えば、Lotus Notes クローラーによってクロールされる文書のアクセス制御に、hrDeptName フィールドを使用するように管理者が指定します。また、そのフィールドが文書に存在しない場合や、そのフィールドにセキュリティ・データが含まれない場合には、文書のアクセス制御に hrgroup1 と hrgroup2 という 2 つのユーザー定義トークンを使用するように管理者が指定することもできます。

セキュリティ・トークンは、クローラーの構成ファイルを通してクローラーから使用できるようになります。各文書に対して、クローラーはセキュリティ・トークン値をメタデータとして提供します。索引作成コンポーネントは、セキュリティ・トークンを読み取り、それを索引内のその文書に関する登録情報に適用します。ネイティブ・データ・ソースの管理者がアクセス制御リストを更新した場合、その更新されたセキュリティ制御は、次に主索引または差分索引の作成が行われたときに有効になります。

規則を Java クラスでエンコードすることにより、セキュリティ・トークンの値の判別にカスタム・ビジネス・ルールを適用することができます。クローラー・プロパティを構成するときに、文書をクロールする際にクローラーに使用させたいプラグインの名前を指定します。プラグインが追加するセキュリティ・トークンは、索引に保管され、文書へのアクセス制御に使用されます。

## 検索アプリケーションによるセキュリティ・トークンの使用方法

文書が適切にフィルター処理されるように検索時にセキュリティ・トークンを提供するの、クライアント検索アプリケーションの責任です。セキュリティ・トークンが提供されないと、デフォルトのパブリック・トークンが検索処理中に自動的に適用されます。



エンタープライズ・サーチのサンプル検索アプリケーションでは、文書レベルのセキュリティを実装する方法の例を示しています。この例では、(クロールされた文書のフィールドからセキュリティ・トークンを抽出するのではなく) 管理者がセキュリティ・トークン値を文書のグループに割り当てたことを想定しています。検索アプリケーションは、ユーザーのログイン ID を使用して、ユーザーがアクセスできる文書を判別します。この検索アプリケーションでは、実際のユーザー ID を使用するのではなく、ユーザーが属するグループ ID を使用しています。グループ ID をセキュリティ・トークンとして使用すると、グループに対してユーザーを追加や削除することができ、索引を再作成する必要がありません。

管理者によって文書セットに割り当てられたセキュリティ・トークンは、有効なオペレーティング・システム・グループ ID を表します。クロール・スペース内の文書ごとに異なるグループ ID が割り当てられます。次に例を示します。

Document1 から 5: セキュリティ・トークン = Group1  
Document6 から 10: セキュリティ・トークン = Group2

## 照会処理中の現行のクレデンシャルの検証

コレクションの作成時にそのコレクションのセキュリティを有効にした場合、特定のタイプのドメインで、ユーザーが照会を実行したときにユーザーの現行のクレデンシャルを検証することができます。

照会の応答を返す前に、検索サーバーはネイティブ・リポジトリとやりとりをして、ユーザーの現行のアクセス権を検証してから、ユーザーがアクセス権を持っていない文書はすべて検索結果に表示しないように除去します。

次のタイプのクローラーを構成するときに、ユーザーのクレデンシャルをネイティブのリポジトリが管理する現行のアクセス制御と比較して検証するオプションを選択できます。文書がクロールされて索引付けされた後、エンタープライズ・サーチ ID 管理コンポーネントを使用して、セキュア・コレクションを検索しようとするユーザーを検証します。

- Content Edition クローラー (Documentum、FileNet Panagon Content Services、Hummingbird DM、Portal Document Manager、および SharePoint リポジトリ・タイプのみ)
- DB2 Content Manager クローラー
- Lotus Domino Document Manager クローラー
- Lotus Notes クローラー
- QuickPlace クローラー
- Windows ファイル・システム・クローラー

次のタイプのクローラーの場合、現行ユーザーのクレデンシャルは、WebSphere Portal の検索ポートレットを使用してエンタープライズ・サーチ・コレクションを検索する場合に検証されます。

- Web Content Management クローラー
- WebSphere Portal クローラー

### 関連概念

314 ページの『Lotus Domino 文書に対する文書レベルのセキュリティの施行』

318 ページの『Windows ファイル・システム文書に対する文書レベルのセキュリティの施行』

### 関連タスク

315 ページの『ユーザーのクレデンシャルを検証するための Lotus Domino トラストッド・サーバーの構成』

## エンタープライズ・サーチ ID 管理

複数のユーザー・クレデンシャルの管理は、エンタープライズでよく問題となります。エンタープライズ・サーチ・システムでは、この問題を解決するために、オプションの ID 管理コンポーネントを提供しています。

### ヒント:

エンタープライズ・サーチの ID 管理コンポーネントのセットアップ方法の詳細な例については、IBM Redbook、IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios を参照してください。

エンタープライズにある情報は、さまざまな状態や形式で存在します。その情報は、エンタープライズ全体に配布して、そのタスクに対して最も適切で使用可能なソフトウェアで管理することができます。例えば、エンタープライズ・ユーザーは、SQL アプリケーションを使用してリレーショナル・データベースにアクセスしたり、文書管理システムを使用して仕事に関連する文書にアクセスしたりします。

こうしたリポジトリ内の機密情報へのアクセス制御は、通常、管理ソフトウェアにより実行されます。ユーザーは、ユーザー ID とパスワードの組み合わせで、ホスト・システムに対して身元確認を行います。システムによって認証された後は、管理ソフトウェアが、定義されているユーザーのアクセス権を基に、ユーザーがどの文書について表示または操作を許可されているかを制御します。

ユーザーが、各リポジトリに関連付けられた異なるユーザー ID とパスワードを持つことはよくあります。ユーザーは、元のエンタープライズ・リポジトリに対する身元確認を求められた場合と同様にして、現行のクレデンシャルの検証を必要とする、エンタープライズ・サーチ・コレクション内の文書を表示する前に、クレデンシャルを提供する必要があります。複数の ID を持つユーザーは、各 ID に対応するクレデンシャルを提示しなければなりません。

管理コンソールで ID 管理にエンタープライズ・サーチを使用するように指定した場合、検索サーバーは、次の方法を使用して、照会処理中にユーザーの現行のクレデンシャルを検証することができます。

- 検索アプリケーションが、さまざまなドメインへのアクセスに必要なクレデンシャルをユーザー・プロファイルに登録するためのプロンプトをユーザーに出すことができます。プロファイルは、セキュア・データ・ストアに暗号化されて保管されており、セキュア・ドメインの検索を可能にします。現行のクレデンシャルの検証を必要とするドメインに対してクレデンシャルが指定されないと、そのドメインからの文書は検索結果から除外されます。
- コレクション内の文書がシングル・サインオン (SSO) セキュリティのサポートを提供するクローラーによってクロールされ、文書のアクセス制御に SSO セキュリティを使用すると指定した場合、システムは検索セッション中に SSO セ

セキュリティー方式を使用してユーザーを認証します。ユーザーは、クレデンシャルを指定するプロファイルを作成したり、セキュア・ドメインの検索でユーザー ID とパスワードを提供したりする必要はありません。

照会の実行時に現行のクレデンシャルの検証を必要とするコレクションをユーザーが検索すると、システムはプロファイルまたは SSO セキュリティー方式を使用して、文書へのアクセスを拒否または許可します。

## ユーザーのグループ情報のリトリブ

ユーザーのクレデンシャルを検証するために、ID 管理コンポーネントは、各ユーザーの ID についてユーザーのグループ情報をリトリブして、この情報をユーザー・セキュリティー・コンテキスト (USC) ストリングに追加する必要があります。このグループ情報は、エンタープライズ・サーチ索引に保管されたアクセス制御データや、SSO 認証データに従って結果をフィルターに掛ける際に使用します。ID 管理コンポーネントは、このために、SSO トークンを使用するか、あるいは、ユーザーのクレデンシャルを使用し、バックエンド・システムに接続してユーザーが属するグループを要求します。

管理コンソールで ID 管理のオプションを構成する際に、このグループ情報のリフレッシュ頻度を指定することができます。ユーザーが検索アプリケーションにログインするたびに新しいグループ・データを抽出することも、あるいは、3 日ごとなど、定期的にグループ・データを抽出することもできます。

## ID 管理コンポーネントを使用しないセキュリティー

企業によっては、エンタープライズ・サーチ ID 管理コンポーネントによるユーザー・コミュニティの複数 ID の管理を望まないことがあります。エンタープライズ・サーチ管理コンソールで ID 管理コンポーネントを使用不可にした場合、検索アプリケーションの責任で、ユーザー・セキュリティー・コンテキスト・ストリングを生成することになります。USC ストリングは、生成された後、各照会に ACL 制約値を設定する際に使用されます。次に例を示します。

```
Query q = factory.createQuery("IBM");
q.setACLConstraints("User's Security Context in XML");
```

**ヒント:** 独自の ID 管理機能の作成を支援するために、検索および索引 API (SI-API) の拡張により、ID 管理データベースのプログラマチックな制御が提供されます。この API を使用すると、USC を Java オブジェクトと共に生成でき、その後、XML ストリングが自動的に作成されます。

XML 照会ストリングは、次の形式でなければなりません。ここで、... には、完全形式の XML ストリングが含まれます。

```
@SecurityContext::'...'
```

XML ストリングのフォーマットは、次のとおりです。

```
<identities id="login_UserName">
  <ssoToken>token_value</ssoToken>
  <identity id="security_domain">
    <type>Notes</type>
    <username>domain_userName</username>
    <password encrypt="no">domain_userPW</password>
  </identity>
</identities>
```

```

    <group id="g1" />
    <group id="g2" />
  </groups>
  <properties>
    <property name="property_name">property_value</property>
    ...
  </properties>
</identity>
...
</identities>

```

### **identities**

id 属性の値は、システムへのログイン時にユーザーが提供するユーザー ID です。

### **ssoToken**

オプション: ブラウザー・セッションの間にユーザーに対して作成される Lightweight Third-Party Authentication (LTPA) トークンを指定します。ターゲット・ドメインが SSO 使用可能で、クローラーが SSO セキュリティーを使用するように構成されている場合にのみ、このパラメーターが使用されます。

### **identity**

特定データ・ソースに対するユーザーのクレデンシャルを含みます。id 属性の値は、ユーザーのクレデンシャルを保管するドメインです (Lotus Domino の場合、これは Lotus Domino ドメイン名です)。

### **type**

データのタイプを識別し、クローラー・タイプ (Lotus Notes、DB2、Exchange Server など) に対応します。

### **username**

ドメインの検索に使用されるユーザー名を指定します。

### **password**

指定されたユーザー名のパスワードを指定します。encrypt 属性は no に設定されなければなりません (エンタープライズ・サーチは、ID 管理コンポーネントの外部で暗号化方式を提供しません)。

### **groups**

ユーザーが属するグループ名を指定します。各グループ名ごとに、別の group エレメントが使用されます。

### **properties**

クローラーの作成に使用された管理者 ID と暗号化されたパスワードや、ソースに SSO が使用可能かなど、接続固有のプロパティのリストを指定します。

#### **property\_name**

プロパティの名前。

#### **property\_value**

プロパティの値。

## **ユーザー・プロファイルを使用したユーザー検証**

検索アプリケーションは、さまざまなドメインへのアクセスに必要なクレデンシャルをユーザー・プロファイルに登録するためのプロンプトをユーザーに出すことができます。

照会が実行されたときにユーザーのクレデンシャルの検証が必要なドメインを検索する場合、ユーザーは、ドメインへのログインで使用するクレデンシャルを検索アプリケーションに提供しなければなりません。エンタープライズ・サーチ ID 管理を使用して、ユーザーは、任意の数のドメインに対するクレデンシャルをユーザー・プロファイルに保管することができます。クレデンシャルは暗号化されて、エンタープライズ・サーチ・システム内に安全に保管されます。

現行のクレデンシャルの検証を必要とするドメインに対してクレデンシャルが指定されないと、そのドメインからの文書は検索結果から除外されます。

ユーザーは、ユーザー・プロファイルを作成して、検索アプリケーションの使用中に自分のクレデンシャルを登録することができます。エンタープライズ・サーチのサンプル検索アプリケーションでは、この機能を「**個人用プロファイル**」オプションで提供しています。カスタム検索アプリケーションでは、この機能の実装が異なることがあります。

コレクションには、さまざまな多くのタイプのソースからの文書が含まれます。例えば、Windows ファイル・システムといくつかの Lotus Notes データベースからクロールされた文書が、コレクションに含まれることがあります。ID 管理コンポーネントは、異なるタイプのソースを区別し、検証を必要とするドメインへのアクセスに必要なクレデンシャルのみを求めます。

デフォルトで、各クレデンシャルは検索に対して使用可能になっており、セキュア・ドメインに対応するユーザー ID とパスワードの提供をユーザーに要求します。ユーザーが特定ドメインのユーザー ID またはパスワードを忘れた場合、チェック・ボックスを選択解除することによって、ドメインを検索に対して使用不可にすることができます。ドメインを使用不可にすると、そのドメインのセキュアな文書が結果セットに戻されなくなります。

プロファイルを作成した後、ユーザーは検索要求を実行することができます。ID 管理コンポーネントは、後続の検索要求で使用される、ユーザーのセキュリティー・コンテキスト (USC) スtringの作成に必要な情報を保持しています。ID 管理コンポーネントを使用しない場合は、現行のクレデンシャルの検証を必要とするドメインをユーザーが照会するときに、検索アプリケーションが USC Stringを提供する必要があります。

ユーザーが次にエンタープライズ・サーチ・コレクションを検索したとき、ID 管理コンポーネントはクレデンシャルの検証処理を繰り返しますが、今回はユーザーのプロファイルを検出することができます。何も変更されていないければ、ユーザーは、検索要求を実行可能なところに自動的に位置付けられ、プロファイルの作成は求められません。

ID 管理コンポーネントがいずれかのユーザーのクレデンシャルで変更を検出した場合、検索アプリケーションがアクセスされたときに、プロファイル・ページがユーザーに自動的に示されます。例えば、検索に対して使用可能になっているドメインのパスワードが変更された場合や、認証を必要とするドメインがコレクションに追加された場合に、これが行われます。

ユーザーは、プロファイル更新の推奨を無視しても構いませんが、そうすると、これらの文書が検索結果から除外されることとなります。

エンタープライズ・サーチで提供されるサンプル検索アプリケーションでは、ツールバーで「個人用プロフィール」を選択し、いつでもプロフィールを更新できます。

## SSO セキュリティーを使用したユーザー検証

コレクション内の文書がシングル・サインオン (SSO) セキュリティーのサポートを提供するクローラーによってクロールされる場合、ID 管理オプションの構成時に、文書のアクセス制御に SSO セキュリティーを使用することを指定できます。

## SSO の使用可能化

シングル・サインオン認証を使用すると、ユーザーは、1 回認証を行った後、クレデンシャルの再提示を求められることなく、多くのリソースへのアクセスを獲得できるようになります。エンタープライズ・サーチ・システムでは、SSO 認証により、セキュア・コレクションの文書のアクセスに指定する多くのユーザー名やパスワードの管理負担が緩和されます。

IBM WebSphere Application Server と Lotus Domino は、Lightweight Third-Party Authentication (LTPA) として知られる SSO の方式をサポートします。ユーザーは、いずれかの製品にアクセスすると、ユーザー名とパスワードによる認証を求められます。このユーザー名とパスワードは、両者の製品が共用する LDAP リポジトリに照らして検証されます。ユーザーが認証されると、LTPA トークンを含むセッション Cookie が作成されます。その後は、認証構成が同じサーバー上の他のリソースに、クレデンシャルの再指定を求められることなくアクセスすることができます。このトークンは、そのブラウザー・セッションが有効な限り持続します。

エンタープライズ・サーチ・コレクションで使用するために SSO サポートを使用可能にするには、次のようにします。

- WebSphere Application Server グローバル・セキュリティと有効な LDAP レジストリーが、エンタープライズ・サーチの検索サーバーで使用可能になっていることを確認します。LDAP レジストリーには、WebSphere Application Server でサポートされる任意の有効な LDAP 製品が可能です。
- WebSphere 認証メカニズムが、LTPA のアクティブな認証メカニズムを使用するように構成されていることを確認します。LTPA の構成時には、your.server.com といった有効で柔軟なドメイン名を指定します。
- LTPA 鍵が、WebSphere Application Server からエクスポートされ、LTPA のサポートを使用可能にする同じドメイン上の他の製品にインポートされていることを確認します。

上記のセキュリティ構成が適切に動作していることをブラウザーで確認した後、エンタープライズ・サーチ管理コンソールを使用して、SSO 認証をサポートするクローラーを構成します。

## SSO と ID 管理

現行のクレデンシャルの検証を必要とするコレクションをユーザーが検索すると、システムは SSO セキュリティー方式を使用して、文書へのアクセスを拒否または許可します。SSO 認証をサポートするソースの検索時に、ユーザーはクレデンシャルを求められません。ID 管理コンポーネントは、以下の条件がすべて当てはまる場合に使用されます。

- WebSphere Application Server とターゲット・ドメインで SSO が適切に使用可能になっている。
- 検索アプリケーションが検索できるコレクションの少なくとも 1 つで、セキュリティが有効になっている。
- エンタープライズ・サーチ管理コンソールで、ID 管理コンポーネントと SSO セキュリティーを使用するオプションが使用可能になっている。
- 以下のクローラー・タイプを構成したときに、SSO セキュリティーを使用するオプションと、文書レベルのセキュリティを施行するオプション (アクセス制御の索引付け、照会処理中の現行のクレデンシャルの検証など) が選択されている。
  - Content Edition (Portal Document Manager リポジトリにのみ使用可能)
  - Lotus Domino Document Manager (DIOP プロトコルを使用するクローラーにのみ使用可能)
  - Lotus Notes (DIOP プロトコルを使用するクローラーにのみ使用可能)
  - QuickPlace (DIOP プロトコルを使用するクローラーにのみ使用可能)

### 検索ポートレットのセキュリティ

エンタープライズ・サーチの検索ポートレットを使用して WebSphere Portal 内からコレクションを検索する場合は、シード・リスト、Web Content Management、および WebSphere Portal クローラーでクロールされる文書についてのセキュリティも提供されます。

このタイプのソースに対してセキュア検索がサポートされるのは、WebSphere Portal 外で実行する検索アプリケーションではなく、ポートレットを使用する場合のみです。WebSphere Portal 内では、ユーザー・クレデンシャルは検索ポートレットを使用して取得されます。ユーザーが WebSphere Portal にログインすると、すべての検索要求にユーザーのセキュリティ・データ (ユーザー名、グループ・メンバーシップなど) が組み込まれます。この情報は常に使用できるので、LTPA トークン・ベースの SSO は不要です。

しかし、別の製品を使用して WebSphere Portal サーバーのサイトおよび文書を保護する場合は、クローラーの構成時に SSO オプションを指定する必要があります。例えば、IBM Tivoli Access Manager WebSEAL や CA SiteMinder SSO Agent for PeopleSoft などの製品を使用する場合は、クローラーがサーバー上の文書にシングル・サインオンでアクセスできるように、クレデンシャルを指定する必要があります。この場合は、SSO がセキュア検索を使用できるようにするのではなく、SSO が保護されたコンテンツにクローラー・アクセスできるようにします。

### ID 管理の構成

エンタープライズ・サーチの ID 管理コンポーネントを使用して、照会処理中にユーザー・クレデンシャルをどのように検証するかを指定できます。

#### 始める前に

ID 管理オプションを構成するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

#### このタスクについて

照会処理中に現行のクレデンシャルの検証を必要とするコレクションをユーザーが検索すると、エンタープライズ・サーチの ID 管理コンポーネントは、ユーザー・プロファイルまたはシングル・サインオン (SSO) セキュリティー方式を使用して、文書へのアクセスを拒否または許可します。

## 手順

ID 管理を構成するには、次のようにします。

1. 「セキュリティ」をクリックして「セキュリティ」ビューを開きます。
2. 「検索アプリケーション」ページで、「ID 管理の構成」をクリックします。
3. 「ID 管理の構成」ページで、エンタープライズ・サーチ ID 管理コンポーネントを使用して照会処理中にユーザー・クレデンシャルをどのように検証するかを制御するチェック・ボックスを選択します。このチェック・ボックスが選択解除されている場合は、照会処理中に現行のクレデンシャルの検証を必要とするドメインをユーザーが照会するときに、検索アプリケーションがユーザー・セキュリティ・コンテキスト (USC) スtringを提供する必要があります。
4. ID 管理コンポーネントが、WebSphere Application Server ユーザー・レジストリーのグループ・レコードからユーザー・クレデンシャルを抽出する頻度を指定します。クレデンシャル・データのリフレッシュは、ユーザーが検索アプリケーションにアクセスするたびに、または指定した日数の経過後に行うことができます。
5. SSO 認証をサポートするクローラー・タイプの場合、ID 管理コンポーネントが、ユーザーの検証に、ユーザー・プロファイルではなく、SSO セキュリティー・トークンを使用するかどうかを指定します。すべてのクローラー・タイプで SSO 認証を使用するためのチェック・ボックスを選択することも、個々のクローラー・タイプごとにチェック・ボックスを選択することもできます。

**重要:** ID 管理コンポーネントは、WebSphere Application Server とターゲット・ドメインで SSO セキュリティーが適切に構成されている場合にのみ、SSO セキュリティー・メカニズムを使用します。

## アンカー・テキスト分析

コレクション・セキュリティを有効にすると、グローバル分析プロセスは、Web クローラーでクロールされる文書の中のアンカー・テキストの索引作成に、特別な規則を適用します。コレクション・セキュリティを有効にしない場合は、個別の Web クローラーを構成するときに禁止文書へのリンクのアンカー・テキストの索引を作成するかどうか指定することができます。

アンカー・テキストは、リンクの接続先のページを説明した、ハイパーテキスト・リンク内の情報です。例えば、次のリンクで、テキスト「Query Syntax」は、syntax.htm ページに接続するリンクの中のアンカー・テキストです。

```
<a href="../doc/syntax.htm">Query Syntax</a>
```

一般に、Web クローラーは文書の中のリンクをたどって、追加の文書をクロールし、これらのリンクされたページを索引の中に組み込みます。グローバル分析の間、索引は、アンカー・テキストが組み込まれる文書 (ソース文書) だけでなく、ターゲット文書にもアンカー・テキストを関連付ける処理をします。上記の例では、



アンカー・テキスト「Query Syntax」は、ターゲット・ページ `syntax.htm` とアンカー構成を含むソース・ページに関連付けられています。この関連付けにより、ソース文書に現れるテキストを指定した照会で、ターゲット文書がリトリートされるようになります。ただし、ユーザーがソース文書は表示できないがターゲット文書の表示は許可されている場合に、この関連付けにより、セキュリティのリスクが生じます。

コレクションを作成するときにコレクション・セキュリティを有効にすると、アンカー・テキスト処理は使用不可になります。実際にそのテキストが文書または文書のメタデータに出現しない限り、そのアンカー・テキストはもはや文書に索引付けされません。このセキュリティ制御により、ユーザーは、アクセスを許可されていない文書の情報を見ることはありません。文書自体のコンテンツまたはメタデータが照会に一致している場合のみ、文書が検索結果に戻されます。

コレクション・セキュリティを有効にすることで、ユーザーが、ユーザーのクレデンシャルとマッチするセキュリティ・トークンの文書のみを検索できるようにし、Web 文書のセキュリティを向上させることができます。しかし、アンカー・テキストを処理しないことで、照会に関連する可能性のある文書が、一部検索結果に含まれなくなるという可能性があります。

コレクション・セキュリティを有効にしない場合は、Web クローラーの拡張プロパティを構成するときに、禁止文書へのリンクのアンカー・テキストの索引を作成するかどうか指定することができます。

## 禁止文書へのリンク内のアンカー・テキストの索引作成

Web クローラーでのクロールが禁止されている文書へのリンクが文書に含まれている場合、それらのリンク用のアンカー・テキストを索引に保存するかどうかを、Web クローラーの構成時に指定できます。

### 始める前に

アンカー・テキストの索引作成のためのオプションを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、構成したい Web Crawler のコレクション管理者でなければなりません。


### このタスクについて

`robots.txt` ファイルまたは Web 文書のメタデータ内のディレクティブにより、Web Crawler が Web サイトで文書にアクセスできないようにすることができます。Web Crawler がクロールを許可されている文書に禁止文書へのリンクが組み込まれている場合、それらのリンクのアンカー・テキストをどのように扱うかを指定できます。

Web クローラーを構成するときに、禁止文書へのアンカー・テキストの索引を作成するかどうかを指定することができます。最も厳しいセキュリティを実現するためには、禁止文書へのリンク内のアンカー・テキストに索引を付け不要に指定してください。アンカー・テキストの索引を作成しないことで、照会に関連する可能性のある文書が、一部検索結果に含まれなくなるという可能性があります。

### 手順

禁止文書へのリンク内のアンカー・テキストの索引作成を使用可能または使用不可に設定するには、以下のようにします。

1. コレクションを編集し、「クロール」ページで、構成する Web クローラーを指して「 クローラーのプロパティ」をクリックします。
2. 「**拡張 Web クローラー・プロパティの編集**」をクリックします。
3. このクローラーがクロールするすべての文書でアンカー・テキストに索引を付けるには、「**禁止文書へのリンク内のアンカー・テキストの索引作成**」チェック・ボックスを選択します。ユーザーは、それらのページをポイントするリンクのアンカー・テキストに入っているテキストを検索すると、Web Crawler がクローラーを許可されていないページを知ることができます。

禁止文書へのアンカー・テキストを索引から除外するには、このチェック・ボックスを外します。ユーザーは、Web Crawler がクローラーを許可されていないページを知ることができません。アンカー・テキストは、禁止文書と共に、索引から除外されます。

4. 「**OK**」をクリックし、「Web Crawler のプロパティ」ページで、もう一度、「**OK**」をクリックします。
5. 変更を有効にするために、クローラーを停止し、再開します。

以前に索引付けされた文書に変更を適用するには、文書をもう一度索引付けするために、文書の再クロールが必要です。以前のクロールで禁止文書に関する情報が索引に追加されている場合、その情報は索引から除去されます。

---

## エンタープライズ・サーチのセキュリティーの使用可能化

ユーザーがエンタープライズ・サーチ・システムを管理または検索する際にセキュリティーを施行する予定の場合、WebSphere Application Server でグローバル・セキュリティーの構成が必要です。また、エンタープライズ・サーチ構成ファイル、管理コンソール、検索アプリケーションでセキュリティー情報を構成することも必要です。

### ヒント:

文書レベルのセキュリティーが確実に施行されるようにさまざまなクローラーを設定する方法の例を含め、LDAP リポジトリを備えた WebSphere Application Server でグローバル・セキュリティーを使用可能にする方法の詳細な例については、IBM Redbook、IBM OmniFind Enterprise Edition Configuration and Implementation Scenarios を参照してください。

### 手順

エンタープライズ・サーチ・システムのセキュリティーを使用可能にするには、次のようにします。

1. ユーザーの認証に使用するユーザー・レジストリーのタイプを決定します。例えば、多くの WebSphere Application Server 管理者が、Lightweight Directory Access Protocol (LDAP) ユーザー・レジストリーの使用を選択します。

2. エンタープライズ・サーチ管理コンソールで、「**セキュリティ**」を選択し、WebSphere Application Server ユーザー・レジストリー内の少なくとも 1 人のユーザーを「**エンタープライズ・サーチ管理者**」の管理役割に割り当てます。

あるいは、OmniFind Enterprise Edition のインストール時に指定されたエンタープライズ・サーチ管理者 ID を WebSphere Application Server ユーザー・レジストリーに追加します。

**重要:** グローバル・セキュリティが使用可能になった後は、ユーザー・レジストリー内にあってエンタープライズ・サーチ管理役割に割り当てられているユーザー ID のみが、管理コンソールにアクセスしてエンタープライズ・サーチを管理することができます。

3. 『WebSphere Application Server におけるグローバル・セキュリティと LDAP ユーザー・レジストリーの構成』 の手順に従って、グローバル・セキュリティを使用可能にし、ユーザー・レジストリーを構成します。
4. OmniFind Enterprise Edition のインストール後にグローバル・セキュリティを使用可能にした場合、エンタープライズ・サーチ・システムに WebSphere Application Server のユーザー ID とパスワードを提供する必要があります。この情報を提供するには、**eschangewaspw** コマンドを使用します。
  - OmniFind Enterprise Edition をシングル・サーバーにインストールした場合は、307 ページの『シングル・サーバー・エンタープライズ・サーチ・システムに対するセキュリティの使用可能化』 の手順に従います。
  - OmniFind Enterprise Edition を複数のサーバーにインストールした場合は、308 ページの『マルチサーバー・エンタープライズ・サーチ・システムに対するセキュリティの使用可能化』 の手順に従います。
5. クロールおよび検索する予定の文書のタイプに応じたタスクを完了します。詳しくは、310 ページの『セキュリティをサポートするためのクローラーのセットアップ要件』を参照してください。

## WebSphere Application Server におけるグローバル・セキュリティと LDAP ユーザー・レジストリーの構成

OmniFind Enterprise Edition でセキュリティを使用可能にするには、まず、WebSphere Application Server でグローバル・セキュリティを使用可能に設定する必要があります。

### このタスクについて

グローバル・セキュリティの使用可能化の一部として、ユーザー ID を認証するユーザー・レジストリーの構成が必要です。以下のタスクでは、WebSphere Application Server でグローバル・セキュリティを使用可能にすると同時に、Lightweight Directory Access Protocol (LDAP) ユーザー・レジストリーを構成する方法について説明します。WebSphere Application Server によって他のタイプのユーザー・レジストリーがサポートされますが、エンタープライズ・サーチ・ユーザーの認証にローカル・オペレーティング・システムのレジストリーを使用することはできません。ローカル・オペレーティング・システムのレジストリーを使用するに

は、エンタープライズ内の全ユーザーのオペレーティング・システム・ユーザー・アカウントが、エンタープライズ・サーチの検索サーバーに存在しなければなりません。

以下のタスクは、WebSphere Application Server バージョン 6 に基づいています。それより前のバージョンの WebSphere Application Server を使用すると、デフォルト・パスや一部のユーザー・インターフェース・ラベルが異なる場合があります。また、このタスクでは、LDAP レジストリーに IBM Tivoli Directory Server を使用します。別のレジストリー・タイプまたはカスタム・レジストリーを使用する場合は、レジストリーに対して適切な情報を提供してください。

このタスクでは、エンタープライズ・サーチ・システムで使用するためにグローバル・セキュリティーを構成するのに必要なステップを要約しています。詳細な手順については、次の URL にある *WebSphere Application Server, Version 6.0.x Information Center* を参照してください。 <http://publib.boulder.ibm.com/infocenter/wasinfo/v6r0/index.jsp>

## 手順

WebSphere Application Server でグローバル・セキュリティーを使用可能にするには、次のようにします。

1. エンタープライズ・サーチの検索サーバーで、次の URL にアクセスし、WebSphere Application Server 管理コンソールを開きます。ここで、*localhost* は、localhost、または *omnifind.search.xyz.com* などのサーバー名です。

`http://localhost:9060/ibm/console`

2. 「**セキュリティー (Security)**」をクリックし、次に「**グローバル・セキュリティー (Global Security)**」をクリックします。
3. LDAP レジストリーを使用するように WebSphere をセットアップします。
  - a. 「**ユーザー・レジストリー (User registries)**」で「**LDAP**」をクリックします。
  - b. アプリケーション・サーバーの実行に使用されるサーバー・ユーザー ID とパスワードを指定します。
  - c. レジストリー・タイプに IBM Tivoli Directory Server を選択します。
  - d. IP アドレスまたはドメイン・ネーム・サービス (DNS) ホスト名のいずれかで、LDAP サーバー・ホスト名を指定します。デフォルト・ポート番号は 389 です。
  - e. レジストリーを検索する開始点であるベース識別名 (DN) (`ou=sales,o=ibm,c=us` など) を指定します。
  - f. レジストリーが検索されるときに匿名バインディングをサポートしない LDAP サーバーがあるため、アプリケーション・サーバーの DN (`cn=searchuser,o=ibm,c=us` など) を指定し、次に、アプリケーション・サーバーのパスワードを指定します。アプリケーション・サーバーは、この DN とパスワードを使用して、レジストリーにバインドします。
  - g. WebSphere と LDAP の間で Secure Sockets Layer (SSL) 通信を使用する場合、「**SSL 使用可能 (SSL enabled)**」チェック・ボックスを選択します。
  - h. 「**適用**」をクリックし、次に「**OK**」をクリックします。

4. 「一般プロパティ (General Properties)」で「グローバル・セキュリティーを使用可能にする (Enable global security)」と「Java 2 セキュリティーを施行する (Enforce Java 2 security)」のチェック・ボックスを選択します。
5. アクティブな認証メカニズムに Simple WebSphere Authentication Mechanism (SWAM) を選択します。
6. アクティブなユーザー・レジストリーに Lightweight Directory Access Protocol (LDAP) ユーザー・レジストリーを選択します。
7. 「OK」をクリックします。
8. ページ上部の「保存」リンクをクリックします。変更内容の保存のプロンプトが表示されたら、「保存」ボタンをクリックします。
9. ツールバーで「ログアウト」をクリックします。
10. ESSearchServer アプリケーションを停止し、再始動します。

#### AIX、Linux、または Solaris

```
./stopServer.sh ESSearchServer
./startServer.sh ESSearchServer
```

#### Windows

```
stopServer ESSearchServer
startServer ESSearchServer
```

これらのスクリプトは、WAS\_INSTALL\_ROOT/AppServer/bin ディレクトリーにあります。

- WebSphere Application Server バージョン 5 の場合、デフォルトのインストール・パスは、AIX システムでは /usr/WebSphere、Linux または Solaris システムでは /opt/WebSphere、Windows システムでは C:\Program Files\WebSphere です。
  - WebSphere Application Server バージョン 6 の場合、デフォルトのインストール・パスは、AIX システムでは /usr/IBM/WebSphere、Linux または Solaris システムでは /opt/IBM/WebSphere、Windows システムでは C:\Program Files\IBM\WebSphere です。
11. WebSphere Application Server 管理コンソールを再始動します。
  12. サーバーはこれでセキュア・モードで開始しているため、コンソールへのログインには、LDAP ユーザー・レジストリーの構成時に指定したサーバー・ユーザー ID とパスワードを入力してください (ステップ 3b (306 ページ) を参照)。

## シングル・サーバー・エンタープライズ・サーチ・システムに対するセキュリティーの使用可能化

OmniFind Enterprise Edition のインストール後に WebSphere Application Server のグローバル・セキュリティーを使用可能にする場合は、**eschangewaspw** コマンドを使用して、WebSphere Application Server ユーザーのパスワードで、エンタープライズ・サーチ構成ファイル **es.cfg** を更新する必要があります。

始める前に

ESSearchApplication アプリケーションの config.properties ファイルで有効な WebSphere Application Server ユーザー名とパスワードを指定していることを確認します。このファイルのデフォルトの場所は、ES\_INSTALL\_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war/WEB-INF です。

### このタスクについて

**eschangewaspw** コマンドは、パスワードを es.cfg ファイルに保管する前に暗号化します。

### 手順

既存のシングル・サーバー・エンタープライズ・サーチ・システムがグローバル・セキュリティーを使用できるようにするには、次のようにします。

1. エンタープライズ・サーチ・サーバーで、エンタープライズ・サーチ管理者としてログインし、エンタープライズ・サーチ・システムを停止します。

```
esadmin system stopall
```

2. ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルの WASUser 項目で有効な WebSphere Application Server ユーザー名を指定していることを確認します。
3. 以下のスクリプトを実行します。ここで、*WAS\_password* は、ES\_NODE\_ROOT/nodeinfo/es.cfg ファイルに指定された WebSphere Application Server ユーザー名のパスワードです (ステップ 2 を参照)。

#### AIX、Linux、または Solaris

```
eschangewaspw.sh WAS_password
```

#### Windows

```
eschangewaspw WAS_password
```

4. Windows で、「コントロール パネル」 → 「管理ツール」 → 「サービス」を選択し、WebSphere Application Server と ESSearchServer サービスに同じ WebSphere Application Server ユーザー名とパスワードを追加します。
5. エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```

## マルチサーバー・エンタープライズ・サーチ・システムに対するセキュリティーの使用可能化

OmniFind Enterprise Edition のインストール後に WebSphere Application Server のグローバル・セキュリティーを使用可能にする場合は、**eschangewaspw** コマンドを使用して、WebSphere Application Server ユーザーのパスワードで、エンタープライズ・サーチ構成ファイル es.cfg を更新する必要があります。

### 始める前に

ESSearchApplication アプリケーションの config.properties ファイルで有効な WebSphere Application Server ユーザー名とパスワードを指定していることを確認します。このファイルのデフォルトの場所は、検索サーバーの ES\_INSTALL\_ROOT/installedApps/ESSearchApplication.ear/ESSearchApplication.war /WEB-INF です。

## このタスクについて

**eschangewaspw** コマンドは、パスワードを **es.cfg** ファイルに保管する前に暗号化します。

### 手順

既存のマルチサーバー・エンタープライズ・サーチ・システムがグローバル・セキュリティーを使用できるようにするには、次のようにします。

1. エンタープライズ・サーチ索引サーバーで次のステップを行います。
  - a. エンタープライズ・サーチ管理者でログインし、エンタープライズ・サーチ・システムを停止します。

```
esadmin system stopall
```

- b. **ES\_NODE\_ROOT/nodeinfo/es.cfg** ファイルの **WASUser** 項目で有効な WebSphere Application Server ユーザー名を指定していることを確認します。
- c. 以下のスクリプトを実行します。ここで、**WAS\_password** は、**ES\_NODE\_ROOT/nodeinfo/es.cfg** ファイルに指定された WebSphere Application Server ユーザーのパスワードです (ステップ 1b を参照)。

#### AIX、Linux、または Solaris

```
eschangewaspw.sh WAS_password
```

#### Windows

```
eschangewaspw WAS_password
```

2. 2 番目の検索サーバー (2 サーバー構成の場合) で、またはクローラー・サーバーと検索サーバーの両方 (4 サーバー構成の場合) で、次のステップを行います。

- a. エンタープライズ・サーチ管理者としてログインします。
- b. 以下のスクリプトを実行します。ここで、**WAS\_password** は、**ES\_NODE\_ROOT/nodeinfo/es.cfg** ファイルに指定された WebSphere Application Server ユーザーのパスワードです (ステップ 1b を参照)。

#### AIX、Linux、または Solaris

```
eschangewaspw.sh WAS_password
```

#### Windows コマンド・プロンプト

```
eschangewaspw WAS_password
```

3. Windows で、「コントロール パネル」 → 「管理ツール」 → 「サービス」を選択し、WebSphere Application Server と ESSearchServer サービスに同じ WebSphere Application Server ユーザー名とパスワードを追加します。
4. エンタープライズ・サーチ索引サーバーで、エンタープライズ・サーチ・システムを再始動します。

```
esadmin system startall
```

---

## セキュリティをサポートするためのクローラーのセットアップ要件

文書レベルのセキュリティを施行するための情報を収集するために、クローラーには、ネイティブ・セキュリティ・データへのアクセス許可が必要です。一部のデータ・タイプでは、セキュアな環境を構成するために追加の手順が必要になります。

表 7. セキュリティをサポートするためのクローラーのセットアップ要件

---

### Content Edition クローラー

直接モードでリポジトリにアクセスするクローラー 関連トピック:

を作成する前に、WebSphere Information Integrator Content Edition システムが直接モードで稼働するように構成し、クローラー・サーバーのコネクタを構成します。

- 50 ページの『Content Edition リポジトリへの直接モード・アクセス』
- 52 ページの『WebSphere II Content Edition リポジトリへのサーバー・モード・アクセス』

サーバー・モードでリポジトリにアクセスするクローラーを作成する前に、クローラー・サーバーを構成するためのスクリプト (AIX、Linux、または Solaris では `escrvbr.sh`、Windows では `escrvbr.vbs`) を実行します。

- 52 ページの『UNIX での WebSphere II Content Edition 用クローラー・サーバーの構成』
- 53 ページの『Windows での WebSphere II Content Edition 用クローラー・サーバーの構成』

クローラーを構成する際に、クロールされる各リポジトリにクローラーがアクセスするためのユーザー ID とパスワードを指定します。必要に応じて、クロール・スペース内の各リポジトリごとに、異なるユーザー ID とパスワードを指定することができます。

---

### DB2 クローラー

クローラーを作成する前に、クローラー・サーバーを構成するためのスクリプト (AIX、Linux、または Solaris では `escredb2.sh`、Windows では `escredb2.vbs`) を実行します。

- 関連トピック:
- 57 ページの『UNIX 上での DB2 クローラー用のクローラー・サーバーの構成』
  - 58 ページの『Windows 上での DB2 クローラー用のクローラー・サーバーの構成』

リモートまたはアンカタログ・データベースをクロールするクローラーを構成する場合、ターゲット・データベース・サーバー上の各データベースをクロールするためのユーザー ID とパスワードを指定します。必要に応じて、クロール・スペース内の各データベースごとに、異なるユーザー ID とパスワードを指定することができます。

---

### DB2 Content Manager クローラー

クローラーを作成する前に、クローラー・サーバーを構成するためのスクリプト (AIX、Linux、または Solaris では `escrcm.sh`、Windows では `escrcm.vbs`) を実行します。

- 関連トピック:
- 66 ページの『UNIX での DB2 Content Manager クローラー用クローラー・サーバーの構成』
  - 68 ページの『Windows での DB2 Content Manager クローラー用クローラー・サーバーの構成』

クローラーを構成する際に、クロールされる各サーバーにクローラーがアクセスするためのユーザー ID とパスワードを指定します。必要に応じて、クロール・スペース内の各サーバーごとに、異なるユーザー ID とパスワードを指定することができます。

---

### Lotus Domino Document Manager、Lotus Notes、および QuickPlace クローラー

---



表7. セキュリティーをサポートするためのクローラーのセットアップ要件 (続き)

<p>Lotus Notes リモート・プロシージャー・コール (NRPC) プロトコルを使用する Lotus Domino サーバーをクローリングするには、次のようにします。</p> <ul style="list-style-type: none"> <li>• AIX システムでは、I/O 完了ポート・モジュールがクローラー・サーバーにインストールされており、使用可能であることを確認します。</li> <li>• クローラーを作成する前に、クローラー・サーバーを構成するためのスクリプト (AIX、Linux、または Solaris では <code>escrnote.sh</code>、Windows では <code>escrnote.vbs</code>) を実行します。</li> <li>• Lotus Domino サーバーが、エンタープライズ・サーチ・クローラー・サーバーにインストールされていなければなりません。この Lotus Domino サーバーは、クローリングされる Lotus Domino ドメインのメンバーである必要があります。</li> <li>• ユーザーが検索要求を実行したときに現行のユーザー・クレデンシャルを検証するには、クローリングされる Lotus Domino サーバーが、Lotus Domino トラストド・サーバーとして構成されていなければなりません。</li> <li>• クローラーを構成する際に、サーバーへのアクセスを許可された Lotus Notes ユーザー ID ファイルのパス (<code>c:\Program Files\lotus\notes\data\name.id</code> や <code>/local/notesdata/name.id</code> など) と、この ID ファイルのパスワードを指定します。</li> </ul>	<p>関連トピック:</p> <ul style="list-style-type: none"> <li>• 92 ページの『Lotus Domino ソースをクローリングするための AIX での I/O 完了ポートの構成』</li> <li>• 87 ページの『Lotus Domino ソースをクローリングするための UNIX でのクローラー・サーバーの構成』</li> <li>• 89 ページの『Lotus Domino ソースをクローリングするための Windows でのクローラー・サーバーの構成』</li> <li>• 315 ページの『ユーザーのクレデンシャルを検証するための Lotus Domino トラストド・サーバーの構成』</li> <li>• 91 ページの『DIIOP プロトコルを使用するサーバーの構成』</li> <li>• 316 ページの『ローカル・ユーザー・セキュリティを使用するための QuickPlace サーバーの構成』</li> <li>• 318 ページの『QuickPlace での Directory Assistance の構成』</li> </ul>
<p>Lotus Domino Internet Inter-ORB Protocol (DIIOP) を使用する Lotus Domino サーバーをクローリングするには、次のようにします。</p> <ul style="list-style-type: none"> <li>• AIX システムでは、I/O 完了ポート・モジュールがクローラー・サーバーにインストールされており、使用可能であることを確認します。</li> <li>• このプロトコルを使用できるようにクローラー・サーバーを構成します。</li> <li>• クローラーを構成する際に、サーバーへのアクセスを許可された完全修飾 Lotus Notes ユーザー ID (User Name/Any Town/My Company など) と、このユーザー ID のパスワードを指定します。</li> </ul>	
<p>QuickPlace サーバーをクローリングするには、使用するセキュリティのタイプに応じて、ローカル・ユーザー・セキュリティまたは Directory Assistance をサポートするように QuickPlace サーバーを構成する必要があります。</p>	
<hr/> <p><b>Exchange Server クローラー</b></p> <hr/>	

表7. セキュリティーをサポートするためのクローラーのセットアップ要件 (続き)

---

クローラーを構成する際に、クローラされる Exchange Server のパブリック・フォルダーへのアクセスを許可されたユーザー ID と、このユーザー ID のパスワードを指定します。	関連トピック: <ul style="list-style-type: none"><li>• 314 ページの『セキュア Exchange Server 文書へのアクセスの検証』</li></ul>
------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------

---

データのクローラ時にクローラーが Exchange Server 鍵管理と Secure Sockets Layer (SSL) プロトコルを使用するためには、鍵ストア・ファイルへの完全修飾パスと、クローラーがこのファイルにアクセスするためのパスワードも指定します。鍵ストア・ファイルは、エンタープライズ・サーチ・クローラー・サーバー上に存在しなければなりません。

---

#### JDBC データベース・クローラー

---

クローラーを構成する際に、ターゲット・データベース内の表をクローラするためのユーザー ID とパスワードを指定できます。必要に応じて、クローラ・スペース内の各データベースごとに、異なるユーザー ID とパスワードを指定することができます。

---

#### NNTP クローラー

---

クローラされる NNTP サーバーは、クローラー・サーバーによるデータの読み取りを許可していなければなりません。

---

#### UNIX ファイル・システム・クローラー

---

クローラされる AIX、Linux、および Solaris のサブディレクトリーは、クローラー・サーバーによるデータの読み取りを許可していなければなりません。

---

#### Web クローラー

---

表7. セキュリティーをサポートするためのクローラーのセットアップ要件 (続き)

---

Web クローラーは、ロボット排除プロトコルに従います。Web サーバーが、サーバー・ディレクトリーの最上位レベルに robots.txt ファイルを含む場合、クローラーはそのファイルを分析し、そのサーバー上の Web サイトのクローリングを許可されている場合のみ、そのサイトをクローリングします。このプロトコルについては、<http://www.robotstxt.org/wc/exclusion.html> を参照してください。

関連トピック:

- 114 ページの『HTTP 基本認証によって保護された Web サイト』
- 115 ページの『フォーム・ベースの認証によって保護された Web サイト』

Web クローラーを構成する場合:

- クローラーのユーザー・エージェント名を指定してください。クローリングされるサーバーの robots.txt ファイルの規則では、この名前を指定してアクセスを許可または拒否します。
- オプション: Web サーバーが HTTP 基本認証を使用して Web サイトへのアクセスを制限している場合、Web クローラーがパスワード保護されたページにアクセスするための認証のクレデンシャルを指定できます。
- オプション: Web サーバーが HTML フォームを使用して Web サイトへのアクセスを制限する場合、Web クローラーがパスワード保護されたページにアクセスするための認証のクレデンシャルを指定できます。

---

#### シード・リスト、Web Content Management、および WebSphere Portal クローラー

クローラーを作成する前に、エンタープライズ・サーチを WebSphere Portal サーバーと統合するためのセットアップ・スクリプトを実行する必要があります。WebSphere Portal のバージョンごとに異なるスクリプトが提供されています。

関連トピック:

- 382 ページの『エンタープライズ・サーチの WebSphere Portal との統合のセットアップ・スクリプト』

クローラーを構成する際、クローリングされるサーバーからクローラーがページをリトリートするための完全修飾識別名 (DN)

(uid=admin,cn=RegularEmployees,ou=Software Group,o=IBM,c=US など) を指定し、この DN のパスワードを指定します。この DN は、WebSphere Portal に構成された DN と一致する必要があります。

指定したユーザー DN の許可が、WebSphere Portal の Portal Access Control (PAC) コンポーネントで定義されていることを確認してください。クローラーは PAC を使用して、クローリングする文書のアクセス制御データを取得します。

---

#### Windows ファイル・システム・クローラー

表7. セキュリティーをサポートするためのクローラーのセットアップ要件 (続き)

クローラーされるサブディレクトリーは、クローラー・

サーバーによるデータの読み取りを許可してなければなりません。リモート・ファイル・システムをクローラーするクローラーを構成する際に、クローラーがリモート・データにアクセスするためのユーザー ID を指定し、このユーザー ID のパスワードを指定します。

- 関連トピック:
- 318 ページの『Windows ファイル・システム文書に対する文書レベルのセキュリティーの施行』
  - 321 ページの『Windows トラステッド・ドメインのセキュア検索』

ユーザーが検索要求を実行したときに現行のユーザー・クレデンシャルを検証するには、ドメイン・アカウントが正しく構成されていることを確認してください。ローカル・コンピューター上でクローラーされたファイルのドメイン・アカウントをセットアップするための要件は、リモート Windows サーバー上でクローラーされたファイルのための要件とは異なります。

## セキュア Exchange Server 文書へのアクセスの検証

ファイアウォールによって保護された文書を Exchange Server クローラーを使用してクローラーするには、クローラー・サーバーが Microsoft Exchange Server パブリック・フォルダー・サーバーにアクセスできることを検証する必要があります。

### このタスクについて

クローラー・サーバーがセキュア Exchange Server サーバーにアクセスできない場合、サーバーから HTTP コード 501 (未実装) を受け取ります。予期しない HTTP 応答を受け取ったことを示すメッセージも表示されます。

### 手順

ファイアウォールで保護された文書にクローラー・サーバーがアクセスできることを検証するには、次のようにします。

1. クローラー・サーバー上で Web ブラウザーを起動します。
2. クローラーしたい Exchange Server パブリック・フォルダー・サーバーの URL に移動します。例: <http://exchange.yourCompany.com/public/>
3. Exchange Server ページをオープンできることを検証します。

Exchange Server サーバーにアクセスできない場合、社内のサーバー管理者に連絡してください。

## Lotus Domino 文書に対する文書レベルのセキュリティーの施行

クローラーされる Lotus Domino サーバーが、Lotus Notes リモート・プロシージャ・コール (NRPC) プロトコルを使用する場合、文書レベルのアクセス制御を施行するように、クローラー・サーバーを構成する必要があります。

NRPC プロトコルを使用する Lotus Domino サーバー上の文書の文書レベルのセキュリティーを施行するには、クローラー・サーバーに Lotus Domino サーバーをインストールする必要があります。この Lotus Domino サーバーは、ご使用の Lotus

Domino ドメインのメンバーでなければなりません。Lotus Domino サーバーをインストールおよび構成するには、Lotus Domino 資料の中の指示に従ってください。

セキュア・コレクションを検索するユーザーが、検索条件に一致した文書の表示を許可されているかどうかを検索サーバーで検査できるように、以下のタスクも実行する必要があります。

- 『ユーザーのクレデンシャルを検証するための Lotus Domino トラステッド・サーバーの構成』
- 305 ページの『WebSphere Application Server におけるグローバル・セキュリティと LDAP ユーザー・レジストリーの構成』

#### 関連概念

295 ページの『照会処理中の現行のクレデンシャルの検証』

83 ページの『Lotus Notes クローラー』

## ユーザーのクレデンシャルを検証するための Lotus Domino トラステッド・サーバーの構成

Lotus Notes リモート・プロシージャ・コール (NRPC) プロトコルを使用する Lotus Notes クローラーによってクロールされた文書にセキュリティを施行するには、クロールされる Lotus Domino サーバーを Lotus Domino トラステッド・サーバーとして構成する必要があります。

### 始める前に

リモート・データベースを検索する際に文書レベルのセキュリティを施行したい場合にのみ、この手順を実行する必要があります。クローラー・サーバーのローカル・データベースを検索するには、この手順は必要ありません。

トラステッド・サーバーを構成するためには、クローラーに Lotus Domino サーバーをインストールしなければなりません。この Lotus Domino サーバーは、ご使用の Lotus Domino ドメインのメンバーでなければなりません。

### このタスクについて

Lotus Notes クローラーに対して文書レベルのセキュリティ・オプションを構成するときは、ユーザーが照会を実行したときにそのユーザーの現行のクレデンシャルを検証することでアクセス制御を施行するかどうかを指定します。このタイプのセキュリティを施行する場合は、クロールされる Lotus Domino サーバーを Lotus Domino トラステッド・サーバーにしなければなりません。

ユーザーの現行のクレデンシャルを検証する必要があるドメインをユーザーが検索するとき、トラステッド・サーバーは、Lotus Domino サーバー ID がコンテキストを現行のユーザー ID に切り替えられるようにします。Lotus Domino データベースは、現行のユーザーがそのデータベースをオープンしたかのようにオープンされ、そのユーザーに関するすべてのデータベース・アクセス制御リスト情報が施行されます。

この方法でコンテキストを切り替える機能は、一般的には、ローカル Lotus Domino サーバーの data ディレクトリーに保管されているデータベースでのみ使用可能です。Lotus Domino バージョン 6.5.1 から、この機能はトラステッド・サーバーに

よって提供されています。トラステッド・サーバーを構成するには、機密の操作 (例えば、データベースがリモート・コンピューターからアクセスされたときに別のユーザーとして動作するなど) を行う際にどの Lotus Domino サーバーを信用すべきかを、Lotus Domino 管理者が指定します。

## 手順

トラステッド・サーバーを構成するには、Lotus Notes クローラーによってクロールされるすべての Lotus Domino サーバー上で次のステップを実行します。

1. Lotus Domino サーバーで、Lotus Domino ドメイン管理者 ID ファイルを使用して、Lotus Domino Administrator クライアントを開きます。
2. 「ファイル」をクリックし、「サーバーを開く」を選択します。
3. トラステッド・サーバー機能を使用可能にする Lotus Domino サーバーの名前を入力します。
4. 「構成」タブを選択します。
5. 「サーバー」オブジェクトを展開し、「現行サーバー」文書を選択して、「サーバーの編集」をクリックします。
6. 「セキュリティー」タブを選択し、文書の最後までスクロールし、「トラステッド・サーバー」の項目を見つけて、下矢印をクリックします。
7. 次のオプションのいずれかを指定します。

### LocalDomainServers

Lotus Domino ドメイン内のすべてのサーバーをトラステッド・サーバーと見なす場合は、このオプションを選択します。

#### *server\_name*

トラステッド・サーバーとしてクロールおよび検索可能な Lotus Domino サーバーの名前を指定します。

クロールしたい Lotus Domino サーバーが別の Lotus Domino ドメインにある場合、サーバー名を指定するか、あるいは **OtherDomainServers** グループを選択する必要があります。エンタープライズ・サーチ Lotus Domino サーバー ID ファイルを他の Lotus Domino ドメインとクロス認証するためには、さらに Lotus Domino 用の手順を実行する必要があります。これらの手順について詳しくは、Lotus Domino サーバーの資料を参照してください。

8. 「保存してクローズ」をクリックして、変更内容を保存します。
9. トラステッド・サーバーとして作動するようにしたリモート Lotus Domino サーバーを停止および再始動します。

### 関連概念

295 ページの『照会処理中の現行のクレデンシャルの検証』

83 ページの『Lotus Notes クローラー』

## ローカル・ユーザー・セキュリティーを使用するための QuickPlace サーバーの構成

ローカル・ユーザー・オプションを使用してセキュリティーを実装するように QuickPlace クローラーを構成する予定の場合、クローラーを作成する前に、Lotus QuickPlace サーバー上に Lotus Domino Directory を構成する必要があります。

## このタスクについて

QuickPlace クローラーを構成するときに、文書レベルのセキュリティーを施行するために使用するクローラーのセキュリティー・モードを選択します。「ローカル・ユーザー」モードを選択する場合、すべてのローカル・ユーザー ID およびローカル・グループが Domino Directory (Domino Directory 階層は QuickPlace 階層に対応していなければなりません) に登録されていることを確認する必要があります。

また、クローラーが使用するよう指定するユーザー ID とパスワードが Domino Directory に登録されていて、クロール対象データベースの読み取り許可があることも確認する必要があります。

QuickPlace を使用するには、ユーザー名のみが必要です。しかし、QuickPlace ソースをクロールするには、完全に展開されたユーザー ID が必要です。展開されたユーザー ID のフォーマットは次のとおりです。

`username/placename/QP/domainname`

この手順を使用して、完全に展開されたバージョンのユーザー ID を決定し、このユーザー ID が QuickPlace データベースの読み取りを認可されていることを確認し、ユーザー ID を Domino Directory に追加します。Domino Directory には、QuickPlace データベースのクロールに使用されるユーザー ID と、すべての QuickPlace ローカル・ユーザーおよびグループが含まれている必要があります (Domino Directory 階層は QuickPlace 階層に対応していなければなりません)。

## 手順

ローカル・ユーザー・セキュリティーを使用するために QuickPlace サーバーを構成するには、次のようにします。

1. ユーザー ID 許可を確認します。
  - a. QuickPlace サーバーで Server 文書を開きます。
  - b. 「ファイル」ページを開き、クロールするデータベースのアクセス制御リスト (ACL) を開きます。
  - c. クローラーが使用するよう構成するローカル・ユーザー ID が ACL に存在することと、このユーザー ID がデータベースの読み取り許可を持っていることを確認します。

このユーザー ID は、ステップ 2 に説明がある完全に展開された形式で指定する必要があります。

2. ユーザーを Domino Directory に追加します。
  - a. QuickPlace サーバーで Server 文書を開きます。
  - b. 「ユーザーとグループ」ページのユーザー・ツリー項目に、ステップ 1 で確認した、完全に展開されたユーザー ID を追加します。
  - c. 「インターネット・パスワード」フィールドに、このユーザー ID のパスワードを指定します。

## QuickPlace での Directory Assistance の構成

LDAP ディレクトリーを使用してセキュリティを実装するように QuickPlace クローラーを構成する予定の場合、クローラーを作成する前に、Lotus QuickPlace サーバー上に Directory Assistance データベースを作成する必要があります。

### 制約事項

クロール対象にする QuickPlace サーバーは、DIIOP タスクおよび HTTP タスクを実行中でなければなりません。

### 手順

QuickPlace サーバー上に LDAP Directory Assistance を構成するには、次のようにします。

1. Directory Assistance データベースを作成します。
  - a. QuickPlace サーバーで Server 文書を開きます。
  - b. **Directory Assistance(6)** テンプレートを使用してデータベースを作成します。このテンプレートは、サーバー上にあります。
  - c. データベース内に文書を作成するため「**Directory Assistance の追加**」をクリックします。
  - d. 「基本」タブを開き、「**DomainType**」フィールドで「**LDAP**」を選択します。
  - e. 「命名コンテキスト」タブを開き、「**クレデンシャルに対してトラステッド**」チェック・ボックスが選択されていることを確認します。
  - f. 「LDAP」タブで LDAP サーバーに関する情報を指定します。
  - g. Server 文書を保存してクローズします。
2. 次のように、Directory Assistance データベースを使用するための QuickPlace サーバーの構成を行います。
  - a. QuickPlace サーバーで Server 文書を開きます。
  - b. 「基本」タブを開き、「**Directory Assistance データベース名**」フィールドに、ステップ 1 で作成したデータベースの名前を指定します。
  - c. Server 文書を保存してクローズします。

これで、QuickPlace サーバーは、2 次 Domino ディレクトリーとして LDAP サーバーを使用できるようになります。

## Windows ファイル・システム文書に対する文書レベルのセキュリティの施行

Windows ファイル・システム・クローラーでクロールされた文書をユーザーが検索するときに現行のクレデンシャルを検証するには、クローラー・サーバーおよび Microsoft Windows サーバーの両方でドメイン・アカウント情報を構成する必要があります。

Windows ファイル・システム クローラーを構成するとき、ローカル・コンピューター上のサブディレクトリーまたはリモート・コンピューター上のサブディレクトリーのどちらをクロールするかを指定します。コレクションでセキュリティが有



効になっている場合、クロールされたサブディレクトリー内の文書へのアクセスを制御するためのオプションも指定できます。

ユーザーが照会を実行したときにそのユーザーの現行のクレデンシャルを検証することでアクセス制御を施行することを選択した場合、ドメイン・アカウントが正しく構成されていることを確認してください。ローカル・コンピューター上でクロールされたファイルのドメイン・アカウントをセットアップするための要件は、リモート Windows サーバー上でクロールされたファイルのための要件とは異なります。

**重要:** 以下の条件が 2 つとも当てはまる場合、ユーザー・クレデンシャルは照会処理中に検証できません。

- クロールされる Windows サーバーがドメインのメンバーでない。
- クロールされるディレクトリーが、`¥サーバー名¥ホスト名` などのリモート・ディレクトリーである。

## ローカル・アクセス制御データでの検証

現行のユーザー・クレデンシャルを検証するために、システムはローカル・ユーザー・アカウント情報とドメイン・アカウント情報の両方を使用します (コンピューターが Windows ドメインに属している場合)。照会処理中にクレデンシャルを検証するには、検索対象の文書のセキュリティ情報に両方のユーザー名をリストしておく必要があります。

### ローカル・アカウント

ローカル・アカウントの場合、ユーザー名は次のフォーマットです。

`COMPUTER_NAME¥USERNAME`

ログインするには、ユーザーはユーザー名のみを指定します。しかし、正しく指定された Windows ユーザー権限割り当てではフルネームが使用されません。例えば、ローカル・アカウントのユーザー名が `abcuser` である場合、アカウントのフルネームとしては `WINSERVER1¥abcuser` などが考えられます。

ユーザーが検索アプリケーションを使用して、ローカル・システム上のセキュア文書の検索用プロファイルを構成するときは、Windows にログインする際に使用するユーザー名 (例えば、`abcuser`) を指定しなければなりません。

### ドメイン・アカウント

ドメイン・アカウントの場合、ユーザー名は次のフォーマットです。

`DOMAIN_NAME¥USERNAME`

ログインするために、ユーザーはこの情報を次のフォーマットで指定します。

`USERNAME@DOMAIN_NAME`

例えば、ファイルのユーザー権限割り当てを構成して、ドメイン `WIN1¥abcuser` を選択する場合、アカウントは `abcuser@win1.company.com` として表示されます。

ユーザーが検索アプリケーションを使用して、セキュア・ドメイン内の文書を検索できるようにするプロファイルを構成するときは、Windows にログインする際に使用するユーザー名 (例えば、abcuser@win1.company.com) を指定しなければなりません。

ローカル・コンピューター上で現行のクレデンシャルの検証を施行するためには、クローラー・サーバーによって使用されるユーザー・アカウントに次の Windows ユーザー権限が必要です。ユーザー権限を割り当てるには、Windows 管理ツールを使用します (「管理ツール」 → 「ローカル・セキュリティ・ポリシー」 → 「ローカル・ポリシー」 → 「ローカル・ユーザー権限の割り当て」)。

- クローラー・サーバーを実行するユーザー ID には、「オペレーティング・システムの一部として動作 (Act as part of the operating system)」の権限が必要です。この権限は、OmniFind Enterprise Edition がインストールされるときに、クローラー・サーバー上のエンタープライズ・サーチ管理ユーザー用に構成されます。
- ユーザーには、「ローカルでログオン」のユーザー権限が必要です。

## リモート・ドメイン・アクセス制御データでの検証

Windows オペレーティング・システムでは、`¥¥servername` で始まるディレクトリーがリモート・ディレクトリーと見なされます。次に例を示します。

```
¥¥software¥utilities¥IBM
```

リモート・ディレクトリーにアクセスするには、ユーザーは、ユーザー名を次のフォーマットで指定します。

```
USERNAME@DOMAIN NAME
```

ユーザーが検索アプリケーションを使用して、リモート・システム上のセキュア文書を検索できるようにするプロファイルを構成するときは、リモート Windows システムにアクセスする際に使用するユーザー名 (例えば、abcuser@win1.company.com) を指定しなければなりません。

リモート・コンピューター上で現行のクレデンシャルの検証を施行するためには、ユーザー・アカウントに次の Windows ユーザー権限が必要です。ユーザーに権限を割り当てるには、「管理ツール」 → 「ドメイン セキュリティー ポリシー」により、Windows 管理ツールを使用します。

- 検索対象のクローラー・サーバーおよび Windows サーバーは、同じドメインのメンバーでなければなりません。
- クローラー・サーバーを実行するユーザー ID には、「オペレーティング・システムの一部として動作 (Act as part of the operating system)」の権限が必要です。この権限は、OmniFind Enterprise Edition がインストールされるときに、クローラー・サーバー上のエンタープライズ・サーチ管理ユーザー用に構成されます。
- ユーザーには、「バッチ・ジョブとしてログオン」のユーザー権限が必要です。

### 関連概念

295 ページの『照会処理中の現行のクレデンシャルの検証』

129 ページの『Windows ファイル・システム・クローラー』

321 ページの『Windows トラストド・ドメインのセキュア検索』

## Windows トラステッド・ドメインのセキュア検索

リモート Windows ファイル・システムに文書レベルのセキュリティーを施行するために、エンタープライズ・サーチ・システムはトラステッド・ドメイン全体のアクセス制御リスト (ACL) 検証をサポートします。

### クローラーの構成

トラステッド・ドメインをサポートするように Windows ファイル・システム クローラーを構成するには、新しい構成ファイルにオプションを指定する必要があります。エンタープライズ・サーチ管理コンソールでこの機能を構成することはサポートされていません。

1. Windows トラステッド・ドメイン全体での文書レベルのセキュリティーをサポートするには、次のファイルを編集します。

```
ES_NODE_ROOT/master_config/session_ID/winfscrawler_ext.xml
```

**ヒント:** 構成する Windows ファイル・システム クローラーのセッション ID を判別するには、エンタープライズ・サーチ管理コンソールでクローラーをモニターするか、**esadmin report collections** コマンドを使用します。

2. Windows ドメイン・ネームと Active Directory の NETBIOS 名を指定します。次に例を示します。

```
<ExtendedProperties>
  <SetAttribute XPath="/Crawler/DataSources/Server/Target"
    Name="Domain">jk.enterprises.com
  </SetAttribute>
  <SetAttribute XPath="/Crawler/DataSources/Server/Target"
    Name="NetBIOSDomain">JKE1
  </SetAttribute>
</ExtendedProperties>
```

3. 変更を有効にするために、クローラーを停止して再開します。

### 制約事項

- 文書には、複数の Windows ドメインの ACL を含むことはできません。ドメイン・ユーザーおよびグループは、コレクションごとに 1 つのドメインに属している必要があります。
- リモート・ファイル・システム・アクセス検証をサポートするには、Windows が、同一 Windows ドメインかまたはトラステッド Windows ドメインで稼働している必要があります。
- Windows ファイル・システム・クローラーは、クロール対象の Windows サーバーに関連した Active Directory の NETBIOS 名を読み取り、この NETBIOS 名を使用して ACL ファイルをフィルターします。クローラー・サーバーが結合する Active Directory は、ユーザー・アカウントとグループ・アカウントを定義する他の Active Directory を信頼します。
- クローラーがリモート Windows サーバーにアクセスするために使用するユーザー・アカウントは、アクセス制御を施行し、検証する Windows ドメインに属している必要があります。
- Windows オペレーティング・システムでは、1 つのファイル・サーバー上のネットワーク・フォルダーに接続できるのは 1 つのアカウントのみです。他のアカウントは、同じファイル・サーバーに同時に接続できません。したがって、別のク

ローラーが別のコレクション内に存在する場合でも、そのクローラーが同じ Windows サーバーをクロールするための別のアカウントを構成することはできません。

#### 関連概念

318 ページの『Windows ファイル・システム文書に対する文書レベルのセキュリティの施行』

---

## エンタープライズ・サーチのセキュリティの使用不可化

WebSphere Application Server でエンタープライズ・サーチ・アプリケーションのセキュリティを使用不可にすることができます。文書レベルのセキュリティ制御を前に構成している場合、その制御を無視するように指定できます。セキュリティ設定は、検索結果で縮小表示された結果がどう表示されるかにも影響します。

### WebSphere Application Server 内でのエンタープライズ・アプリケーションのセキュリティの使用不可化

ユーザー認証を必要とするエンタープライズ・サーチ・アクティビティを制御するために、WebSphere Application Server 内の個々のエンタープライズ・アプリケーションについてグローバル・セキュリティを使用不可にすることができます。

#### このタスクについて

OmniFind Enterprise Edition インストール・プログラムは、次の 3 つのエンタープライズ・アプリケーションを WebSphere Application Server にデプロイします。

- ESAdmin アプリケーション。エンタープライズ・サーチ管理コンソールのインターフェイスが含まれています。
- ESSearchApplication アプリケーション。サンプル検索アプリケーションのインターフェイスが含まれています。
- ESSearchServer アプリケーション。エンタープライズ・サーチ SIAPI インプリメンテーションのすべての遠隔通信を提供し、SIAPI インターフェイスが検索サーバーと通信できるようにします。

デフォルトで、3 つのエンタープライズ・アプリケーションはすべて、WebSphere Application Server グローバル・セキュリティをサポートします。これらのアプリケーションは、グローバル・セキュリティが使用可能であることを検出すると、受信したすべての要求の認証を始めます。

組織によっては、特定のエンタープライズ・アプリケーションについてセキュリティを使用可能または使用不可にしたい場合があります。例えば、エンタープライズ・サーチ管理コンソールにアクセスするすべてのユーザーを認証したいが、SIAPI インターフェイスまたはサンプル検索アプリケーションを使用するユーザーは認証したくない場合があります。

#### 手順

エンタープライズ・アプリケーションについてセキュリティを使用不可にするには、次のようにします。

1. 検索サーバー上で、WebSphere Application Server 管理コンソールを開始します。

管理コンソールは、以下の方法で開くことができます。

- Windows の「スタート」メニューを使用して、該当するプログラムを選択します。
  - WebSphere Application Server バージョン 5 の場合、Web ブラウザーを開き、`http://hostname:port/admin` にアクセスします。ここで、`hostname` は検索サーバーのホスト名であり、`port` は WebSphere Application Server 管理コンソールのポート番号です。通常、管理コンソールのポートは 9090 です。
  - WebSphere Application Server バージョン 6 の場合、Web ブラウザーを開き、`http://hostname:port/ibm/console` にアクセスします。ここで、`hostname` は検索サーバーのホスト名であり、`port` は WebSphere Application Server 管理コンソールのポート番号です。通常、管理コンソールのポートは 9060 です。
2. ユーザー ID とパスワードの入力を求めるプロンプトが出たら、WebSphere Application Server でグローバル・セキュリティを使用可能にしたときに指定した管理者 ID とパスワードを入力します。
  3. 管理コンソールにログインしたら、「アプリケーション」をクリックし、さらに「エンタープライズ・アプリケーション」をクリックします。
  4. セキュリティーを使用不可にしたいエンタープライズ・アプリケーションの名前の横にあるチェック・ボックスを選択します。
  5. スクロールダウンして、「ユーザー/グループへのセキュリティ役割のマッピング」リンクをクリックします。
  6. **AllAuthenticated** 役割を見つけて、「全利用者?」欄の下のチェック・ボックスを選択します。
  7. 「OK」をクリックします。
  8. 「保存」リンクをクリックして、変更を保存します。
  9. WebSphere Network Deployment を使用している場合は、「ノードとの変更の同期化」チェック・ボックスを選択します。
  10. 「保存」をクリックします。
  11. ESSearchServer アプリケーションを停止し、再始動します。

#### AIX、Linux、または Solaris

```
./stopServer.sh ESSearchServer  
./startServer.sh ESSearchServer
```

#### Windows

```
stopServer ESSearchServer  
startServer ESSearchServer
```

これらのスクリプトは、`WAS_INSTALL_ROOT/AppServer/bin` ディレクトリーにあります。

- WebSphere Application Server バージョン 5 の場合、デフォルトのインストール・パスは、AIX システムでは /usr/WebSphere、Linux または Solaris システムでは /opt/WebSphere、Windows システムでは C:%Program Files%WebSphere です。
- WebSphere Application Server バージョン 6 の場合、デフォルトのインストール・パスは、AIX システムでは /usr/IBM/WebSphere、Linux または Solaris システムでは /opt/IBM/WebSphere、Windows システムでは C:%Program Files%IBM%WebSphere です。

## 文書レベルのセキュリティを無効にする

アクセス制御が索引内の文書と関連付けられているかどうかに関係なく、ユーザーがコレクションを検索できるようにすることができます。現行のクレデンシャルの検証をサポートするクローラーでは、照会処理中に現行のアクセス制御を検証することなく、ユーザーにコレクションを検索させることができます。

### 始める前に

コレクションのすべての文書に文書レベルのセキュリティを有効または無効にするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 制約事項

文書レベルのセキュリティ・オプションを指定できるのは、コレクションが作成されたときにセキュリティがコレクションに対して有効な場合だけです。

### このタスクについて

クローラーは、クロール時に文書にセキュリティ・トークンを関連付けるように構成することができます。検索アプリケーションは、索引に格納されているこれらのトークンを使用して、ユーザーがコレクションを検索する際にアクセス制御を施行できます。クローラーの中には、ユーザーが照会を実行したときに、ネイティブ・リポジトリ内の、文書に関連付けられた現行のアクセス制御を検証するように指定できるものもあります。

これらのセキュリティ制限を除去するために、照会を使って渡されるセキュリティ・トークンを検索サーバーが無視するように指定できます。クレデンシャルを現行のアクセス制御と比較することなく、ユーザーが文書を照会できるようにすることもできます。

新規コレクションをテストする場合や、検索アプリケーションの問題をトラブルシューティングする必要がある場合に、文書レベルのセキュリティを一時的に無効にしたいことがあります。

### 手順

文書レベルのアクセス制御を無効にするには、次のようにします。

1. コレクションを編集し、「一般」ページを選択して、「**文書レベルのセキュリティの有効/無効**」をクリックします。

2. クローラーが文書と関連付けたセキュリティー・トークンを、ユーザーがコレクションを照会するときを使用したくない場合、「すべての文書に対する文書レベルのセキュリティー」ページで、「索引内の文書レベルのアクセス制御を無視する」チェック・ボックスを選択します。

クローラーは文書にセキュリティー・トークンを追加し続けますが、検索サーバーはトークンを無視し、ユーザーが以前に保護された文書を検索できるようにします。

3. ユーザーが照会を実行したときに、ネイティブ・リポジトリ内の、文書に関連付けられた現行のアクセス制御を検証したくない場合、「照会処理中に現行のクレデンシャルの検証を行わない (Do not validate current credentials during query processing)」チェック・ボックスを選択します。このチェック・ボックスは、この機能をサポートするクローラーによってクロールされた文書に対してのみ選択可能です。

このチェック・ボックスを選択しても、その他の文書レベルのセキュリティー・オプションは有効なままです。例えば、クローラーの構成時にアクセス制御を索引内に保管するオプションを指定した場合、「索引内の文書レベルのアクセス制御を無視する」チェック・ボックスも同様に選択しないと、それらのセキュリティー制御が引き続き適用されます。

## 縮小表示の検索結果に対するセキュリティーの使用不可化

コレクション・セキュリティーが有効になっている場合、同じサイトからの検索結果は、照会処理中にユーザー・クレデンシャルを検証しないと指定しない限り、検索結果で縮小表示できません。

### 始める前に

コレクションのすべての文書に現行のクレデンシャルの検証を使用可能または使用不可にするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 制約事項

文書レベルのセキュリティー・オプションを指定できるのは、コレクションが作成されたときにセキュリティーがコレクションに対して有効な場合だけです。

### このタスクについて

現行のクレデンシャルを検証する際には、各文書のソースが検査され、検証のためにルーティングされる可能性があります。同じ URI 接頭部を持つ文書や以前に構成した縮小表示の URI グループに属する文書を、検索結果で縮小表示する場合は、現行のクレデンシャルの検証を使用不可にする必要があります。

### 手順

検索結果で文書を縮小表示するために現行のクレデンシャルの検証を使用不可にするには、次のようにします。

1. コレクションを編集し、「一般」ページを選択して、「文書レベルのセキュリティーの有効/無効」をクリックします。

2. 「すべての文書に対する文書レベルのセキュリティー」ページで「**照会処理中に  
現行のクレデンシャルを検証しない**」チェック・ボックスを選択します。
3. コレクションをモニターして、「検索」ページを選択し、検索サーバー・プロセスを一度停止してから再始動します。

ユーザーがコレクションを照会すると、同じ URI 接頭部を持つ文書や縮小表示を構成されたサイトに属する文書が、検索結果で縮小表示されます。サンプル検索アプリケーションでは、「**同じソースの結果をさらに表示**」リンクをクリックすることにより、縮小表示された結果を表示できます。



---

## エンタープライズ・サーチ・システムの開始と停止

コレクションを作成後、データのクロール、構文解析、索引作成を行うためのサーバーを開始する必要があります (検索サーバーは自動的に開始されます)。コレクションに変更を加えた後は、サーバーを停止して再始動してください。

ほとんどのエンタープライズ・サーチ・サーバーは、連続して、または指定されたスケジュールに従って稼働できます。例えば、主索引と差分索引の作成スケジュールを指定できます。エンタープライズ・サーチ・システムを開始した後は、構成設定を変更する (カテゴリの更新や検索キャッシュのサイズ増加など) 場合以外、通常、サーバー・プロセスを停止し再始動する必要はありません。

コレクションの内容を変更した場合、あるいは、エンタープライズのソースからクローラーがデータを収集する方法の規則を変更した場合、それらの変更を有効にするために、クローラーの停止と再始動が必要です。クロール規則を変更しなければ、Web クローラーは連続して実行され、他のクローラーは指定したスケジュールに従って実行されます。

索引サーバーと管理コンソールを使用できないときに検索サーバーの可用性を高めるために、コレクションの検索サーバーをスタンドアロン・モードで開始するためのコマンドを指定できます。

---

## エンタープライズ・サーチ・システムの開始

ユーザーがコレクションを検索できるようにするために、システム・プロセスを開始し、その後、コレクションのクロール、構文解析、索引付けを行うサーバーを開始する必要があります (検索サーバーは自動的に開始されます)。

### 始める前に

クロールしたいデータ・ソースを構成し、そのデータが構文解析、索引付け、検索される方法について、オプションを指定します。例えば、ユーザーが検索結果にカテゴリ詳細を表示できるようにしたい場合は、パーサーを開始する前にカテゴリを構成します。

エンタープライズ・サーチ・サーバーを開始するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者、またはそのコレクションを管理する権限をもつオペレーターである必要があります。

エンタープライズ・サーチ・サーバーの開始は、正しい順序で行う必要があります。例えば、クローラーを開始してデータのクロールを行った後でないと、クロールされたデータの索引付けはできません。

### 制約事項

エンタープライズ・サーチ・システムを開始するには、ローカル認証で認証可能なユーザー・アカウントを使用する必要があります。Andrew File System (AFS<sup>®</sup>) アカウントでシステムを開始しようとすると、エラーが発生します。

## 手順

エンタープライズ・サーチ・システムを開始するには、次のようにします。

1. 2 サーバーまたは 4 サーバー構成でエンタープライズ・サーチを使用する場合、エンタープライズ・サーチ管理者でログインし、各サーバーで共通通信層 (CCL) を開始します。

### AIX、Linux、または Solaris

```
startccl.sh -bg
```

### Windows コマンド・プロンプト



```
startccl
```

### Windows サービス管理ツール

バックグラウンドで CCL を開始するには、次のようにします。


- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」により、Windows サービスを起動します。
  - b. 「**IBM OmniFind Enterprise Edition**」を右クリックして、「開始」をクリックします。
2. エンタープライズ・サーチ・システム・コンポーネントを開始します。
    - a. 任意のエンタープライズ・サーチ・サーバーでエンタープライズ・サーチ管理者としてログインします。
    - b. すべてのシステム・コンポーネントを開始します。 `esadmin system startall`

このコマンドは、次のプロセスとアプリケーションを開始します。


- Web サーバー (マルチサーバー構成では、両方の検索サーバーで Web サーバーが開始されます)
  - WebSphere Application Server の ESSearchServer および ESAdmin アプリケーション (マルチサーバー構成では、両方の検索サーバーでアプリケーションが開始されます)
  - 索引サーバーの ESAdmin セッション
  - コマンドを実行したコンピューター上の CCL (CCL がまだ稼働していない場合)
  - エンタープライズ・サーチのデータベース・ネットワーク・サーバー
  - エンタープライズ・サーチ・インフォメーション・センター (マルチサーバー構成では、両方の検索サーバーでインフォメーション・センターが開始されます)
3. エンタープライズ・サーチ管理コンソールを始動して、エンタープライズ・サーチ管理者としてログインします。管理役割を使用した場合、開始したいコレクションに対する権限をもつコレクション管理者またはオペレーターとしてログインできます。
  4. 「コレクション」ビューで、管理するコレクションを指して、「 モニター」をクリックします。
  5. 「クローल」ページで、開始する各クローラーごとに、「 開始」をクリックします。

- Web クローラーを開始すると、クローラーは即時にデータのクロールを始めます。これらのタイプのクローラーは、連続して稼働し、Web 文書のクロールと再クロールを行ないます。
- 他のクローラー・タイプのいずれかを開始すると、クローラー・セッションが開始します。クローラーは、スケジュールされた日時にクロールを開始します。クローラーをスケジュールしていない場合、またはクローラーを早く開始したい場合は、クローラーをモニターして、クロールしたい各データ・ソースの開始アイコンをクリックします。



クローラーが開始した後は、それを続けて実行させておくことができます。クローラーをスケジュールしていれば、スケジュールされた日時に再び実行されます。

6. データのクロールが終わったら、「解析」ページを開き、「 **開始**」をクリックしてパーサーを開始します。

パーサーは続けて実行させておくことができます。通常、データの構文解析方法(カテゴリーまたは XML フィールド・マッピングの更新)を変更しない限り、パーサーを停止する必要はありません。

7. オプション: 索引付けがスケジュールされた日時に開始するのを待たずに、索引付けプロセスを強制始動するには、「索引」ページを開き、「主」領域で「 **開始**」をクリックします。

索引付けプロセスは続けて実行させておくことができます。索引は、スケジュールされた日時に作成されます。

**ヒント:** 検索サーバーは自動的に始動します。それは、続けて実行させておくことができます。通常、検索キャッシュまたは文書サマリーの設定を変更しない限り、検索サーバーを停止する必要はありません。検索サーバーを再始動するには、「検索」ページを開き、「 **停止**」をクリックしてから「 **開始**」をクリックします。

管理コンソールを使用できないときに検索サーバーの可用性を高めるために、コレクションの検索サーバーをスタンドアロン・モードで開始するためのコマンドを指定できます。索引サーバーが稼働していなければ、管理コンソールは使用できません。

#### 関連概念

288 ページの『管理役割』

#### 関連タスク

21 ページの『管理コンソールへのログイン』

331 ページの『開始または停止対象のコンポーネントの制御』

332 ページの『スタンドアロン・モードでの検索サーバーの管理』

## エンタープライズ・サーチ・システムの停止

エンタープライズ・サーチ・コンポーネントの構成に変更を加えた場合や、問題のトラブルシューティングが必要な場合などには、エンタープライズ・サーチ・サーバーを一度停止してから再始動する必要があります。

### 始める前に


エンタープライズ・サーチ・サーバーを停止するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者、またはそのコレクションを管理する権限をもつオペレーターである必要があります。

### このタスクについて

エンタープライズ・サーチ・サーバーは、他のサーバーとは無関係に停止することができます。例えば、クローラーの構成を変更し、それを組み込むためにクローラーを一度停止してから再始動する場合、検索サーバーを停止してから再始動する必要はありません。

### 手順

エンタープライズ・サーチ・サーバーを停止するには、次のようにします。

1. 「コレクション」ビューで、管理するコレクションを指して、「 モニター」をクリックします。
2. 「クロール」ページで、管理したいクローラーを見つけ、停止または休止します。


クロール・スペースまたはクローラーのプロパティを変更する場合は、変更を組み込むために、クローラーを一度停止してから再始動してください。クロール・スペースを変更し、既に索引作成された文書に変更を適用したい場合は、文書の再クロールも必要です。

**ヒント:** プロセスがバックグラウンドでまだ実行されているにもかかわらず、要求された操作がタイムアウトになったというメッセージが表示される場合があります。タスクが完了したかどうかを判断するには、管理コンソールで「リフレッシュ」をクリックします (Web ブラウザーの「リフレッシュ」はクリックしないでください)。プロセスは、クローラーの状況アイコンが停止したことを示したら完了です。

3. 「解析」ページで、「 停止」をクリックして、パーサーを停止します。

データを構文解析するための規則を変更する場合は、変更を組み込むために、パーサーを一度停止してから再始動してください。変更は、新たにクロールされた文書にのみ適用されます。既に索引に入っている文書に変更を適用する場合は、フル・クロールを開始して、すべての文書を再クロールする必要があります。これにより、それらに対して再び解析と索引付けができるようになります。

4. 「索引」ページで「 停止」をクリックして作成中の索引を停止します。

また、索引キューをモニター中にも、索引作成を停止できます。これを行うには、ツールバーで「システム」を選択し、「索引」ページを開いて、作成を停止する索引について「 停止」をクリックします。

5. 「検索」ページで「 停止」をクリックして、検索サーバーを停止します。通常、検索サーバーを一度停止してから再始動する必要があるのは、検索キャッシュまたは文書サマリーの設定を変更した場合だけです。
6. 個別サーバーではなく、エンタープライズ・サーチ・システムを停止するには、次のようにします。
  - a. 任意のエンタープライズ・サーチ・サーバーでエンタープライズ・サーチ管理者としてログインします。
  - b. すべてのシステム・コンポーネントを停止します。 `esadmin system stopall`

このコマンドは、次のプロセスとアプリケーションを停止します。

- Web サーバー (マルチサーバー構成では、両方の検索サーバーで Web サーバーが停止されます)
- WebSphere Application Server の ESSearchServer および ESAdmin アプリケーション (マルチサーバー構成では、両方の検索サーバーでアプリケーションが停止されます)
- 索引サーバーの ESAdmin セッション
- コマンドを実行したコンピューター上のエンタープライズ・サーチの共通通信層 (CCL)
- エンタープライズ・サーチのデータベース・ネットワーク・サーバー
- エンタープライズ・サーチ・インフォメーション・センター (マルチサーバー構成では、両方の検索サーバーでインフォメーション・センターが停止されます)

#### 関連タスク

21 ページの『管理コンソールへのログイン』

『開始または停止対象のコンポーネントの制御』

332 ページの『スタンドアロン・モードでの検索サーバーの管理』

---

## 開始または停止対象のコンポーネントの制御

`esadmin system startall` および `esadmin system stopall` コマンドによって開始または停止されるコンポーネントを制御できます。

#### このタスクについて

`ES_INSTALL_ROOT/default_config/AutoRunComponents.properties` ファイルには、`esadmin system startall` および `esadmin system stopall` コマンドによって開始または停止できるエンタープライズ・サーチ・コンポーネントのリストが含まれます。デフォルトでは、リストされたすべてのコンポーネントがこれらのコマンドによって開始および停止されます。

プロパティ・ファイルを編集して、特定のコンポーネントを開始または停止させないようにすることができます。

#### 手順

エンタープライズ・サーチ・システムを始動または停止したときに、開始または停止されるコンポーネントを指定するには、次のようにします。

1. **esadmin system startall** または **esadmin system stopall** コマンドを実行するサーバーに、エンタープライズ・サーチ管理者としてログインします。
2. `ES_INSTALL_ROOT/default_config/AutoRunComponents.properties` ファイルを編集します。
3. コンポーネントが開始されないようにするには、対象のコンポーネントに、`Component.startable.component_ID=false` のフィールドを追加します。ここで、`component_ID` は、開始させないコンポーネントです。
4. コンポーネントが停止されないようにするには、対象のコンポーネントに、`Component.stopable.component_ID=false` のフィールドを追加します。ここで、`component_ID` は停止させないコンポーネントです。
5. ファイルを保存して、終了します。

この次に **esadmin system startall** または **esadmin system stopall** コマンドを使用すると、変更したコンポーネントは、プロパティ・ファイルに加えた変更に従って、開始または停止されません。

**例:** この例で、HTTP サーバー は **esadmin system startall** コマンドで開始されますが (デフォルトの設定)、**esadmin system stopall** コマンドでは停止されません (例の中の太字で表された行による制御のため)。

```
#####
# コンポーネント 3 の詳細
#####
Component.name.3=IBM HTTP server
Component.impl.class.3=com.ibm.es.control.util.component.impl.HTTPControlImpl
Component.nodes.3=search
# デフォルトでは、すべてのコンポーネントが開始の対象
Component.stopable.3=false
```

#### 関連タスク

327 ページの『エンタープライズ・サーチ・システムの開始』

330 ページの『エンタープライズ・サーチ・システムの停止』

---

## スタンドアロン・モードでの検索サーバーの管理

検索サーバーの高可用性を確保するために、たとえ索引サーバーが稼働していなくても個々のコレクションの検索サーバーを開始できます。

### 制約事項

検索サーバーを停止および開始するには、エンタープライズ・サーチ管理者でなければなりません。

スタンドアロン・モードの検索サーバーを開始または停止する機能は、エンタープライズ・サーチ管理コンソールからは使用できません。索引サーバーが稼働していなければ、管理コンソールにアクセスできません。

スタンドアロン・モードで検索サーバーを開始または停止するには、その前に、コレクションのクローラー、パーサー、索引および検索サーバーがすべて、少なくとも一度開始されたことを確認する必要があります。これは、必要なファイルが検索サーバーで確実に同期するために必要です。

コレクションの文書レベルのセキュリティーが有効である場合は、クローラー・サーバーが開始していることも確認してください。これは、文書レベルのセキュリティー管理が確実に実施されるようにするために必要です。クローラー・サーバーが稼働していなければ、検索結果には認証を必要としない文書のみが戻されます。

## このタスクについて

索引サーバーが使用できない場合に、検索サーバーをスタンドアロン・モードを開始することで、ユーザーは確実にシステムの検索を継続できます。検索サーバーを開始および停止するコマンドは、マルチサーバー・インストールのどのエンタープライズ・サーチ・サーバーからも実行できます。コマンドは、すべての使用可能な検索サーバー上の指定されたコレクションの検索サーバーを開始または停止しようとします。

## 手順

スタンドアロン・モードで検索サーバーを開始または停止するには、次のようにします。

1. 索引サーバーが稼働していないときにコレクションの検索サーバーを開始するには:

- a. クローラー・サーバーおよび検索サーバーに、エンタープライズ・サーチ管理者としてログインし、共通通信層 (CCL) サービスを開始します。

### AIX、Linux、または Solaris

```
startccl.sh -bg
```

### Windows コマンド・プロンプト

```
startccl
```

### Windows サービス管理ツール

バックグラウンドで CCL を開始するには、次のようにします。

- 1) 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」 により、Windows サービスを起動します。
  - 2) 「**IBM OmniFind Enterprise Edition**」を右クリックして、「開始」をクリックします。
- b. 検索サーバーで、**startServer** スクリプト (このスクリプトは `WAS_INSTALL_ROOT/AppServer/bin` ディレクトリに存在します) を実行して、WebSphere Application Server で **ESSearchServer** アプリケーションを開始します。

### AIX、Linux、または Solaris

```
./startServer.sh ESSearchServer
```

### Windows

```
startServer ESSearchServer
```

- c. 次のコマンドを入力します。ここで、**コレクション ID** は開始する検索サーバーを所有するコレクションを示します。

```
esadmin startSearch -cid コレクション ID
```

2. 索引サーバーが稼働していないときにコレクションの検索サーバーを停止するには:

- a. 任意のエンタープライズ・サーチ・サーバーでエンタープライズ・サーチ管理者としてログインします。
- b. 次のコマンドを入力します。ここで、コレクション ID は停止する検索サーバーを所有するコレクションを示します。

```
esadmin stopSearch -cid コレクション ID
```

#### 関連タスク

327 ページの『エンタープライズ・サーチ・システムの開始』

330 ページの『エンタープライズ・サーチ・システムの停止』



---

## エンタープライズ・サーチ・アクティビティのモニター

システム・アクティビティおよびコレクション・アクティビティをモニターすると、さまざまなプロセスの状況の表示、問題発生の可能性の監視、または、パフォーマンス改善のための構成設定の調整を行うことができます。

エンタープライズ・サーチ管理コンソールを使用すれば、システムをモニターし、必要に応じて操作を調整することができます。主要アクティビティ（クロール、構文解析、索引作成、検索）については、詳細な統計データを参照することができます。統計データには、平均応答時間や、セッション内にクロールまたは索引付けされた文書数などの進行状況情報があります。

ほとんどのアクティビティは、アイコンをクリックするだけで、開始および停止することができます。この開始/停止操作により、アクティビティを一時停止して、構成の変更または問題のトラブルシューティングを行い、アクティビティを進められる状況になった時点で処理を再開することができます。

### 関連タスク

327 ページの『エンタープライズ・サーチ・システムの開始』

330 ページの『エンタープライズ・サーチ・システムの停止』

---

## 単一コレクション内の文書数の見積もり

エンタープライズ・サーチ・コレクションを作成または編集するときは、保持しておきたいコレクションの文書数の見積もりを用意します。システムは、この数値を使用して、コレクションに必要なメモリーとディスク・リソースを見積もりますが、コレクションのサイズを制限することはありません。

### 始める前に

コレクションの見積もりサイズを変更するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

コレクションが大きくなって、見積もったサイズに達しても、システムは索引への文書追加を停止しません。コレクションについてアラートを構成し、索引内の文書の数が限度を超えたとき通知するオプションを選択した場合、デフォルトの限度は、コレクション内の見積もり文書数に指定した値と一致します。システムは、この見積もりと指定したアラートしきい値パーセントをモニターし、コレクションに構成された最大文書数に達しそうになると、E メールを送信します。

### 手順

コレクションの可能なサイズを見積もるには、次のようにします。

1. コレクションを編集、「一般」ページを選択、そして「一般オプションの構成」をクリックします。

2. 「見積もり文書数」フィールドに、コレクションの成長が見込まれる大きさを表す数値を入力します。 デフォルト値は 1,000,000 文書です。

---

## コレクションのモニター


コレクション内の各コンポーネントの状態に関する一般情報を表示するか、または、個々のコンポーネントおよび URI に関する詳細情報を表示するオプションを選択することができます。


### 始める前に


すべてのエンタープライズ・サーチ管理ユーザーは、コレクションをモニターできます。コンポーネントの開始または停止を行うか、スケジュールを使用可能/使用不可にするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順

コレクションをモニターするには、次のようにします。

1. 「コレクション」ビューで、モニターするコレクションを指して、「 **モニター**」をクリックします。各コレクション・コンポーネントの現在の状況が表示されます。

**ヒント:** コレクションを編集中であり、既に「一般」ページが表示されている場合は、「 **モニター**」をクリックすれば、コレクションをモニターするためのビューに移動できます。

2. URI に関する詳細情報を表示するには、「 **URI の詳細**」をクリックします。

例えば、特定の URI が索引中にあるかどうかや、その URI が含まれる索引が検索サーバーにコピーされたかどうかを見ることができます。

3. 個々のコンポーネントをモニターし、そのコンポーネントのアクティビティに関する詳細な統計を表示するには、「**状況**」アイコンをクリックします。

---

## URI についての詳細表示

URI に関する詳細情報を見ることができます。この URI で表される文書がどのようにクロール、索引付け、検索されるのかについての現在情報および履歴情報を知ることができます。

### 始める前に

URI レポートを表示する要求、またはレポートを E メール・アドレスへ送信する要求を実行する前に、情報の発信元になるコンポーネントがアクティブであることを確認してください。例えば、文書がどのようにクロール、索引付け、検索されるのかについての詳細情報を表示する場合には、Web クローラー、索引サーバー、検索サーバーが稼働中であることを確認してください。ドロップされた文書をトラッキングするには、文書トラッキングのロギング・オプションが構成されていることを確認してください。

### このタスクについて


URI に関する情報の収集は、時間のかかるプロセスです。情報を要求した後、情報が表示されるのを待つというオプションを選択できます。より効率的なオプションは、指定した E メール・アドレスにレポートを送信するようにするオプションです。


レポートを受信できるようにするには、その前に、エンタープライズ・サーチに対してメール・サーバーの情報が構成されていることを確認する必要があります。この情報は、「システム」ビューの「ログ」ページで E メール・オプションを構成するときに指定します。


索引サーバーと検索サーバーは、すべての URI に関する情報（例えば、ある URI が索引にあるかどうか、それが検索サーバーにコピーされたかどうか）を提供できます。文書がどのようにクロールされたかに関する情報を表示するには、Web クローラーによってクロールされた文書の URI を指定する必要があります。

## 手順

URI に関する詳細を表示するには、次のようにします。

1. 「コレクション」ビューで、モニターするコレクションを指して、「 モニター」をクリックします。

**ヒント:** コレクションを編集中であり、既に「一般」ページが表示されている場合は、「 モニター」をクリックすれば、コレクションをモニターするためのビューに移動できます。

2. 「 URI の詳細」をクリックします。
3. 「URI の詳細」ページで、情報を見たい URI を入力します。
4. 見たい情報のタイプに合わせてチェック・ボックスを選択します。

### クローラー詳細 (Web クローラーにのみ使用可能)

このチェック・ボックスを選択すると、Web クローラーによって文書がどのようにクロールされたかに関する情報と、クロール・スペース内の現在の状況に関する情報が示されます。

### 索引の詳細

このチェック・ボックスを選択すると、文書が索引付けされたかどうかと、文書が検索サーバーにコピーされたかどうかを示されます。

### 検索の詳細

このチェック・ボックスを選択すると、文書がどのように検索できるかに関する情報と、文書が検索に使用可能かどうかを示されます。

### パーサーによってドロップされた文書

その文書が構文解析中にエンタープライズ・サーチ・システムからドロップされたのかどうか、またその場合のドロップされた理由を参照する場合は、このチェック・ボックスを選択します。

### 索引からドロップされた文書

その文書が索引作成中またはアナライズ中にエンタープライズ・サーチ・システムからドロップされたのかどうか、またその場合のドロップされた理由を参照する場合は、このチェック・ボックスを選択します。

5. レポートが表示されるのを待つ場合は、「レポート表示」をクリックします。

6. 後でレポートを表示できるように E メール・アドレスにレポートを送信するには、「レポートの送信」をクリックします。
  - a. 「詳細 URI レポートの送信」ページで、「通知を受け取る E メール・アドレス」フィールドに、レポートを受け取る E メール・アドレスを入力します。
  - b. 「レポートの送信」をクリックします。

#### 関連概念

355 ページの『文書トラッキング』

#### 関連タスク

356 ページの『ドロップされた文書に関するレポートの表示』

#### 関連資料

135 ページの『エンタープライズ・サーチ索引における URI フォーマット』

---

## クローラーのモニター


コレクション内の各クローラーの状態に関する一般情報を表示するか、または、クローラーのアクティビティに関する詳細情報を表示するオプションを選択することができます。


### 始める前に


使用している管理役割が、コレクションのモニターに限定されている場合は、クローラー統計を表示することはできませんが、クローラーの動作の変更 (クローラーの開始や停止など) はできません。

### 手順


クローラーをモニターするには、次のようにします。

1. 「コレクション」ビューで、モニターするコレクションを指して、「 モニター」をクリックします。
2. 「クロール」ページを開きます。

**ヒント:** コレクションを編集集中であり、既に「クロール」ページが表示されている場合は、「 モニター」をクリックすれば、クローラーをモニターするためのビューに移動できます。

3. クローラーが稼働中または休止していて、クローラーに関する詳細な状況情報を表示する場合は、「 詳細」をクリックします。表示される統計のタイプは、クローラー・タイプによって異なります。

コレクションのプロセス管理を許可する管理役割を持っている場合、クローラー・アクティビティに関する詳細を表示中に、クローラーを開始、停止、休止することができます。クローラーがスケジュール可能である場合、クロールのスケジュールを使用可能または使用不可にすることもできます。

4. クローラーが停止または休止していて、クローラー・セッションを開始する場合は、「 開始」または「再開」をクリックします。

#### Web クローラーの場合:

クローラーが停止されていた場合、クローラーは再びクロールを開始



し、クロール・スペース全体をクロールします。クローラーが休止されていた場合、クローラーは、休止されたときのターゲットの先頭からクロールを再開します。

即時にクローラーにフル・クロールを開始させたい場合、「詳細」アイコンをクリックして、その後、「再フル・クロールの開始」アイコンをクリックします。クローラーは、最後にクロールしてから変更が行われなかったページも含めて、クロール・スペース全体のクロールを開始します。すべての文書を再クロールしたい場合があります。例えば、文書の構文解析の規則を変更して、以前に索引作成した文書にその規則を適用したい場合などです。

#### 他のすべてのクローラー・タイプの場合:

クローラーが停止されていた場合、クローラーは、スケジュールされた日時にクロールを開始します。クローラーがデータ・ソースをクロールする最初の時、クローラーはフル・クロールを実行します。クロールが繰り返しスケジュールされている場合、クローラーはデータ・ソースに対するすべての更新 (文書の追加、削除、および変更) か、または文書の追加および変更のみのどちらかをクロールします。クローラー・スケジュールでクロールのタイプを構成します。

クローラー・スケジュールがまだ行われていない場合、またはクローラーを早く開始したい場合は、「詳細」アイコンをクリックしてください。その後、クロール・スペースの詳細領域で、開始したいクロールのタイプのアイコン「フル・クロール」、「すべて更新」、または「新規および変更文書のみ」をクリックします。クロールしたい各データ・ソース (サーバー、データベース、サブフォルダーなど) の該当する開始アイコンをクリックする必要があります。

- 稼働中のクローラーを停止する場合、「 停止」または「 休止」をクリックします。クローラーは、再始動または再開されるまで、データのクロールを停止します。

休止されたクローラーを再開する場合、クローラーは、休止されたターゲットの先頭でクロールを開始します。例えば、DB2 クローラーは、停止されたときにクロール中だった表の先頭行からクロールを再開します。

---

## Web クローラーのアクティビティ詳細の表示


Web クローラーのアクティビティに関する詳細を表示することによって、全体的なパフォーマンスを査定し、必要に応じて、Web クローラーのプロパティおよびクロール・スペース定義を調整できます。


### 始める前に


すべてのエンタープライズ・サーチ管理ユーザーは、クローラーのアクティビティをモニターできます。クローラーの開始または停止を行うには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順

Web クローラーのアクティビティーに関する詳細を表示するには、次のようにします。

1. 「コレクション」ビューで、モニターする Web クローラーを所有するコレクションを指して、「 **モニター**」をクリックします。
2. 「クロール」ページを開きます。

**ヒント:** コレクションを編集集中であり、既に「クロール」ページが表示されている場合は、「 **モニター**」をクリックすれば、クローラーをモニターするためのビューに移動できます。

3. モニターする Web クローラーが稼働中であるか休止している場合は、「 **詳細**」をクリックします。
4. Web クローラーの詳細ページで、クローラーの現在と過去のアクティビティーに関する詳細な統計を表示するため、以下のオプションを確認または選択します。
  - 「**スレッドの詳細**」をクリックすると、アクティブに Web サイトをクロールしているスレッドの数と、非アクティブ状態のスレッドの数が示されます。
  - 「**アクティブ・サイト**」をクリックすると、クローラーがアクティブにクロール中の Web サイトに関する情報が示されます。
  - 「**最近クロールされた URL**」をクリックします。この情報は、クローラーが最近クロールした対象を示します。ビューをリフレッシュしてもリスト中の項目が変わらない場合、発生しているクロールはありません。
  - 「**クローラー履歴**」をクリックすると、クローラーの過去のアクティビティーに関するレポートが表示されます。
  - 「**URL 状況**」領域で、情報を表示する URL を入力します。
    - a. 「**URL の詳細**」をクリックして、URL の状況情報を表示します。前にクロールされた URL のみの URL の詳細を要求できます。
    - b. 「**サイトの詳細**」をクリックして、その URL が属する Web サイトに関するレポートに組み込む情報を指定します。前にクロールされた Web サイト、またはまだクロールされていない Web サイトのサイトの詳細を要求できます。

例えば、このオプションを使用して、ある URL がクロール・スペース内にあるかどうか、それがクロールされたのか、または、単に検出されただけなのか、再度クロールされるべきなのはいつかを見たり、Web サイトに対する最新クロール試行に関する情報を見ることができます。また、Web サイトの robots.txt ファイルの内容を表示するように求めることもできます。これは、そのサイトがクロールされていない理由を判別する際に役立つことがあります。

## Web クローラーのスレッドの詳細

Web クローラーをモニターすることによって、アクティブに Web サイトをクロールしているスレッドの数と、非アクティブ状態のスレッドの数を見ることができます。

コレクションのモニター中に Web クローラーに関する詳細を表示すると、クローラー・スレッドの状況を見ることができます。一般的な状態には次のものがあります。

**待機中** スレッドにはクロールすべき URL がないことを示します。この状態は、スレッドがクロールを完了し、クローラーがそれ以上のクロールすべき URL を見つける速度が不十分な場合に発生します。例えば、クローラーのプロパティの設定で、クローラーが同じサイトから別のページをリトリブできるまでに待たなければならない時間が長すぎる場合には、URL が供給される速度が十分ではないことがあります。

#### フェッチ中

スレッドが Web サイトからページをダウンロード中であることを示します。

**完了** スレッドが、クロールしたページをクローラーの残りに送信中であるが、別の URL をクロールする準備はまだできていないことを示します。

**中断** クローラーが休止していることを示します。

理想的なのは、すべてのスレッドが常にページをフェッチ中であることです。スレッドが頻繁に完了状態になる場合は、データベースにスループットの問題があると考えられます。

スレッドが頻繁に待機状態になる場合は、クローラーのプロパティの「**アクティブ・ホストの最大数**」フィールドに指定された値を検討してください。この値が小さい場合、スレッドを継続的にビジーにするにはクロール・スペース内のサイト数が十分でないか、クロールに適格な URL の数が十分でないことが考えられます。低アクティビティを引き起こす条件として、DNS ルックアップ失敗とロボット・ルックアップ失敗があります。

## Web クローラー・アクティブ・サイト

Web クローラーをモニターして、クローラーがアクティブにクロール中の Web サイトに関する情報を見ることができます。

コレクションのモニター中に Web クローラーに関する詳細を表示すると、アクティブ・サイトに関する統計を見ることができます。統計には、以下の情報が示されます。

- 現時点でクロールのためにクローラーが内部データベースからメモリーへ移した URL の数。
- これまでにクローラーがクロールを試行した URL の数。
- サイトが非アクティブにされてクローラーのこの回の反復用のメモリーから除去されるまでの残り時間。
- これまでにサイトがメモリー内にあった時間。

この情報は、クローラーに対して構成されたクロール規則をクローラーが進んでいくのに従って刻々と変化します。アクティブな URL 数が、クローラー・メモリーのプロパティの「**アクティブ・ホストの最大数**」フィールドに構成された値に近いことが理想的です。

アクティブにされた URL 数がゼロに近い場合、クローラーは適格 URL を検出していません。そのような低いアクティビティを引き起こす可能性のある条件には、DNS ルックアップの失敗、ネットワーク接続性の問題、データベース・エラー、クロール・スペース定義の問題があります。次に例を示します。

- 長時間にわたってメモリー内に多数のサイトがあり、クロールされた URL が少数である場合、ネットワーク接続性の問題がないか調べてください。
- 十分な数のサイトがリストにある場合、クロール・スペース定義の問題、または DNS ルックアップの問題がないか調べてください。
- 適切な速度でサイトがクロールされているが、多数の URL がクロールされずにサイトがメモリーから出ている場合、クローラー・メモリーのプロパティを編集し、「各ホストがアクティブでいられる時間」フィールド内のタイムアウト値を調整して、サイトをメモリーに保持する時間を長くしてください。

## Web クローラーのクロール速度

Web クローラーをモニターして、クローラーが Web サイトからページをダウンロードしている速度に関する情報を見ることができます。

コレクションのモニター中に Web クローラーに関する詳細を表示すると、クローラーがデータをクロールする速さ (クロール速度) に関する統計を見ることができます。また、現行セッションの開始以降にクローラーがクロールした URL の数に関する統計も見ることができます。

クロール速度は、1 秒あたりにクロールされているページ数です。この数は、Web クローラーに対して構成できるいくつかのプロパティと関連しています。

- クローラー・スレッドの数
- アクティブ・サイトの数
- クローラーが同じ Web サーバーから別のページをリトリブできるまでに待たなければならない時間。

クローラーがクローラー・スレッドごとに 1 つのアクティブ・サイトを持っている場合で、クローラーが 2 秒間待った後でないと同じ Web サーバーから別のページをリトリブできない場合、クローラーは 2 秒あたりに、スレッド当たり 1 ページより速くクロールすることはできません。例えば、クローラーがデフォルト数 (200) のスレッドを使用する場合、クローラーは 200 スレッドで 1 秒あたり 100 ページをクロールできます。

クローラー・スレッド数の倍のアクティブ・サイトがある場合で、クローラーが 2 秒間待った後でないと同じ Web サーバーから別のページをリトリブできないとすると、クローラーは 1 秒にスレッド当たり 1 ページの速度に達すると考えられます。しかし、そうすると、ネットワークのダウンロード速度およびデータベース・スループットが速度を限定する要因になってきます。良好なクローラー・パフォーマンスが示されるのは、クロール速度が、クローラー・スレッド数、アクティブ・サイト数、およびクローラー待ち時間と調和している場合です。

Web クローラーのパフォーマンスをモニターするときに検討が必要なもう 1 つの因子は、現行クローラー・セッションの開始以降にクローラーがクロールした URL の数です。この数を、クローラーが稼働した合計時間で除算した値が、長期間の平均スループットです。この数値が増加しない場合、クローラーは完了しているか、



進行できないかのいずれかです。例えば、ネットワーク接続エラー、データベース・エラー、DNS ルックアップ失敗などがあると、クローラーの進行が妨げられます。

## Web クローラー・レポートの作成

Web クローラーの過去のアクティビティーに関するレポートを表示することによって、全体的なパフォーマンスを査定し、必要に応じて、Web クローラーのプロパティーおよびクロール・スペース定義を調整できます。

### 始める前に

コレクションのモニター限定の管理役割を使用している場合、クローラー統計を表示することができ、クローラーのアクティビティーに関するレポートを作成することもできますが、クローラーの動作の変更 (クローラーの開始や停止など) はできません。


### このタスクについて


さまざまなタイプのレポートがあり、Web クローラーのアクティビティーに関する情報が提供されます。レポートのタイプによっては、クローラーの内部データベースから収集できる限りの速度で情報が戻されるものもあります。サイト・レポートおよび HTTP 状況コード・レポートは、作成に時間を要します。これらのタイプのレポートを作成する場合、結果がエンタープライズ・サーチ管理コンソールに戻されるのを待つ代わりに、レポートを受け取る E メール・アドレスを指定できます。


レポートでの統計の解釈方法については、Web クローラーのモニター中およびレポートの作成中に「ヘルプ」をクリックしてください。

### 手順

Web クローラー・レポートを作成するには、次のようにします。

1. 「コレクション」ビューで、モニターする Web クローラーを所有するコレクションを指して、「 **モニター**」をクリックします。
2. 「クロール」ページを開きます。

**ヒント:** コレクションを編集中であり、既に「クロール」ページが表示されている場合は、「 **モニター**」をクリックすれば、クローラーをモニターするためのビューに移動できます。

3. レポートを作成する Web クローラーが稼働中であるか休止している場合、「 **詳細**」をクリックします。
4. Web クローラーの詳細ページで、作成したいレポートのタイプに応じてオプションを選択します。
  - クローラーと、クローラーが検出またはクロールするすべてのサイトに関するレポートを作成するには、「クローラー状況サマリー」領域で「クローラー履歴」をクリックします。
  - 「URL 状況」領域で、レポートを作成する特定サイトの URL を指定し、「サイトの詳細」をクリックします。

5. クローラー履歴レポートとサイト・レポートの両方について、レポートに入れる各統計のチェック・ボックスを選択し、「レポートの表示」をクリックできません。

これらのタイプの統計について、クローラーは、内部データベースから情報をリトリブできる限りの速度で、レポートを管理コンソールに戻します。

6. クローラー履歴レポートを作成している場合に、サイト・レポートを作成するためのオプションを指定し、「レポートの実行」をクリックすることができます。

このレポートは、選択した統計が組み込まれて作成され、指定するファイルに保存されます (このファイルの名前は絶対名でなければなりません)。レポート作成後に E メールを受け取ることを指定できます。

7. クローラー履歴レポートを作成している場合に、HTTP 状況コード・レポートを作成するためのオプションを指定し、「レポートの実行」をクリックすることができます。

このレポートには、サイトごとの HTTP 状況コードの数の分布についての情報が示されます。このレポートは、指定するファイルに保存されます (このファイルの名前は絶対名でなければなりません)。レポート作成後に E メールを受け取ることを指定できます。

このレポートを使用すると、4xx 状況コード (ページが見つからなかったことを示す)、5xx 状況コード (サーバーの問題を示す)、6xx 状況コード (接続での問題を示す) などを多く戻しているサイトを確認できます。

このレポートは、クローラーがある程度の時間アクティブであった場合に最も利用価値があります (例えば、数週間にわたってアクティブであったクローラーの場合など)。このレポートは、消失したサイト、新着サイト、大量の URL が含まれるサイト (Lotus Notes データベースの重複クローリングを表している可能性があります)、HTTP サーバーによる再帰的ファイル・システムのサービスを受けているサイトを識別するのに役立ちます。多数の HTTP 状況コードが付いたサイトが索引に寄与していない場合、それらのサイトをクローリング・スペースから除去することによって、クローラーのパフォーマンスを向上させることができます。

## Web クローラーに戻される HTTP 状況コード

Web クローラーをモニターするとき、クローリングするページからクローラーが受け取る HTTP 状況コードに関する情報を表示することができます。

### 表のサマリー

Web クローラー履歴をモニターするか、または特定 URL の状況をモニターするとき、クローラーに戻された HTTP 状況コードに関する情報を表示することができます。この情報を使用して、クローリング・スペースの管理およびクローラー・パフォーマンスの最適化を行えます。例えば、ある URL に対して多数の HTTP 状況コードをクローラーが受け取り、それらの状況コードが、その場所にあるページをクローリングできないことを示すものである場合、その URL をクローリング・スペースから除去することによってパフォーマンスを改善できます。

以下の表に、HTTP 状況コードと、Web クローラーがそれらをどのように解釈するの  
 のかを示します。100 から 505 までの値は標準 HTTP 状況コードです (詳しく  
 は、<http://www.w3.org/Protocols/rfc2616/rfc2616.html> を参照してください)。それ以外  
 の HTTP 状況コードは、エンタープライズ・サーチおよび Web クローラー専用の  
 コードです。

表 8. Web クローラーからの HTTP 状況コード

コード	説明	コード	説明	コード	説明	コード	説明
NULL	未クロール	400	不正な要求	500	内部サーバー・エラー	693	選択失敗 (URLFetcher)
100	継続	401	無許可	501	未実装	694	書き込みエラー (URLFetcher)
101	プロトコル切り替え	402	支払要	502	不正なゲートウェイ	695	不完全なブロック・ヘッダー (URLFetcher)
200	成功	403	禁止	503	サービス利用不可	699	予期しないエラー (URLFetcher)
201	作成されました	404	検出されません	504	ゲートウェイのタイムアウト	700	構文解析エラー (ヘッダー・エンドなし)
202	受け入れられました	405	許可されないメソッド	505	サポートされない HTTP バージョン	710	構文解析エラー (ヘッダー)
203	非認定情報	406	受け入れ不能	611	読み取りエラー	720	構文解析エラー (HTTP コードなし)
204	コンテンツなし	407	プロキシ認証が必要	612	接続エラー	730	構文解析エラー (本文)
205	コンテンツのリセット	408	要求タイムアウト	613	読み取りタイムアウト	740 または 4044	robots.txt ファイルにより除外
206	部分的コンテンツ	409	競合	614	SSL ハンドシェイク失敗	741	ロボットが一時的に利用不可
300	多肢選択	410	もう存在しない	615	他の読み取りエラー	760	クロール・スペース定義により除外
301	恒久的に移動	411	長さが必要	616	FBA 異常	761	ローカル・クロール・スペースにより不許可。グローバルにより許可
302	検出	412	前提条件失敗	617	エンコード・エラー	770	不正なプロトコルまたは非標準システム・ポート
303	他を参照	413	要求エンティティが大きすぎる	618	リダイレクト URL がいないリダイレクト	780	ファイル・タイプ排他により除外

表 8. Web クローラーからの HTTP 状況コード (続き)

コード	説明	コード	説明	コード	説明	コード	説明
304	未変更	414	要求 URI が長すぎる	680	DNS ルックアップ失敗	786	無効な URL
305	プロキシ使用	415	サポートされないメディア・タイプ	690	誤った形式の URL	2004	索引 META タグなし
306	(未使用)	416	要求範囲が不適合	691	接続喪失 (URLFetcher)	3020	ソフト・リダイレクト
307	一時的リダイレクト	417	予想失敗	692	書き込みタイムアウト (URLFetcher)		

## 表の注

### 4xx 状況コード

400 (不正な要求) コードは、めったにありません。HTTP 状況コードの規格によると、4xx コードは、クライアント (クローラー) が失敗したことを示すものとされています。しかし、サーバーに問題があるか、クローラーがリンクとして受け取った URL に問題があるのが普通です。例えば、一部の Web サーバーは、サイト・ルート (例えば、<http://xyz.ibm.com/././sales>) からナビゲートしようとする URL を許容しません。他の Web サーバーでは、このような上方ナビゲーションでの問題はなく、クローラーが既にルートにある場合は親ディレクトリー記号 (..) を無視します。

一部のサーバーはサイト・ルートに対する要求をエラーとして扱い、古くなったリンクが、もう認識または実装されていない操作を要求する可能性があります。もう提供されていないページが求められると、その要求はもはや有効とは見なされないため、アプリケーション・サーバーは例外をスローし、それにより、Web サーバーは HTTP 状況コード 400 を戻します。

**615** Web サイトからデータをダウンロードするクローラー・サーバーで予期しない例外が検出されたことを示します。このタイプの状況コードが多数ある場合、クローラーで問題があることを示している可能性があります。

### 61x 状況コード

615 を除く 61x 状況コードは、タイムアウトなど、クローラーにおいて予想される問題を示します。以下の状況コードについては、修正処置が必要な場合があります。

#### 611、612、613

サイトが低速であるか、ネットワーク・パフォーマンスが不十分であることがこの問題の原因である可能性があります。

**611** クローラーが文書をリトリートしたときにエラーが発生したことを示します。

**612** クローラーが Web サーバーに接続しようとしたときにエラーが発生したことを示します。

**613** クローラーが文書をリトリートしている間にタイムアウトが発生したことを示します。

- 614** クローラーがセキュア (HTTPS) サイトをクロールできないことを示します。これらのサイトがアクセス可能であるという確信がある場合、証明書がクローラー・サーバーおよびターゲット Web サーバー上で正しくセットアップされていることを確認してください。例えば、認識された認証局 (CA) によって認証済みのサイトの場合、クローラーが使用するトラスト・ストアに新しい CA を追加することができます。

また、クロールしようとしているサイトで自己署名証明書がどのように構成されているかも確認してください。クローラーは、自己署名証明書を受け入れるように構成されています。サイトによっては、ルート URL (例えば、<http://sales.ibm.com/>) に対して自己署名証明書を作成し、その後、サブドメイン (例えば、<http://internal.sales.ibm.com/>) でその証明書の使用を試みます。クローラーは、このような方法で使用される証明書を受け入れられません。クローラーが自己署名証明書を受け入れるのは、サブジェクトのドメイン名 (sales.ibm.com) および証明書の署名者が、要求されているページのドメイン名と一致する場合のみです。

- 616** フォーム・ベースの認証 (FBA) のログイン・フォームが、再認証後にダウンロードにまだ出現することを示します。

FBA 構成ファイルで提供される情報 (ログイン・フォームと、ユーザー名、パスワードなどの認証データ) がクローラーの認証に失敗した場合、状況コード 616 が、フォーム・ベースの認証に依存するすべてのページに割り当てられます。管理者は、FBA 構成が機能していない理由を調べて判別します。

- 617** エンコード・ストリング (文字セット) が正しくないか、または文書に無効なバイトが含まれているために、文書のバイト・コンテンツからストリングを作成できないことを表します。

- 618** クローラーが以下の HTTP 状況コードを受け取ったときに、リダイレクト URL が無効であることを示します。HTTP 応答ヘッダーのロケーションが無効である可能性があります。

301 恒久的に移動  
302 検出

- 680** おそらくネットワーク・アクセスの問題のため、クローラーが、クロール・スペース内でホストの IP アドレスを入手できなかったことを示します。このタイプのエラーは、クローラーが、いくつかの URL をクロールできなかったのではなく、サイト全体をクロールできないことを示します。このタイプの状況コードが多数ある場合、スループットが大幅に減少します。

#### 69x 状況コード

状況コード 690 から 699 は、クローラーの永続データベースに記録されません。これらのコードは、リモート・ホストからのダウンロードの実際の結果ではなく、クローラー内部の一時的な状態 (別のコンポーネントが結果を待機または送信しているときにコンポーネントがシャットダウンしたなど) を反映した結果を表します。これらの状況コードは一部のログに出現しますが、永続レコードには出現しません。そのため、選択セットの値には使用しないでください。

## 7xx 状況コード

7xx コードは、ほとんどの場合、クローラ・スペースにおける規則が原因です。

### 710 - 730

問題があるためにクローラーがダウンロードを完全にできないこと、または、クローラーがサイトで無効な HTML データを検出したことを示します。これらのタイプの状況コードが多数ある場合、エンタープライズ・サーチのサポート担当者に支援を依頼してください。

### 740 または 4044

サイトの robots.txt ファイル内の制限によって文書が除外されているため、ファイルのコンテンツを索引付けできないことを示します。

**740** 除外された文書をポイントするアンカー・リンクを索引に含められることを示します。

**4044** 除外された文書をポイントする文書内のアンカー・リンクが索引でも除外されることを示します。

**741** サイトにはクローラを許可する robots.txt ファイルがあるが、ダウンロードが失敗したことを示します。URL をクローラできない状態が繰り返される場合、その URL はクローラ・スペースから除去されます。このタイプの状況コードが多数ある場合、ターゲット・サイトが利用不可であるのは一時的なものか、永久的なものかを確認してください。ターゲット・サイトがもう使用可能でない場合、それをクローラ・スペースから除去してください。

これ以外の 7xx 状況コードは、多くが、クローラーがしばらく稼働した後でクローラ・スペースに変更を加えた場合に発生するものです。これらの状況コードは通常、対応が必要な問題を示すものではありません。

**3020** 状況コードが 200 の文書に、ユーザー・エージェントを別の URL に向けるロケーション・ヘッダーが含まれていることを表します。

---

## パーサーのモニター

パーサーによって分析される文書に関する情報を表示する必要がある場合に、それらがエンタープライズ・サーチ索引に追加される前にパーサーをモニターします。オプションを使用して、統計を検討し、パーサーのアクティビティを管理できます。

### 始める前に

管理役割がコレクションのモニターに限定されている場合、パーサーの状況を表示することはできますが、パーサーを開始したり、停止したりすることはできません。

### このタスクについて


パーサーの詳細をモニターしているとき、特定の瞬間に合わせて構文解析アクティビティに関する統計を提供するパーサー・アクティビティのスナップショットが表示されます。統計により、クローラーが終了し、現在、構文解析中または構文解析を待っている文書の数や、構文解析が終了し、索引への格納を待っている文書の数が示されます。


パーサーがアクティブであれば、メッセージにより、パーサーの状態に関するその他の情報が提供されます。次に例を示します。


- パーサーは、アクティブに文書を構文解析しています。
- パーサーはアイドル中です。パーサーは、さらに多くの文書が構文解析可能になるまでスリープ状態です。エラーが発生すると、パーサーは再始動を待機します。使用可能なパーサー・サービスがなくなると、パーサーは再始動します (例えば、自動再始動は、パーサー・サービスへの接続が確立できない時、またはパーサー Java 仮想マシンのすべてが他のコレクションのためビジーである場合に発生します)。
- パーサーは休止しています (例えば、パーサーは、索引作成が完了するまで休止します)。

## 手順


コレクションのパーサーをモニターするには、次のようにします。

1. 「コレクション」ビューで、モニターするコレクションを指して、「 モニター」をクリックします。
2. 「解析」ページを開きます。

**ヒント:** コレクションを編集集中であり、既に「解析」ページが表示されている場合は、「 モニター」をクリックすれば、コレクションをモニターするためのビューに移動できます。

3. パーサーが稼働中で、構文解析アクティビティに関する詳細な状況情報を表示する場合、「 詳細」をクリックします。

コレクションのプロセス管理を許可する管理役割を持っている場合、構文解析アクティビティに関する詳細を表示中にパーサーを開始および停止することができます。

4. パーサーが停止されていて、それを開始する場合は、「 開始」をクリックします。

最初にコレクションを作成するときには、クローラーがデータのクローラーを開始した後で、パーサーを開始してください。そうすれば、パーサーが分析とカテゴリ化を行うデータが必ずあることとなります。構文解析規則を変更しない限り、パーサーは続けて実行させておくことができます。

5. パーサーが稼働中で、それを停止する場合は、「 停止」をクリックします。

構文解析規則を変更する場合、パーサーを一度停止してから再始動する必要があります。例えば、パーサー構成を変更した場合、パーサーを停止してから再始動した後でないと、変更が有効になりません。

## コレクションの索引アクティビティのモニター


作成中の索引の進行を知る必要がある場合、索引スケジュールを使用可能/使用不可にする必要がある場合、または、索引付けアクティビティの開始/停止を行う必要がある場合に、コレクションの索引をモニターします。


### 始める前に




すべてのエンタープライズ・サーチ管理ユーザーは、索引アクティビティをモニターできます。索引作成の開始または停止を行うか、索引スケジュールを使用可能/使用不可にするには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順


コレクションの索引をモニターするには、次のようにします。

1. 「コレクション」ビューで、モニターするコレクションを指して、「 **モニター**」をクリックします。
2. 「索引」ページを開きます。

**ヒント:** コレクションを編集集中であり、既に「索引」ページが表示されている場合は、「 **モニター**」をクリックすれば、コレクションをモニターするためのビューに移動できます。

3. 索引がスケジュール済みの場合にスケジュールされた日時に索引が作成されないようにするには、「 **スケジュールを使用不可にする**」をクリックします。その索引は、スケジュールを使用可能にするか、索引作成プロセスを開始するまで、作成されません。
4. 索引がスケジュール済みであるが、索引作成のスケジュールが使用不可にされている場合は、「 **スケジュールを使用可能にする**」をクリックします。索引は、索引付けスケジュールに指定した日時に、作成のためにキューに入れられません。
5. 索引が停止されていて、それを開始する場合は、「 **開始**」をクリックします。

通常、索引付けは定期的なスケジュールに従って行われます。索引が作成されているときに停止した場合、または、索引のスケジュールを使用不可にした場合、「**開始**」をクリックして索引作成を強制開始できます。

6. 索引作成がアクティブであり、それを停止する場合、「 **停止**」をクリックします。

例えば、差分索引作成を停止する必要がある場合として、コレクション内で使用されるカテゴリー化のタイプを変更した後に主索引作成を強制する場合などが考えられます。

7. 索引作成中にエラーが発生した場合は、「 **エラー**」をクリックします。

「ログ・ファイルのコンテンツ」ページが表示されて、索引作成エラーに関する追加情報を表示できます。そのページで、個々のエラー・メッセージを選択して、問題に関する詳細を参照することができます。



---

## エンタープライズ・サーチの索引キューのモニター

索引キュー内のすべての索引作成の状況の表示、作成中の索引の停止、またはキューからの索引の削除を行うことができます。

### 始める前に

索引キューを管理するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

### このタスクについて

複数の索引を同時に作成することができますが、一度にキューに入れられるコレクションは 1 つだけです。システムの索引オプションを構成するとき、いくつかの索引がキューおよび索引作成用リソースを同時に共用できるのかを指定します。

### 手順


索引キューをモニターするには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「索引」ページを選択します。


索引キュー内に索引があるコレクションのリストが表示されます。それぞれの索引ごとに、作成中の索引のタイプ (差分またはフル)、索引が索引キューに入れられた時刻、索引作成が始まった時刻 (作成が進行中である場合) が表示されます。


3. 個々の索引を管理するには、「状況」アイコンをクリックします。

例えば、ある索引がどの程度まで完了しているかを見たり、索引内の文書数を見たり、索引スケジュールを使用不可にすることができます。

4. 作成している索引を停止するには、「 停止」をクリックします。

例えば、カテゴリ規則を変更した場合、主索引の作成を強制的に開始できるように差分索引の作成を停止することがあります。

索引作成を停止した後に開始するには、スケジュールされている次の開始日時に索引が索引キューに入るのを待つか、「状況」アイコンをクリックして索引をモニターしてから、「 開始」をクリックして索引作成を開始します。

5. 索引キューから索引を除去するには、「 除去」をクリックします。

---

## 検索サーバーのモニター


ある特定のコレクションについて検索サーバーのアクティビティーに関する詳細な状況情報を表示するか、または、エンタープライズ・サーチ・システム全体について検索サーバーの詳細な状況情報を表示することができます。


### 始める前に



すべてのエンタープライズ・サーチ管理ユーザーは、管理を認可されているコレクションの検索サーバーをモニターできます。エンタープライズ・サーチ・システムのすべての検索サーバーをモニターするには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。

検索サーバーの開始または停止を行うには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

#### 手順

1. 単一のコレクションの検索サーバーをモニターするには、次のようにします。
  - a. 「コレクション」ビューで、モニターするコレクションを指して、「 **モニター**」をクリックします。
  - b. 「検索」ページを開きます。

**ヒント:** コレクションを編集集中であり、既に「検索」ページが表示されている場合は、「 **モニター**」をクリックすれば、コレクションをモニターするためのビューに移動できます。

2. エンタープライズ・サーチ・システムのすべての検索サーバーをモニターするには、次のようにします。
  - a. 「システム」をクリックして「システム」ビューを開きます。
  - b. 「検索」ページを選択します。
3. 検索サーバーが停止されていて、それを開始する場合は、「 **開始**」をクリックします。
4. 検索サーバーが稼働中で、それを停止する場合は、「 **停止**」をクリックします。

検索キャッシュを使用可能または使用不可にした、検索キャッシュのサイズを変更した、またはクイック・リンクを変更した場合は、変更を有効にするために、検索サーバーを停止して、再始動する必要があります。

5. 検索サーバーが検索要求の処理に費やす時間のサマリーを表示するには、「**応答時間履歴**」をクリックします。

このレポートには、ある特定の日に検索要求への応答に検索サーバーが費やした平均時間がミリ秒で示されます。

平均応答時間は、システムのパフォーマンスの指標であり、サービス品質に対応します。応答時間の増加は、システムに負荷がかかっていることを示します。例えば、検索対象のコレクションの数およびコレクション・サイズがシステムの負担になっていることが考えられます。

6. 最も頻繁に実行されている照会のリストを表示するには、「**高頻度の照会**」をクリックします。

このレポートでは、最も頻繁に実行された上位 50 の照会におけるキーワードと、ある特定の照会をユーザーが実行した回数が示されます。

高頻度の照会を検討することによって、クイック・リンクの候補を識別できます。クイック・リンクを作成すれば、多くのユーザーのために検索品質を高めることができます。関連の高い文書が常に検索結果に戻されることを保証できます。

また、これらの照会に答えるようなリソースへのエンタープライズ・ポータルからのリンクを作成するとい場合もあります。例えば、交際費についての情報をユーザーが頻繁に検索している場合に、イントラネット・ホーム・ページに、交際費手続きに関するページへのリンクを組み込みます。

7. 最近実行された照会のリストを表示するには、「**最新の照会**」をクリックします。

このレポートでは、最近実行された 50 個の照会におけるキーワードが示されます。

最新の照会を検討することによって、組織における現在の傾向と緊急事態を識別できます。例えば、何らかのトピックに対する関心が急に高まっていることなどが分かります。そのような関心の高まりは、そのトピック用のクイック・リンクの必要性を示していたり、そのトピックを他の方法 (エンタープライズ・ポータルでのリンクの提供など) でユーザーが使用できるようにする必要があることを示している可能性があります。

## 照会統計の計算方法の変更

よく使用される照会や最近の照会の数の計算方法を変更できます。

### このタスクについて

検索サーバーをモニターする場合は、最もよく使用された上位 50 の照会のリストや、最近に処理された上位 50 の照会のリストを表示するオプションを選択できます。デフォルトの検索サーバー構成では、同等の照会ストリングで結果の範囲設定が異なる照会は、独立した照会としてカウントされます。したがって、例えば、結果の次のページを表示するオプションをユーザーがクリックすると、照会カウントが増えます。

システムによる照会の統計の計算方法を変更するには、検索サーバーの `runtime-generic.properties` ファイルを編集します。**distinctRecentQueryCheck** パラメーターに `true` を設定すると、結果の初期ページに戻す照会のみが独立した照会としてカウントされます。

索引サーバーのプロパティを編集するには、エンタープライズ・サーチ管理者としてログインする必要があります。検索サーバーの開始または停止を行うには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。



### 手順

照会の統計をシステムで計算する方法を変更するには、次の手順で行います。

1. エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。マルチサーバー構成では、検索サーバーにログインします。

2. 次のファイルを編集します。ここで、*collection\_ID* は変更するコレクションの ID で、*node\_ID* は変更する検索サーバーを示します。

```
ES_NODE_ROOT/master_config/collection_ID.runtime.node_ID/runtime-  
generic.properties
```

3. **distinctRecentQueryCheck** パラメーターの値を `true` に設定し、ファイルを保存して閉じます。
4. マルチサーバー構成では、2 番目の検索サーバーで前の手順を繰り返します。
5. 追加コレクションの照会の統計動作を変更する場合は、変更するコレクションごとに前の手順を繰り返します。
6. 変更を有効にするために、エンタープライズ・サーチ管理コンソールにログインして検索サーバーを再始動します。
  - a. 「システム」をクリックして「システム」ビューを開きます。
  - b. 「検索」ページを選択します。
  - c. 変更した検索サーバーごとに、「 停止」をクリックします。
  - d. 変更した検索サーバーごとに、「 開始」をクリックします。

---

## データ・リスナーのモニター

データ・リスナーをモニターしてその状況を確認し、クライアント・データ・リスナー・アプリケーションのアクティビティーに関する詳細を表示します。


### 始める前に

データ・リスナーをモニターするには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。


**重要:** データ・リスナーは、将来のリリースではサポートされなくなります。エンタープライズ・サーチのクライアント・アプリケーションを開発するには、データ・リスナー API に代わって検索および索引 (SI-API) API を使用してください。下の説明は、以前にデータ・リスナー用アプリケーションを作成したユーザー向けです。

### 手順

データ・リスナーをモニターするには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「データ・リスナー」ページで、状況アイコンを表示して、データ・リスナーがアクティブであるのか、停止しているのかを確認します。
3. データ・リスナーが稼働中で、クライアント・アプリケーションのアクティビティーに関する詳細な状況情報を表示する場合は、「 詳細」をクリックします。

「データ・リスナー詳細」ページの状況アイコンは、データ・リスナーが稼働中か、停止しているかを示します。統計により、処理を待っている要求の数、クライアント・アプリケーション要求に対して作動している各スレッドの現在の状態、指定のスレッド状態についてアクティブなスレッドの数が示されます。

4. データ・リスナーのポート番号を変更した場合、または、一時ストレージに保持できる文書の最大数を変更した場合は、「 再始動」をクリックします。

データ・リスナーは、エンタープライズ・サーチ・システムが開始されたときに開始されます。これらの構成オプションの一方または両方を変更しない限り、データ・リスナーの再始動は必要ありません。

---

## 文書トラッキング

文書は、処理の様々な段階でシステムからドロップすることができます。文書がドロップされた時、およびドロップの原因となった問題の内容を確認する場合に、このオプションを指定できます。

文書の構文解析を妨げるエラーをパーサーが検出した場合、ドロップされた文書に関して、理由コードの付いたメッセージがログに記録されます。(このタイプのエラーによって、旧バージョンの文書が索引から除去されることはありません。)

索引作成の段階で、文書がドロップされることがあり、この情報もまたログに記録されます。例えば、URI および URI パターンは、明示的に削除できます。ある文書をクロールし、クロールしたそのクローラーを後で削除してしまうということがあります。ソース文書がもはや存在しないか (その文書にはネガティブ HTTP コードが関連付けられます)、またはその文書に関連付けられた HTTP コードが認識されないこととなります。また、グローバル分析を要求する文書にランク情報が欠如していると、文書がドロップされるということもあります。

クロールされた文書がわかっているが、その文書が索引に表示されない場合、エンタープライズ・サーチ管理コンソールを使用して、システム全体にわたる文書のフローをトラックすることができます。いつ、どこで、なぜ、その文書がドロップされたのかを示す詳細レポートを表示することができます。例えば、そのレポートは、グローバル分析中に文書が予期せずにドロップされたことを示したり、または管理者によって URI が索引から除去されたことを示したりすることができます。

### 関連タスク

336 ページの『URI についての詳細表示』

## 文書トラッキングに対するログ・ファイルの構成

いつ、どこで、およびなぜ、文書がシステムからドロップされたかを判別する場合、ドロップされた文書に関する情報をトラッキングするようにログ・ファイルを構成することができます。

### 始める前に

ドロップされた文書をトラッキングするためのオプションを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### このタスクについて

ログ・ファイルが消費するディスク・スペースが大きくなりすぎるのを防止するために、システムはログ・ファイルを循環させ、現在日付が変わるたびに新しいログ・ファイルを開始します。1 つのログ・ファイルが許容最大サイズに達し、日付

が変わっていない場合、システムは新しいログ・ファイルを作成します。ログ・ファイルの最大数に達した場合、新しいログ・ファイルを作成できるように、最も古いログ・ファイルが廃棄されます。

### 手順

文書トラッキングのログ・ファイルを構成するには、次のようにします。

1. コレクションを編集、「ログ」ページを選択、そして「**文書トラッキングの構成**」をクリックします。
2. 「文書トラッキング」ページで、文書トラッキングのチェック・ボックスが選択されていることを確認します。
3. システムからドロップされた文書に関する情報をログに記録するのに使用するログ・ファイルの数を指定します。これらのログ・ファイルは、文書がドロップされる可能性のあるすべてのセッションで共用されます。

## ドロップされた文書に関するレポートの表示

エンタープライズ・サーチ・システムからドロップされた文書に関する詳細情報を表示することができます。この情報は、コレクションに対して文書トラッキングが使用可能になっている場合にのみ有効です。

### 始める前に

ドロップされた文書に関するレポートを表示する要求を実行する、または E メール・アドレスにレポートを送信する前に、情報の発信元になるセッションがアクティブであることを確認してください。例えば、構文解析中または索引作成中にドロップされた文書に関して確認するには、コレクションに対するパーサーまたは索引セッションが開始していることを確認してください。

レポートを受信できるようにするには、その前に、メール・サーバーがエンタープライズ・サーチに対して確実に構成されているようにしてください。この情報は、「システム」ビューの「ログ」ページで E メール・オプションを構成するときに指定します。


### このタスクについて


ドロップされた文書に関する情報の収集は、時間のかかるプロセスです。情報を表示する場合に、その情報が表示されるのを待つというオプションを選択できます。より効率的なオプションは、指定した E メール・アドレスにレポートを送信するようにするオプションです。


文書がドロップされた場合、文書がドロップされた日時、エラーの重大度レベル、問題が発生したコンポーネントとセッション、およびエラー・メッセージが、そのレポートによって表示されます。

### 手順

ドロップされた文書に関する詳細を表示するには、次のようにします。

1. 「コレクション」ビューで、モニターするコレクションを指して、「 **モニター**」をクリックします。

**ヒント:** コレクションを編集中であり、既に「一般」ページが表示されている場合は、「 **モニター**」をクリックすれば、コレクションをモニターするためのビューに移動できます。

2. 「 **URI の詳細**」をクリックします。
3. 「URI の詳細」ページで、情報を見たい URI を入力します。
4. 見たい情報のタイプに合わせてチェック・ボックスを選択します。

#### パーサーによってドロップされた文書

その文書が構文解析中にドロップされたのかどうか、またその場合のドロップされた理由を参照する場合は、このチェック・ボックスを選択します。

#### 索引からドロップされた文書

その文書が索引作成中またはアナライズ中にドロップされたのかどうか、またその場合のドロップされた理由を参照する場合は、このチェック・ボックスを選択します。

5. どのようにレポートを表示するかを指定します。
  - レポートが表示されるのを待つ場合は、「**レポート表示**」をクリックします。
  - 後でレポートを表示できるように E メール・アドレスにレポートを送信するには、「**レポートの送信**」をクリックします。

「詳細 URI レポートの送信」ページで、「**通知を受け取る E メール・アドレス**」フィールドに、レポートを受け取る E メール・アドレスを入力して、「**レポートの送信**」をクリックします。

#### 関連タスク

336 ページの『URI についての詳細表示』

#### 関連資料

135 ページの『エンタープライズ・サーチ索引における URI フォーマット』

## ドロップされた文書に関するログ・ファイルの表示


エンタープライズ・サーチ・システムからドロップされた文書に関するログ・メッセージを表示することができます。この情報は、コレクションに対して文書トラッキングが使用可能になっている場合にのみ有効です。


#### このタスクについて


ドロップされた文書に関するレポートを表示するには、文書の URI を知る必要があります。ドロップされた文書のログ・ファイルを表示すると、文書がドロップされた日時、エラーの重大度レベル、問題が発生したコンポーネントとセッション、詳細なエラー・メッセージが表示されます。

#### 手順

ドロップされた文書に関するログ・ファイルを表示するには、次のようにします。

1. 「コレクション」ビューで、モニターするコレクションを指して、「 **モニター**」をクリックします。

**ヒント:** コレクションを編集集中であり、既に「一般」ページが表示されている場合は、「 モニター」をクリックすれば、コレクションをモニターするためのビューに移動できます。


2. 「 ドロップされた文書のログ・ファイル」をクリックします。
3. 「ドロップされた文書のログ・ファイル」ページで、表示するログ・ファイルを選択します。各ログ・ファイルの名前は、文書がパーサーによってドロップされたのか (pd)、あるいは索引作成の間にドロップされたのか (in) を示し、ファイルが作成された日付も含まれます。同じ日に同じタイプのログ・ファイルが複数作成される場合は、数字のサフィックスによって、その日のファイルの作成順が示されます。次に例を示します。

```
dropped_doc_in_20060525.log  
dropped_doc_pd_20060524.log (この日の最新の項目が含まれます)  
dropped_doc_pd_20060524.log.1  
dropped_doc_pd_20060524.log.2 (この日の最も古い項目が含まれます)
```

4. 「ログの表示」をクリックします。

「ログ・ファイルのコンテンツ」ページの各メッセージごとに、メッセージが発行された日時、メッセージ重大度レベル、メッセージを発行したセッションの名前、およびメッセージ ID とエラー・テキストが表示されます。

ボタンをクリックすることで、ログ・ファイルの先頭ページ、最終ページ、直前のページ、または次のページに移動することができます。ページ番号を指定することで、そのページに直接移動することもできます。

5. メッセージについての詳細情報を表示する場合は、「 詳細」をクリックします。

「ログ・メッセージの詳細」ページに、メッセージを出したエンタープライズ・サーチ・サーバーのホスト名、エラーを引き起こしたファイルの名前、エラーが発生した関数名および行番号、プロセス ID、スレッド ID が表示されます。

移動のボタンをクリックして、ログ・ファイル内の前後のメッセージに移動できます。



---

## ログ・ファイルおよびアラート

コレクションおよびシステムに対してログを記録するときのメッセージのタイプを選択し、ログ・ファイルの作成と表示のオプション、アラート受信のオプション、およびメッセージに関する E メール受信のオプションを指定することができます。

通常運用では、エンタープライズ・サーチ・コンポーネントは、メッセージを共通のログ・ファイルに書き込みます。このログ・ファイルは、索引サーバーの `ES_NODE_ROOT/logs` ディレクトリーにあります。管理コンソールを使用して、この共通ログ・データを表示することができます。

ネットワーク通信障害などの問題が発生すると、当該コンポーネントは、そのコンポーネントがインストールされているサーバーの `logs` ディレクトリーにログ・メッセージを書き込みます。これらのローカル・ログ・ファイルを表示するには、UNIX システムの `tail` ユーティリティーなど当該コンピューターのファイル・ビューアーを使用してください。このタイプのログ・ファイルの表示に管理コンソールを使用することはできません。

ログ・ファイルの構成時には、ログに記録するメッセージのタイプ (エラー・メッセージ、警告メッセージなど) の選択、古くなったログ・ファイルを廃棄して新規のログ・ファイル用のスペースを確保する場合の廃棄頻度の指定、ログ・ファイルの最大サイズの指定、およびメッセージ言語の選択を行うことができます。また、特定の事象が発生した場合や、特定のメッセージ、あるいは特定のタイプのメッセージがログに記録された場合に、必ず E メールを受信するようにするオプションを指定することもできます。

ログ・ファイルをモニターする場合、開きたいログ・ファイルを選択できます。ログ・ファイルの内容をフィルターに掛けて、特定の重大度レベル (エラー・メッセージのみなど) のメッセージ、または特定のエンタープライズ・サーチ・セッションが生成したメッセージだけを表示するように設定することができます。ログ・ファイルを表示する場合、個別のメッセージの詳細を表示することができます。例えば、メッセージを生成した機能名、その他問題を修正する際に役立つ情報を必要に応じて確認することができます。

### 関連概念

 [エンタープライズ・サーチのメッセージ](#)

---

## アラート

エンタープライズ・サーチでは、特定のイベントの発生が検出された場合に、メッセージをログ・ファイルに書き込むように構成することができます。

イベントによって起動され、作成されるメッセージ (アラートと呼ばれます) は、特定のリソースでフリー・スペースがなくなりかけているなど、対処が必要と思われる状態の発生を管理者に通知します。エンタープライズ・サーチでアラートを構成

する場合、システムでモニターする条件を指定します。その条件に該当する状態が発生した場合、システムによって自動的にログ・ファイルにメッセージが書き込まれます。

特定の状態について、直接管理者に通知されるようにしたい場合は、モニターによる特定のメッセージがログに記録されたときに E メールを送信するオプションを指定できます。

アラートは、コレクション・レベルのイベント、およびシステム・レベルで発生したイベントに対して構成できます。コレクション・レベルでは、システムは、次のような構成が可能です。

- 各クローラーがクロールする文書数をモニターし、指定された最大文書数に近くなった場合にアラート・メッセージを発行する。
- コレクションの索引に追加される文書数をモニターし、指定された最大文書数に近くなった場合にアラート・メッセージを発行する。
- 検索要求に対する応答時間が指定した限度を超える場合に通知する。

システム・レベルでは、エンタープライズ・サーチ・サーバーのディスク・スペースをモニターし、フリー・スペースの量が少なくなった場合にアラート・メッセージを発行する構成を設定することができます。

## コレクション・レベル・アラートの構成

アラートを構成することによって、特定のコレクション・レベルのイベントが発生すると必ずログ・ファイルにメッセージが書き込まれるようにすることができます。また、これらのイベントに関するメッセージがログに記録されると、E メールを受信できます。

### 始める前に

コレクションのアラートを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。

### 手順

コレクション・レベルのアラートを構成するには、次のようにします。

1. コレクションを編集します。「ログ」ページを選択し、「アラートの構成」をクリックします。
2. システムが、それぞれのクローラーがクロールしている文書の数モニターするようにしたければ、次のステップを行います。
  - a. 「クローラーがクロールした文書数が最大許容値の一定比率に達した場合」チェック・ボックスを選択します。
  - b. 「比率 (%)」フィールドに、メッセージをログに記録する時点のパーセンテージを指定します。この数値は、クローラーがクロールできる最大文書数に対するパーセンテージで指定します (クローラー・プロパティを構成するときに、「クロールする最大文書数」を指定します)。デフォルト値は 90 % です。

それぞれのクローラーに異なる制限を構成できるので、クローラーごとに別個のメッセージがログに記録されます。例えば、デフォルトのアラートしきい値を使用し、DB2 クローラーのクロール可能文書数を 2,000,000 文書に設定し、Lotus Notes クローラーのクロール可能文書数を 1,000,000 文書に設定した場合、DB2 クローラーが 1,800,000 文書をクロールした時点でメッセージがログに記録され、Lotus Notes クローラーが 900,000 文書をクロールした時点で別のメッセージがログに記録されます。

3. システムが、索引に追加されている文書の数モニターするようにしたければ、次のステップを行います。
  - a. 「**コレクション内の文書数が見積もり数の一定比率に達した場合**」チェック・ボックスを選択します。
  - b. 「**比率 (%)**」フィールドに、メッセージをログに記録する時点のパーセンテージを指定します。この数値は、コレクションが保持すると想定される見積もり文書数のパーセンテージで指定します。デフォルト値は 85 % です。

「**限度**」フィールドに、コレクションの現在の見積もり数が表示されます。この値を変更するには、コレクションの「**一般**」ページを開き、一般オプションを構成するオプションを選択して、「**見積もり文書数**」フィールドに新しい値を指定します。

**重要:** この限度、およびあるコレクションのために構成する文書の見積もり数は、そのコレクションの拡大をモニターするためにのみ使用されます。それらの数値は、索引が拡大できる大きさの絶対的な制限を押しつけるものではありません。

4. システムが、検索要求の応答に必要な時間が限界を超えた場合に通知するようにしたければ、次のステップを行います。
  - a. 「**検索応答時間が限度を超えた場合**」チェック・ボックスを選択します。
  - b. 「**限度**」フィールドに、最長の検索応答時間として許容できる秒数を入力します。

この数を超えると、システムは、このイベントについてのログ・メッセージを書き込みます。例えば、デフォルト値のままにした場合、システムは、検索サーバーの検索要求に回答する時間が平均 5 秒以上かかると、ログ・メッセージを作成します。

標準的な応答時間は、0.5 秒以内です。平均が 1 秒を超える場合、オペレーティング・システムにパフォーマンスを改善する調整が必要であるか、あるいは、検索サーバーの構成設定に問題があることを示しています。例えば、検索キャッシュに割り振るスペースの量を増やすことも考えられます。

5. 「**OK**」をクリックします。

これらのイベントについてのメッセージをシステムがログに記録したときに、E メールを受信するようにしたければ、「**ログ**」ページを開き、「**メッセージに対する E メール・オプションの構成**」をクリックして、E メール・アドレスを指定します。使用可能にしたアラートのメッセージ ID は、E メールが送信されるメッセージ ID のリストに自動的に追加されます。

E メールを受信できるようにするには、その前に、ユーザーのメール・サーバーの情報が構成されていることを確認する必要があります。これを行うには、エンタープライズ・サーチ管理者は、ツールバーの「システム」を選択し、「ログ」ページを開いて、「メッセージに対する E メール・オプションの構成」をクリックする必要があります。

#### 関連タスク

365 ページの『ログ・メッセージについての E メール受信』

## システム・レベル・アラートの構成


アラートを構成することによって、特定のシステム・レベルのイベントが発生すると必ずログ・ファイルにメッセージが書き込まれるようにすることができます。また、これらのイベントに関するメッセージがログに記録されると、E メールを受信できます。

#### 始める前に

システム・レベルのアラートを構成するには、エンタープライズ・サーチ管理者でなければなりません。

#### 手順

システム・レベルのアラートを構成するには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「 編集」をクリックして、システム編集ビューに切り替えます。
3. 「ログ」ページで、「アラートの構成」をクリックします。
4. システムが、エンタープライズ・サーチ・サーバーで使用可能なフリー・スペースの量をモニターするようにしたい場合は、「ファイル・システムの使用可能スペース量が全体スペースの一定比率に達した場合」チェック・ボックスを選択します。
5. 「比率 (%)」フィールドに、システムが、サーバーのフリー・スペースの量がなくなったことを通知する時点のパーセンテージを指定します。この数値は、ファイル・システムの合計スペースに対するパーセンテージで指定します。デフォルト値は 80 % です。

使用しているエンタープライズ・サーチ・システムが複数のサーバー上にセットアップされている場合、システムは、サーバーごとに別個のログ・メッセージを作成します。例えば、あるメッセージはクローラー・サーバー上のスペースがなくなったことを通知し、また別のメッセージは、索引および検索サーバーのスペース制約について通知します。

6. 「OK」をクリックします。

このイベントについてのメッセージをシステムがログに記録したときに、E メールを受信するようにしたければ、「ログ」ページを開き、「メッセージに対する E メール・オプションの構成」をクリックして、E メール・アドレスと、メール・サーバーについての情報を指定します。

#### 関連タスク

365 ページの『ログ・メッセージについての E メール受信』

---

## ログ・ファイルの構成

ログに記録したいメッセージのタイプを指定し、ログ・ファイル作成のオプションを指定できます。

### 始める前に

コレクション・レベルのログ・ファイルを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。システム・レベルのログ・ファイルを構成するには、エンタープライズ・サーチ管理者でなければなりません。

### このタスクについて

ログ・ファイルが消費するディスク・スペースが大きくなりすぎるのを防止するために、システムはログ・ファイルを循環させ、現在日付が変わるたびに新しいログ・ファイルを開始します。1つのログ・ファイルが許容最大サイズに達し、日付が変わっていない場合、システムは新しいログ・ファイルを作成します。ログ・ファイルの最大数に達した場合、新しいログ・ファイルを作成できるように、最も古いログ・ファイルが廃棄されます。

ログに記録されたメッセージについての E メールを受信するには、まず E メールが配信される方法に関する情報を指定します。次に、どのメッセージについて E メールを受信したいのかを指定します。

### 手順

エンタープライズ・サーチ・ログ・ファイルを構成するには、次のようにします。

1. システム・レベルのログ・ファイルの作成と循環についてのオプションを構成したい場合、次のようにします。
  - a. 「システム」をクリックして「システム」ビューを開きます。
  - b. 「 編集」をクリックして、システム編集ビューに切り替えます。
  - c. 「ログ」ページで、「ログ・ファイル・オプションの構成」をクリックします。「システム・レベルのログ・ファイル・オプション」ページが表示されます。
2. コレクション・レベルのログ・ファイルの作成と循環についてのオプションを構成したい場合、次のようにします。
  - a. 「コレクション」ビューで、オプションを指定するコレクションを指して、「 編集」をクリックします。
  - b. 「ログ」ページで、「ログ・ファイル・オプションの構成」をクリックします。「コレクション・レベルのログ・ファイル・オプション」ページが表示されます。
3. 「ログ対象の情報タイプ」フィールドで、ログに記録したいメッセージのタイプを選択します。

### エラー・メッセージのみ

エラー・メッセージは、望ましくない状態または予期しない動作が発生したこと、処理が継続できないことを示します。何らかの処置を実行して問題を解決する必要があります。

### エラーおよび警告メッセージ

警告メッセージは、起こりうる競合または不整合を示しますが、そのために処理が停止するわけではありません。このオプションがデフォルトです。

### すべてのメッセージ

情報メッセージは、システムまたは現行タスクの一般情報を示し、修正処置は何も必要ありません。

**重要:** このオプションを選択すると、システム・パフォーマンスに悪影響を与えることがあります。すべてのメッセージのログ記録は、問題をトラブルシューティングする必要がある場合、または IBM ソフトウェア・サポートから要求された場合にのみ行なってください。

4. 「各ログ・ファイルの最大サイズ」フィールドに、各ログ・ファイルの最大サイズを M バイトで入力します。デフォルト値は 10MB です。

ログ・ファイルがこのサイズに達すると、ログ・ファイルの許容最大数に達するまで新しいログ・ファイルが作成されます。ログ・ファイルを比較的小さいサイズにしておく、表示して調べるときの効率がよくなります。

5. 「ログ・ファイルの最大数」フィールドに、作成するログ・ファイルの最大数を入力します。デフォルト値は 16 です。

古いログ・メッセージも確実に参照できるようにしたい場合は、この数値を大きくしてください。関心があるのは最近のメッセージであり、長期間のアクティビティ履歴を保持する必要がない場合は、この値を減らしてください。

6. 「デフォルト・ロケール」フィールドで、ログ・メッセージに使用したい言語を選択します。デフォルト値は英語です。
7. 「OK」をクリックします。
8. 変更を有効にするために、次のコマンドを入力して、エンタープライズ・サーチ・システムを一度停止してから再始動します。

```
esdmin system stopall  
esdmin system startall
```

## SMTP サーバー情報の構成

エンタープライズ・サーチのアクティビティに関する E メールを受信できるようにするには、その前に、SMTP (Simple Mail Transfer Protocol) サーバーの情報を構成しておく必要があります。

### 始める前に

SMTP サーバーの情報を構成するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。


### このタスクについて

いくつかのエンタープライズ・サーチ管理機能を使用して、E メールを受信できます。これらの機能のいずれかから E メールを受信するには、その前に、SMTP サーバーの情報を指定する必要があります。

- コレクション・レベルのアラートまたはシステム・レベルのアラートを構成すると、これらのメッセージがログに記録されるかどうかに関係なく、E メールを受信できます。また、モニター対象のイベントに起因するメッセージだけでなく、他のメッセージがログに記録されたときに E メールを受信することもできます。
- 索引内の URI、またはエンタープライズ・サーチ・システムからドロップされた文書に関して詳しい情報を表示したい場合、レポートを E メールで受け取ることができます。
- Web クローラーをモニターし、Web クローラーの履歴レポートを作成するよう指定した場合、レポートの作成後に E メールで通知されます。

### 手順

SMTP サーバーの情報を構成するには、次のようにします。

1. 「システム」をクリックして「システム」ビューを開きます。
2. 「 編集」をクリックして、システム編集ビューに切り替えます。
3. 「ログ」ページで、「メッセージに対する E メール・オプションの構成」をクリックします。
4. 「システム・メッセージに対する E メール・オプション」ページで、「E メール配信に使用する SMTP メール・サーバー」フィールドに、使用したい SMTP サーバーの完全修飾ホスト名または IP アドレスを入力します。

システムは、このサーバーを使用して、指定したアドレスに E メールを送信します。

5. 「Eメールの確認頻度」フィールドに、適格メッセージのチェックとそれらに関する Eメールの送信をシステムが行う頻度を指定します。

システムは、1 つの特定の E メール・アドレス向けのメッセージはすべて結合して 1 つのメッセージにし、指定した頻度でそのメッセージを送信します。

6. 「OK」をクリックします。

## ログ・メッセージについての Eメールの受信

特定のメッセージまたはメッセージ・タイプがログに記録されたら Eメールを受信するというオプションを指定できます。

### 始める前に

システム・レベル・メッセージの Eメール・オプションを構成するには、エンタープライズ・サーチ管理者役割のメンバーである必要があります。コレクション・レベル・メッセージの Eメール・オプションを構成するには、エンタープライズ・サーチ管理者役割のメンバーであるか、そのコレクションのコレクション管理者である必要があります。


Eメールを受信できるようにするには、最初に、Eメールが送達できるように SMTP (Simple Mail Transfer Protocol) サーバーの情報を構成しておく必要があります。

### このタスクについて

アラートの構成では、特定のイベントが発生したらメッセージをログに記録するというオプションを選択できます。それらのオプションを使用可能にした場合、それらのメッセージがログに記録されたら自動的に E メールを受信するというオプションを構成できます。また、イベントに起因するメッセージだけでなく、他のメッセージがログに記録されたときに E メールを受信するというオプションを指定することもできます。

## 手順

メッセージについての E メール・オプションを構成するには、次のようにします。

1. システム・メッセージに関して E メールを受け取りたい場合は、次のようにします。
  - a. 「システム」をクリックして「システム」ビューを開きます。
  - b. 「 編集」をクリックして、システム編集ビューに切り替えます。
  - c. 「ログ」ページで、「メッセージに対する E メール・オプションの構成」をクリックします。
  - d. 「システム・メッセージの E メール・オプション」ページで、「システム・レベル・メッセージに関する E メールを送信する」チェック・ボックスを選択します。
  - e. 「E メール受信用 E メール・アドレス」フィールドに、1 つ以上の E メール・アドレスを入力します。通常、エンタープライズ・サーチ管理者は、システム・メッセージに関する情報を受信する必要があります。


各アドレスはコンマで区切ってください。次に例を示します。

```
steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com
```

- f. ログに記録されるすべてのエラー・メッセージについて E メールを受信したい場合は、「全エラー・メッセージに関する E メールを送信する」チェック・ボックスを選択します。
- g. 特定のシステム・レベル・メッセージがログに記録されたときにのみ E メールを受信したい場合は、「特定のメッセージに関する E メールを送信する」領域でそれらのメッセージのメッセージ ID を入力します。1 行につき 1 つのメッセージ ID を入力してください。次に例を示します。

```
FFQC4819E  
FFQ00005E
```

デフォルトでは、いくつかのメッセージ ID がリストされます (これらのメッセージの説明については、「ヘルプ」をクリックしてください)。

- h. 「OK」をクリックします。
2. コレクションのメッセージに関して E メールを受け取りたい場合は、次のようにします。
    - a. 「コレクション」をクリックして「コレクション」ビューを開きます。
    - b. コレクションのリストで、構成するコレクションを指して、「 編集」をクリックします。
    - c. 「ログ」ページで、「メッセージに対する E メール・オプションの構成」をクリックします。



- d. 「コレクション・メッセージの E メール・オプション」 ページで、「コレクション・レベル・メッセージに関する E メールを送信する」 チェック・ボックスを選択します。
- e. 「E メール受信用 E メール・アドレス」 フィールドに、1 つ以上の E メール・アドレスを入力します。通常、コレクション管理者は、コレクション・レベルのメッセージに関する情報を受信する必要があります。

各アドレスはコンマで区切ってください。次に例を示します。

steinbeck@us.ibm.com, yeats@ireland.ibm.com, dante@it.ibm.com

- f. ログに記録されるすべてのエラー・メッセージについて E メールを受信したい場合は、「全エラー・メッセージに関する E メールを送信する」 チェック・ボックスを選択します。
- g. 特定のコレクション・レベル・メッセージがログに記録されたときのみ E メールを受信したい場合は、「特定のメッセージに関する E メールを送信する」 領域でそれらのメッセージのメッセージ ID を入力します。1 行につき 1 つのメッセージ ID を入力してください。次に例を示します。

FFQC4819E  
FFQ00005E

デフォルトでは、いくつかのメッセージ ID がリストされます (これらのメッセージの説明については、「ヘルプ」をクリックしてください)。

- h. 「OK」 をクリックします。

#### 関連概念

 エンタープライズ・サーチのメッセージ

#### 関連タスク

360 ページの『コレクション・レベル・アラートの構成』

362 ページの『システム・レベル・アラートの構成』

## 照会ログのサイズの変更

照会処理のために作成したログ・ファイルのサイズを増大または減少するには、構成ファイルを編集します。エンタープライズ・サーチ管理コンソールでは、このタスクのサポートはありません。

### このタスクについて

照会処理時に、ログ・データが *collection\_ID\_OmniFindQueryLog\_date.log* ファイルに書き込まれます。ここで、*collection\_ID* は構成するコレクションを示し、*date* はログ・ファイルが作成される日付です。ログ・ファイルが新たに作成される前に、ログ記録するデータ量に応じて、このログ・ファイルのサイズを増大または減少できます。

### 手順

照会処理のログ・ファイルのサイズを変更するには、次の手順で行います。

1. エンタープライズ・サーチ管理者としてログインします。マルチサーバー・エンタープライズ・サーチ・システムでは、索引サーバーにログインします。

2. ES\_ROOT\_NODE/master\_config/collection\_ID.runtime.nodel/runtime-generic.properties ファイルを開きます。
3. **MaxFileSize** プロパティを検索します。ログ・ファイルのサイズを増大または減少するために、この値を増大または減少させて、変更内容を保存します。
4. シングル・サーバー・エンタープライズ・サーチ・システムの場合は、次のようにします。
  - a. エンタープライズ・サーチ管理コンソールを使用して、変更した停止をモニターし、検索サーバーを停止します。
  - b. WebSphere Application Server の ESSearchServer アプリケーションを再始動します。
  - c. 管理コンソールで、停止した検索サーバーを再始動します。
  - d. 新規ブラウザで検索アプリケーションを開きます。
5. マルチサーバー・エンタープライズ・サーチ・システムの場合は、次のようにします。
  - a. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。
  - b. 以下のコマンドを入力して、エンタープライズ・サーチ・システムを再始動します。

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

---


## ログ・ファイルの表示


システムおよびコレクション・コンポーネントが共通ログ・ファイルに書き込むログ・メッセージを表示できます。また、フィルターを指定して、特定の重大度レベルのメッセージ、および特定のエンタープライズ・サーチ・セッションのメッセージのみを表示することもできます。

### 始める前に

すべてのエンタープライズ・サーチ管理ユーザーは、管理を認可されているコレクションのログ・ファイルを表示できます。システム・レベルのログ・ファイルを表示するには、エンタープライズ・サーチ管理者役割のメンバーであるか、または「システム」ツールバーへのアクセス許可を持っている必要があります。

### 手順

1. 単一のコレクションのログ・ファイルを表示するには、次のようにします。
  - a. 「**コレクション**」をクリックして「コレクション」ビューを開きます。
  - b. コレクションのリストで、表示するコレクションを指して、「 **モニター**」をクリックし、「ログ」ページを開きます。

**ヒント:** コレクションを編集途中であり、既に「ログ」ページが表示されている場合は、「 **モニター**」をクリックすれば、コレクションをモニターするためのビューに移動できます。

2. システム・レベルのログ・ファイルを表示するには、次のようにします。
  - a. 「**システム**」をクリックして「システム」ビューを開きます。

- b. 「ログ」 ページを選択します。
3. 「ログ・ファイル」 フィールドで、表示したいログ・ファイルを選択します。


各ログ・ファイルの名前には、ログ・ファイルのタイプ (システムまたはコレクション名など) とファイルの作成日が含まれています。同じ日に同じタイプのログ・ファイルが複数作成された場合は、数字のサフィックスによって、ファイルの作成順が示されます。次に例を示します。

```
log_file_type_20060526.log (この日の最新の項目が含まれます)
log_file_type_20060526.log.1
log_file_type_20060526.log.2 (この日の最も古い項目が含まれます)
log_file_type_20060525.log (この日の最新の項目が含まれます)
log_file_type_20060525.log.1
log_file_type_20060525.log.2
log_file_type_20060525.log.3 (この日の最も古い項目が含まれます)
```

4. 特定の重大度レベルのメッセージのみを表示するには、「重大度」フィールドで適切なチェック・ボックスを選択します。
5. 特定のセッションからのメッセージのみを表示するには、「セッション」フィールドで適切なチェック・ボックスを選択します。
6. 「ログの表示」をクリックします。

「ログ・ファイルのコンテンツ」ページの各メッセージごとに、メッセージが発行された日時、メッセージ重大度レベル、メッセージを発行したセッションの名前、およびメッセージ ID とエラー・テキストが表示されます。

ボタンをクリックすることで、ログ・ファイルの先頭ページ、最終ページ、直前のページ、または次のページに移動することができます。ページ番号を指定することで、そのページに直接移動することもできます。

7. メッセージについての詳細情報を表示する場合は、「 詳細」をクリックします。

「ログ・メッセージの詳細」ページに、メッセージを出したエンタープライズ・サーチ・サーバーのホスト名、エラーを引き起こしたファイルの名前、エラーが発生した関数名および行番号、プロセス ID、スレッド ID が表示されます。

移動のボタンをクリックして、ログ・ファイル内の前後のメッセージに移動できます。



---

## エンタープライズ・サーチ・システムのバックアップとリストア

バックアップ・スクリプトおよびリストア・スクリプトを使用して、エンタープライズ・サーチ・システムをバックアップおよびリストアすることができます。

### スクリプトでバックアップされるもの

これらのスクリプトは、次のファイルをバックアップおよびリストアします。

- ES\_NODE\_ROOT/master\_config ディレクトリーの構成ファイル
- クローラーのデータベース・ファイル。データ・ソースが最後にいつクローラされたかなどの、すべてのクローラー・メタデータを含みます。
- ES\_NODE\_ROOT/data ディレクトリー内のすべてのファイル
- デフォルト以外のデータ・ディレクトリーで構成されたコレクションの索引ファイル

### バックアップ・ディレクトリーの構造

バックアップ・スクリプトは、スクリプトの実行時に指定したディレクトリーの下に以下のサブディレクトリーを作成します。エンタープライズ・サーチ管理者 ID は、指定したディレクトリーに対して書き込み許可を持っている必要があります。

#### master\_config

ES\_NODE\_ROOT/master\_config ディレクトリーの構成ファイルが保管されます。

#### database

クローラー・サーバーのデータベース・ファイルが保管されます。

**data** 索引サーバーの索引ファイルが保管されます。

### 使用法のガイドライン

- あるコンピューターからデータをバックアップし、それを別のコンピューターにリストアすることができます。ただし、
  - あるバージョンの OmniFind Enterprise Edition からバックアップされたファイルを、異なるバージョンの OmniFind Enterprise Edition が稼働するシステムにリストアすることはできません。
  - 同数以上のエンタープライズ・サーチ・サーバーがあるシステムにデータをリストアしてください。例えば、1 台のサーバーで稼働するエンタープライズ・サーチ・システムをバックアップする場合は、2 台または 4 台のエンタープライズ・サーチ・サーバーを使用するシステムにデータをリストアできます。4 サーバー・システムからバックアップされたデータは、2 台のサーバーまたは 1 台のサーバーを使用するシステムにリストアできません。
  - あるオペレーティング・システムからバックアップされたファイルを、異なるオペレーティング・システムが稼働するシステムにリストアすることはできません。例えば、過去にエンタープライズ・サーチを AIX システムにインストール

ールしたが、今はエンタープライズ・サーチを Linux で実行したいという場合は、新しいエンタープライズ・サーチ・システムを Linux サーバーにインストールする必要があります。

- 最新の索引付けされたデータがバックアップされるようにするために、バックアップを開始する前に、主索引を作成してください。
- インストール・ディレクトリー (ES\_INSTALL\_ROOT)、データ・ディレクトリー (ES\_NODE\_ROOT)、エンタープライズ・サーチ管理者 ID およびパスワードの設定はすべて、バックアップされるシステムと、データのリストア先のシステムで同じでなければなりません。
- マルチサーバー構成の場合、システムのバックアップおよびリストアは、エンタープライズ・サーチ索引サーバーから行ってください。すべてのクローラー・データがクローラー・サーバー上のデータベースにあるため、スクリプトは、リモート・コマンドを実行してクローラー・データをバックアップおよびリストアします。
- エンタープライズ・サーチ・システム・ファイルを別のディレクトリーにバックアップする場合、十分なディスク・スペースを確保する必要があります。バックアップ・スクリプトおよびリストア・スクリプトでは、ファイルのチェックは行いません。
- バックアップ・スクリプトおよびリストア・スクリプトの実行中は、すべてのシステム・セッションが停止されます。誤ったシステム情報または矛盾するシステム情報が表示されるのを回避するため、スクリプトの実行中は、エンタープライズ・サーチ管理コンソールを使用しないでください。
- 回復不能エラーによるシステム障害が発生した場合は、OmniFind Enterprise Edition を再インストールし、リストア・スクリプトを実行する必要があります。

---

## エンタープライズ・サーチ・システムのバックアップ

エンタープライズ・サーチ・システムのバックアップは、AIX、Linux、または Solaris の場合は `esbackup.sh` スクリプトを、Microsoft Windows の場合は `esbackup.bat` スクリプトを使用して行います。

### 制約事項

エンタープライズ・サーチ管理者 ID には、バックアップ・スクリプトを実行するときに指定するディレクトリーへの書き込み権限が必要です。

バックアップ・スクリプトおよびリストア・スクリプトの実行中は、すべてのシステム・セッションが停止されます。誤ったシステム情報または矛盾するシステム情報が表示されるのを回避するため、スクリプトの実行中は、エンタープライズ・サーチ管理コンソールを使用しないでください。

**重要:** Ctrl キーを押しながら C を押してバックアップ・スクリプトを中断すると、システムは不整合な状態に入ります。次のコマンドを入力して、すべてのサービス・セッションと、バックアップ処理中に停止された実行中セッションを開始してください。

```
esadmin system startall
```

### 手順

エンタープライズ・サーチ・システムをバックアップするには、次のようにします。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. エンタープライズ・サーチの共通通信層 (CCL) が稼働していない場合は、それを始動します。

#### AIX、Linux、または Solaris

```
startccl.sh -bg
```

#### Windows コマンド・プロンプト

```
startccl
```

#### Windows サービス管理ツール

バックグラウンドで CCL を開始するには、次のようにします。

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」により、Windows サービスを起動します。
  - b. 「IBM OmniFind Enterprise Edition」を右クリックし、「開始」をクリックします。
3. マルチサーバー構成の場合は、CCL が各サーバーで開始されていることを確認してください。必要に応じて前のステップを繰り返してください。
  4. 次のようにしてバックアップを開始します。ここで、*backup\_directory* はバックアップされるデータのディレクトリーです。

#### AIX、Linux、または Solaris

```
esbackup.sh backup_directory
```

#### Windows コマンド・プロンプト

```
esbackup.bat backup_directory
```

#### 関連資料

409 ページの『エンタープライズ・サーチ・コマンド、戻りコード、およびセッション ID』

---

## エンタープライズ・サーチ・システムのリストア

OmniFind Enterprise Edition を再インストールした後、`esrestore.sh` スクリプト (AIX、Linux、Solaris の場合) または `esrestore.bat` スクリプト (Microsoft Windows の場合) を使用して、エンタープライズ・サーチ・システムをリストアすることができます。

#### 制約事項

バックアップ・スクリプトおよびリストア・スクリプトの実行中は、すべてのシステム・セッションが停止されます。誤ったシステム情報または矛盾するシステム情報が表示されるのを回避するため、スクリプトの実行中は、エンタープライズ・サーチ管理コンソールを使用しないでください。

あるバージョンの OmniFind Enterprise Edition からバックアップされたファイルを、異なるバージョンの OmniFind Enterprise Edition が稼働するシステムにリストア

アすることはできません。さらに、データのリストア先システムには、データがバックアップされたシステムと同数以上のエンタープライズ・サーチ・サーバーが必要です。

## 手順

エンタープライズ・サーチ・システムをリストアするには、次のようにします。

1. 索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
2. エンタープライズ・サーチの共通通信層 (CCL) が稼働していない場合は、それを始動します。

### AIX、Linux、または Solaris

```
startccl.sh -bg
```

### Windows コマンド・プロンプト

```
startccl
```

### Windows サービス管理ツール

バックグラウンドで CCL を開始するには、次のようにします。

- a. 「スタート」 → 「プログラム」 → 「管理ツール」 → 「サービス」により、Windows サービスを起動します。
  - b. 「**IBM OmniFind Enterprise Edition**」を右クリックして、「開始」をクリックします。
3. マルチサーバー構成の場合は、CCL が各サーバーで開始されていることを確認してください。必要に応じて前のステップを繰り返してください。
  4. コントローラーを停止します。

```
esadmin stop
```

5. エンタープライズ・サーチのデータをリストアします。ここで、*backup\_directory* はファイルをバックアップしたディレクトリーです。

### AIX、Linux、または Solaris

```
esrestore.sh backup_directory
```

### Windows コマンド・プロンプト

```
esrestore.bat backup_directory
```

### 関連資料

409 ページの『エンタープライズ・サーチ・コマンド、戻りコード、およびセッション ID』

---

## コレクション構成のエクスポートとインポート

コレクション構成のエクスポートとインポートは個別に行うことができます。エクスポートおよびインポートされるのは、コレクション・データではなく、コレクション構成ファイルのみです。

エンタープライズ・サーチ・システムからコレクションをエクスポートして、そのコレクションを別のエンタープライズ・サーチ・システムにインポートすることができます。ただし、コレクションのエクスポートとインポートを行えるのは、同じ



バージョンのエンタープライズ・サーチを実行しているシステム間に限られます。例えば、バージョン 8.4 のシステムからエクスポートしたコレクションをバージョン 8.4 やバージョン 8.4.0.150 のシステムにインポートすることは可能ですが、バージョン 8.3 のシステムからエクスポートしたコレクションをバージョン 8.4 のシステムにインポートすることはできません。

コレクションをエクスポートした場合、そのコレクションは、同じオペレーティング・システムを使用しているエンタープライズ・サーチ・システムにのみインポートできます。例えば、Linux にインストールされているエンタープライズ・サーチ・システムからエクスポートしたコレクションを、Windows にインストールされているエンタープライズ・サーチ・システムにインポートすることはできません。

## コレクションのエクスポート

コレクションをエクスポートするには、一度に 1 つのコレクションをエクスポートする **esadmin export** コマンドを使用します。エンタープライズ・サーチ管理コンソールでは、コレクションのエクスポートはサポートされません。

1. エンタープライズ・サーチ管理者としてログインします。複数サーバー構成では、索引サーバーでログインする必要があります。(他のエンタープライズ・サーチ・サーバーからコマンドを実行しようとする、コマンドは失敗します。)
2. 以下のコマンドを入力します。

```
esadmin export -help
```

```
esadmin export -cid コレクション ID [-fname エクスポートするファイル名] [-verbose]
```

ここで、それぞれ次のことを表します。

### **-help**

このコマンドのヘルプ情報を提供します。

### **-cid コレクション ID**

エクスポートするコレクションのコレクション ID を指定します。

**ヒント:** エクスポートするコレクションのコレクション ID を調べるには、エンタープライズ・サーチ管理コンソールまたは **esadmin report collections** コマンドを使用できます。

### **-fname エクスポートするファイル名**

エクスポート・ファイルのパスを指定します。ファイル名が絶対ファイル名でない場合は、ES\_NODE\_ROOT/dump ディレクトリーを指定したものと見なされます。このオプションを省略すると、次の命名規則に従うファイルが ES\_NODE\_ROOT/dump ディレクトリーに作成されます。

```
_export_yyyyMMdd_HHmssz.zip
```

各要素の意味:

*yyyyMMdd*

エクスポート・コマンドが実行される現在の年、月、および日です。

*HHmss*

エクスポート・コマンドが実行される現在の時、分、および秒です。

- z エクスポート・コマンドが実行される GMT からの時間帯オフセットです。例えば、太平洋標準時は GMT -0800 です。

**-verbose**

エクスポート操作のトラブルシューティングに役立つ情報を表示します。

## コレクションのインポート

コレクションをインポートするには、一度に 1 つのコレクションをインポートする **esadmin import** コマンドを使用します。エンタープライズ・サーチ管理コンソールでは、コレクションのインポートはサポートされません。

1. エンタープライズ・サーチ管理者としてログインします。複数サーバー構成では、索引サーバーでログインする必要があります。(他のエンタープライズ・サーチ・サーバーからコマンドを実行しようとする、コマンドは失敗します。)
2. 以下のコマンドを入力します。

```
esadmin import -help
esadmin import -fname インポートするファイル名
    [-cid 新規コレクション ID]
    [-name 新規コレクション ID]
    [-colDataDir 新規コレクション・データ・ディレクトリー]
    [-force]
    [-verbose]
```

ここで、それぞれ次のことを表します。

**-help**

このコマンドのヘルプ情報を提供します。

**-fname** インポートするファイル名

インポート・ファイルのパスを指定します。ファイル名が絶対ファイル名でない場合は、ES\_NODE\_ROOT/dump ディレクトリーを指定したものと見なされます。

**-cid** 新規コレクション ID

コレクションを別の ID でインポートする必要がある場合は、新しいコレクション ID を指定します。

**ヒント:** インポートするコレクションのコレクション ID を調べるには、エンタープライズ・サーチ管理コンソールまたは **esadmin report collections** コマンドを使用できます。

**-name** *new\_collection\_name*

コレクションを別の名前でもインポートする必要がある場合は、新しいコレクション名を指定します。

**-colDataDir** *new\_collection\_data\_directory*

コレクションのデータ・ディレクトリーを指定します。省略した場合は、デフォルトのディレクトリーが割り当てられます。

**-force**

既にターゲット・システムに存在するコレクションのインポートを強制します。システムは、インポートされたコレクションのコレクション ID を使用して、ターゲット・システムで置き換えるコレクションを決定します。

### 重要:

- コレクションをインポートした後、同じコレクション ID を持つコレクションを再度インポートすることはできません。どうしてもしたい場合は、このオプションを使用して、コレクションのインポートを強制します。
- 強制インポートされたコレクションは、テキスト分析エンジン、辞書、およびクローラー・プラグインを保存しません。これらは、インポートされたコレクションに含まれている情報で上書きされます。
- コレクションに関連付けられているクローラーはすべて除去され、インポート・ファイルに指定されているクローラーで置き換えられます。クローラーが除去されたので、まだ解析されていないクローラーのメタデータおよび文書もすべて除去されます。コレクションがインポートされたならば、すべての文書を再クロールして、インポートされたクローラー用にこのデータを再生成する必要があります。

### -verbose

インポート操作のトラブルシューティングに役立つ情報を表示します。

## 使用法のガイドライン

### テキスト分析エンジンと辞書

エクスポート・システムでコレクションに関連付けられているテキスト分析エンジンと辞書 (同義語辞書、ストップワード辞書、およびランキング調整ワード辞書) は、コレクション構成データと一緒にエクスポートされません。テキスト分析エンジンと辞書への名前関連付けのみがエクスポートされます。

インポート時に、同じ名前を持つテキスト分析エンジンまたは辞書がターゲット・システムに存在すると、それがインポートされたコレクションに関連付けられます。存在しない場合は、関連付けを確立できなかったことを示す警告メッセージが表示されます。このため、テキスト分析エンジンと辞書をターゲット・システムにアップロードし、インポートされたコレクションで使用された名前と同じものを使用する必要があります。

同じ名前を持つテキスト分析エンジンまたは辞書がターゲット・システムに存在しない場合は、関連付けが断ち切られます。関連付けが断ち切られると、ターゲット・システムではコレクションが正しく機能しますが、コレクションは対応するテキスト分析エンジンも辞書も使用しなくなります。

### メモリー・モデル

エクスポート・システム用に構成されたメモリー・モデルは保存されません。ターゲット・システム用に構成されたメモリー・モデルは構成されたままとなります。より小さいメモリー・モデル用に構成されたシステムにコレクションをインポートしようとする、警告メッセージが表示されます。コレクションは期待どおりに作動しないことがあり、より小規模のシステムのリソース使用に重大な影響を及ぼす可能性があります。問題を避けるためには、エクスポート・システムと同じメモリー・モデルか、より大きいメモリー・モデルを使用しているシステムにコレクションをインポートしてください。

### クローラー

クローラーの前提条件がターゲット・システムにインストールされて構成されていることを確認する必要があります。

インポートされたクローラーは、クローラーがクロールするデータ・ソースにアクセスできるまで機能しません。例えば、ローカル・ファイル・システムをクロールするファイル・システム・クローラーがコレクションに含まれていると、同じディレクトリー構造がターゲット・システムに存在するようになるまでクローラーはクロールできません。

クローラー・プラグインはエクスポートされません。コレクションをインポートすると、警告メッセージが表示され、その後インポートは完了まで続行されます。コレクションをインポートしたならば、クローラーのプロパティを更新し、クローラーのプラグインを関連付けてください。クローラーが正しく機能するようにクローラーのプラグインがターゲット・システムにデプロイされたことを確認する必要があります。

#### **関連資料**

409 ページの『エンタープライズ・サーチ・コマンド、戻りコード、およびセッション ID』

---

## Lotus Notes バージョン 8 との統合

IBM Lotus Notes バージョン 8 クライアント・デプロイメントの検索機能を拡張するには、OmniFind Enterprise Edition Lotus Notes 検索プラグインをデプロイします。このプラグインによって、Lotus Notes クライアントの検索バーからエンタープライズ・サーチ・コレクションを検索できるようになります。

エンタープライズ・サーチ機能と Lotus Notes を統合するには、プラグイン更新サイトを作成する必要があります。その後、プラグインをローカルの Lotus Notes バージョン 8 クライアント・システムにインストールできます。

---

### エンタープライズ・サーチ・プラグイン更新サイトの作成

エンタープライズ・サーチと Lotus Notes バージョン 8 を統合するには、エンタープライズ・サーチ Eclipse 更新サイトを組織の Web サーバーにデプロイする必要があります。更新サイトによって、ユーザーがエンタープライズ・サーチ・プラグインをローカルの Lotus Notes クライアント・インストール・システムにデプロイできます。

#### このタスクについて

エンタープライズ・サーチ更新サイトの作成に必要なファイルは、`com.ibm.es.notes.search.plugin_8.4.0.150.zip` ファイルにあります。

#### 手順

エンタープライズ・サーチ Eclipse 更新サイトを作成するには、次のようにします。

1. `com.ibm.es.notes.search.plugin_8.4.0.150.zip` ファイルを組織の Web サーバーにコピーします。
2. 共用ディレクトリーのサブディレクトリー内の zip ファイルの内容を Web サーバーのインストール・パスにアンパックします。

例えば、更新サイトをエンタープライズ・サーチに使用する Web サーバーにデプロイするには、zip ファイルの内容を `C:\Program Files\IBM\HTTP Server\htdocs\en_US\omnifind` ディレクトリーにアンパックします。

---

### Lotus Notes バージョン 8 クライアントへのエンタープライズ・サーチ・プラグインのインストール

照会処理にエンタープライズ・サーチ機能を使用するために、Lotus Notes 検索バーにエンタープライズ・サーチ・プラグインを追加できます。

#### このタスクについて

Lotus Notes クライアントの右上隅に、検索バーがあります。検索バーには、基本 Lotus Notes クライアント・インストール・システムで提供されるいくつかの検索プ

ログインがあります。エンタープライズ・サーチ・プラグインをインストールすると、検索バーからエンタープライズ・サーチ・コレクションを検索できます。

## 手順

エンタープライズ・サーチ・プラグインを Lotus Notes クライアントにインストールするには、次の手順で行います。

1. NOTES\_INSTALL\_ROOT¥notes.ini ファイルを開きます。NOTES\_INSTALL\_ROOT は通常は c:¥lotus¥notes です。
2. ファイルの末尾に、OMNIFIND\_ENTERPRISE\_EDITION\_SEARCH\_SERVER\_URL というプロパティを追加して、その値に http://ホスト名:ポート/ESSearchApplication/search.do?q= を指定します。ここで、「ホスト名」は、エンタープライズ・サーチの検索サーバーのホスト名で、ポートは Web サーバーのポートです。
3. ファイルを保存して閉じます。
4. Lotus Notes クライアント・アプリケーションを開きます。
5. 「ファイル」 → 「アプリケーション」 → 「インストール」を選択します。
6. 「インストールする新機能の選択」ラジオ・ボタンを選択します。
7. 「リモート・ロケーションの追加」をクリックします。
8. この新しいロケーションに名前を指定します。例えば、「エンタープライズ・サーチ」とします。URL フィールドに、エンタープライズ・サーチ更新サイトのルート URL を入力します。例えば、  
com.ibm.es.notes.search.plugin\_8.4.0.150.zip ファイルをルート IBM HTTP Server English ディレクトリーの omnifind サブディレクトリーにアンパックする場合は、http://ホスト名:port/omnifind/ と指定します。
9. 「完了」をクリックして、新規リモート・サイトを保存します。
10. サイト名の横のチェック・ボックスを選択して「次へ」をクリックします。
11. ご使用条件を確認し、「使用条件の条項に同意します」を選択して「次へ」をクリックします。
12. 「完了」をクリックして、プラグインをインストールします。未署名プラグインをインストールしているという警告が表示されたら、「このプラグインをインストール」ラジオ・ボタンを選択して「OK」をクリックします。
13. プロンプトが出されたら、「はい」をクリックして、Lotus Notes クライアントを再始動します。

これで Lotus Notes 検索バーの横のメニューをクリックして、検索オプションとして「**OmniFind Enterprise Edition**」を選択できます。照会条件を入力して「検索」ボタンをクリックすると、要求がエンタープライズ・サーチ・サーバーに転送され、結果が Lotus Notes の Web ブラウザー・ウィンドウに表示されます。

---

## WebSphere Portal との統合

IBM WebSphere Portal の検索機能は、WebSphere Portal にエンタープライズ・サーチ・ポートレットの検索ポートレットをデプロイすることや、WebSphere Portal がデフォルト検索エンジンとして検索ポートレットを使用するように構成することによって拡張することができます。

### 統合の設定ポイント

OmniFind Enterprise Edition インストール・プログラムには、エンタープライズ・サーチを WebSphere Portal と統合するためのセットアップ・スクリプトが用意されています。これらのスクリプトを実行した後、エンタープライズ・サーチ・システムはいくつかの方法で WebSphere Portal と統合することができます。

#### エンタープライズ・サーチの検索ポートレット

WebSphere Portal は、アプリケーション、コンテンツ、プロセス、人と相互通信を行うための単一のアクセス・ポイントをユーザーに提供します。

WebSphere Portal フレームワークにより、ポートレットと呼ばれる新規のアプリケーションを、ポータル内の他のアプリケーションに影響を与えることなく統合し、デプロイすることができます。

エンタープライズ・サーチの検索ポートレットを WebSphere Portal にデプロイすれば、WebSphere Portal インターフェースを使用してエンタープライズ・サーチ・コレクションを検索し、検索結果を処理することができます。WebSphere Portal の構成設定により、エンタープライズ・サーチ・ポートレットのルック・アンド・フィールは WebSphere Portal 環境の他のポートレットと同じ状態に保たれます。

#### WebSphere Portal 検索センター

WebSphere Portal 検索センターは、WebSphere Portal で検索が可能になっているすべてのソースを検索するときを中心となる起点です。検索センターおよびユニバーサル・サーチ・ポートレットを使用すれば、WebSphere Portal コンテンツや、検索センターに登録されたその他のコレクションを検索することができます。

エンタープライズ・サーチを WebSphere Portal バージョン 5.1 と統合するセットアップ・スクリプトを実行した場合、エンタープライズ・サーチ・ページが検索センター・インターフェースのページに追加されます。このページを選択してエンタープライズ・サーチ・コレクションだけを検索することも、エンタープライズ・サーチ・コレクションと検索センターで利用できる他のコレクションを検索する照会を入力することもできます。

エンタープライズ・サーチを WebSphere Portal バージョン 6 と統合するセットアップ・スクリプトを実行した場合、エンタープライズ・サーチ機能がフェデレーテッド・サービスとして統合され、これを使用してエンタープライズ・サーチ・コレクションと検索センターで利用できる他のコレクションを検索することができます。

#### WebSphere Portal 検索バー

すべての WebSphere Portal インターフェース・テーマの右上隅に、検索バ

ーがあります。このバーのデフォルト動作では、すべての検索要求をデフォルトの検索センター検索エンジンに送信します。より強力なエンタープライズ・サーチ機能を照会処理に使用するために、このデフォルト動作を変更して、すべての検索要求がエンタープライズ・サーチの検索ポートレットにリダイレクトされるようにできます。

### **WebSphere Portal および Web Content Management クローラー**

WebSphere Portal サイトおよび IBM Workplace Web Content Management サイトをエンタープライズ・サーチ索引に組み込むには、エンタープライズ・サーチ管理コンソールを使用して WebSphere Portal および Web Content Management クローラーを構成します。その後、エンタープライズ・サーチ・ポートレットまたは検索アプリケーションを使用して、索引付けされたコンテンツを検索します。

WebSphere Portal クローラーは、WebSphere Portal バージョン 5.1 と WebSphere Portal バージョン 6 のサイトをクロールできます。Web Content Management クローラーは、WebSphere Portal バージョン 6 サーバーのサイトをクロールできます。

### **IBM Lotus Quickr 文書**

エンタープライズ・サーチ索引に IBM Lotus Quickr 文書を含める場合は、エンタープライズ・サーチ管理コンソールを使用して、シード・リスト クローラーを構成できます。その後、WebSphere Portal のエンタープライズ・サーチ・ポートレットまたはスタンドアロン検索アプリケーションを使用して、索引付けされたコンテンツを検索できます。

シード・リスト クローラーは、WebSphere Portal バージョン 6 サーバー上の Lotus Quickr コンテンツ (文書) ライブラリーをクロールできます。

### **統合による利点**

エンタープライズ・サーチによって、より広いデータ・ソース・タイプの検索がサポートされ、WebSphere Portal 検索環境が拡張されます。エンタープライズ・サーチの検索ポートレットを使用することにより、Web サイトに加え、エンタープライズ・サーチ・システムでサポートされる他のすべてのデータ・ソース・タイプも検索することができます。

また、エンタープライズ・サーチには、スケーラビリティの面でも利点があります。ポータル検索エンジンは、1 つのサーバーで十分に検索負荷をサポートできる中小規模の企業に便利です。エンタープライズ・レベルの容量をサポートするために、エンタープライズ・サーチのワークロードは、複数のサーバーに分散し、2 つのサーバーが検索および取り出し処理のサポートを提供するようにすることができます。

---

## **エンタープライズ・サーチの WebSphere Portal との統合のセットアップ・スクリプト**

エンタープライズ・サーチを IBM WebSphere Portal と統合するには、OmniFind Enterprise Edition インストール・プログラムで提供されるセットアップ・スクリプトを実行することができます。



ご使用バージョンの WebSphere Portal 用のセットアップ・スクリプトが入った JAR ファイルを、エンタープライズ・サーチ・サーバーから WebSphere Portal がインストールされているサーバーにコピーしてください。セットアップ・スクリプトは、以下を行います。

- WebSphere Portal 内でエンタープライズ・サーチを使用できるようにする EAR ファイルをデプロイし、 WebSphere Portal と IBM Workplace Web Content Management のコンテンツをエンタープライズ・サーチ・コレクションに追加するためのクローラーを作成します。
- エンタープライズ・サーチ・ポートレットに必要な WAR ファイルをデプロイします。
- WebSphere Portal でページを作成し、エンタープライズ・サーチ・ポートレット・ファイルをそのページに割り当てます。
- 必要なすべての JAR ファイルを WebSphere Portal インストール・ディレクトリにコピーします (インストール・ディレクトリに既にあった JAR ファイルは、エンタープライズ・サーチに使用される JAR ファイルがコピーされる前にバックアップされます)。
- WebSphere Information Integrator Content Edition が Portal Document Manager 文書を検索するための統合ポイントを提供します。

スクリプトを実行した後、 WebSphere Portal 管理インターフェースを使用して、検索ポートレットのプロパティを更新し、エンタープライズ・サーチの検索サーバーに関する情報を指定してください。

## 使用法のガイドライン

- スクリプトでは、エンタープライズ・サーチと WebSphere Portal 間のすべての統合ポイントをセットアップします。例えば、ポートレットを選択的にインストールして、 WebSphere Portal および Web Content Management クローラーをサポートする EAR ファイルをインストールしないということとはできません。
- WebSphere Information Integrator Content Edition をセットアップしていないで、後で Portal Document Manager 文書を検索するポートレットを使用することにした場合は、 WebSphere Portal からエンタープライズ・サーチを除去するスクリプトを実行する必要があります。その後、セットアップ・スクリプトを再び実行し、 WebSphere Information Integrator Content Edition インストール・パスを指定できます。
- スクリプトでは、WebSphere Portal を停止して再始動します。ポータル・サービスを利用できないことでユーザー・コミュニティに影響がないように、スクリプトは、通常の運用時間後に実行したほうがよいでしょう。
- セットアップ・スクリプトの実行中にエラーが発生した場合は、セットアップ・スクリプトを再実行します。最初の試行で正常に完了したタスクはエラーを報告する可能性がありますが、セットアップ処理は続行し、残りのタスクを完了します。
- セットアップ・スクリプトの実行後にエンタープライズ・サーチ・ポートレット・ページに初めてアクセスした場合、このポートレット用の Java Server Pages (JSP ファイル) をシステムがコンパイルする必要があるため、ページの表示に時間がかかることがあります。

---

## WebSphere Portal バージョン 5.1 でのエンタープライズ・サーチのセットアップ

エンタープライズ・サーチ・システムを WebSphere Portal バージョン 5.1.0 以降と統合するには、`wp5_install` スクリプトを使用します。

### このタスクについて

エンタープライズ・サーチと WebSphere Portal の統合に必要なファイルは、`es.wp5.install.jar` ファイルで提供されています。このファイルをアンパックすると、次のファイルが解凍されます。

- `ESSearchPortlet.war`
- `ESSearchAdapterPortlet.war`
- `ESSearchAdapter.ear`
- `ESPACServer.ear`
- `esapi.jar`
- `siapi.jar`
- `es.security.jar`
- インストールに必要なスクリプト、バッチ、XML、および JACL ファイル

### 手順

エンタープライズ・サーチを WebSphere Portal バージョン 5.1 システムと統合するには、次のようにします。

1. `es.wp5.install.jar` ファイルをエンタープライズ・サーチ・サーバーから WebSphere Portal サーバーにコピーし、Java **JAR** コマンド (または **TAR** コマンド) を使用してファイルをアンパックします。
2. オプション: WebSphere Portal Document Manager (PDM) との統合をサポートしたい場合は、次のいずれかのステップを行います。
  - WebSphere Information Integrator Content Edition のインストール・プログラムを実行し、コネクタのみのインストールを実行するオプションを選択して、PDM コネクタを WebSphere Portal サーバーにインストールします。
  - WebSphere Information Integrator Content Edition のディレクトリ構造を WebSphere Portal サーバーに作成し、以下のファイルを既存の WebSphere Information Integrator Content Edition インストール済み環境から WebSphere Portal サーバーにコピーします。ここで、`CE_ROOT` はルートの WebSphere Information Integrator Content Edition インストール・ディレクトリを指定します。

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

3. コマンド・プロンプトで、**wp5\_install.bat** コマンド (Windows の場合) または **wp5\_install.sh** コマンド (AIX、Linux、または Solaris の場合) を実行します。以下の例では、読みやすくするためにパラメーターを別々の行に示しています。パラメーターはコマンドで次のように指定してください。

```
wp5_install.bat
-WASDir "C:\Program Files\WebSphere\AppServer"
-WASUser wpsbind -WASPassword wpsbind
-WPSDir "C:\Program Files\WebSphere\PortalServer"
-WPSUser wpsadmin -WSPassword wpsadmin
-WPSHost "portalserver.ibm.com:9081"
-IICEDir "C:\IICE"
```

#### **WASDir**

WebSphere Application Server のインストール・ディレクトリーの完全修飾パス。

#### **WASUser**

WebSphere Application Server 管理ユーザーのユーザー名。グローバル・セキュリティが WebSphere Application Server で使用可能になっている場合のみ必要です。

#### **WASPassword**

WebSphere Application Server 管理ユーザーが指定された場合に、そのパスワード。

#### **WPSDir**

WebSphere Portal インストール・ディレクトリーの完全修飾パス。

#### **WPSUser**

WebSphere Portal 管理ユーザーのユーザー名。

#### **WSPassword**

指定された WebSphere Portal 管理ユーザーのパスワード。

#### **WPSHost**

WebSphere Portal サーバーのホスト名とポート番号。

#### **IICEDir**

WebSphere Information Integrator Content Edition インストール・ディレクトリーの完全修飾パス。前に Portal Document Manager コネクタを WebSphere Portal サーバーにセットアップした場合にのみ必要です。

4. スクリプトを実行 (WebSphere Portal は停止され、再始動される) した後、次のようにして、検索サーバーを特定するようにエンタープライズ・サーチ・ポートレットを更新します。
  - a. Portal 管理者 ID およびパスワードで WebSphere Portal にログインします。
  - b. 右上隅にある「管理」をクリックします。
  - c. 左側のナビゲーション領域で「ポートレット管理」をクリックし、「ポートレット」をクリックします。
  - d. 「検索方法」オプションを「タイトルの内容」に変更します。
  - e. 「検索」フィールドに「エンタープライズ・サーチ」と入力し、「検索」ボタンをクリックします。
  - f. 右に新しいアイコンが表示されたら、レンチのアイコンをクリックして、エンタープライズ・サーチの検索ポートレットを構成します。
  - g. ポートレット・パラメーターのリストで、以下のパラメーターを変更します。

**hostname**

エンタープライズ・サーチ用の検索サーバーの完全修飾ホスト名を指定します。

**port**

エンタープライズ・サーチ用の検索サーバーで WebSphere Application Server が使用するポート番号を指定します。デフォルト値は 80 です (SSL 通信のデフォルト値は 443 です)。

**username**

検索サーバー上の WebSphere Application Server でグローバル・セキュリティが使用可能になっている場合、WebSphere Application Server ユーザー・レジストリーで有効なユーザー名を指定します。

**password**

WebSphere Application Server ユーザー名を指定した場合に、対応するパスワードを指定します。

**protocol**

WebSphere Portal と検索サーバー間の通信に使用されるプロトコルを指定します。デフォルトは HTTP です。SSL を使用する場合は、HTTPS を指定します。

**trustStore**

SSL を使用する場合、SSL 証明書ストアの完全修飾パス (ファイル名を含む) を指定します。

**trustPassword**

SSL を使用する場合、指定された trustStore ファイルのパスワードを指定します。

**ssoCookieName**

シングル・サインオン (SSO) トークン・ストリングを含む Cookie の名前を指定します。デフォルト値は LtpaToken です。

**proxyHost**

エンタープライズ・サーチの検索サーバーへのアクセスにプロキシ・サーバーが必要な場合、プロキシ・サーバーの完全修飾ホスト名を指定します。

**proxyPort**

プロキシ・サーバーを指定した場合、プロキシ・サーバーのポート番号を指定します。

**proxyUser**

プロキシ・サーバーが基本認証を必要とする場合、プロキシ・サーバーへのログインに使用するユーザー名を指定します。

**proxyPassword**

プロキシ・サーバーのユーザー名を指定した場合に、対応するパスワードを指定します。

- h. 「OK」をクリックして、変更内容を保存します。

## WebSphere Portal バージョン 5.1 検索バーによるエンタープライズ・サーチ使用の構成

検索バーでユーザーが照会を実行したときに、デフォルトの WebSphere Portal 検索エンジンではなく、エンタープライズ・サーチを使用するように、WebSphere Portal バージョン 5.1.0 以降を構成することができます。

### 始める前に

検索要求をエンタープライズ・サーチにリダイレクトするには、その前に、**wps\_install** セットアップ・スクリプトを実行して、エンタープライズ・サーチを WebSphere Portal と統合する必要があります。また、エンタープライズ・サーチ・ポートレットのパラメーターを更新して、ホスト名、ポート、エンタープライズ・サーチの検索サーバーに関する他の情報を指定する必要があります。

### このタスクについて

すべての WebSphere Portal インターフェース・テーマの右上隅に、検索バーがあります。このバーのデフォルト動作では、すべての検索要求を検索センター・ポートレットに送信します。より強力なエンタープライズ・サーチ機能を照会処理に使用するために、このデフォルト動作を変更して、すべての検索要求がエンタープライズ・サーチの検索ポートレットにリダイレクトされるようにできます。

検索バーをリダイレクトすると、変更は、エンタープライズ・サーチの検索ポートレットと同じ WebSphere Portal テーマを使用するページに影響を与えます。これらのページは、SearchBarInclude.jsp ファイルを呼び出さなければなりません。別のテーマを使用するページや、SearchBarInclude.jsp ファイルを呼び出さないページは、デフォルトの検索センター・ポートレットを引き続き使用します。

このタスクの完了後は、変更の取り消し (例えば、元の SearchBarInclude.jsp ファイルを復元します) を行わない限り、検索センターを使用できません。

### 手順

WebSphere Portal の検索バーで照会を実行したときにエンタープライズ・サーチ・ポートレットを使用するには、以下のようにします。

1. WebSphere Portal サーバーで、`WAS_INSTALL_ROOT/AppServer/installedApps/eswin1/wps.ear/wps.war/themes/html` ディレクトリーに移動します。
2. SearchBarInclude.jsp ファイルをコピーして名前変更する (例えば、SearchBarInclude.jsp.BACKUP) ことによって、このファイルのバックアップを作成します。
3. SearchBarInclude.jsp ファイルを編集して、`<wps:if loggedIn="yes">` と `</wps:if>` ステートメントの間の内容を以下のテキストで置換します。 form エレメントの `action=` 属性の、`localhost:9081` を WebSphere Portal サーバーのホスト名およびポート番号で置換します。

```
<wps:if loggedIn="yes">
<wps-internal:adminlinkinfo name="<%=AdminUniqueNamesMappingService.SEARCH_CENTER%>">
<td nowrap="nowrap" align="right" valign="middle" class="wpsLinkBar">
<form style="margin: 0px;" method="post"
action="http://localhost:9081/wps/omnifind/porta1SearchBar51.jsp">
<table border="0" cellpadding="0" cellspacing="0">
<tr>
```

```

<td valign="middle" nowrap="">&nbsp;   
<label for="q"><wps:text key="search.theme.searchbox" bundle="nls.engine"/>
</label></td>
<td valign="middle" style="padding: 0px 4px 0px 4px;">
<input class="wpsEditField" name="q" type="text">
</input></td>
<td valign="middle">
<input valign="middle" alt='<wps:text key="search.theme.searchbox.alttext"
bundle="nls.engine"/>' src="%= wpsBaseUrl %>/images/icons/Search.gif" type="image">
</input></td>
</tr></table>
</form>
</td>
</wps-internal:adminlinkinfo>
</wps:if>

```

4. Default.jsp ファイルを開き、このファイルを保存します。このステップは、ファイルが確実に再コンパイルされるようにファイルの変更日付を更新するもので、デフォルトの WebSphere Portal テーマでなく独自のテーマを使用する場合にはオプションです。
5. AdminLinkBarInclude.jsp ファイルを開き、このファイルを保存します。このステップは、ファイルが確実に再コンパイルされるようにファイルの変更日付を更新するもので、デフォルトの WebSphere Portal テーマでなく独自のテーマを使用する場合にはオプションです。
6. WebSphere Portal アプリケーション・サーバー・インスタンスを停止して再始動します。

## WebSphere Portal バージョン 5.1 からのエンタープライズ・サーチの除去

WebSphere Portal バージョン 5.1.0 以降のシステムからエンタープライズ・サーチを除去するには、wp5\_uninstall スクリプトを使用します。

### このタスクについて

WebSphere Portal からエンタープライズ・サーチを除去すると、セットアップ処理の中でエンタープライズ・サーチ・ポートレットに指定したポートレット・パラメーターは保存されません。

スクリプトを開始すると、スクリプトは WebSphere Portal サーバーを停止します。エンタープライズ・サーチ・ソフトウェアが除去された後、スクリプトは WebSphere Portal サーバーを再始動します。

### 手順

WebSphere Portal バージョン 5.1 システムからエンタープライズ・サーチを除去するには、次のようにします。

コマンド・プロンプトで、**wp5\_uninstall.bat** コマンド (Windows の場合) または **wp5\_uninstall.sh** コマンド (AIX、Linux、または Solaris の場合) を実行します。以下の例では、読みやすくするためにパラメーターを別々の行に示しています。パラメーターはコマンドで次のように指定してください。

```
wp5_uninstall.bat
-WASDir "C:\Program Files\WebSphere\AppServer"
-WASUser wpsbind -WASPassword wpsbind
-WPSDir "C:\Program Files\WebSphere\PortalServer"
-WPSUser wpsadmin -WSPassword wpsadmin
-WPSHost "portalserver.ibm.com:9081"
```

**WASDir**

WebSphere Application Server のインストール・ディレクトリーの完全修飾パス。

**WASUser**

WebSphere Application Server 管理ユーザーのユーザー名。グローバル・セキュリティが WebSphere Application Server で使用可能になっている場合にのみ必要です。

**WASPassword**

WebSphere Application Server 管理ユーザーが指定された場合に、そのパスワード。

**WPSDir**

WebSphere Portal インストール・ディレクトリーの完全修飾パス。

**WPSUser**

WebSphere Portal 管理ユーザーのユーザー名。

**WSPassword**

指定された WebSphere Portal 管理ユーザーのパスワード。

**WPSHost**

WebSphere Portal サーバーのホスト名とポート番号。

---

## WebSphere Portal バージョン 6 でのエンタープライズ・サーチのセットアップ

エンタープライズ・サーチ・システムを WebSphere Portal バージョン 6 と統合するには、wp6\_install スクリプトを使用します。

### このタスクについて

エンタープライズ・サーチと WebSphere Portal の統合に必要なファイルは、es.wp6.install.jar ファイルで提供されています。このファイルをアンパックすると、次のファイルが解凍されます。

- ESSearchPortlet.war
- ESPACServer.ear
- esapi.jar
- es.search.provider.jar
- es.security.jar
- 検索プロバイダー結果ページで使用される検索アプリケーション・ソース・タイプ・アイコン
- インストールに必要なスクリプト、バッチ、XML、および JACL ファイル

### 手順

エンタープライズ・サーチを WebSphere Portal バージョン 6 システムと統合するには、次のようにします。

1. `es.wp6.install.jar` ファイルをエンタープライズ・サーチ・サーバーから WebSphere Portal サーバーにコピーし、Java **JAR** コマンド (または **TAR** コマンド) を使用してファイルをアンパックします。
2. オプション: WebSphere Portal Document Manager (PDM) との統合をサポートしたい場合は、次のいずれかのステップを行います。
  - WebSphere Information Integrator Content Edition のインストール・プログラムを実行し、コネクタのみのインストールを実行するオプションを選択して、PDM コネクタを WebSphere Portal サーバーにインストールします。
  - WebSphere Information Integrator Content Edition のディレクトリ構造を WebSphere Portal サーバーに作成し、以下のファイルを既存の WebSphere Information Integrator Content Edition インストール済み環境から WebSphere Portal サーバーにコピーします。ここで、`CE_ROOT` はルートの WebSphere Information Integrator Content Edition インストール・ディレクトリを指定します。

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

3. コマンド・プロンプトで、**wp6\_install.bat** コマンド (Windows の場合) または **wp6\_install.sh** コマンド (AIX、Linux、または Solaris の場合) を実行します。以下の例では、読みやすくするためにオプションを別々の行に示しています。オプションはコマンドで次のように指定してください。

```
wp6_install.bat
-WSPProfileDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥profiles¥¥wp_profile"
-WASDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥AppServer"
-WASUser wpsbind -WASPassword wpsbind
-WPSDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥PortalServer"
-WPSUser wpsadmin -WSPassword wpsadmin
-WPSHost "portalserver.ibm.com:9081"
-IICEDir "C:¥¥IICE"
```

#### **WSPProfileDir**

WebSphere Portal プロファイル・ディレクトリの完全修飾パス。デフォルト・パスは、AIX システムでは `/usr/IBM/WebSphere/AppServer/profiles/wp_profile`、Linux または Solaris システムでは `/opt/IBM/WebSphere/AppServer/profiles/wp_profile`、Windows システムでは `C:¥Program Files¥IBM¥WebSphere¥profiles¥wp_profile` です。

#### **WASDir**

WebSphere Application Server のルート・ディレクトリの完全修飾パス。AIX、Linux、および Solaris システムでのみ必要です。デフォルトのルート・ディレクトリ・パスは、AIX システムで `/usr/IBM/WebSphere/AppServer`、Linux または Solaris システムで `/opt/IBM/WebSphere/AppServer`、Windows システムで `C:¥Program Files¥IBM¥WebSphere¥AppServer` です。



**WASUser**

WebSphere Application Server 管理ユーザーのユーザー名。グローバル・セキュリティが WebSphere Application Server で使用可能になっている場合のみ必要です。

**WASPassword**

WebSphere Application Server 管理ユーザーが指定された場合に、そのパスワード。

**WPSDir**

WebSphere Portal インストール・ディレクトリーの完全修飾パス。

**WPSUser**

WebSphere Portal 管理ユーザーのユーザー名。

**WPSPassword**

指定された WebSphere Portal 管理ユーザーのパスワード。

**WPSHost**

WebSphere Portal サーバーのホスト名とポート番号。

**IICEDir**

WebSphere Information Integrator Content Edition インストール・ディレクトリーの完全修飾パス。前に Portal Document Manager コネクタを WebSphere Portal サーバーにセットアップした場合にのみ必要です。

4. スクリプトを実行した後、WebSphere Portal を停止して再始動し、エンタープライズ・サーチ・ポートレットを更新して検索サーバーを識別します。
  - a. Portal 管理者 ID およびパスワードで WebSphere Portal にログインします。
  - b. 左下隅にある「管理」をクリックします。
  - c. 左側のナビゲーション領域で「ポートレット管理」をクリックし、「ポートレット」をクリックします。
  - d. 「検索方法」オプションを「タイトルの内容」に変更します。
  - e. 「検索」フィールドに「エンタープライズ・サーチ」と入力し、「検索」ボタンをクリックします。
  - f. 右に新しいアイコンが表示されたら、レンチのアイコンをクリックして、エンタープライズ・サーチの検索ポートレットを構成します。
  - g. ポートレット・パラメーターのリストで、以下のパラメーターを変更します。

**hostname**

エンタープライズ・サーチ用の検索サーバーの完全修飾ホスト名を指定します。

**port**

エンタープライズ・サーチ用の検索サーバーで WebSphere Application Server が使用するポート番号を指定します。デフォルト値は 80 です (SSL 通信のデフォルト値は 443 です)。

**username**

検索サーバー上の WebSphere Application Server でグローバル・セキュリティが使用可能になっている場合、WebSphere Application Server ユーザー・レジストリーで有効なユーザー名を指定します。

**password**

WebSphere Application Server ユーザー名を指定した場合に、対応するパスワードを指定します。

**protocol**

WebSphere Portal と検索サーバー間の通信に使用されるプロトコルを指定します。デフォルトは HTTP です。SSL を使用する場合は、HTTPS を指定します。

**trustStore**

SSL を使用する場合、SSL 証明書ストアの完全修飾パス (ファイル名を含む) を指定します。

**trustPassword**

SSL を使用する場合、指定された trustStore ファイルのパスワードを指定します。

**ssoCookieName**

シングル・サインオン (SSO) トークン・ストリングを含む Cookie の名前を指定します。デフォルト値は LtpaToken です。

**proxyHost**

エンタープライズ・サーチの検索サーバーへのアクセスにプロキシ・サーバーが必要な場合、プロキシ・サーバーの完全修飾ホスト名を指定します。

**proxyPort**

プロキシ・サーバーを指定した場合、プロキシ・サーバーのポート番号を指定します。

**proxyUser**

プロキシ・サーバーが基本認証を必要とする場合、プロキシ・サーバーへのログインに使用するユーザー名を指定します。

**proxyPassword**

プロキシ・サーバーのユーザー名を指定した場合に、対応するパスワードを指定します。

- h. 「OK」をクリックして、変更内容を保存します。

## WebSphere Portal バージョン 6 検索センターのエンタープライズ・サーチに対する構成

WebSphere Portal 検索センターでユーザーが照会を実行したときに、エンタープライズ・サーチ・コレクションを検索するように WebSphere Portal バージョン 6 を構成することができます。

### 制約事項

検索されるエンタープライズ・サーチ・コレクションがセキュアな場合、ユーザーは、エンタープライズ・サーチの検索ポートレットを実行してユーザー・プロファイルを構成する必要があります。プロファイルは、暗号化されてセキュアなエンタープライズ・サーチ・ストアに保管されています。ユーザーが WebSphere Portal 検索センターからセキュアなコレクションを検索する照会を実行するには、その前に、プロファイルが存在しなければなりません。

## このタスクについて

WebSphere Portal バージョン 6 の検索センターは、複数のコレクションにまたがるフェデレーテッド・サーチ機能をサポートします。コレクションは、ポータル文書ライブラリーや Portal Content (ページおよびポートレット) など、さまざまなタイプのコンテンツを含むことができます。エンタープライズ・サーチを WebSphere Portal と統合するセットアップ・スクリプトを実行した後、検索センターがエンタープライズ・サーチ・コレクションも検索するように構成することができます。

### 手順

検索センターがエンタープライズ・サーチ・コレクションを検索するように構成するには、次のようにします。

1. Portal 管理者 ID およびパスワードで WebSphere Portal にログインします。
2. 左下隅にある「管理」をクリックします。
3. 左側のナビゲーション領域で「検索管理」をクリックしてから、「検索の管理」をクリックします。
4. 「検索サービス」をクリックし、「新規検索サービス」をクリックします。
5. 「検索サービスのインプリメンテーション」フィールドで「エンタープライズ・サーチ検索サービス」を選択し、サービスに使用する名前を「サービス名」テキスト・ボックスに入力します。
6. パラメーターのリストで、以下のパラメーターを変更します。

#### hostname

エンタープライズ・サーチ用の検索サーバーの完全修飾ホスト名を指定します。

**port** エンタープライズ・サーチ用の検索サーバーで WebSphere Application Server が使用するポート番号を指定します。デフォルト値は 80 です (SSL 通信のデフォルト値は 443 です)。

#### username

検索サーバー上の WebSphere Application Server でグローバル・セキュリティティが使用可能になっている場合、WebSphere Application Server ユーザー・レジストリーで有効なユーザー名を指定します。

#### password

WebSphere Application Server ユーザー名を指定した場合に、対応するパスワードを指定します。

#### protocol

WebSphere Portal と検索サーバー間の通信に使用されるプロトコルを指定します。デフォルトは HTTP です。SSL を使用する場合は、HTTPS を指定します。

#### trustStore

SSL を使用する場合、SSL 証明書ストアの完全修飾パス (ファイル名を含む) を指定します。

#### trustPassword

SSL を使用する場合、指定された trustStore ファイルのパスワードを指定します。

**ssoCookieName**

シングル・サインオン (SSO) トークン・ストリングを含む Cookie の名前を指定します。デフォルト値は LtpaToken です。

**proxyHost**

エンタープライズ・サーチの検索サーバーへのアクセスにプロキシ・サーバーが必要な場合、プロキシ・サーバーの完全修飾ホスト名を指定します。

**proxyPort**

プロキシ・サーバーを指定した場合、プロキシ・サーバーのポート番号を指定します。

**proxyUser**

プロキシ・サーバーが基本認証を必要とする場合、プロキシ・サーバーへのログインに使用するユーザー名を指定します。

**proxyPassword**

プロキシ・サーバーのユーザー名を指定した場合に、対応するパスワードを指定します。

7. 「OK」をクリックして、変更内容を保存します。

## WebSphere Portal バージョン 6 検索バーによるエンタープライズ・サーチ使用の構成

検索バーでユーザーが照会を実行したときに、デフォルトの WebSphere Portal 検索エンジンではなく、エンタープライズ・サーチを使用するように、WebSphere Portal バージョン 6 を構成することができます。

### 始める前に

検索要求をエンタープライズ・サーチにリダイレクトするには、その前に、**wp6\_install** セットアップ・スクリプトを実行して、エンタープライズ・サーチを WebSphere Portal と統合する必要があります。また、エンタープライズ・サーチ・ポートレットのパラメーターを更新して、ホスト名、ポート、エンタープライズ・サーチの検索サーバーに関する他の情報を指定する必要があります。

### このタスクについて

すべての WebSphere Portal インターフェース・テーマの右上隅に、検索バーがあります。このバーのデフォルト動作では、すべての検索要求を検索センター・ポートレットに送信します。より強力なエンタープライズ・サーチ機能を照会処理に使用するために、このデフォルト動作を変更して、すべての検索要求がエンタープライズ・サーチの検索ポートレットにリダイレクトされるようになります。

検索バーをリダイレクトすると、変更は、エンタープライズ・サーチの検索ポートレットと同じ WebSphere Portal テーマを使用するページに影響を与えます。これらのページは、`banner_searchControl.jspf` ファイルを呼び出さなければなりません。別のテーマを使用するページや、`banner_searchControl.jspf` ファイルを呼び出さないページは、デフォルトの検索センター・ポートレットを引き続き使用します。

このタスクの完了後は、変更の取り消し (例えば、元の banner\_searchControl.jspf ファイルを復元します) を行わない限り、検索センターを使用できません。

## 手順

WebSphere Portal の検索バーで照会を実行したときにエンタープライズ・サーチ・ポートレットを使用するには、以下のようにします。

1. WebSphere Portal アプリケーション・サーバー・インスタンスを停止します。
2. WebSphere Portal サーバーで、WPS\_PROFILE\_ROOT/installedApps/ノード名/wps.ear/wps.war/themes/html/現在のテーマ名 ディレクトリーに移動します。ここで、ノード名は WebSphere Portal サーバーのノード名で、現在のテーマ名は WebSphere Portal サーバーに現在適用されているテーマです。WebSphere Portal サーバーのデフォルトのテーマ名は IBM です。
3. banner\_searchControl.jspf ファイルをコピーして名前変更する (例えば、banner\_searchControl.jsp.BACKUP) ことによって、このファイルのバックアップを作成します。
4. banner\_searchControl.jspf ファイルを編集して、内容を以下のテキストで置換します。form エLEMENTの action= 属性の、localhost:10038 を WebSphere Portal サーバーのホスト名およびポート番号で置換します。

```
<%@ taglib uri="/WEB-INF/tld/SearchMenuControl.tld" prefix="searchmenu" %>
<%String ic = (bidiImageRTL == null) ? "icons/scope_search_submit.gif" :
  "icons/scope_search_submit"+bidiImageRTL+".gif";%>

<searchmenu:adminlinkinfo name="SEARCH_CENTER">
<div class="searchControl">
  <form name="SearchForm" style="margin: 0px;" method="GET"
    action="http://localhost:10038/wps/omnifind/portalSearchBar.jsp">
  <table border="0" cellpadding="0" cellspacing="0">
  <tr>
  <td><span class="wpsInstructionText">
    <portal-fmt:text key="search.theme.control.label" bundle="nls.engine"/></span></td>
  <td valign="middle" style="padding: 0px 4px 0px 4px;">
    <input type="text" name="q"></input></td>
  <td valign="middle"><input tabIndex="4" valign="middle"
    title="<portal-fmt:text key='search.theme.searchresultsicon.alttext' bundle='nls.engine' />"
    alt="<portal-fmt:text key='search.theme.searchresultsicon.alttext' bundle='nls.engine' />"
    src="<portal-logic:urlFindInTheme file=">" />" type="image"></input></td>
  </tr>
  </table>
  </form>
  </div>
</searchmenu:adminlinkinfo>
```

5. banner.jspf ファイルを開き、このファイルを保存します。このステップは、ファイルが確実に再コンパイルされるようにファイルの変更日付を更新するもので、デフォルトの WebSphere Portal テーマでなく独自のテーマを使用する場合にはオプションです。
6. Default.jsp ファイルを開き、このファイルを保存します。
7. WebSphere Portal アプリケーション・サーバー・インスタンスを再始動します。

# Lotus Quickr 用のエンタープライズ・サーチ・ポートレットのセットアップ

エンタープライズ・サーチ・ポートレットを WebSphere Portal バージョン 6 にセットアップして、Lotus Quickr ソースを検索することができます。

## 始める前に

**wp6\_install.bat** コマンド (Windows) または **wp6\_install.sh** コマンド (AIX、Linux、または Solaris) を実行して、WebSphere Portal バージョン 6 にエンタープライズ・サーチをセットアップする手順に従います。

## 手順

WebSphere Portal バージョン 6 にエンタープライズ・サーチ・ポートレットをセットアップして Lotus Quickr ソースを検索するには、次の手順で行います。

1. WebSphere Portal 構成内のポートレット・パラメーターを更新します。
  - a. Lotus Quickr 管理者 ID およびパスワードで WebSphere Portal にログインします。
  - b. 「**サイト管理**」をクリックして、「**拡張管理**」をクリックします。
  - c. 左側のナビゲーション領域で「**ポートレット管理**」をクリックし、「**ポートレット**」をクリックします。
  - d. 「**検索方法**」オプションを「**タイトルの内容**」に変更します。
  - e. 「**検索**」フィールドに「**エンタープライズ・サーチ**」と入力し、「**検索**」ボタンをクリックします。
  - f. 右に新しいアイコンが表示されたら、レンチのアイコンをクリックして、エンタープライズ・サーチの検索ポートレットを構成します。
  - g. ポートレット・パラメーターのリストで、以下のパラメーターを変更します。

### hostname

エンタープライズ・サーチ用の検索サーバーの完全修飾ホスト名を指定します。

### port

エンタープライズ・サーチ用の検索サーバーで WebSphere Application Server が使用するポート番号を指定します。デフォルト値は 80 です (SSL 通信のデフォルト値は 443 です)。

### username

検索サーバー上の WebSphere Application Server でグローバル・セキュリティが使用可能になっている場合、WebSphere Application Server ユーザー・レジストリーで有効なユーザー名を指定します。

### password

WebSphere Application Server ユーザー名を指定した場合に、対応するパスワードを指定します。

### protocol

WebSphere Portal と検索サーバー間の通信に使用されるプロトコルを指定します。デフォルトは HTTP です。SSL を使用する場合は、HTTPS を指定します。

**trustStore**

SSL を使用する場合、SSL 証明書ストアの完全修飾パス (ファイル名を含む) を指定します。

**trustPassword**

SSL を使用する場合、指定された `trustStore` ファイルのパスワードを指定します。

**ssoCookieName**

シングル・サインオン (SSO) トークン・ストリングを含む Cookie の名前を指定します。デフォルト値は `LtpaToken` です。

**proxyHost**

エンタープライズ・サーチの検索サーバーへのアクセスにプロキシ・サーバーが必要な場合、プロキシ・サーバーの完全修飾ホスト名を指定します。

**proxyPort**

プロキシ・サーバーを指定した場合、プロキシ・サーバーのポート番号を指定します。

**proxyUser**

プロキシ・サーバーが基本認証を必要とする場合、プロキシ・サーバーへのログインに使用するユーザー名を指定します。

**proxyPassword**

プロキシ・サーバーのユーザー名を指定した場合に、対応するパスワードを指定します。

- h. 「OK」をクリックして、変更内容を保存します。
2. ポートレットを設定後、それにアクセスするには、次のようにします。
  - a. Lotus Quickr サーバーにログインします。
  - b. ブラウザー・ウィンドウで、URL を次に変更します。 `http://host_name:port/lotus/myquickr/ESSearchPortlet`

## WebSphere Portal バージョン 6 からのエンタープライズ・サーチの除去

WebSphere Portal バージョン 6 システムからエンタープライズ・サーチを除去するには、`wp6_uninstall` スクリプトを使用します。

### このタスクについて

WebSphere Portal からエンタープライズ・サーチを除去すると、セットアップ処理の中でエンタープライズ・サーチ・ポートレットに指定したポートレット・パラメーターは保存されません。

スクリプトを開始すると、スクリプトは WebSphere Portal サーバーを停止します。エンタープライズ・サーチ・ソフトウェアが除去された後、スクリプトは WebSphere Portal サーバーを再始動します。

### 手順

WebSphere Portal バージョン 6 システムからエンタープライズ・サーチを除去するには、次のようにします。

コマンド・プロンプトで、**wp6\_uninstall.bat** コマンド (Windows の場合) または **wp6\_uninstall.sh** コマンド (AIX、Linux、または Solaris の場合) を実行します。以下の例では、読みやすくするためにパラメーターを別々の行に示しています。パラメーターはコマンドで次のように指定してください。

```
wp6_uninstall.bat
-WSPProfileDir "C:\Program Files\IBM\WebSphere\AppServer\profiles\wp_profile"
-WASDir "C:\Program Files\IBM\WebSphere\AppServer"
-WASUser wpsbind -WASPassword wpsbind
-WPSDir "C:\Program Files\IBM\WebSphere\PortalServer"
-WPSUser wpsadmin -WSPassword wpsadmin
-WPSHost "portalserver.ibm.com:9081"
```

#### **WSPProfileDir**

WebSphere Portal プロファイル・ディレクトリーの完全修飾パス。

#### **WASDir**

WebSphere Application Server のルート・ディレクトリーの完全修飾パス。  
AIX、Linux、および Solaris システムでのみ必要です。

#### **WASUser**

WebSphere Application Server 管理ユーザーのユーザー名。グローバル・セキュリティが WebSphere Application Server で使用可能になっている場合にのみ必要です。

#### **WASPassword**

WebSphere Application Server 管理ユーザーが指定された場合に、そのパスワード。

#### **WPSDir**

WebSphere Portal インストール・ディレクトリーの完全修飾パス。

#### **WPSUser**

WebSphere Portal 管理ユーザーのユーザー名。

#### **WSPassword**

指定された WebSphere Portal 管理ユーザーのパスワード。

#### **WPSHost**

WebSphere Portal サーバーのホスト名とポート番号。

---

## **WebSphere Portal クラスタ化システムとのエンタープライズ・サーチの統合**

WebSphere Portal バージョン 6 クラスタ化システムで稼働するようにエンタープライズ・サーチ・ポートレットをセットアップできます。

### **使用法のガイドライン**

- エンタープライズ・サーチのセットアップ・スクリプトを実行する前に、WebSphere Application Server Network Deployment Manager が稼働していて、クラスタ内の各ノードが実行していることを確認してください。
- スクリプトでは、エンタープライズ・サーチと WebSphere Portal 間のすべての統合ポイントをセットアップします。例えば、ポートレットを選択的にインストー



ルして、 WebSphere Portal および Web Content Management クローラーをサポートする EAR ファイルをインストールしないこととはできません。

- このスクリプトは、クラスター内の WebSphere Portal サーバーのすべてのインスタンスを停止して再始動します。ポータル・サービスを利用できないことでユーザー・コミュニティーに影響がないように、スクリプトは、通常の営業時間後に実行する必要がある場合があります。
- セットアップ・スクリプトの実行中にエラーが発生した場合は、セットアップ・スクリプトを再実行します。最初の試行で正常に完了したタスクはエラーを報告する可能性があります。セットアップ処理は続行し、残りのタスクを完了します。
- セットアップ・スクリプトの実行後にエンタープライズ・サーチ・ポートレット・ページに初めてアクセスした場合、このポートレット用の Java Server Pages (JSP ファイル) をシステムがコンパイルする必要があるため、ページの表示に時間がかかることがあります。

## WebSphere Portal クラスター化システムでのエンタープライズ・サーチのセットアップ

エンタープライズ・サーチ・システムを WebSphere Portal バージョン 6 クラスター化システムと統合するには、wp6\_cluster\_install スクリプトを使用します。

### このタスクについて

エンタープライズ・サーチと WebSphere Portal の統合に必要なファイルは、es.wp6.install.jar ファイルで提供されています。このファイルをアンパックすると、次のファイルが解凍されます。

- ESSearchPortlet.war
- ESPACServer.ear
- esapi.jar
- es.search.provider.jar
- es.security.jar
- インストールに必要なスクリプト、バッチ、XML、および JAACL ファイル

### 手順

エンタープライズ・サーチを WebSphere Portal バージョン 6 クラスター化システムと統合するには、次の手順で行います。

1. es.wp6.install.jar ファイルをエンタープライズ・サーチ・サーバーから WebSphere Portal がインストールされているクラスター内の各ノードにコピーし、Java **JAR** コマンド (または **TAR** コマンド) を使用してファイルをアンパックします。
2. オプション: WebSphere Portal Document Manager (PDM) との統合をサポートする場合は、クラスター内の各ノードで次のいずれかのステップを実行します。
  - WebSphere Information Integrator Content Edition のインストール・プログラムを実行し、コネクタのみのインストールを実行するオプションを選択して、PDM コネクタを WebSphere Portal サーバーにインストールします。

- WebSphere Information Integrator Content Edition のディレクトリー構造を WebSphere Portal サーバーに作成し、以下のファイルを既存の WebSphere Information Integrator Content Edition インストール済み環境から WebSphere Portal サーバーにコピーします。ここで、CE\_ROOT はルートの WebSphere Information Integrator Content Edition インストール・ディレクトリーを指定します。

```
CE_ROOT/lib/vbr.jar
CE_ROOT/ejb/vbr_pdm.jar
CE_ROOT/war/services.war
CE_ROOT/vbr_services.properties
```

3. コマンド・プロンプトで、WebSphere Portal がインストールされているクラスター内の各ノードで、**wp6\_cluster\_copyFiles.bat** コマンド (Windows) または **wp6\_cluster\_copyFiles.sh** コマンド (AIX、Linux、または Solaris) を実行します。以下の例では、読みやすくするためにオプションを別々の行に示しています。オプションはコマンドで指定する必要があります。

```
wp6_cluster_copyFiles.bat
-WPSDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥PortalServer"
-WSPProfileDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥AppServer¥¥profiles¥¥wp_profile"

wp6_cluster_copyFiles.sh
-WASDir /opt/IBM/WebSphere/AppServer
-WPSDir /opt/IBM/WebSphere/PortalServer
-WSPProfileDir /opt/IBM/WebSphere/AppServer/profiles/wp_profile
```

4. コマンド・プロンプトで、**wp6\_cluster\_install.bat** コマンド (Windows) または **wp6\_cluster\_install.sh** コマンド (AIX、Linux、または Solaris) を実行します。以下の例では、読みやすくするためにオプションを別々の行に示しています。オプションはコマンドで次のように指定してください。

```
wp6_cluster_install.bat
-WPSClusterName MyCluster
-WSPProfileDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥profiles¥¥wp_profile"
-WASDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥AppServer"
-WASUser wpsbind -WASPassword wpsbind
-WPSDir "C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥PortalServer"
-WPSUser wpsadmin -WSPPassword wpsadmin
-WPSHost "portalserver.ibm.com"
-webServerName webserver1
-webServerNodeName node1
-IICEDir "C:\\Program Files\\IBM\\Content Edition"
```

#### WPSClusterName

WebSphere Portal がインストールされているクラスターの名前。

#### WSPProfileDir

WebSphere Portal プロファイル・ディレクトリーの完全修飾パス。デフォルトのパスは、/usr/IBM/WebSphere/AppServer/profiles/wp\_profile (AIX システム)、/opt/IBM/WebSphere/AppServer/profiles/wp\_profile (Linux または Solaris システム)、および C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥profiles¥¥wp\_profile (Windows システム) です。

#### WASDir

WebSphere Application Server のルート・ディレクトリーの完全修飾パス。AIX、Linux、および Solaris システムでのみ必要です。デフォルトのルー

ト・ディレクトリー・パスは、/usr/IBM/WebSphere/AppServer (AIX システム)、/opt/IBM/WebSphere/AppServer (Linux または Solaris システム) です。

**WASUser**

WebSphere Application Server 管理ユーザーのユーザー名。グローバル・セキュリティが WebSphere Application Server で使用可能になっている場合にのみ必要です。

**WASPassword**

WebSphere Application Server 管理ユーザーが指定された場合に、そのパスワード。

**WPSDir**

WebSphere Portal インストール・ディレクトリーの完全修飾パス。

**WPSUser**

WebSphere Portal 管理ユーザーのユーザー名。

**WPSPassword**

指定された WebSphere Portal 管理ユーザーのパスワード。

**WPSHost**

WebSphere Portal サーバーのホスト名とポート番号。

**webServerName**

WebSphere Portal が属する Web サーバー定義の名前。

**webServerNodeName**

Web サーバー定義が属する WebSphere Application Server ノードの名前。

**IICEDir**

WebSphere Information Integrator Content Edition インストール・ディレクトリーの完全修飾パス。前に Portal Document Manager コネクタを WebSphere Portal サーバーにセットアップした場合にのみ必要です。

5. スクリプトが完了したら、Web ブラウザーを開いて、Network Deployment サーバー上の WebSphere 管理コンソールにログインします。アドレスは通常は、http://hostname:9060/ibm/console です。
6. 「サーバー」セクションを拡張して、「Web サーバー」を選択します。
7. Web サーバーの横の「選択」ボックスを選択して、「プラグインを生成」ボタンを選択します。
8. Web サーバーの横の「選択」ボックスを選択して、「プラグインを伝搬」ボタンを選択します。
9. 管理コンソールからログアウトします。
10. エンタープライズ・サーチ・ポートレットを更新して、検索サーバーを識別します。
  - a. Portal 管理者 ID およびパスワードで WebSphere Portal にログインします。
  - b. 左下隅にある「管理」をクリックします。
  - c. 左側のナビゲーション領域で「ポートレット管理」をクリックし、「ポートレット」をクリックします。
  - d. 「検索方法」オプションを「タイトルの内容」に変更します。

- e. 「検索」フィールドに「エンタープライズ・サーチ」と入力し、「検索」ボタンをクリックします。
- f. 右に新しいアイコンが表示されたら、レンチのアイコンをクリックして、エンタープライズ・サーチの検索ポートレットを構成します。
- g. ポートレット・パラメーターのリストで、以下のパラメーターを変更します。

**hostname**

エンタープライズ・サーチ用の検索サーバーの完全修飾ホスト名を指定します。

**port** エンタープライズ・サーチ用の検索サーバーで WebSphere Application Server が使用するポート番号を指定します。デフォルト値は 80 です (SSL 通信のデフォルト値は 443 です)。

**username**

検索サーバー上の WebSphere Application Server でグローバル・セキュリティが使用可能になっている場合、WebSphere Application Server ユーザー・レジストリーで有効なユーザー名を指定します。

**password**

WebSphere Application Server ユーザー名を指定した場合に、対応するパスワードを指定します。

**protocol**

WebSphere Portal と検索サーバー間の通信に使用されるプロトコルを指定します。デフォルトは HTTP です。SSL を使用する場合は、HTTPS を指定します。

**trustStore**

SSL を使用する場合、SSL 証明書ストアの完全修飾パス (ファイル名を含む) を指定します。

**trustPassword**

SSL を使用する場合、指定された trustStore ファイルのパスワードを指定します。

**ssoCookieName**

シングル・サインオン (SSO) トークン・ストリングを含む Cookie の名前を指定します。デフォルト値は LtpaToken です。

**proxyHost**

エンタープライズ・サーチの検索サーバーへのアクセスにプロキシ・サーバーが必要な場合、プロキシ・サーバーの完全修飾ホスト名を指定します。

**proxyPort**

プロキシ・サーバーを指定した場合、プロキシ・サーバーのポート番号を指定します。

**proxyUser**

プロキシ・サーバーが基本認証を必要とする場合、プロキシ・サーバーへのログインに使用するユーザー名を指定します。

### proxyPassword

プロキシ・サーバーのユーザー名を指定した場合に、対応するパスワードを指定します。

- h. 「OK」をクリックして、変更内容を保存します。

## WebSphere Portal クラスタ化システムからのエンタープライズ・サーチの除去

WebSphere Portal バージョン 6 クラスタ化システムからエンタープライズ・サーチを除去するには、wp6\_cluster\_uninstall スクリプトを使用します。

### このタスクについて

WebSphere Portal からエンタープライズ・サーチを除去すると、セットアップ処理の中でエンタープライズ・サーチ・ポートレットに指定したポートレット・パラメーターは保存されません。

スクリプトを開始すると、スクリプトは WebSphere Portal サーバーを停止します。エンタープライズ・サーチ・ソフトウェアが除去された後、スクリプトは WebSphere Portal サーバーを再始動します。

### 手順

WebSphere Portal バージョン 6 クラスタ化システムからエンタープライズ・サーチを除去するには、次の手順で行います。

1. コマンド・プロンプトで、**wp6\_cluster\_uninstall.bat** コマンド (Windows) または **wp6\_cluster\_uninstall.sh** コマンド (AIX、Linux、または Solaris) をクラスター内のいずれかのノードで実行します。以下の例では、読みやすくするためにパラメーターを別々の行に示しています。パラメーターはコマンドで次のように指定してください。

```
wp6_cluster_uninstall.bat
-WPSClusterName MyCluster
-WSPProfileDir "C:\\Program Files\\IBM\\WebSphere\\profiles\\wp_profile"
-WASDir "C:\\Program Files\\IBM\\WebSphere\\AppServer"
-WASUser wpsbind
-WASPassword wpsbind
-WPSDir "C:\\Program Files\\IBM\\WebSphere\\PortalServer"
-WPSUser wpsadmin
-WPSPassword wpsadmin
-WPSHost "portalserver.ibm.com"
-webServerName webserver1
-webServerNodeName node1
```

### WPSClusterName

WebSphere Portal がインストールされているクラスターの名前。

### WSPProfileDir

WebSphere Portal プロファイル・ディレクトリーの完全修飾パス。デフォルトのパスは、/usr/IBM/WebSphere/AppServer/profiles/wp\_profile (AIX システム)、/opt/IBM/WebSphere/AppServer/profiles/wp\_profile (Linux または Solaris システム)、および C:¥¥Program Files¥¥IBM¥¥WebSphere¥¥profiles¥¥wp\_profile (Windows システム) です。

### WASDir

WebSphere Application Server のルート・ディレクトリーの完全修飾パス。

AIX、Linux、および Solaris システムでのみ必要です。デフォルトのルート・ディレクトリー・パスは、/usr/IBM/WebSphere/AppServer (AIX システム)、/opt/IBM/WebSphere/AppServer (Linux または Solaris システム) です。

**WASUser**

WebSphere Application Server 管理ユーザーのユーザー名。グローバル・セキュリティが WebSphere Application Server で使用可能になっている場合のみ必要です。

**WASPassword**

WebSphere Application Server 管理ユーザーが指定された場合に、そのパスワード。

**WPSDir**

WebSphere Portal インストール・ディレクトリーの完全修飾パス。

**WPSUser**

WebSphere Portal 管理ユーザーのユーザー名。

**WPSPassword**

指定された WebSphere Portal 管理ユーザーのパスワード。

**WPSHost**

WebSphere Portal サーバーのホスト名とポート番号。

**webServerName**

WebSphere Portal が属する Web サーバー定義の名前。

**webServerNodeName**

Web サーバー定義が属する WebSphere Application Server ノードの名前。

2. スクリプトが完了したら、Web ブラウザーを開いて、Network Deployment サーバー上の WebSphere 管理コンソールにログインします。アドレスは通常は、`http://hostname:9060/ibm/console` です。
3. 「サーバー」セクションを拡張して、「**Web サーバー**」を選択します。
4. Web サーバーの横の「**選択**」ボックスを選択して、「**プラグインを生成**」ボタンを選択します。
5. Web サーバーの横の「**選択**」ボックスを選択して、「**プラグインを伝搬**」ボタンを選択します。
6. 管理コンソールからログアウトします。

---

## WebSphere Portal からエンタープライズ・サーチへの移行

エンタープライズ・サーチでは、IBM WebSphere Portal からエンタープライズ・サーチにコレクションとルール・ベースの分類構造を移行するときに使用する移行ウィザードが用意されています。

エンタープライズ・サーチでは、分類構造はカテゴリ・ツリー と言います。分類構造を移行した後は、エンタープライズ・サーチ管理コンソールを使用して、カテゴリ・ツリーとカテゴリ規則を編集します。コレクションを移行した後は、管理コンソールを使用してコレクションを管理します。

分類構造およびコレクションを移行する場合、エンタープライズ・サーチ索引サーバーで移行ウィザードを実行します。

---

## WebSphere Portal からのコレクションの移行

WebSphere Portal からエンタープライズ・サーチにコレクションとルール・ベースの分類構造を移行するには、WebSphere Portal にコレクションを準備し、移行ウィザードを使用してそれらを移行します。

### 始める前に

分類構造とコレクションを移行する予定の場合は、この手順でコレクションを移行する前に、分類構造ファイルを移行してください。この方法により、移行されたカテゴリ化規則が、移行されたコレクションと動作するようになります。

### 手順

WebSphere Portal からエンタープライズ・サーチにコレクションを移行する (また、オプションで分類構造を移行する) には、次のようにします。

1. WebSphere ポータル検索エンジンで、移行するコレクションのすべてのクローラー・プロセスを停止し、保留中のすべての文書を承認または拒否します。(エンタープライズ・サーチでは、保留中の文書という概念はサポートしません。)
2. 移行する各コレクションについて、ポータル検索エンジン・ポートレットを使用してその設定を XML ファイルにエクスポートします。
3. エンタープライズ・サーチ索引サーバーが別のサーバーにインストールされている場合は、エクスポートした XML ファイルを索引サーバーにコピーします。
4. エンタープライズ・サーチ索引サーバーで、エンタープライズ・サーチ管理者としてログインします。このユーザー ID は、OmniFind Enterprise Edition のインストール時に指定されたものです。
5. エンタープライズ・サーチのインストール・ディレクトリーに移動します。

```
UNIX: cd $ES_INSTALL_ROOT/bin
```

```
Windows: cd %ES_INSTALL_ROOT%\bin
```

6. セキュリティーを有効にしたコレクションを移行するには、次のコマンドを入力して移行ウィザードを開始し、「次へ」をクリックします。

UNIX: ./eswpsmigrate.sh

Windows: eswpsmigrate.bat

7. 移行するコレクションに対してコレクション・レベルのセキュリティーを無効にするには、次のコマンドを入力して移行ウィザードを開始し、「次へ」をクリックします。

UNIX: ./eswpsmigrate.sh disable.security

Windows: eswpsmigrate.bat disable.security

8. 「WebSphere Portal のポータル検索エンジンから検索設定を移行する」を選択し、「次へ」をクリックします。
9. エクスポートされたポータル検索エンジン構成ファイルがあるディレクトリーを参照し、移行したいファイルを選択して「次へ」をクリックします。選択された構成ファイルが分析、検証されます。
10. コレクションごとに次の情報を入力し、「次へ」をクリックすると、コレクションのエンタープライズ・サーチへの移行が開始されます。
  - エンタープライズ・サーチで使用するコレクションの名前。
  - コレクションで文書の重要度を判断するための基準。この静的ランキング要素は、なし、文書日付を基準にする、Web 文書への他の Web 文書からのリンク数を基準にする、のいずれかにすることができます。
  - このコレクションに使用するカテゴリー化のタイプ。なしを指定すると、分類構造の情報がエンタープライズ・サーチに移行されません。ルール・ベースのカテゴリーを選択した場合、分類構造がコレクションと共にエンタープライズ・サーチに移行されます。

移行時にエラーが発生した場合、移行ウィザードがインストールされているディレクトリーにある MigrationWizard.log ファイルを参照してください。

移行されたコレクションに追加設定を構成するには、エンタープライズ・サーチ管理コンソールを使用することができます。

**要件:** 移行したコレクションの Web クローラー・プロパティーを構成するとき、クローラーに関するコメントを受信する E メール・アドレスとユーザー・エージェント名を指定する必要があります (Web クローラー・プロパティーの構成中に不明な点があったら、「ヘルプ」をクリックしてください)。

11. エンタープライズ・サーチ管理コンソールから、移行済みコレクションに対してクロール、構文解析、および索引作成プロセスを開始します。
12. 移行済みコレクションがエンタープライズ・サーチで検索できるとわかったら、ポータル検索エンジンにある元のコレクションを削除します。
13. オプション: WebSphere Portal 管理者が、WebSphere Portal のポータルから移行したコレクションをユーザーが検索できるようにしたい場合は、以下のステップを実行します。
  - a. エンタープライズ・サーチ・ポートレットを WebSphere Portal システムにデプロイします。



WebSphere Portal サーバー・クラスターでは、この作業は WebSphere Application Server Deployment Manager がインストールされているサーバーに対して行う必要があります。 WebSphere Portal サーバー・クラスターの他のサーバーには、この Deployment Manager によってエンタープライズ・サーチ・ポートレットが配布されます。

- b. エンタープライズ・サーチ・ポートレットを該当するポータル・ページに追加します。

WebSphere Portal では、検索ポートレットのアクセス制御が、特定ページおよびポートレットに対するアクセス可能度によってモデル化されています。コレクション設定は移行されますが、ポートレットは WebSphere Portal サーバー管理者によって手動で配置する必要があります。

---

## 移行されたコレクションの設定

IBM WebSphere Portal からコレクションを移行すると、移行ウィザードにより、コレクションおよびクローラーのデフォルト設定が作成されます。

ポータル検索エンジン・コレクションとエンタープライズ・サーチ・コレクションに同じ設定がある場合、ウィザードでは、コレクションのエンタープライズ・サーチへの移行時にポータル検索エンジンの設定を使用します。エンタープライズ・サーチだけにある設定については、コレクションの移行時に指定した設定、または、エンタープライズ・サーチのコレクションに対するデフォルト設定がウィザードによって使用されます。

### ポータル検索エンジンとエンタープライズ・サーチの両方にある設定

移行ウィザードは、移行対象の各コレクションについて、以下の設定を移行します。

- ポータル検索エンジン・コレクション内のポータル検索エンジン・サイト
- コレクション言語
- 分類構造 (カテゴリー・ツリー)、およびエンタープライズ・サーチ・コレクションでルール・ベースのカテゴリーを使用する場合はルール・ベース・カテゴリーの規則

コレクション内の各ポータル検索エンジン・サイトは、エンタープライズ・サーチ Web クローラーに統合されます。移行ウィザードは、以下のクローラー設定を移行します。

- 開始 URL
- 並行クロール・プロセスの数
- クロールの深さ
- 文書検索のタイムアウト値 (秒)
- デフォルトの文字セット
- Web サイトをクロールするための規則 (組み込む、または除外する)

## エンタープライズ・サーチだけにある設定

コレクションを移行する場合、そのコレクションに関する情報を指定します。移行ウィザードでは、それらの設定を移行し、コレクションに対するエンタープライズ・サーチのデフォルト設定を使用して、移行した各コレクションを構成します。

エンタープライズ・サーチ管理コンソールを使用すれば、このコレクションおよび Web クローラー構成を変更することができます。括弧内に示された値は、移行されたデータに対するデフォルト設定です。

- コレクション名
- 文書の静的ランキング計画
- 使用されるカテゴリー化のタイプ (ルール・ベースまたはなし)
- 検索キャッシュを使用するかどうか、および検索キャッシュが保持できる照会応答数 (使用する、5000)
- 検索応答時間をモニターし、限度を超えたらアラートを発行するかどうか (モニターする、限度 5 秒)
- アクセス制御を使用するかどうか (しない)
- 差分索引の作成スケジュール
- 主索引の作成スケジュール
- ログの詳細レベル (すべてのメッセージ)

また、移行ウィザードでは、各 Web クローラーに対して以下の設定を作成します。

- クローラー名
- クローラーの説明
- 最大ページ長
- 文書セキュリティ設定
- データ・ソース・タイプで該当する場合、クローラーする必要がある文書の MIME タイプ

新規に移行された Web クローラーを開始する前に、クローラー・プロパティーとクローラー・スペース設定を全部よく検討し、必要な値がすべて指定されていることを確認してください (必要フィールドは、赤いアスタリスクでマークされています)。特に、クローラーに関するコメントを受信する E メール・アドレスとクローラーのユーザー・エージェント名を指定していることを確認してください。 Web クローラー・プロパティーの構成中に不明な点があったら、「ヘルプ」をクリックしてください。

---

## 移行ウィザードのログ・ファイル

移行ウィザードでは、すべてのメッセージを、移行ウィザードがインストールされているディレクトリーにある WpsMigratorLog.log ファイルに書き込みます。

移行済みの各コレクションについて、WpsMigratorLog.log ログ・ファイルには、WebSphere ポータル検索エンジンから読み込まれたすべての設定値、およびそれらの設定値がエンタープライズ・サーチ・コレクションにインポートされた場所が格納されています。

---

## エンタープライズ・サーチ・コマンド、戻りコード、およびセッション ID

コマンドを使用して、問題の診断、システムの別の部分の状況の判別、セッションの開始および停止、あるいはシステムの始動および停止を行うことができます。

複数のサーバーがインストールされている環境では、システム内のどのサーバーからでもコマンドを実行することができます。しかしながら、索引サーバーからコマンドを実行することをお勧めします。索引サーバー (コントローラー・サーバー) は、システム内にある他のすべてのサーバーの情報にアクセスすることができます。

大部分のコマンドは、以下のようなフォーマットです。

```
esadmin command_name arguments  
esadmin session_ID action -option
```

すべてのコマンドについて詳しく知りたい場合は、`esadmin help` と入力してください。特定のコマンドについて詳しく知りたい場合は、`esadmin action help` と入力してください。

### エンタープライズ・サーチ esadmin コマンド

次のコマンドを 1 行に入力します。

表 9. エンタープライズ・サーチ **esadmin** コマンド

コマンド	説明
<code>esadmin system startall</code>	<p>すべてのエンタープライズ・サーチ・サーバー上のエンタープライズ・サーチ・コンポーネントを開始します。それには、Web サーバー、ESSearchServer アプリケーション、検索サーバー上のインフォメーション・センター、クローラー・サーバー上のクローラー・セッション、索引サーバー上の索引セッションが含まれます。ローカル・サーバー上でのみ、共通通信層 (CCL) を開始します。CCL を更新するには、リモートの各エンタープライズ・サーチ・サーバーで、CCL を手動で一度停止してから再始動する必要があります。</p> <p><b>コマンドの例:</b></p> <pre>esadmin system startall</pre>

表9. エンタープライズ・サーチ **esadmin** コマンド (続き)

コマンド	説明
esadmin system stopall	<p>すべてのエンタープライズ・サーチ・サーバー上のエンタープライズ・サーチ・コンポーネントを停止します。それには、インフォメーション・センター、ESSearchServer アプリケーション、検索サーバー上の Web サーバー、クローラー・サーバー上のクローラー・セッション、索引サーバー上の索引セッションが含まれます。ローカル・サーバー上でのみ、CCL を停止します。CCL を更新するには、リモートの各エンタープライズ・サーチ・サーバーで、CCL を手動で一度停止してから再始動する必要があります。</p> <p><b>コマンドの例:</b></p> <pre>esadmin system stopall</pre>
esadmin system checkall	<p>すべてのエンタープライズ・サーチ・サーバー上にあるすべてのエンタープライズ・サーチ・コンポーネントの状況を検査します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin system checkall</pre>
esadmin crawler_session_id start	<p>クローラー・セッションを開始します。このコマンドでは、ロール・アクティビティーは開始されません。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col1.WEB1.esadmin start</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5310I WEBCrawler1 (sid: col1.WEB1.esadmin) は実行されていません。 FFQC5314I 次の結果が発生しました: 0</pre>
esadmin crawler_session_id startCrawl	<p>クローリングを開始します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col3.DB21.esadmin startCrawl</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) は既に実行中です。PID: 23650 FFQC5314I 次の結果が発生しました: 0</pre>
esadmin crawler_session_id pause	<p>クローリングを一時停止します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col3.DB21.esadmin pause</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) は既に実行中です。PID: 23650 FFQC5314I 次の結果が発生しました: 0</pre>

表 9. エンタープライズ・サーチ **esadmin** コマンド (続き)

コマンド	説明
<code>esadmin crawler_session_id resume</code>	<p>クローラを再開します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col3.DB21.esadmin resume</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) は既に実行中です。PID: 23650 FFQC5314I 次の結果が発生しました: 0</pre>
<code>esadmin crawler_session_id stopCrawl</code>	<p>クローラを停止します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col3.DB21.esadmin stopCrawl</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) は既に実行中です。PID: 23650 FFQC5314I 次の結果が発生しました: 0</pre>
<code>esadmin crawler_session_id stop</code>	<p>クローラー・セッションを停止します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col3.DB21.esadmin stop</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) は既に実行中です。PID: 23650 FFQC5314I 次の結果が発生しました: 0</pre>
<code>esadmin crawler_session_id getCrawlerStatus</code>	<p>クローラーの状況を取得します。戻される情報は、そのクローラーが Web クローラーであるか、あるいはその他すべてのデータ・ソース用のクローラーであるかによって決まります。</p> <p><b>Web クローラーの場合の例:</b></p> <pre>esadmin col1.WEB1.esadmin getCrawlerStatus</pre> <p><b>Web クローラーの場合の戻りコードおよびメッセージの例:</b></p> <pre>FFQC5303I WebCrawler1 (sid: col1.WEB1.esadmin) は既に実行中です。PID: 23650</pre> <p><b>Web 以外のクローラーの場合の例:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlerStatus</pre> <p><b>Web 以外のクローラーの場合の戻りコードおよびメッセージの例:</b></p> <pre>FFQC5303I db2crawler (sid: db2col.DB2_96945) は既に実行中です。PID: 5936</pre> <p>戻される状況メッセージについて詳しくは、414 ページの『状況コマンドの詳細情報』を参照してください。</p>

表 9. エンタープライズ・サーチ **esadmin** コマンド (続き)

コマンド	説明
<pre>esadmin dscrawler_session_id getCrawlSpaceStatus</pre>	<p>Web 以外のクローラーのクロール・スペース状況概要を取得します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatus</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) は既に実行中です。PID: 23650</pre> <p>Web クローラーのクロール・スペース状況概要を取得します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col11.WEB1.esadmin getCrawlStatus</pre> <p>戻される状況メッセージについて詳しくは、414 ページの『状況コマンドの詳細情報』を参照してください。</p>
<pre>esadmin dscrawler_session_id getCrawlSpaceStatusDetail -ts target_server_id</pre> <pre>esadmin webcrawler_session_id getCrawlDetailsPerSite -url string -selections num -threshold num</pre>	<p>Web 以外のクローラーのクロール・スペース状況詳細を取得します。ターゲット・サーバーのオプションを指定しないと、すべてのターゲット・サーバーのデータが戻されます。例えば、DB2 クローラーが FOUNTAIN および SAMPLE データベースをクロールする場合、ターゲット・サーバー・オプションを指定しないと、FOUNTAIN および SAMPLE データベースのすべての表の状況が戻されます。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col3.DB21.esadmin getCrawlSpaceStatusDetail -ts FOUNTAIN</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I DB2Crawler1 (sid: col3.DB21.esadmin) は既に実行中です。PID: 23650</pre> <p>Web クローラーのクロール・スペース状況詳細を取得します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin col11.WEB1.esadmin getCrawlDetailsPerSite</pre> <p>戻される状況メッセージについて詳しくは、414 ページの『状況コマンドの詳細情報』を参照してください。</p>
<pre>esadmin monitor getCollectionParserMonitorStatus -cid collection_ID</pre>	<p>パーサーの状況を取得します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin monitor getCollectionParserMonitorStatus -cid col1</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Monitor (node1) (sid: monitor) は既に実行中です。PID: 12543</pre> <p>戻される状況メッセージについて詳しくは、414 ページの『状況コマンドの詳細情報』を参照してください。</p>

表 9. エンタープライズ・サーチ **esadmin** コマンド (続き)

コマンド	説明
<code>esadmin startMain -cid コレクション ID</code>	<p>主索引の作成を開始します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin startMain -cid coll</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) は既に実行中です。PID: 25917 FFQC5314I 次の結果が発生しました: 1117671147056</pre>
<code>esadmin startDelta -cid コレクション ID</code>	<p>差分索引の作成を開始します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin startDelta -cid coll</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) は既に実行中です。PID: 4548 FFQC5314I 次の結果が発生しました: 1117670603408</pre>
<code>esadmin controller startIndexBuild -cid コレクション ID -buildType type -detectChanges</code>	<p>主または差分索引作成を開始して、索引に適用する必要がある変更が検出されない場合にのみ、作成を行うことを指定します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin controller startIndexBuild -cid col_1 -buildType main -detectChanges</pre>
<code>esadmin monitor getCollectionIndexMonitorStatus -cid コレクション ID -buildType [main   delta] -numrecords lastNrecords</code>	<p>主索引または差分索引の作成の状況を取得します。オプション <code>numrecords</code> を指定すると、最後の <i>N</i> 件の索引作成状況レコードが表示されます。 <code>numrecords</code> を省略した場合は、最後の 20 個の索引作成状況が戻されます。</p> <p><b>コマンドの例:</b></p> <pre>esadmin monitor getCollectionIndexMonitorStatus -cid coll -buildType main -numrecords 4</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Monitor (node1) (sid: monitor) は既に実行中です。PID: 12649</pre> <p>戻される状況メッセージについて詳しくは、414 ページの『状況コマンドの詳細情報』を参照してください。</p>
<code>esadmin startSearch -cid コレクション ID</code>	<p>検索サーバー・プロセスを開始します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin startSearch -cid coll</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) は既に実行中です。PID: 25917 FFQC5314I 次の結果が発生しました: 0</pre>

表 9. エンタープライズ・サーチ **esadmin** コマンド (続き)

コマンド	説明
<code>esadmin stopSearch -cid コレクション ID</code>	<p>検索サーバー・プロセスを停止します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin stopSearch -cid coll</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Controller (node1) (sid: controller) は既に実行中です。PID: 15292 FFQC5314I 次の結果が発生しました: 0</pre>
<pre>esadmin monitor getCollectionSearchMonitorStatus -cid コレクション ID  esadmin searchmanager session_id getStatus -cid コレクション ID</pre>	<p>検索サーバーの状況を取得します。</p> <p><b>コマンドの例:</b></p> <pre>esadmin monitor getCollectionSearchMonitorStatus -cid coll</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Monitor (node1) (sid: monitor) は既に実行中です。PID: 12649</pre> <p>所定の検索サーバー上のコレクションに関する詳細な検索索引状況情報を戻します。検索サーバーごとに、検索マネージャー・セッションが 1 つあります。特定の検索サーバー上の検索索引をモニターおよび操作するのは、各検索マネージャー・セッションの役目です。</p> <p><b>コマンドの例:</b></p> <pre>esadmin searchmanager.node1 getStatus -cid coll</pre> <p><b>メッセージおよび戻りコードの例:</b></p> <pre>FFQC5303I Search Manager (node1) (sid: searchmanager.node1) は既に実行中です。PID: 15711 FFQC5314I 次の結果が発生しました: PID=18390 CacheHits=3 QueryRate=1 Port=44008 SessionId=coll.runtime.node1 CacheHitRate=0.333 ResponseTime=70 Status=1 SessionName=coll.runtime.node1.1</pre> <p>戻される状況メッセージについて詳しくは、『状況コマンドの詳細情報』を参照してください。</p>

## 状況コマンドの詳細情報

コマンドの中には、広範囲にわたる情報を戻せるものがあります。このセクションでは、クローラー状況およびクロール・スペース状況に関して戻ることができる情報について説明します。409 ページの『エンタープライズ・サーチ **esadmin** コマンド』の表で、各 **esadmin** コマンドから戻される情報を示しています。このセクションでは、以下のコマンドから戻される情報について説明します。



- Web クローラー状況
- Web 以外のクローラー状況
- Web クローラーのクロール・スペース状況
- Web 以外のクローラーのクロール・スペース状況
- Web クローラーの詳細なクロール・スペース状況
- Web 以外のクローラーの詳細なクロール・スペース状況
- パーサー状況
- 索引作成状況
- 検索サーバー状況
- 詳細な検索サーバー状況

**Web クローラー状況:** Web クローラー状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。 Web クローラー状況コマンドによって、以下の情報を得ることができます。

```
FFQC5314I 次の結果が発生しました: <?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>
```

以下の表で、Web クローラー状況コマンドによって戻される各 XML エlement およびその属性について説明します。

表 10. Web クローラー状況情報

エレメント	属性	説明
CrawlerStatus	<ul style="list-style-type: none"> <li>• CrawlerThreadStateDist</li> <li>• ActiveBucketList</li> <li>• CrawlRate</li> <li>• RecentlyCrawledURLList</li> <li>• NumURLsThisSession</li> </ul>	クローラー状況。

表 10. Web クローラー状況情報 (続き)

エレメント	属性	説明
CrawlerRunLevel Value	<ul style="list-style-type: none"> <li>・ スtring (英語) 「Not started」: クローラー・セッションが存在します。しかし、文書処理するための開始メッセージはまだ受け取っていません。</li> <li>・ 『Started』: クローラーは開始しています。</li> <li>・ 『Running』: クローラーは初期化および開始処理を終えて、アクティブな状態でクロールしています。</li> <li>・ 『Paused』: クローラーは、アクティブ・クロールを中断し、しかし終了はしないように指示されました。</li> <li>・ 『Stopping』: クローラーはストップ信号を受け取り、停止しようとしています。</li> <li>・ 『Error』: クローラーはリカバリー不能な状態に陥っており、クロールを再開するにはクローラーを停止および再開する必要があります。</li> </ul>	クローラーが何をしているかという情報。
CrawlerThreadState State	String (英語)	クローラー・スレッドのアクティビティ。このフィールドは、スレッド (1 つ以上) が何をしているかということを示します。
ActiveBucket	<ul style="list-style-type: none"> <li>・ URL: String (URL 仕様) URL をクロールしているプロトコル、ホスト、およびポート。</li> <li>・ NumActURLs: 整数 (正数) クロール用にバケットが使用可能になった (アクティブになった) ときの、そのバケット内の URL の数。</li> <li>・ NumProcURLs: 整数 (負数以外) それまでに処理された (クロールまたは拒否された) バケットの URL の数。</li> <li>・ TimeRem: 整数 バケットがタイムアウトになるまでの残りの秒数。</li> <li>・ Duration: 整数 (負数以外) バケットがアクティブになってからの秒数。</li> </ul>	指定された Web サイトの現在のアクティビティ。
CrawlRate	値: 整数 (負数以外)  1 秒あたりにクロールされたページ (全バケットを結合したもの)。	クローラーのスループット測定。

表 10. Web クローラー状況情報 (続き)

エレメント	属性	説明
RecentlyCrawledURL	URL: ストリング (URL 仕様)	最近クロールされたページ。
	クロールされたプロトコル、ホスト、ポート、およびファイルを示すストリング。	
NumURLsThisSession	値: 整数 (負数以外)	クローラー (プロセス) のこのインスタンスがクロールを開始してからクロールされた URL の数。

**Web 以外のクローラー状況:** Web 以外のクローラー状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。 **getCrawlerStatus** コマンドによって、Web クローラー以外のクローラーについて次の情報を戻すことができます。

```
FFQC5314I 次の結果が発生しました: <?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>
```

以下の表で、Web クローラー以外の各エンタープライズ・サーチ・クローラーでの、XML エレメントおよび属性について説明します。この情報は、クローラー状況コマンドで戻されるものです。

表 11. NNTP、DB2、JDBC データベース、および Lotus Notes クローラーのクローラー状況情報

エレメントおよび属性の名前	DB2 および JDBC データ		
	NNTP クローラー	ベース・クローラー	Lotus Notes クローラー
Status	状況 (0、1、2、-1)	状況 (0、1、2、-1)	状況 (0、1、2、-1)
StatusMessage	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー
NumberOfServers	クロール・スペース内の NNTP サーバーの数。	クロール・スペース内のデータベースの数。	クロール・スペース内のデータベースの数。
NumberOfCompletedServers	クロールされた NNTP サーバーの数。	クロールされたデータベースの数。	クロールされたデータベースの数。
NumberOfTargets	クロール・スペース内の新規グループの数。	クロール・スペース内のデータベースの数。	クロール・スペース内のビューおよびフォルダーの数。
NumberOfCompletedTargets	クロールされた新規グループの数。	クロールされた表の数。	クロールされたビューおよびフォルダーの数。
NumberOfCompletedRecords	クロールされた項目の数。	クロールされたレコードの数。	クロールされた文書の数。
RunningThreads	クローラー・スレッドの数。	クローラー・スレッドの数。	クローラー・スレッドの数。

表 12. Exchange Server、DB2 Content Manager、および Content Edition クローラーのクローラー状況情報

エレメントおよび属性の名前	Exchange Server クローラ ー	DB2 Content Manager ク ローラー	Content Edition クローラ ー
Status	状況 (0、1、2、-1)	状況 (0、1、2、-1)	状況 (0、1、2、-1)
StatusMessage	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー
NumberOfServers	クロール・スペース内の Exchange Server サーバーの数。	クロール・スペース内の Content Manager サーバーの数。	クロール・スペース内のリポジトリの数。
NumberOfCompletedServers	クロールされた Exchange Server サーバーの数。	クロールされた Content Manager サーバーの数。	クロールされたりリポジトリの数。
NumberOfTargets	クロール・スペース内のサブフォルダーの数。	クロール・スペース内の項目タイプの数。	クロール・スペース内のクラスの数。
NumberOfCompletedTargets	クロールされたサブフォルダーの数。	クロールされた項目タイプの数。	クロールされた項目クラスの数。
NumberOfCompletedRecords	クロールされた文書の数。	クロールされた文書の数。	クロールされた文書の数。
RunningThreads	クローラー・スレッドの数。	クローラー・スレッドの数。	クローラー・スレッドの数。

表 13. QuickPlace、Lotus Domino Document Manager、UNIX ファイル・システム、および Windows ファイル・システム・クローラーのクローラー状況情報

エレメントおよび属性の名前	QuickPlace クローラー	Lotus Domino Document Manager クローラー	UNIX および Windows ファイル・システム・クローラー
Status	状況 (0、1、2、-1)	状況 (0、1、2、-1)	状況 (0、1、2、-1)
StatusMessage	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー
NumberOfServers	クロール・スペース内のプレースの数。	クロール・スペース内のライブラリーの数。	固定値 1。
NumberOfCompletedServers	クロールされたプレースの数。	クロールされたライブラリーの数。	すべてのサブディレクトリーがクロールされた場合、0 または 1。
NumberOfTargets	クロール・スペース内のプレース・データベースおよびルーム・データベースの数。	クロール・スペース内のキャビネットの数。	クロール・スペース内のサブディレクトリーの数。
NumberOfCompletedTargets	クロールされたプレース・データベースおよびルーム・データベースの数。	クロールされたキャビネットの数。	クロールされたサブディレクトリーの数。
NumberOfCompletedRecords	クロールされた文書の数。	クロールされた文書の数。	クロールされたファイルの数。
RunningThreads	クローラー・スレッドの数。	クローラー・スレッドの数。	クローラー・スレッドの数。

表 14. WebSphere Portal および Web Content Management クローラーのクローラー状況情報

エレメントおよび属性の名前	WebSphere Portal クローラー	Web Content Management クローラー
Status	状況 (0、1、2、-1)	状況 (0、1、2、-1)
StatusMessage	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー	状況: 0 - アイドル、1 - 実行中、2 - 休止、-1 - エラー
NumberOfServers	クロール・スペース内のサーバーの数。	クロール・スペース内のサイトの数。
NumberOfCompletedServers	クロールされたサーバーの数。	クロールされたサイトの数。
NumberOfTargets	クロール・スペース内のサーバーの数。	クロール・スペース内のサイトの数。
NumberOfCompletedTargets	クロールされたサーバーの数。	クロールされたサイトの数。
NumberOfCompletedRecords	クロールされた文書の数。	クロールされた文書の数。
RunningThreads	クローラー・スレッドの数。	クローラー・スレッドの数。

**Web クローラーのクロール・スペース状況:** Web クローラーのクロール・スペース状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。Web クロール・スペース状況コマンドによって、以下の情報を得ることができます。

表 15. Web クローラーのクロール・スペース状況コマンドの選択マスク値

マスク・ビット	選択
1	ロー・データ・ストア内のページの数。
2	検出されたサイトの数。
4	DNS のあるサイトの数。
8	DNS のないサイトの数。
16	検出された URL の数。
32	保存された固有ページの数。
64	クロールされた URL の数。
128	クロールされていない URL の数。
256	期限切れの URL の数。
512	HTTP 状況コードの分布。

すべての値が、現行の内部データベースを使用するすべてのセッションでの累積合計を表します。

```
<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"/>
  </HTTPCodeDist>
</CrawlStatus>
```

```

    <HTTPCode Code="404" Count ="10000"/>
    <HTTPCode Code="780" Count="20000"/>
  </HTTPCode Code="780" Count="20000">
<?CrawlStatus>

```

戻りデータには、以下のエレメントのいくつかまたはすべてが含まれます (無いこともあります)。

表 16. Web クローラーのクロール・スペース状況情報

エレメント	属性	説明
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>	クロールの累積状態に関して迅速に取得できる情報 (すべてのセッション)。
NumPagesInRDS	値: 負でない整数  ロー・データ・ストア (RDS) のステージング・エリア内に現在入っているページの数 (このクローラーのみ)。	ロー・データ・ストア (RDS) がどのくらいフルになっているか (このクローラーのコンテンツリビューションのみ)。
NumSitesDiscovered	値: 負でない整数  クロールによって検出されたホストの数 (または、シード)。	クロールすべきドメインの、クローラーの適用範囲の指標 (ホスト・カウント)。
NumSitesWithDNS	値: 負でない整数  IP アドレスに関連付けられたホストの数 (バックグラウンドのクローラーによって解決されたもの)。	URL 内で DNS 名によって検出されるホストの IP アドレスを、クローラーがどのくらい効率的に取得できるかの指標。
NumSitesWithoutDNS	値: 負でない整数  IP アドレスに関連付けられていないホストの数 (バックグラウンドのクローラーによって解決されたもの)。	URL 内で DNS 名によって検出されるホストの IP アドレスを、クローラーがどのくらい効率的に取得できるかの指標。
NumURLsDiscovered	値: 負でない整数  クローラーによって訪問された固有の URL の数。	クロールすべきドメインの、クローラーの適用範囲の指標 (URL カウント)。
NumUniquePagesSaved	値: 負でない整数  その他のエンタープライズ・サーチ・コンポーネントによってさらに処理するために RDS に書き込まれた固有のページの数。	索引のサイズに対する、このクローラーのコンテンツリビューション。

表 16. Web クローラーのクロール・スペース状況情報 (続き)

エレメント	属性	説明
NumURLsCrawled	値: 負でない整数  クローラーによってクロールされた固有の URL の数。	データを処理するための、クローラーの機能の指標 (エンドツーエンド)。この数は、RDS に書き込まれるページの数によって異なります。それは、クロールされたすべてのページが RDS に書き込まれるわけではないからです。
NumURLsOverdue	値: 負でない整数  再クロールする対象として適格な固有の URL の数。	Web スペースをトラバースするための、クローラーの機能の指標。

**Web 以外のクローラーのクロール・スペース状況:** Web 以外のクローラーのクロール・スペース状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。 **getCrawlSpaceStatus** コマンドによって、Web クローラー以外のクローラーについて次の情報を戻すことができます。

```
FFQC5314I 次の結果が発生しました: <?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>
```

以下の表で、Web クローラー以外の各エンタープライズ・サーチ・クローラーでの、XML エレメントおよび属性について説明します。この情報は、クロール・スペース状況コマンドで戻されるものです。Notes クローラーの場合、集約レベルが 0 のとき、Server@Name は サーバー名 + データベース名 です。集約レベルが 1 のとき、Server@Name は サーバー名 + ディレクトリー名 です。

表 17. NNTP、DB2、JDBC データベース、および Lotus Notes クローラーのクロール・スペース状況情報

エレメントおよび属性の名前	DB2 および JDBC データ		
	NNTP クローラー	ベース・クローラー	Lotus Notes クローラー
Server@Name	ニュース・サーバー名	データベース名	データベース名またはディレクトリー名

表 17. NNTP、DB2、JDBC データベース、および Lotus Notes クローラーのクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	NNTP クローラー	DB2 および JDBC データベース・クローラー	Lotus Notes クローラー
Server/Status	状況: (0、1、2、3、4、5、-1) ・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	状況 (0、1、2、3、4、5、-1) ・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	状況 (0、1、2、3、4、5、-1) ・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー
Server/StatusMessage	・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー
Server/NumberOfTargets	クロール・スペース内の新規グループの数。	クロール・スペース内のデータベースの数。	クロール・スペース内のビュー、およびフォルダーまたはディレクトリーの数。
Server/NumberOfCompletedTargets	クロールされた新規グループの数。	クロールされた表の数。	クロールされたビュー、およびフォルダーまたはディレクトリーの数。
Server/NumberOfErrors	エラーの数。	エラーの数。	エラーの数。
Server/StartTime	適用される場合、開始日時。	適用される場合、開始日時。	適用される場合、開始日時。
Server/EndTime	適用される場合、終了日時。	適用される場合、終了日時。	適用される場合、終了日時。



表 17. NNTP、DB2、JDBC データベース、および Lotus Notes クローラーのクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	NNTP クローラー	DB2 および JDBC データベース・クローラー	Lotus Notes クローラー
Server/ScheduleConfigured	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>
Server/ScheduleTime	適用される場合、スケジュール日時。	適用される場合、スケジュール日時。	適用される場合、スケジュール日時。
Server/TotalTime	適用される場合、合計時間。	適用される場合、合計時間。	適用される場合、合計時間。
Server/AggregationLevel	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。	0、1: <ul style="list-style-type: none"> <li>0: Notes クローラーは、通常モードで文書をクロールします。(Notes クローラー以外のクローラーでは、常に 0 が戻されます。)</li> <li>1: Notes クローラーは、ディレクトリー・モードで文書をクロールします。</li> </ul>

表 18. Exchange Server、DB2 Content Manager、および Content Edition クローラーのクロール・スペース状況情報

エレメントおよび属性の名前	Exchange Server クローラー	DB2 Content Manager クローラー	Content Edition クローラー
Server@Name	Exchange Server サーバー名	DB2 Content Manager サーバー	リポジトリ名

表 18. Exchange Server、DB2 Content Manager、および Content Edition クローラーのクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	Exchange Server クローラー	DB2 Content Manager クローラー	Content Edition クローラー
Server/Status	状況 (0、1、2、3、4、5、-1) <ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• 5: スケジュールされている</li> <li>• -1: エラー</li> </ul>	状況 (0、1、2、3、4、5、-1) <ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• 5: スケジュールされている</li> <li>• -1: エラー</li> </ul>	状況 (0、1、2、3、4、5、-1) <ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• 5: スケジュールされている</li> <li>• -1: エラー</li> </ul>
Server/StatusMessage	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• 5: スケジュールされている</li> <li>• -1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• 5: スケジュールされている</li> <li>• -1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• 5: スケジュールされている</li> <li>• -1: エラー</li> </ul>
Server/NumberOfTargets	クロール・スペース内のサブフォルダーの数。	クロール・スペース内の項目タイプの数。	クロール・スペース内の項目クラスの数。
Server/NumberOfCompletedTargets	クロールされたサブフォルダーの数。	クロールされた項目タイプの数。	クロールされた項目クラスの数。
Server/NumberOfErrors	エラーの数。	エラーの数。	エラーの数。
Server/StartTime	適用される場合、開始日時。	適用される場合、開始日時。	適用される場合、開始日時。
Server/EndTime	適用される場合、終了日時。	適用される場合、終了日時。	適用される場合、終了日時。

表 18. Exchange Server、DB2 Content Manager、および Content Edition クローラーのクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	Exchange Server クローラー	DB2 Content Manager クローラー	Content Edition クローラー
Server/ScheduleConfigured	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>
Server/ScheduleTime	適用される場合、スケジュール日時。	適用される場合、スケジュール日時。	適用される場合、スケジュール日時。
Server/TotalTime	適用される場合、合計時間。	適用される場合、合計時間。	適用される場合、合計時間。
Server/AggregationLevel	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。

表 19. QuickPlace、Domino Document Manager、UNIX ファイル・システム、および Windows ファイル・システム・クローラーのクロール・スペース状況情報

エレメントおよび属性の名前	QuickPlace クローラー	Lotus Domino Document Manager クローラー	UNIX および Windows ファイル・システム・クローラー
Server@Name	ブレース・ディレクトリー	ライブラリー・データベース	ローカル・ホストの固定値

表 19. QuickPlace、Domino Document Manager、UNIX ファイル・システム、および Windows ファイル・システム・クローラーのクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	QuickPlace クローラー	Lotus Domino Document Manager クローラー	UNIX および Windows ファイル・システム・クローラー
Server/Status	状況 (0、1、2、3、4、5、-1) ・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	状況 (0、1、2、3、4、5、-1) ・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	状況 (0、1、2、3、4、5、-1) ・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー
Server/StatusMessage	・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー	・ 0: クロールされていない ・ 1: クロール中 ・ 2: 完了 (スケジュールされていない) ・ 3: 待機中 ・ 4: 休止されている ・ 5: スケジュールされている ・ -1: エラー
Server/NumberOfTargets	クロール・スペース内のプレース・データベースおよびルーム・データベースの数。	クロール・スペース内のキャビネットの数。	クロール・スペース内のサブディレクトリーの数。
Server/NumberOfCompletedTargets	クロールされたプレース・データベースおよびルーム・データベースの数。	クロールされたキャビネットの数。	クロール・スペース内のサブディレクトリーの数。
Server/NumberOfErrors	エラーの数。	エラーの数。	エラーの数。
Server/StartTime	適用される場合、開始日時。	適用される場合、開始日時。	適用される場合、開始日時。
Server/EndTime	適用される場合、終了日時。	適用される場合、終了日時。	適用される場合、終了日時。

表 19. QuickPlace、Domino Document Manager、UNIX ファイル・システム、および Windows ファイル・システム・クローラーのクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	QuickPlace クローラー	Lotus Domino Document Manager クローラー	UNIX および Windows ファイル・システム・クローラー
Server/ScheduleConfigured	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジューリングされるように構成されていません。</li> <li>1: クローラーが、スケジューリングされるように構成されています。しかし、スケジューリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジューリングされるように構成されています。そして、スケジューリングはそのセッションで使用可能になっています。</li> </ul>
Server/ScheduleTime	適用される場合、スケジュール日時。	適用される場合、スケジュール日時。	適用される場合、スケジュール日時。
Server/TotalTime	適用される場合、合計時間。	適用される場合、合計時間。	適用される場合、合計時間。
Server/AggregationLevel	0: クローラーは、通常モードで文書をクロールしません。	0: クローラーは、通常モードで文書をクロールしません。	0: クローラーは、通常モードで文書をクロールしません。

表 20. WebSphere Portal および Web Content Management クローラーのクロール・スペース状況情報

エレメントおよび属性の名前	WebSphere Portal クローラー	Web Content Management クローラー
Server@Name	WebSphere Portal サーバー	Web Content Management 検索のシード URL
Server/Status	状況 (0、1、2、3、4、5、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>5: スケジュールされている</li> <li>-1: エラー</li> </ul>	状況 (0、1、2、3、4、5、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>5: スケジュールされている</li> <li>-1: エラー</li> </ul>

表 20. WebSphere Portal および Web Content Management クローラーのクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	WebSphere Portal クローラー	Web Content Management クローラー
Server/StatusMessage	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>5: スケジュールされている</li> <li>-1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>5: スケジュールされている</li> <li>-1: エラー</li> </ul>
Server/NumberOfTargets	クロール・スペース内のサーバーの数。	クロール・スペース内のサイトの数。
Server/NumberOf CompletedTargets	クロールされたサーバーの数。	クロールされたサイトの数。
Server/NumberOfErrors	エラーの数。	エラーの数。
Server/StartTime	適用される場合、開始日時。	適用される場合、開始日時。
Server/EndTime	適用される場合、終了日時。	適用される場合、終了日時。
Server/ScheduleConfigured	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジュールリングされるように構成されていません。</li> <li>1: クローラーが、スケジュールリングされるように構成されています。しかし、スケジュールリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジュールリングされるように構成されています。そして、スケジュールリングはそのセッションで使用可能になっています。</li> </ul>	0、1、2 <ul style="list-style-type: none"> <li>0: クローラーが、クローラー構成ファイルに従って、スケジュールリングされるように構成されていません。</li> <li>1: クローラーが、スケジュールリングされるように構成されています。しかし、スケジュールリングはそのセッションで使用不可になっていました。</li> <li>2: クローラーが、スケジュールリングされるように構成されています。そして、スケジュールリングはそのセッションで使用可能になっています。</li> </ul>
Server/ScheduleTime	適用される場合、スケジュール日時。	適用される場合、スケジュール日時。
Server/TotalTime	適用される場合、合計時間。	適用される場合、合計時間。
Server/AggregationLevel	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。

**Web クローラーの詳細なクロール・スペース状況:** Web クローラーの詳細なクロール・スペース状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。詳細なクロール・スペース状況コマンドによって、以下の情報を得ることができます。

表 21. Web クローラーの詳細なクロール・スペース状況コマンドの選択マスク値

マスク・ビット	選択
1	ロー・データ・ストア内のページの数。

表 21. Web クローラーの詳細なクロール・スペース状況コマンドの選択マスク値 (続き)

マスク・ビット	選択
2	検出されたサイトの数。
4	DNS のあるサイトの数。
8	DNS のないサイトの数。
16	検出された URL の数。
32	保存された固有ページの数。
64	クロールされた URL の数。
128	クロールされていない URL の数。
256	期限切れの URL の数。
512	HTTP 状況コードの分布。

戻り情報の例:

```
<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>
```

以下の表で、Web クローラーの詳細なクロール・スペース状況として戻される各フィールドについて説明します。

表 22. Web クローラーの詳細なクロール・スペース状況情報

エレメント	属性	説明
CrawlDetailsPerSite	<ul style="list-style-type: none"> <li>LastActivationTime:</li> <li>LastActivationDuration:</li> <li>IPAddressList:</li> <li>RobotsContent:</li> <li>HTTPCodeDist:</li> </ul>	あるサイトの詳細な状態に関して迅速に取得できる情報。
Site	URL	サイト・ルート・ページの URL。
NumURLsDiscovered	Value	サイトから検出された URL の数。
NumURLsOverdue	Value	サイトから再クロールする対象として適格な URL の数。

表 22. Web クローラーの詳細なクロール・スペース状況情報 (続き)

エレメント	属性	説明
NumURLsCrawled	Value	サイトに応じてクロールされた URL の数。
NumURLsUncrawled	Value	サイトに応じてまだクロールされていない URL の数。
NumURLsOverdueBy	Threshold、Value: 整数 (正数または負数)  Value は、再クロールの対象として適格な URL の数を示します。Threshold には、URL が再クロールされるまでの待ち時間を指定します。Threshold は、現在時刻からの経過秒数で示されます。Threshold が負の場合、URL の再クロールが待ちになっていることを示します。Threshold が正の場合、URL の再クロールの時期がまだで、この後待ちになることを示します。	少なくとも示された秒数前に再クロールの対象として適格であった URL の数、あるいは、示された秒数後に再クロールの対象として適格になる URL の数。
NumURLsActivated	Value	このサイトの最後のスキャン中にメモリー内に入れられ、クローラー・スレッドから使用できるようになった URL の数。
LastActivationTime	Value	このサイトの URL が最後にメモリー内に入れられたエポック以降、経過した秒数。
LastActivationDuration	Value	このサイトの URL が最後にメモリーに入れられ、クローラー・スレッドから使用できるようになっていた秒数。
IPAddressList	IPAddress	このサイトのサーバー・ホスト用の、認識されているすべての IP アドレス。
IPAddress	Value	サイトのサーバー・ホスト用の、IPv4 ドット表記アドレス。
RobotsContent	テキスト	ロボット・ファイルにテキストが存在している場合、そのテキスト。
HTTPCodeDist	HTTPCode	このサイトの試行されたダウンロードからの HTTP コードの分布。
HTTPCode	Code: 整数  HTTP 状況コードまたは別の内部コード。	このサイトのクロール中に、特定の HTTP 状況コードが発生した回数。

**Web 以外のクローラーの詳細なクロール・スペース状況:** Web 以外のクローラーの詳細なクロール・スペース状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。 `getCrawlSpaceStatusDetail` コマンドによって、Web クローラー以外のクローラーについて次の情報を戻すことができます。



```

FFQC5314I 次の結果が発生しました: <?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>

```

表 23. NNTP、DB2、JDBC データベース、および Lotus Notes クローラーの詳細なクロール・スペース状況情報

エレメントおよび属性の名前	NNTP クローラー	DB2 および JDBC データベース・クローラー	Lotus Notes クローラー
Target@Name	ニュースグループ名	表名	ビューまたはフォルダー名
Target@CrawlType	適用されない。	0、1 (DB2); 0 (JDBC データベース) <ul style="list-style-type: none"> <li>0: アクティブ・クロール (通常)</li> <li>1: パッシブ・クロール (DB2 Event Publishing)</li> </ul>	0
Target/Status	状況: (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	状況: (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	状況: (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>
Target/NumberOfRecords	サーバー上の最後の項目番号。	クロールされたレコードの数。	クロールされた文書の数。
Target/NumberOfCompletedRecords	クロールされた項目の数。	クロールされたレコードの数。	クロールされた文書の数。
Target/NumberOfInsertedRecords	新しく通知された項目の数。	挿入されたレコードの数。	挿入されたレコードの数。
Target/NumberOfUpdatedRecords	適用されない。	更新されたレコードの数。	更新されたレコードの数。

表 23. NNTP、DB2、JDBC データベース、および Lotus Notes クローラーの詳細なクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	NNTP クローラー	DB2 および JDBC データベース・クローラー	Lotus Notes クローラー
Target/NumberOfDeletedRecords	適用されない。	削除されたレコードの数。	削除されたレコードの数。
Target/StartTime	クローラーが最後に開始された日時。	クローラーが最後に開始された日時。	クローラーが最後に開始された日時。
Target/EndTime	クロールが完了した日時。	クロールが完了した日時。	クロールが完了した日時。
Target/TotalTime	クローラーがクロールに費やした時間の合計。	クローラーがクロールに費やした時間の合計。	クローラーがクロールに費やした時間の合計。
Target/AggregationLevel	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。	0、1: <ul style="list-style-type: none"> <li>0: クローラーは、通常モードで文書をクロールします。</li> <li>クローラーは、ディレクトリー・モードで文書をクロールします。</li> </ul>
Target/LastUpdatedTime	適用されない。	最終更新日時: <ul style="list-style-type: none"> <li>0: アクティブ・クロール (通常)</li> <li>1: パッシブ・クロール (DB2 Event Publishing)</li> </ul>	適用されない。
Target/LastResetTime	適用されない。	統計の最終リセット日時: <ul style="list-style-type: none"> <li>0: アクティブ・クロール (通常)</li> <li>1: パッシブ・クロール (DB2 Event Publishing)</li> </ul>	適用されない。

表 24. Exchange Server、DB2 Content Manager、および Content Edition クローラーの詳細なクロール・スペース状況情報

エレメントおよび属性の名前	Exchange Server クローラー	DB2 Content Manager クローラー	Content Edition クローラー
Target@Name	サブフォルダー名	項目タイプ名	項目クラス名
Target@CrawlType	0	0	0
Target/Status	状況 (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	状況 (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	状況 (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>

表 24. Exchange Server、DB2 Content Manager、および Content Edition クローラーの詳細なクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	Exchange Server クローラー	DB2 Content Manager クローラー	Content Edition クローラー
Target/StatusMessage	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• -1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• -1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• -1: エラー</li> </ul>
Target/NumberOf Records	適用されない。	適用されない。	適用されない。
Target/NumberOf CompletedRecords	クロールされた文書の数。	クロールされた文書の数。	クロールされた文書の数。
Target/NumberOf InsertedRecords	挿入されたレコードの数。	挿入されたレコードの数。	挿入されたレコードの数。
Target/NumberOf UpdatedRecords	適用されない。	更新されたレコードの数。	更新されたレコードの数。
Target/NumberOf DeletedRecords	適用されない。	削除されたレコードの数。	削除されたレコードの数。
Target/StartTime	クローラーが最後に開始された日時。	クローラーが最後に開始された日時。	クローラーが最後に開始された日時。
Target/EndTime	クロールが完了した日時。	クロールが完了した日時。	クロールが完了した日時。
Target/TotalTime	クローラーがクロールに費やした時間の合計。	クローラーがクロールに費やした時間の合計。	クローラーがクロールに費やした時間の合計。
Target/AggregationLevel	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。
Target/LastUpdateTime	適用されない。	適用されない。	適用されない。
Target/LastResetTime	適用されない。	適用されない。	適用されない。

表 25. QuickPlace、Lotus Domino Document Manager、UNIX ファイル・システム、および Windows ファイル・システム・クローラーの詳細なクロール・スペース状況情報

エレメントおよび属性の名前	QuickPlace クローラー	Lotus Domino Document Manager クローラー	UNIX および Windows ファイル・システム・クローラー
Target@Name	プレイス・データベース名またはルーム・データベース名	キャビネット・データベース名	サブディレクトリー名
Target@CrawlType	0	0	0

表 25. QuickPlace、Lotus Domino Document Manager、UNIX ファイル・システム、および Windows ファイル・システム・クローラーの詳細なクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	QuickPlace クローラー	Lotus Domino Document Manager クローラー	UNIX および Windows ファイル・システム・クローラー
Target/Status	状況 (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	状況 (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	状況 (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>0: クロールされていない</li> <li>1: クロール中</li> <li>2: 完了 (スケジュールされていない)</li> <li>3: 待機中</li> <li>4: 休止されている</li> <li>-1: エラー</li> </ul>
Target/NumberOfRecords	適用されない。	適用されない。	適用されない。
Target/NumberOfCompletedRecords	クロールされた文書の数。	クロールされた文書の数。	クロールされたファイルの数。
Target/NumberOfInsertedRecords	挿入されたレコードの数。	挿入されたレコードの数。	挿入されたレコードの数。
Target/NumberOfUpdatedRecords	更新されたレコードの数。	更新されたレコードの数。	更新されたレコードの数。
Target/NumberOfDeletedRecords	削除されたレコードの数。	削除されたレコードの数。	削除されたレコードの数。
Target/StartTime	クローラーが最後に開始された日時。	クローラーが最後に開始された日時。	クローラーが最後に開始された日時。
Target/EndTime	クロールが完了した日時。	クロールが完了した日時。	クロールが完了した日時。
Target/TotalTime	クローラーがクロールに費やした時間の合計。	クローラーがクロールに費やした時間の合計。	クローラーがクロールに費やした時間の合計。
Target/AggregationLevel	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。
Target/LastUpdatedTime	適用されない。	適用されない。	適用されない。
Target/LastResetTime	適用されない。	適用されない。	適用されない。

表 26. WebSphere Portal および Web Content Management クローラーの詳細なクロール・スペース状況情報

エレメントおよび属性の名前	WebSphere Portal クローラー	Web Content Management クローラー
Target@Name	WebSphere Portal サーバー名	サイトを表す検索シード URL
Target@CrawlType	0	0

表 26. WebSphere Portal および Web Content Management クローラーの詳細なクロール・スペース状況情報 (続き)

エレメントおよび属性の名前	WebSphere Portal クローラー	Web Content Management クローラー
Target/Status	状況: (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• -1: エラー</li> </ul>	状況: (0、1、2、3、4、-1) <ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• -1: エラー</li> </ul>
Target/StatusMessage	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• -1: エラー</li> </ul>	<ul style="list-style-type: none"> <li>• 0: クロールされていない</li> <li>• 1: クロール中</li> <li>• 2: 完了 (スケジュールされていない)</li> <li>• 3: 待機中</li> <li>• 4: 休止されている</li> <li>• -1: エラー</li> </ul>
Target/NumberOf Records	適用されない。	適用されない。
Target/NumberOf CompletedRecords	クロールされたレコードの総数。	クロールされたレコードの総数。
Target/NumberOf InsertedRecords	挿入されたレコードの数。	挿入されたレコードの数。
Target/NumberOf UpdatedRecords	更新されたレコードの数。	更新されたレコードの数。
Target/NumberOf DeletedRecords	削除されたレコードの数。	削除されたレコードの数。
Target/StartTime	クローラーが最後に開始された日時。	クローラーが最後に開始された日時。
Target/EndTime	クロールが完了した日時。	クロールが完了した日時。
Target/TotalTime	クローラーがクロールに費やした時間の合計。	クローラーがクロールに費やした時間の合計。
Target/AggregationLevel	0: クローラーは、通常モードで文書をクロールします。	0: クローラーは、通常モードで文書をクロールします。
Target/LastUpdatedTime	適用されない。	適用されない。
Target/LastResetTime	適用されない。	適用されない。

**パーサー状況:** パーサー状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。パーサー状況コマンドによって、以下の情報を得ることができます。

```
FFQC5314I 次の結果が発生しました:
<Monitor Type="Parser">
<ParserStatus>
  <Status>1</Status>
  <State>Parsing</State>
  <NumberOfDocsToBeIndexed>231974</NumberOfDocsToBeIndexed>
  <ParseRate>0</ParseRate>
  <ParseRateMBPerHour>0</ParseRateMBPerHour>
  <NumberOfCpmThreads>3</NumberOfCpmThreads>
  <ParserServiceSession>parserservice.1</ParserServiceSession>
</ParserStatus>
<CrawlerStatus>
```

```

<Name>WEBCrawler1</Name>
<Crawlerid>col1.WEB1.esadmin</Crawlerid>
<Type>WEB</Type>
<ParserStatus>1</ParserStatus>
<NumberOfDocsAlreadyParsed>29</NumberOfDocsAlreadyParsed>
</CrawlerStatus>
</Monitor>

```

以下の表で、パーサー状況コマンドによって戻される情報の XML エlement について説明します。

表 27. パーサー状況コマンドのElement

Element	説明
Status	<ul style="list-style-type: none"> <li>• 0: このコレクションのためのパーサー・セッションが停止しています。</li> <li>• 1: このコレクションのためのパーサー・セッションが稼働中です。</li> </ul>
State	<p>起こりうる状態は、Initializing (初期化中)、Idle (アイドル)、Restart (再始動)、Parsing (構文解析中)、Stopped (停止)、Paused (休止)、Resuming (再開)、NoParserServiceIsAvailable (使用可能なパーサー・サービスがない) です。</p> <p>Initializing (初期化中) という状況は、パーサーが開始され、その状態を初期化していることを示します。</p> <p>Idle (アイドル) という状況は、パーサーが、このコレクション内のクローラーからさらに多くの文書が到着するのを待って、N 分間スリープしていることを示します。デフォルトのスリープ時間は 300 秒です。</p> <p>Restart (再始動) という状況は、パーサーが、構文解析/トークン化 JVM が再始動するのを待機していることを示します。構文解析/トークン化 JVM は別個のセッションで実行されます。これは、文書が最終的に処理される場所です。</p> <p>Parsing (構文解析中) という状況は、パーサーが文書を処理していることを示します。</p> <p>Paused (休止) という状況は、パーサーが、このコレクション用の索引作成セッションによって休止されたことを示します。</p> <p>Resuming (再開) という状況は、パーサーが、このコレクション用の索引作成セッションによって、Paused (休止) 状態から Parsing (構文解析中) 状態に変更されたことを示します。</p> <p>NoParserServiceIsAvailable (使用可能なパーサー・サービスがない) という状況は、このコレクション用の文書を処理するために使用可能な構文解析/トークン化 JVM がないことを示します。この状況は、すべての構文解析/トークン化 JVM が、他のコレクションで使用されていることを意味します。</p>
NumberOfDocsToBeIndexed	このコレクション用のストア内の文書の数。この数には、次回の索引作成から削除するようマークされている文書も含まれます。
ParseRate	構文解析速度 (1 秒あたりの文書数)。
ParseRateMBPerHour	構文解析速度 (1 時間あたりの MB)。
NumberOfCpmThreads	このコレクション用の文書を処理するために構文解析/トークン化 JVM が使用する CPM スレッドの数。

表 27. パーサー状況コマンドの要素 (続き)

要素	説明
ParserServiceSession	このコレクション用の文書処理している構文解析/トークン化 JVM の名前。このフィールドは、パーサーが Parsing (構文解析中) 状態になっている場合にのみ使用可能になります。
Name	クローラーの名前。
Crawlerid	システムによってこのクローラーに作成された ID。
Type	クローラーのタイプ (Web、NNTP、DB2、など)。
ParserStatus	<ul style="list-style-type: none"> <li>0: このクローラーからの文書は構文解析されていません (パーサー・セッションが停止しています)。</li> <li>1: このクローラーからの文書は構文解析されています (パーサー・セッションが稼働中です)。</li> </ul>
NumberofDocsAlreadyParsed	このクローラーからの構文解析された文書の数。

**索引作成状況:** 索引作成状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。索引作成状況コマンドによって、以下の情報を得ることができます。

```
<?xml version="1.0"?>
<Monitor Type="MainIndexHistory" Count="1">
  <IndexStatus Id="1">
    <StartTime>1131987633901</StartTime>
    <Progress>0</Progress>
    <CurrentPhase>0</CurrentPhase>
    <TotalPhase>3</TotalPhase>
    <IndexCopyTime>49822</IndexCopyTime>
    <CurrentServer>0</CurrentServer>
    <TotalServer>0</TotalServer>
    <IndexBuildTime>46158</IndexBuildTime>
    <Status>0</Status>
    <JobID>1131987633899</JobID>
    <MessagesAvailable>>false</MessagesAvailable>
    <StopTime>1131987734199</StopTime>
    <TotalTime>100298</TotalTime>
    <NumberOfDocuments>43</NumberOfDocuments>
  </IndexStatus>
  <CurrentIndexWildcardSupport/>
  <NextIndexWildcardSupport Type="None" Limit="0"/>
  <ScheduleStatus>
    <Status>1</Status>
    <ScheduledTime Enabled="false"></ScheduledTime>
  </ScheduleStatus>
</Monitor>
```

以下の表で、索引作成状況コマンドによって戻される情報の各 XML エlement について説明します。

表 28. 索引作成状況コマンドの要素

要素	説明
IndexStatusId	索引状況 ID。
StartTime	この索引作成が開始された時刻を、1970 年からの秒数で表した時間。この時間が表す現在時刻を計算するには、January 1, 1970 %2B StartTime という式を使用します。エポック時間については、 <a href="http://en.wikipedia.org/wiki/Unix_epoch">http://en.wikipedia.org/wiki/Unix_epoch</a> を参照してください。

表 28. 索引作成状況コマンドの要素 (続き)

要素	説明
Progress	この索引作成の完了したパーセンテージ。
CurrentPhase	<ul style="list-style-type: none"> <li>1: 保管再書き込みフェーズ</li> <li>2: グローバル分析フェーズ</li> <li>3: 索引作成フェーズ</li> </ul>
TotalPhase	この索引作成のフェーズの数。現在この値は 3 です。
IndexCopyProgress	索引コピーの完了したパーセンテージ。索引コピー処理では、作成された索引を索引作成サーバーから検索サーバーにコピーします。
CurrentServer	索引コピーで、索引がコピーされる先の検索サーバー。
TotalServer	索引をコピーする先の検索サーバーの数。
IndexCopyTime	すべての検索サーバーに索引をコピーするための合計時間。
IndexBuildTime	索引作成のすべてのフェーズの合計時間
Status	<ul style="list-style-type: none"> <li>0: 索引作成とコピー</li> <li>-1: 索引作成要求の失敗</li> <li>1: 索引作成、コピー、あるいはその両方が進行中</li> </ul>
JobID	索引作成の各要求に関連付けられた固有 ID。
MessagesAvailable	(障害時に) エラー・メッセージがあるかどうかを示すブール値。
StopTime	索引作成 (すべてのフェーズ) および索引コピーの終了日時。
TotalTime	開始日時と停止日時の間の期間。
NumberOfDocuments	索引内の文書の数。
CurrentIndexWildCardSupport	次の索引作成に使用すべきワイルドカード設定。取りうる値は、None、QueryExpansion、または IndexExpansion です。
ScheduleStatus	<ul style="list-style-type: none"> <li>0: このコレクションおよび索引タイプのためのスケジュールが使用可能になっていない場合。</li> <li>1: このコレクションおよび索引タイプのためのスケジュールが使用可能になっている場合。</li> </ul>
ScheduledTimeEnabled	このコレクションと索引タイプのために次に索引作成が実行される時刻を、1970 年からの秒数で表した時間。この時間が表す現在時刻を計算するには、January 1, 1970 %2B ScheduledTimeEnabled という式を使用します。エポック時間については、 <a href="http://en.wikipedia.org/wiki/Unix_epoch">http://en.wikipedia.org/wiki/Unix_epoch</a> を参照してください。

**検索サーバー状況:** 検索サーバー状況を取得するコマンドを実行すると、XML 文書フォーマットで情報が戻されます。検索サーバー状況コマンドによって、以下の情報を得ることができます。

```
FFQC5314I 次の結果が発生しました: <?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>
```

以下の表で、検索サーバー状況コマンドによって戻される情報の XML エlement について説明します。



表 29. 検索サーバー状況コマンドのエレメント

エレメント	説明
SearchStatusName	このコレクション用の検索索引のモニターおよび保守を行っている検索マネージャー・セッションの名前および ID。
HostName	検索索引が実行されているサーバーのホスト名。
Status	<ul style="list-style-type: none"> <li>0: このコレクション用の検索索引が実行されていない場合。</li> <li>1: このコレクション用の検索索引が実行されている場合。</li> </ul>

**詳細な検索サーバー状況:** 検索サーバー状況を戻すコマンドによって、以下の情報を得ることができます。

```
FFQC5303I Search Manager (node1) (sid: searchmanager.node1)
は既に実行中です。PID: 15711
FFQC5314I 次の結果が発生しました: PID=18390
CacheHits=3
QueryRate=1
Port=44008
SessionId=coll.runtime.node1
CacheHitRate=0.333
ResponseTime=70
Status=1
SessionName=coll.runtime.node1.1
```

以下の表で、詳細な検索サーバー状況コマンドによって戻される情報の項目について説明します。

表 30. 詳細な検索サーバー状況コマンドの項目

項目	説明
CacheHits	検索キャッシュからリトリートされる結果の数。
QueryRate	最後の時間間隔で受信した照会の数。デフォルトでは、時間間隔は 5 分です。
Port	照会を listen または受信するために検索索引が使用するポート番号。
SessionId	このコレクションの検索索引のセッション ID。
CacheHitRate	検索キャッシュからリトリートされる結果の数 (すべての検索結果に対するパーセンテージ)。
ResponseTime	指定した時間間隔 (デフォルトは 5 分) での平均応答時間 (ミリ秒)。
Status	<ul style="list-style-type: none"> <li>0: このコレクション用の検索索引が実行されていない場合。</li> <li>1: このコレクション用の検索索引が実行されている場合。</li> </ul>
SessionName	このコレクションの検索索引のセッション名。

## esadmin コマンドの戻りコード

**esadmin** コマンドから戻るコードは、以下のとおりです。

表 31. **esadmin** コマンドの戻りコード

コード	名前	説明
0	CODE_ERROR_NONE	コマンドが正常に完了しました。
102	CODE_ERROR_INSTANTIATION_EXCEPTION	コマンド・ハンドラーをインスタンス化するときにエラーが発生しました。

表 31. esadmin コマンドの戻りコード (続き)

コード	名前	説明
103	CODE_ERROR_ACCESS_EXCEPTION	コマンド・ハンドラーをインスタンス化するときに、正しくないアクセスによるエラーが発生しました。
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	

## セッション ID の取得

esadmin check コマンドを使用して、エンタープライズ・サーチ・コンポーネントおよびそれらに対応するセッション ID のリストを表示することができます。以下の表で、共通セッション、それらの ID、それらが存在しているサーバー、およびセッションの状態のリストを示します。

表 32. セッション名、起点サーバー、セッション ID、およびセッション状態の例

セッション	セッションが実行されているサーバー	セッション ID	セッションの状態
configmanager	索引サーバー	10433	開始済み
controller	索引サーバー	10464	開始済み
customcommunication	索引サーバー	適用されない	適用されない
discovery	索引サーバー	10649	開始済み
monitor	索引サーバー	10682	開始済み
parserservice	索引サーバー	10718	開始済み
resource.node1	索引サーバー	10759	開始済み
samplecpp	索引サーバー	10827	開始済み
sampletest	索引サーバー	10857	開始済み
scheduler	索引サーバー	10889	開始済み
searchmanager.node1	索引サーバー	10927	開始済み
utilities.node1	索引サーバー	10384	開始済み

### 関連概念

335 ページの『エンタープライズ・サーチ・アクティビティのモニター』

371 ページの『エンタープライズ・サーチ・システムのバックアップとリストア』

 エンタープライズ・サーチのメッセージ

 エンタープライズ・サーチのメッセージ

### 関連タスク

338 ページの『クローラーのモニター』

- 327 ページの『エンタープライズ・サーチ・システムの開始』
- 330 ページの『エンタープライズ・サーチ・システムの停止』
- 332 ページの『スタンドアロン・モードでの検索サーバーの管理』



---

## エンタープライズ・サーチにおける大/小文字の区別

照会構文、クイック・リンク、フィールド名など、エンタープライズ・サーチのコンポーネントでは、その大/小文字の扱いが異なります。

### 照会構文

以下の場合を除き、検索では大/小文字を区別しません。

#### XML エlement名と属性名

大/小文字を区別します。語と属性値は、XML 照会でも大/小文字を区別しません。例えば、次の文書で例を示します。

```
<book>
  <Author>
    <Name>Ferdinand</Name>
    <Contact Type="eMail">ferdi@nand.org</Contact Type>
    <Contact Type="Phone">+1 408 876 4242</Contact Type>
  </Author>
</book>
```

以下の照会は、この文書を戻しません。

- @xmlxp::'author[Name ftcontains ("Ferdinand")]'
- @xmlxp::'//contact[@type="eMail"]'
- @xmlf2::'<author><name>Ferdinand</name></author>
- @xmlf2::'<CONTACT TYPE="email">ferdi</contact>

一方、以下の照会は、この文書を戻します。

- @xmlxp::'Author[Name ftcontains ("ferdinand")]'
- @xmlxp::'//Contact[@Type="email"]'
- @xmlf2::'<Author><Name>ferdinand</Author><Name>
- @xmlf2::'<Contact Type="email">ferdi</Contact>

#### アクセス制御リスト (ACL)

大/小文字を区別します。

#### docid: および samegroupas: 条件内の URL

大/小文字を区別します。ただし、site: または url: 条件に含まれる URL の構成部分は大/小文字を区別しません。例えば、http://www.here.com/HR/ という URL の文書では、以下ようになります。

- docid:http://www.here.com/hr という照会は、この文書を戻しません。
- url:hr と url:HERE という照会は、この文書を戻します。
- site:HERE.com という照会は、この文書を戻します。

#### カテゴリー ID と分類構造 ID

大/小文字を区別しません。例えば、taxonomy\_id::category\_id という照会は、taxonomy\_id と category\_id の両者とも大/小文字はどちらでも問題ありません。RuleBased::c42 という照会は、ルール・ベースの分類構造のカ

テゴリー c42 に一致し、rulebased::C42 にも一致します。カテゴリ名では検索できませんが、カテゴリ ID で検索できます。

#### 有効範囲

大/小文字を区別しません。例えば、Scope:RESEARCH と scope::research の両方が、Research という有効範囲から文書を戻します。

#### ワイルドカードの条件

大/小文字を区別しません。例えば、条件 Fer\*n\*d は fer\*n\*d と同等です。

#### フィールド名

大/小文字を区別しません。例えば、

Title:Expenses、TITLE:expenses、title:expenses という照会はすべて同等です。フィールド名はすべて、XML マッピング・ファイルからののものであっても大/小文字を区別しません。ただし、外部ソースは、固有の照会セマンティクスに従って、大/小文字を区別するものとしてフィールド名を扱う可能性があります。

#### クイック・リンク

大/小文字を区別しません。

#### ルール・ベースのカテゴリライザー

URL 規則は大/小文字を区別しますが、文書コンテンツ規則は大/小文字を区別しません。

#### 縮小表示された URI と URI パターン・ベースのランキング調整定義

URI は大/小文字を区別しますが、縮小表示された URI グループ名は大/小文字を区別しません。大/小文字のみが異なる 2 つの縮小表示された URI グループ名を定義することはできません。

#### フィールド名

大/小文字を区別しません。この規則は、照会、XML マッピング、ランキング調整クラス定義、およびフィールド名が指定される他のすべての管理インターフェースのフィールド名に適用されます。インターフェースで大文字または大/小文字混合のフィールド名を構成すると、システムによって小文字にされて、構成を次に表示したときに小文字で表示されることがあります。また、フィールド名は、外部ソースによって、大/小文字を区別して解釈される可能性があります。

**辞書** 大/小文字を区別しません。この規則は、同義語辞書、ストップワード辞書、スペル提案辞書、ランキング調整ワード辞書に適用されます。

#### コレクション名とコレクション ID

大/小文字を区別します。管理または検索の API でコレクション名またはコレクション ID を指定する場合、それは、その名前のコレクションの大/小文字表記に正確に一致しなければなりません。ただし、コレクション ID が大/小文字を区別するといっても、大/小文字のみが異なる 2 つのコレクション ID を指定することはできません。この制約事項は、クローラーとデータ・ソースの名前および ID にも適用されます。

#### 検索アプリケーション

大/小文字を区別します。大/小文字のみが異なる 2 つのアプリケーション名または ID を指定することはできません。

## エンタープライズ・サーチの資料

OmniFind Enterprise Edition の資料は、PDF 形式または HTML 形式で読むことができます。

OmniFind Enterprise Edition のインストール・プログラムは、エンタープライズ・サーチ用資料の HTML バージョンを含むインフォメーション・センターを自動的にインストールします。複数のサーバーをインストールする場合は、インフォメーション・センターが両方の検索サーバーにインストールされます。インフォメーション・センターをインストールしなかった場合は、「ヘルプ」をクリックすると、IBM Web サイトのインフォメーション・センターが開きます。

PDF 文書を参照するには、ES\_INSTALL\_ROOT/docs/locale/pdf に移動します。例えば、英語の資料を見つけるには、ES\_INSTALL\_ROOT/docs/en\_US/pdf に移動します。

使用可能なすべての言語の PDF バージョンの文書にアクセスするには、OmniFind Enterprise Edition Version 8.5 documentationサイトを参照してください。

製品ダウンロード、フィックスパック、技術情報、およびインフォメーション・センターには、OmniFind Enterprise Edition Support サイトからアクセスすることもできます。

以下の表は、使用可能な資料、ファイル名、ロケーションを示します。

表 33. エンタープライズ・サーチの資料

タイトル	ファイル名	場所
インフォメーション・センター		<a href="http://publib.boulder.ibm.com/infocenter/discover/v8r5/">http://publib.boulder.ibm.com/infocenter/discover/v8r5/</a>
エンタープライズ・サーチ インストール・ガイド	iiysi.pdf	ES_INSTALL_ROOT/docs/locale/pdf/
クイック・スタート・ガイド (英語、フランス語、および日 本語ではハードコピー版も用 意されています。)	OmniFindEE850_qsg_ locale を表す 2 文字 .pdf	ES_INSTALL_ROOT/docs/locale/pdf/
エンタープライズ・サーチの 管理	iiysa.pdf	ES_INSTALL_ROOT/docs/locale/pdf/
<i>Programming Guide and API Reference for Enterprise Search</i>	iiysp.pdf	ES_INSTALL_ROOT/docs/en_US/pdf/
トラブルシューティング・ガ イドおよびメッセージ	iiysm.pdf	ES_INSTALL_ROOT/docs/locale/pdf/
テキスト分析機能ガイド	iiyst.pdf	ES_INSTALL_ROOT/docs/locale/pdf/
Google デスクトップ検索用 プラグイン	iiysg.pdf	ES_INSTALL_ROOT/docs/locale/pdf/





---

## アクセシビリティ機能

アクセシビリティ機能は、運動障害や視覚障害といった身体的障害を持つユーザーが IT 製品を快適に使用できるように支援するものです。

IBM は、年齢や能力にかかわらず誰もが使用できる製品を提供するように努めております。

### アクセシビリティ機能

OmniFind Enterprise Edition における主要なアクセシビリティ機能は次のとおりです。

- キーボードのみの操作
- スクリーン・リーダー (読み上げソフトウェア) が通常使用するインターフェース

OmniFind Enterprise Edition のインフォメーション・センターと関連資料はアクセシビリティ対応になっています。このインフォメーション・センターのアクセシビリティ機能の説明は [http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/topic/com.ibm.classify.nav.doc/dochome/accessibility\\_info.htm](http://publib.boulder.ibm.com/infocenter/discover/v8r5m0/topic/com.ibm.classify.nav.doc/dochome/accessibility_info.htm) にあります。

### キーボード・ナビゲーション

この製品では、標準的な Microsoft Windows ナビゲーション・キーを使用します。

さらに、以下のキーボード・ショートカットを使用して、OmniFind Enterprise Edition インストール・プログラム内をナビゲートすることができます。

表 34. インストール・プログラム用キーボード・ショートカット

アクション	ショートカット
ラジオ・ボタンの強調表示	矢印キー
ラジオ・ボタンの選択	Tab キー
プッシュボタンの強調表示	Tab キー
プッシュボタンの選択	Enter キー
次のウィンドウまたは前のウィンドウへ移動、またはキャンセル	Tab キーを押してプッシュボタンを強調表示し、Enter キーを押す
アクティブ・ウィンドウを非アクティブにする	Ctrl + Alt + Esc

### インターフェース情報

管理コンソール、サンプル検索アプリケーション、および検索アプリケーション・カスタマイザーのユーザー・インターフェースは、Microsoft Internet Explorer または Mozilla FireFox で表示できる、ブラウザー・ベースのインターフェースです。ブラウザーのキーボード・ショートカットのリストおよび他のアクセシビリティ機能については、Internet Explorer または FireFox のオンライン・ヘルプを参照してください。

## 関連するアクセシビリティ情報

Adobe Acrobat Reader を使用すれば、OmniFind Enterprise Edition の資料を、Adobe PDF で表示できます。これらの PDF は、本製品と同梱の CD に収録されていますが、<http://www.ibm.com/support/docview.wss?rs=63&uid=swg27010938> でアクセスすることもできます。

## IBM とアクセシビリティ

IBM のアクセシビリティに対する取り組みの詳細については、IBM Human Ability and Accessibility Center を参照してください。

---

## エンタープライズ・サーチの用語集

この用語集では、エンタープライズ・サーチのインターフェースおよび資料で使用される用語を定義します。

### アクセス制御リスト (access control list)

コンピューター・セキュリティーにおいて、特定のオブジェクトに関連付けられているリスト。そのオブジェクトにアクセスできるすべてのサブジェクトとそれらのアクセス権限を識別する。

### アノテーター (annotator)

特定の言語分析タスクを実行して、注釈を生成し、記録するソフトウェア・コンポーネント。アノテーターは、分析エンジンにおける分析論理コンポーネントです。

### エスケープ文字 (escape character)

後続の 1 つ以上の文字に対して特殊な意味を抑制または設定する文字。

### エンキュー (enqueue)

メッセージや項目をキューに入れること。

### エンタープライズ・サーチ・ベース・アノテーター (enterprise search base annotators)

エンタープライズ・サーチの中でデフォルトの文書分析の処理に使用される標準テキスト分析エンジンのセット。

### エンタープライズ・サーチ管理者 (enterprise search administrator)

エンタープライズ・サーチ・システム全体を管理できる管理役割。

### オペレーター (operator)

コレクション・レベルのプロセスを監視、開始、停止する権限を持つエンタープライズ・サーチ・ユーザー。

### 改行文字 (newline character)

印刷または表示位置を 1 行下へ移動させる制御文字。

### 開始 Uniform Resource Locator (URL) (start Uniform Resource Locator (URL))

クロールの開始点。

### 概念抽出 (concept extraction)

テキスト文書にある重要な語彙項目 (人、場所、製品など) を識別し、その項目リストを生成するテキスト分析。『テーマ抽出』も参照。

### 外部データ・ソース (external data source)

OmniFind Enterprise Edition によってクロール、解析、または索引付けされていないフェデレーションのデータ・ソース。外部データ・ソースの検索は、それらのデータ・ソースの照会アプリケーション・プログラミング・インターフェースに委任されます。

### 鍵ストア・ファイル (keystore file)

署名者証明書として保管される公開鍵と個人証明書に保管される秘密鍵の両方を含む鍵リング。

### 鍵データベース・ファイル (key database file)

『鍵リング』を参照。

### 鍵リング (key ring)

コンピューター・セキュリティーにおいて、公開鍵、秘密鍵、トラステッド・ルート、および証明書を含むファイル。『鍵ストア・ファイル』も参照。

### カスタム・テキスト分析エンジン (custom text analysis engine)

構成解除情報管理アーキテクチャー (UIMA) Software Development Kit (SDK) の使用によって作成されるテキスト分析エンジンで、標準エンタープライズ・サーチ・テキスト分析エンジン (エンタープライズ・サーチ・ベース・アノテーターとも呼ばれる) のセットに追加できる。『テキスト分析エンジン』も参照。

### カタカナ (Katakana)

2 つの一般的日本語表音文字の 1 つで使用されるシンボルから構成される文字セット。外国語のワードを表音的に書く場合に主に使用されます。

### カテゴリー・ツリー (category tree)

カテゴリーの階層。

### 管理役割 (administrative role)

ユーザーに対してアクセス権限を規定するユーザーの種別。

### 共通通信層 (CCL) (Common Communication Layer)

OmniFind Enterprise Edition の各種コンポーネント (コントローラー、パーサー、クローラー、索引サーバー) を結合する通信インフラストラクチャー。

### 共通分析構造 (common analysis structure)

文書のコンテンツとメタデータ、および、テキスト分析エンジンによって生成されたすべての分析結果を保管する構造。文書分析中のすべてのデータ交換は、共通分析構造を使用して処理されます。

### 共通分析構造コンシューマー (CAS コンシューマー) (common analysis structure consumer (CAS consumer))

共通分析構造に保管された分析結果の最終処理を行うコンシューマー。例えば、コンシューマーは、検索エンジン内の共通分析構造のコンテンツの索引付けを行うか、あるいは、リレーショナル・データベースに特定の分析結果を取り込みます。

### 近接検索 (proximity search)

同一文内や同一段落内など、お互いから一定の距離内に 2 つ以上の一致条件があるとき結果を戻すテキスト検索。

### クイック・リンク (quick link)

URI とキーワード、または URI と句との間の関連。

### クレデンシャル (credential)

認証時に取得される詳細情報で、ユーザー、グループ関連、およびその他のセキュリティー関係識別属性を記述するもの。クレデンシャルは、許可、監査、および委任など、多数のサービスを実行するのに使用できます。例えば、あるユーザーのサインオン情報 (ユーザー ID とパスワード) は、そのユーザーがアカウントにアクセスすることを許可するクレデンシャルです。

**クローラー (crawler)**

データ・ソースから文書をリトリートし、検索索引作成用の情報を収集するソフトウェア・プログラム。

**クロール・スペース (crawl space)**

指定パターン (URL、データベース名、ファイル・システム・パス、ドメイン名、IP アドレスなど) に一致するソースの集合。クローラーはここからの読み取りを行って、索引用の項目をリトリートします。

**言語の識別 (language identification)**

エンタープライズ・サーチにおいて、文書の言語を決定する検索機能。

**言語分析検索 (linguistic search)**

基本型に戻したり (例: *mice* は *mouse* として索引付けされる)、または基本型を使用して拡張したり (複合語のように) した語句を使用して文書を表示、リトリート、索引付けする検索タイプ。

**検索アプリケーション (search application)**

エンタープライズ・サーチにおいて、照会の処理、索引の検索、検索結果の戻し、およびソース文書のリトリートを行うプログラム。

**検索エンジン (search engine)**

検索要求を受け取り、文書リストをユーザーに戻すプログラム。

**検索キャッシュ (search cache)**

以前の検索要求のデータと結果を保持するバッファ。

**検索結果 (search results)**

検索要求に一致する文書のリスト。

**検索索引ファイル (search index files)**

検索エンジンで索引が保管されているファイルのセット。

**検出機能 (discoverer)**

クローラー機能の 1 つで、クローラーが情報検索に使用できるデータソースを判別する機能。

**合字 (ligature)**

2 文字以上を連結することによって 1 文字として表示されるようにしたもの。例えば、ff や ffi は合字として表示できる文字です。

**高頻度ランキング (popular ranking)**

文書の検索頻度に基づいて文書の既存ランキングを上げるランキング・タイプ。

**語のステミング (word stemming)**

言語学的な正規化のプロセス。1 つのワードの異形を一般形に分解する。例えば、*connections*、*connective*、および *connected* のようなワードは *connect* に戻されます。

**コレクション (collection)**

データ・ソースと、そのクロール、解析、索引作成、検索用のオプションの集合。

**サーブレット (servlet)**

Web サーバー上で稼働し、Web クライアント要求に対する応答として動的

コンテンツを生成することにより、サーバーの機能性を拡張する Java プログラム。サーブレットは、一般的に、データベースを Web に接続するのに使用されます。

#### 索引 (index)

『フルテキスト索引』を参照。

#### 索引キュー (index queue)

処理される主索引作成および差分索引作成の要求リスト。

#### 索引付けしないディレクティブ (no-index directive)

Web ページ内のディレクティブで、そのページの内容を索引に含めないようロボット (Web クローラーなど) に指示するもの。

#### 差分索引作成 (delta index build)

エンタープライズ・サーチ・システムにおいて、新しい情報を既存の索引に追加する処理。『主索引作成』と対比。

#### シード・リスト・ページ (seed list page)

WebSphere Portal の中で、ポータルで使用可能なページへのリンクを含む XML ページ。クローラーは、シード・リストを使用してクロールする文書を識別します。シード・リスト・ページには、クロールされた文書と一緒にエンタープライズ・サーチ索引に保管されるメタデータも含まれます。

#### 識別名 (distinguished name)

ディレクトリーのエントリーを一意的に識別する名前。識別名は、コンマで分離された「属性:値 (attribute:value)」ペアで構成されます。また、デジタル証明書のエンティティを一意的に識別する名前/値ペアのセット(例: CN=個人の名前および C=国または地域)。

#### 字句類縁性 (lexical affinity)

文書内で互いに近い意味を持つ検索語間の関係。字句類縁性を使用して、結果の適合度を算出します。

#### 主索引作成 (main index build)

エンタープライズ・サーチにおいて、索引全体を作成する処理。『差分索引作成』と対比。

#### 情報抽出 (information extraction)

概念抽出のタイプの 1 つで、テキスト文書内の重要な語彙項目 (名前、用語、式など) を自動的に認識するもの。

#### 証明書 (certificate)

コンピューター・セキュリティーにおいて、公開鍵を証明書の所有者の ID に結合するデジタル文書で、それによって証明書の所有者を認証済みにすることができるもの。証明書は、認証局によって発行され、その認証局によってデジタル署名されます。

#### 処理エンジン・アーカイブ (processing engine archive)

Unstructured Information Management Architecture (UIMA) 分析エンジン、およびエンタープライズ・サーチのカスタム分析に使用するために要求されるリソースのすべてを含む .pear zip アーカイブ・ファイル。

#### ステミング (stemming)

『語のステミング』を参照。

### ストップワード (stop word)

共通に使用されるワードで、*the*、*an*、または *and* など、検索アプリケーションが無視するもの。

### ストップワードの除去 (stop word removal)

共通ワードを無視して、より関連性のある結果を戻すために、照会からストップワードを除去するプロセス。

### 正規表現アノテーター (regular expression annotator)

文書テキストで検索される厳密パターンを記述する正規表現に基づいて、テキスト文書内の情報エンティティや情報単位 (製品番号など) を検出するソフトウェア・コンポーネント。正規表現の 1 つが文書テキストの一部と一致すると、正規表現アノテーターは、一致の一部または全体を取り入れた、対応する注釈を作成します。これらの注釈の付いたテキストは、索引マッピング・ファイルを使用してエンタープライズ・サーチ索引に保管されるか、データベース・マッピング・ファイルを使用して JDBC 可能データベースに保管されます。

### 静的要約 (static summarization)

要約タイプの 1 つ。検索結果には、文書の指定および保管された要約が含まれる。『動的要約』と対比。

### 静的ランキング (static ranking)

ランキング・タイプの 1 つ。日付や、その文書を指すリンク数など、ランキングされる文書に関する係数でランキングが上がる。『動的ランキング』と対比。

### セキュリティー・トークン (security token)

コレクションの文書へのアクセス許可に使用される ID とセキュリティーに関する情報。データ・ソース・タイプによって、サポートするセキュリティー・トークンのタイプは異なる。例えば、ユーザー役割、ユーザー ID、グループ ID や、コンテンツへのアクセス制御用のその他の情報などがある。

### セグメンテーション (segmentation)

テキストを明確な字句単位に分割すること。非辞書ベースの処理には空白文字と N-gram セグメンテーションが含まれ、辞書ベースのサポートには、ワード、文、段落のセグメンテーションとレンマタイゼーションが含まれます。

### 接語 (clitic)

構文的には分離して機能するが、音声学的には別のワードに接続するワード。接語は、結合されるワードとは、接続して書かれたり、分離して書かれたりします。接語の一般的な例としては、英語における縮小語の終わりの部分が含まれます (*wouldn't* または *you're*)。

### セマンティック検索 (semantic search)

キーワード検索のタイプで、言語分析と文脈分析を統合したもの。『テキスト分析』も参照。

### ソフト・エラー・ページ (soft error page)

Web ページの 1 つのタイプで、要求された Web ページを戻せない理由を説明する情報を提供するもの。例えば、HTTP サーバーは、単純な状況コードを戻す代わりに、状況コードを詳しく説明するページを戻すことができます。

### タイプ・システム (type system)

タイプ・システムは、文書の中でテキスト分析エンジンによって発見される可能性のあるオブジェクト (フィーチャー構造) のタイプを定義する。タイプ・システムは、タイプとフィーチャーに関してすべての可能なフィーチャー構造を定義します。タイプ・システムの中に、異なるタイプをいくつでも定義できます。タイプ・システムはドメインおよびアプリケーションに固有です。

### 注釈 (annotation)

テキストのスパンに関する情報。例えば、注釈は、テキストのスパンが会社名を表すことを指示することもあり得ます。Unstructured Information Management Architecture (UIMA) では、注釈は、特別な種類のフィーチャー構造です。

### データ・ストア (data store)

文書が解析された形式で保持されるデータ構造。

### データ・ソース (data source)

文書をリトリートできるデータ・リポジトリ。Web、リレーショナルおよび非リレーショナル・データベース、およびコンテンツ・マネージメント・システムなど。

### データ・ソース・タイプ (data source type)

データ・アクセス用のプロトコルに応じたデータ・ソースのグループ。

### テーマ抽出 (theme extraction)

概念抽出のタイプの 1 つで、テキスト文書内の重要な語彙項目を自動的に認識して、文書のテーマやトピックを抽出するもの。『概念抽出』も参照。

### テキスト・セグメンテーション (text segmentation)

『セグメンテーション』を参照。

### テキスト・ベースのスコアリング (text-based scoring)

照会内の語に対する文書の適合度を表す整数値を、文書に割り当てるプロセス。整数値が大きいほど、照会への一致が緊密であることを表す。『動的ランキング』も参照。

### テキスト分析 (text analysis)

コレクションのデータの検索性を高めるために、テキストから意味やその他の情報を抽出するプロセス。『セマンティック検索』も参照。

### テキスト分析エンジン (text analysis engine)

テキスト内のコンテキストおよびセマンティック・コンテンツを検索および表すことに関与するソフトウェア・コンポーネント。

### デキュー (dequeue)

キューから項目を除去すること。

### トークナイザー (tokenizer)

テキストをスキャンし、一続きの文字をトークンとして認識できる場合に、それを判別するテキスト・セグメンテーション・プログラム。

### トークン (token)

エンタープライズ・サーチによって索引付けされる基本テキスト単位。トークンは、言語内のワードにすることもできますし、索引付けに適切な、他のテキスト単位にすることもできます。



### トークン化 (tokenization)

入力を解析してトークンを生成する処理。

### 同義語辞書 (synonym dictionary)

ユーザーがコレクションを検索するときに、その照会条件の同義語を検索できるようにする辞書。

### 動的要約 (dynamic summarization)

要約タイプの 1 つ。検索語が強調表示され、検索結果には検索している文書の概念を最もよく表す句が含まれる。『静的要約』と対比。

### 動的ランキング (dynamic ranking)

照会の条件を検索中の文書に関して分析し、結果のランクを決定するランキングのタイプ。『テキスト・ベースのスコアリング』も参照。『静的ランキング』と対比。

### ドキュメント・ハッシュ (shingle)

文から取り出される連続トークン (ワード) のストリング。例えば、「This is a very short sentence.」からの 3 ワード・ドキュメント・ハッシュ (つまり隣り合う 3 ワード) は次のとおりです。

This is a  
is a very  
a very short  
very short sentence

ドキュメント・ハッシュは、統計言語学で使用できます。例えば、2 つの異なるテキストに多くの共通するドキュメント・ハッシュが含まれている場合、これらのテキストはおそらく何らかの形で関連しています。

### 認証局 (certificate authority)

信頼できる第三者機関または企業で、デジタル署名および公開鍵と秘密鍵のペアの作成に使用されるデジタル証明書を発行するもの。認証局は、固有の証明書を付与される個人の身元を保証します。

### パーサー (parser)

エンタープライズ・サーチ・データ・ストアに追加された文書を解釈するプログラム。パーサーは、文書から情報を抽出し、索引付け、検索、取得の準備を行う。

### パーサー・サービス (parser service)

文書コレクション全体にわたって、すべての文書の解析とテキスト分析処理を行うエンタープライズ・サーチ・サービス。少なくとも 1 つのパーサー・サービスが常に実行中です。

### パーサー・ドライバー (parser driver)

エンタープライズ・サーチにおいて、パーサー・サービスに文書を供給するサービス。それぞれのコレクションに 1 つのパーサー・ドライバーがあります。コレクションのパーサー・ドライバー・サービスは、エンタープライズ・サーチ管理コンソールの中のコレクションのパーサーに対応します。

### ハイブリッド検索 (hybrid search)

ブール検索とフリー・テキスト検索を組み合わせたもの。

**発音区別符号 (diacritic)**

文字または文字の組み合わせの音価の変化を示す符号。

**パラメトリック検索 (parametric search)**

指定された範囲内の数値または属性 (日付、整数、その他のデータ・タイプなど) を含むオブジェクトを探す検索のタイプ。

**ブール検索 (Boolean search)**

1 つ以上の検索語が、AND、NOT、OR などの演算子を使って結合された検索。

**ファジー検索 (fuzzy search)**

検索語にスペルが似た語を戻す検索。

**フィーチャー・パス (feature path)**

Unstructured Information Management Architecture (UIMA) フィーチャー構造内のフィーチャーの値にアクセスするのに使用されるパス。

**フィーチャー構造 (feature structure)**

テキスト分析の結果を表す、基礎となるデータ構造。フィーチャー構造は、属性 - 値の構造をしています。各フィーチャー構造は、タイプに属します。すべてのタイプは、Java クラスと非常に類似している、有効なフィーチャーまたは属性の指定されたセットを持ちます。

**フィールド (field)**

特定のカテゴリのデータや制御情報が入力される領域。

**フィールド検索 (fielded search)**

特定のフィールドに限定された照会。

**フェデレーション (federation)**

命名システムを結合する処理。それによって、集合システムは、命名システムをスパンする複合名を処理できます。

**フェデレーテッド・サーチ (federated search)**

複数の検索サービスにわたって検索を可能にし、検索結果の統合化されたりリストを戻す検索機能。

**フリー・テキスト検索 (free text search)**

フリー・フォーム・テキストで検索語を表現した検索。

**フリー・フォーム・テキスト (free-form text)**

語や文から成る非構造化テキスト。

**フルテキスト索引 (full-text index)**

データ項目を参照して、照会条件を含む文書を検索によって見つけれられるようにするデータ構造。

**プレース (place)**

個人やグループが共同作業するために出会うポータルで、可視になる仮想ロケーション。ポータルでは、各ユーザーは、専用作業のための個人用プレースを持ち、個人やグループは、さまざまな共有スペースへのアクセス権を持ちます。そこは、パブリック・プレースにも、制限されたプレースにもなり得ます。『Lotus QuickPlace プレース』をも参照。

**プロキシ・サーバー (proxy server)**

アプリケーションまたは Web サーバーがホストする HTTP Web 要求に対

する中継として動作するサーバー。プロキシ・サーバーは、エンタープライズのコンテンツ・サーバーの代理として動作します。

#### **分析エンジン (analysis engine)**

『テキスト分析エンジン』を参照。

#### **分析結果 (analysis results)**

アナテーターが生成する情報。分析結果は、共通分析構造と呼ばれるデータ構造に書き込まれます。カスタム・テキスト分析エンジン (アナテーター) によって生成された分析結果は、エンタープライズ・サーチ索引に含めることによって、検索可能にできます。

#### **分類構造 (taxonomy)**

類似性に基づいてオブジェクトをグループに分類したもの。エンタープライズ・サーチでは、分類構造によってデータはカテゴリとサブカテゴリに編成される。『カテゴリ・ツリー』も参照。

#### **マスク文字 (masking character)**

検索語の先頭、中間、および末尾にある任意の文字を表す文字。マスク文字は通常、索引で語の異形を検索するために使用される。『ワイルドカード文字』も参照。

#### **末尾の文字 (trailing character)**

ワードにおける最後の位置を保持する文字。

#### **見出し語 (lemma)**

ワードの基本型。見出し語は、チェコ語など、大きく語形変化する言語では重要です。

#### **文字の正規化 (character normalization)**

大文字化や発音区別符号など、文字の異体形式が共通形式に合わせられる処理。

#### **モニター担当者 (monitor)**

コレクション・レベルのプロセスを監視する権限を持つエンタープライズ・サーチ・ユーザー。

#### **ユーザー・エージェント (user agent)**

Web をブラウザし、アクセスしたサイトに自身の情報を残すアプリケーション。エンタープライズ・サーチで、Web クローラーはユーザー・エージェント。

#### **用語加重検索 (weighted term search)**

一定の用語が重視される照会。

#### **要約 (summarization)**

文書の内容を簡潔に説明する非冗長文を検索結果に組み込む処理。『動的的要約』と『静的要約』も参照。

#### **ライブラリー (library)**

他のオブジェクトにディレクトリーとしてサービスを提供するシステム・オブジェクト。『Lotus Domino Document Manager ライブラリー』を参照。

#### **ランキング (ranking)**

照会によって得られた検索結果の各文書に整数値を割り当てること。検索結

果における文書の順序は、照会への適合度に基づいて決まる。ランクが高いほど、緊密な一致を表す。『動的ランキング』と『静的ランキング』も参照。

#### **ランキング調整クラス (boost class)**

検索結果の文書の相対的ランクに影響を与えることのできる仕様を含むオブジェクト。

#### **ランキング調整ワード (boost word)**

検索結果内の文書の相対的ランクに影響を与えることのできるワード。照会処理中に、ワードに事前定義したスコアに従って、ランキング調整ワードを含む文書の重要度を調整することも可能です。

#### **リモート・フェデレーター (Remote Federator)**

検索可能なオブジェクトのセットをフェデレートするサーバー・フェデレーター。

#### **リンク分析 (link analysis)**

文書間のハイパーリンクの分析に基づき、コレクション内のどのページがユーザーにとって重要なかを判別するための方法。

#### **リンクをたどらないディレクティブ (no-follow directive)**

Web ページ内のディレクティブで、そのページで検出されたリンクをたどらないようロボット (Web クローラーなど) に指示するもの。

#### **ルーム (room)**

ユーザーが、他の人々が読む文書を作成し、他の人々からのコメントに回答し、プロジェクトの状況と期限を検討することができるようにするプログラム。ユーザーは、同じルームにいる他の人々とチャットすることもできます。『Lotus QuickPlace ルーム』をも参照。

#### **ルール・ベースのカテゴリ (rule-based category)**

どの文書が、どのカテゴリと関連付けられるかを指定する規則によって作成されるカテゴリ。例えば、一定の語を含む、または含まない文書や一定の Uniform Resource Identifier (URI) パターンに一致する文書を、特定のカテゴリと関連付ける規則を定義する。

#### **レンマタイゼーション (lemmatization)**

語の原形とさまざまな文法的形式とを識別する処理。例えば、「mouse」という語を検索すると語「mice」を含む文書も検出され、「go」という語を検索すると「going」、「gone」、または「went」を含む文書も検出されます。

#### **ロー・データ・ストア (raw data store)**

クローリング済みの文書がパーサーに送られる前に保管されるデータ構造。クローラーはロー・データ・ストアに書き込み、パーサーはロー・データ・ストアから読み取ります。文書が解析されると、ロー・データ・ストアから除去されます。データ・ストアと混同しないでください。

#### **ローカル・フェデレーター (Local Federator)**

エンタープライズ・サーチ・アプリケーションにおいて、検索および索引 API によって作成されるクライアント・オブジェクトで、ユーザーが異種コレクションの集合を検索し、一連の統一された検索結果を取得できるようにするもの。

**ロボット排除プロトコル (Robots Exclusion Protocol)**

サイトのある部分をロボットが訪問しないように、Web サイト管理者が、訪問するロボットに指示できるようにするプロトコル。

**ワイルドカード文字 (wildcard character)**

検索語の先頭、中間、または末尾にある任意の文字を表す文字。

**Document Object Model (DOM)**

XML ファイルなど、構造化文書を、プログラマチックにアクセスおよび更新できるオブジェクトのツリーとして表示するシステム。

**ID 管理 (identity management)**

保護データへのアクセスを制御し、ユーザーがコレクション内の各リポジトリごとにユーザー ID とパスワードを指定しなくてもそのコレクションを検索できるようにするエンタープライズ・サーチ API の集合。

**IP アドレス (IP address)**

ネットワーク上のデバイスあるいは論理装置の固有のアドレスで、IP 標準に従うもの。

**Java Database Connectivity (JDBC)**

Java プラットフォームと広範なデータベースとの間のデータベース依存接続の業界標準。JDBC インターフェースは、SQL ベースのデータベース・アクセスに対する呼び出しレベルの API を提供します。

**Java 仮想マシン (JVM) (Java virtual machine (JVM))**

コンパイル済み Java コード (アプレットおよびアプリケーション) を実行するプロセッサのソフトウェア・インプリメンテーション。

**JavaScript**

ブラウザおよび Web サーバーで使用される Web スクリプト言語。

**JavaServer Pages (JSP)**

動的コンテンツをクライアントに戻すために、Java コードを動的に Web ページ (HTML ファイル) に組み込むことができるようにし、そのページのサービスが提供されるときに実行するサーバー・スクリプト・テクノロジー。

**Lightweight Directory Access Protocol (LDAP)**

オープン・プロトコルの 1 つで、TCP/IP を使用して、X.500 モデルをサポートするディレクトリーにアクセスできるようにするが、より複雑な X.500 ディレクトリー・アクセス・プロトコル (DAP) のリソース要件を負わないもの。例えば、LDAP を使用して、インターネットまたはイントラネット・ディレクトリーで個人、組織その他のリソースを検索できます。

**Lotus Domino Document Manager キャビネット (Domino Document Manager cabinet)**

文書を編成するのに使用される Lotus Domino Document Manager データベース。キャビネットが Lotus Domino データベースを保持します。

**Lotus Domino Document Manager ライブラリー (Domino Document Manager library)**

Lotus Domino Document Manager に対するエントリー・ポイントである Lotus Domino Document Manager データベース。

**Lotus Domino Internet Inter-ORB Protocol (IIOP)**

サーバー上で稼働し、Lotus Domino Object Request Broker と連動して、

Lotus Notes Java クラスを使用して作成される Java アプレットと Lotus Domino サーバーとの間の通信を可能にするサーバー・タスク。ブラウザ・ユーザーおよび Lotus Domino サーバーは、DIOP を使用して通信し、オブジェクト・データを交換します。

**Lotus Notes リモート・プロシージャー・コール (NRPC) (Notes remote procedure call (NRPC))**

すべての Notes-to-Notes 通信に使用される Lotus Notes の通信機構。

**Lotus QuickPlace プレース (Lotus QuickPlace place)**

地理的に分散した参加者が、構造化されてセキュアなワークスペースにおいて、プロジェクトで共同作業し、オンラインで通信することができるようになる、Lotus QuickPlace によって提供される Web の場。

**Lotus QuickPlace ルーム (Lotus QuickPlace room)**

共通の興味、および集会的作業の必要を共有する、許可されたメンバーに制限された Lotus QuickPlace プレースのパーティション化領域。

**MIME タイプ (MIME type)**

インターネット全体にわたって転送されるオブジェクトのタイプを識別するためのインターネット標準。

**N-gram セグメンテーション (n-gram segmentation)**

Unicode ベースの空白文字のセグメンテーションのように、ワードを区切るのにブランク・スペースを使用するよりはむしろ、所与の数の文字の重複シーケンスを単一文字とみなす分析方法。

**Portal Document Manager (PDM)**

チーム・コラボレーションのために、ユーザーに文書の中央リポジトリを 1 つ持たせるもの。管理者は文書を効果的に管理する能力を持ち、ユーザーが情報と対話する方法を制御できます。

**Secure Sockets Layer (SSL)**

通信プライバシーを提供するセキュリティー・プロトコル。SSL により、クライアント/サーバー・アプリケーションは通信中の盗聴、改ざん、およびメッセージ偽造を回避することができます。

**Unicode ベースの空白文字のセグメンテーション (Unicode-based white space segmentation)**

トークンと区切り文字を区別するために Unicode 文字プロパティを使用するトークン化の方式。

**Unstructured Information Management Architecture (UIMA)**

非構造化データの分析用システムをインプリメントするフレームワークを定義する IBM アーキテクチャー。

**URI (Uniform Resource Identifier)**

抽象的または物理的リソースを識別するコンパクトな文字ストリング。

**URL (Uniform Resource Locator)**

インターネットなどのネットワークを使用してアクセスできる情報リソースの固有アドレス。URL には、その情報リソースへのアクセスに使用されるプロトコルの省略名と、そのプロトコルが情報リソースを見つけるために使用する情報とが含まれます。

**Web クローラー (Web crawler)**

クローラーの 1 つのタイプで、Web 文書を検索したり、文書内のリンクをたどったりすることによって Web を探索するもの。

**XML パス言語 (XPath) (XML Path Language (XPath))**

XSLT、XQuery、および XML パーサーなどの XML 関連技術との併用目的で、ソース XML データの各部分を一意的に識別したり処置したりするように設計された言語。XPath は World Wide Web Consortium 標準です。





---

## 特記事項および商標

---

### 特記事項

本書は米国 IBM が提供する製品およびサービスについて作成したものです。

本書に記載の製品、サービス、または機能が日本においては提供されていない場合があります。日本で利用可能な製品、サービス、および機能については、日本 IBM の営業担当員にお尋ねください。本書で IBM 製品、プログラム、またはサービスに言及していても、その IBM 製品、プログラム、またはサービスのみが使用可能であることを意味するものではありません。これらに代えて、IBM の知的所有権を侵害することのない、機能的に同等の製品、プログラム、またはサービスを使用することができます。ただし、IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

IBM は、本書に記載されている内容に関して特許権 (特許出願中のものを含む) を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

〒106-8711  
東京都港区六本木 3-2-12  
日本アイ・ビー・エム株式会社  
法務・知的財産  
知的財産権ライセンス渉外

以下の保証は、国または地域の法律に沿わない場合は、適用されません。IBM およびその直接または間接の子会社は、本書を特定物として現存するままの状態を提供し、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。

この情報には、技術的に不適切な記述や誤植を含む場合があります。本書は定期的に見直され、必要な変更は本書の次版に組み込まれます。IBM は予告なしに、随時、この文書に記載されている製品またはプログラムに対して、改良または変更を行うことがあります。

本書において IBM 以外の Web サイトに言及している場合がありますが、便宜のため記載しただけであり、決してそれらの Web サイトを推奨するものではありません。それらの Web サイトにある資料は、この IBM 製品の資料の一部ではありません。それらの Web サイトは、お客様の責任でご使用ください。

IBM は、お客様が提供するいかなる情報も、お客様に対してなんら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。

本プログラムのライセンス保持者で、(i) 独自に作成したプログラムとその他のプログラム (本プログラムを含む) との間での情報交換、および (ii) 交換された情報の相互利用を可能にすることを目的として、本プログラムに関する情報を必要とする方は、下記に連絡してください。

IBM Corporation  
J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003  
U.S.A.

本プログラムに関する上記の情報は、適切な使用条件の下で使用することができませんが、有償の場合もあります。

本書で説明されているライセンス・プログラムまたはその他のライセンス資料は、IBM 所定のプログラム契約の契約条項、IBM プログラムのご使用条件、またはそれと同等の条項に基づいて、IBM より提供されます。

この文書に含まれるいかなるパフォーマンス・データも、管理環境下で決定されたものです。そのため、他の操作環境で得られた結果は、異なる可能性があります。一部の測定が、開発レベルのシステムで行われた可能性がありますが、その測定値が、一般に利用可能なシステムのものと同じである保証はありません。さらに、一部の測定値が、推定値である可能性があります。実際の結果は、異なる可能性があります。お客様は、お客様の特定の環境に適したデータを確かめる必要があります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者をお願いします。

IBM の将来の方向または意向に関する記述については、予告なしに変更または撤回される場合があります、単に目標を示しているものです。

表示されている IBM の価格は IBM が小売り価格として提示しているもので、現行価格であり、通知なしに変更されるものです。卸価格は、異なる場合があります。

本書はプランニング目的としてのみ記述されています。記述内容は製品が使用可能になる前に変更になる場合があります。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名称はすべて架空のものであり、名称や住所が類似する企業が実在しているとしても、それは偶然にすぎません。

著作権使用許諾:

本書には、様々なオペレーティング・プラットフォームでのプログラミング手法を例示するサンプル・アプリケーション・プログラムがソース言語で掲載されています。

す。お客様は、サンプル・プログラムが書かれているオペレーティング・プラットフォームのアプリケーション・プログラミング・インターフェースに準拠したアプリケーション・プログラムの開発、使用、販売、配布を目的として、いかなる形式においても、IBM に対価を支払うことなくこれを複製し、改変し、配布することができます。このサンプル・プログラムは、あらゆる条件下における完全なテストを経ていません。従って IBM は、これらのサンプル・プログラムについて信頼性、利便性もしくは機能性があることをほのめかしたり、保証することはできません。

それぞれの複製物、サンプル・プログラムのいかなる部分、またはすべての派生的創作物にも、次のように、著作権表示を入れていただく必要があります。

© (お客様の会社名) (西暦年). このコードの一部は、IBM Corp. のサンプル・プログラムから取られています。 © Copyright IBM Corp. \_年を入れる\_. All rights reserved.

この製品には次のものが含まれています。

- Oracle® Outside In Content Access, Copyright © 1992, 2008, Oracle. All rights reserved.
- IBM XSLT Processor Licensed Materials - Property of IBM © Copyright IBM Corp., 1999-2008. All Rights Reserved.

## 商標

IBM の商標については、<http://www.ibm.com/legal/copytrade.shtml> を参照してください。

以下は、それぞれ各社の商標または登録商標です。

Adobe、PostScript は、Adobe Systems Incorporated の米国およびその他の国における登録商標または商標です。

Intel、Intel (ロゴ)、Intel Inside、Intel Inside (ロゴ)、Intel Centrino、Intel Centrino (ロゴ)、Celeron、Intel Xeon、Intel SpeedStep、Itanium、Pentium は、Intel Corporation または子会社の米国およびその他の国における商標または登録商標です。

Java およびすべての Java 関連の商標およびロゴは Sun Microsystems, Inc.の米国およびその他の国における商標です。

Linux は、Linus Torvalds の米国およびその他の国における商標です。

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

UNIX は The Open Group の米国およびその他の国における登録商標です。

他の会社名、製品名およびサービス名等はそれぞれ各社の商標です。



# 索引

日本語、数字、英字、特殊文字の順に配列されています。なお、濁音と半濁音は清音と同等に扱われています。

## [ア行]

アーカイブ・ファイル  
クローラ 134  
サポートされるフォーマット 134  
URI フォーマット 135  
アクセス制御  
  現在のユーザーの検証 295  
  コレクションの無効化 324  
  シングル・サインオン・セキュリティ 300  
  説明 288  
  文書レベルのセキュリティ 294  
  ID 管理 296, 299  
  Lotus Domino での要件 314, 315  
  Windows ファイル・システムでの要件 318  
アクティブ Web サイトのモニター 339, 341  
アナテーター 161  
アプリケーション ID 291  
アラート  
  クローラ済み文書数 360  
  検索応答時間 360  
  コレクション・レベル 335, 360  
  サーバー上のフリー・スペース 362  
  索引付けされた文書 360  
  索引の制限 335  
  システム・レベル 362  
  説明 359  
  E メールでの受信 365  
  E メール・オプション 360, 362  
  SMTP サーバー構成 364  
アンカー・テキスト分析  
  グローバル分析 302  
  コレクション・セキュリティ 302  
  説明 285  
  文書の索引作成 303  
移行  
  コレクション 405  
  ルール・ベースの分類構造 405  
移行ウィザード  
  開始 405  
  コレクション 405  
  説明 405

移行ウィザード (続き)  
  デフォルトのクローラ設定 407  
  デフォルトのコレクション設定 407  
  ルール・ベースの分類構造 405  
  ログ・ファイル 408  
イベント・パブリッシング  
  説明 55  
  AIX オペレーティング・システムでのセットアップ 57  
  DB2 クローラ構成 59, 62  
  Linux オペレーティング・システムでのセットアップ 57  
  Solaris オペレーティング環境でのセットアップ 57  
  Windows でのセットアップ 58  
エラー・メッセージ  
  ドロップされた文書のログ・ファイルの表示 357  
  ログ・ファイルの表示 368  
  E メールでの受信 363, 365  
  SMTP サーバー構成 364  
エンタープライズ・サーチ  
  概要 1  
  管理コンソール 10  
  クローラ・サーバー 5  
  検索アプリケーション 12  
  検索サーバー 9  
  コンポーネント 3  
  索引サーバー 7  
  データ・フロー・ダイアグラム 13  
  パーサー 5  
  API 11  
エンタープライズ・アプリケーション  
  ESAdmin アプリケーション 322  
  ESSearchApplication アプリケーション 322  
  ESSearchServer アプリケーション 322  
エンタープライズ・サーチ  
  管理役割 288  
  クローラ・サーバー 43  
  検索サーバー 215  
  検索サーバーの開始 332  
  検索サーバーの停止 332  
  コマンド 409  
  コレクション・レベルのセキュリティ 290  
  サーバーの開始 327  
  サーバーの停止 327, 330  
  索引サーバー 193  
  セキュリティ 285  
  セッション ID 409

エンタープライズ・サーチ (続き)  
  パーサー 147  
  バックアップ 372  
  バックアップ・スクリプト 371  
  バックアップ・データからのリストア 373  
  文書レベルのセキュリティ 292  
  ポート番号構成 27  
  戻りコード 409  
  モニター 335  
  リストア・スクリプト 371  
  ログ・ファイル 359  
  Lotus Notes の統合 379  
  URI フォーマット 135  
  WebSphere Portal との統合 381  
エンタープライズ・サーチ管理者  
  シングル・サーバーでのパスワードの変更 22  
  説明 288  
  マルチサーバーでのパスワードの変更 24  
  役割の構成 289  
エンタープライズ・サーチのバックアップ 371, 372  
エンタープライズ・サーチのリストア 371  
エンタープライズ・サーチ用 HTML 資料 445  
エンタープライズ・サーチ用 PDF 資料 445  
エンタープライズ・サーチ・サーバー  
  二重 IP サポート 30  
  IP アドレスの変更 29  
  IPv6 プロトコル・サポート 31  
応答時間履歴のモニター 351  
オペレーター  
  説明 288  
  役割の構成 289

## [カ行]

改行文字の除去 190  
開始  
  移行ウィザード 405  
  エンタープライズ・サーチ・サーバー 327  
  クローラ・サーバー 338  
  検索アプリケーション 273  
  検索アプリケーション・カスタマイザー 266  
  検索サーバー 332, 351

- 開始 (続き)
  - 索引作成 350
  - データ・リスナー 354
  - パーサー 348
- 外部クローラー
  - 構成 131
  - データ・リスナー・アプリケーション 131
- 外部ソース
  - アプリケーション ID セキュリティー 291
  - 検索アプリケーションとの関連付け 282
  - 構成 279
  - 説明 279
- 鍵ストア・ファイル 274
- カスタム・テキスト分析
  - 共通分析構造の索引へのマッピング 165
  - 共通分析構造のリレーショナル・データベースへのマッピング 167
  - 索引への分析結果のマッピング 165
  - 説明 161
  - テキスト分析エンジン 162, 163
  - 分析結果のリレーショナル・データベースへのマッピング 167
  - XML エLEMENTのマッピング 164
- カテゴリ
  - カテゴリ化タイプ 151
  - カテゴリ・ツリー 151
  - 検索 148
  - 作成 152
  - サブカテゴリのネスティング 151
  - 説明 148
  - ルール・ベース 149, 151
  - URI フォーマット 135
  - WebSphere Portal からの移行 405
- カテゴリ化タイプ
  - 選択 37, 151
  - ルール・ベース 149
- カテゴリ規則
  - 構成 152
  - 文書コンテンツ 149, 152
  - URI パターン 149, 152
- カテゴリ・ツリー
  - 説明 151
  - WebSphere Portal からの移行 405
- 韓国語
  - 複合語分析 168
  - N-gram セグメンテーション 190
- 完全一致検索フィールド、説明 200
- 管理コンソール
  - インターフェース 17
  - 説明 10
  - タスクの概要 17
  - ログイン 21
- 管理者パスワード
  - シングル・サーバーでの変更 22
  - マルチサーバーでの変更 24
- 管理役割
  - エンタープライズ・サーチ管理者 288, 289
  - オペレーター 288, 289
  - 構成 289
  - コレクション管理者 288, 289
  - 説明 288
  - モニター担当者 288, 289
- キーワード、クイック・リンク 226
- 共通分析構造
  - 索引へのマッピング 165
  - 説明 161
  - リレーショナル・データベースへのマッピング 167
  - XML エLEMENTのマッピング 164
- 許可、説明 288
- クイック・リンク
  - 検索 226
  - 作成 226
  - 説明 226
  - URI フォーマット 135
- 空白文字の除去 190
- クラスター
  - WebSphere Portal 398
- グローバル Web クロール・スペース 119
- グローバル分析
  - アンカー・テキスト分析 285, 302
  - 説明 7
  - 重複文書の検出 205, 285, 290
- クローラー
  - アーカイブ・ファイル 134
  - 外部サポート 131
  - クローラーのプロパティの編集 46
  - クロール・スペースの編集 47
  - 構成の概要 43
  - 異なるクローラー・タイプの組み合わせ 43
  - 削除 47
  - 作成 45
  - シード・リスト 98
  - システム状況 338
  - 初期値 45
  - スケジューリング 43, 48
  - 説明 5
  - データ・リスナー・アプリケーション 131
  - デフォルトの移行設定 407
  - プラグイン 132
  - 文書レベルのセキュリティ 292
  - 文書レベルのセキュリティを有効にする 43
  - ベースとなる値 43
- クローラー (続き)
  - モニター 338
  - Content Edition 49, 50, 52
  - DB2 55
  - DB2 Content Manager 65
  - Domino Document Manager 70
  - Exchange Server 73, 314
  - JDBC データベース 74, 76, 78
  - Lotus Notes 83, 86
  - NNTP 82
  - QuickPlace 94
  - UNIX ファイル・システム 100
  - URI フォーマット 135
  - Web 101
  - Web Content Management 124, 127
  - WebSphere Portal 126, 127
  - Windows ファイル・システム 129
- クローラーのプロパティ
  - 説明 5
  - 編集 46
- クローラー履歴レポート
  - サイト・レポート 343
  - 作成 343
  - 説明 339
  - HTTP 状況コード・レポート 343
- クローラー・サーバー
  - 開始 327, 338
  - 停止 330, 338
- クローラー・タイプ
  - 単一コレクション内に組み合わせる 43
  - ベースとなる値 43
- クローラー・プラグイン 132
- 「クロール」 ページ、説明 17
- クロールされた文書の日付
  - Web クローラーの構成 122
- クロール速度のモニター 342
- クロール・スペース
  - アラート 360
  - 説明 5
  - 編集 47
  - Web クローラーの構成 106
- 言語
  - 検索 187
  - サポートされる 187, 188
  - 自動検出 188
  - 2 文字コード 187
- 言語サポート
  - カスタム・テキスト分析 161
  - 空白文字の除去 190
  - 言語検出 188
  - 言語コード 187
  - コード・ページ検出 189
  - ストップワード辞書 219
  - セマンティック検索 161, 169
  - 同義語辞書 217

- 言語サポート (続き)
  - ネイティブ XML 検索 169
  - ランキング調整ワード辞書 233
  - ロケール 187
  - N-gram セグメンテーション 190
- 検索
  - カテゴリー 148
  - クイック・リンク 226
  - コレクション 200
  - HTML 文書 158, 159
  - XML 文書 154, 155, 164
  - 「検索」ページ、説明 17
- 検索アプリケーション
  - アクセス 273
  - アプリケーション ID 291
  - 開始 273
  - 外部ソースとの関連付け 282
  - カスタマイズ 266
  - カスタム 243
  - コレクションとの関連付け 244
  - コレクション・レベルのセキュリティ
    - ー 291
  - サンプル 243, 245
  - 説明 12
- 検索アプリケーションのカスタマイズ
  - 264, 266
- 検索アプリケーション・カスタマイザー
  - 開始 266
  - config.properties ファイル 266
- 検索応答時間
  - アラート 360
  - モニター 351
- 検索オプション
  - 完全一致 200
  - 検索結果の 200
  - ソート可能フィールド 200
  - パラメトリック検索 200
  - フィールド検索 200
  - フリー・テキスト検索 200
  - 文書コンテンツ 200
- 検索および索引 API 11, 243
- 検索キャッシュ
  - 構成 216
  - 説明 216
- 検索結果
  - グループ化 211, 212
  - サマリー 224
  - サマリーのカスタマイズ 224
  - 事後フィルター 292
  - 事前フィルター 292
  - 縮小表示 211, 212, 325
  - 静的ランキング 230
  - 説明 229
  - テキスト・ベースのスコアリング 229
  - 動的要約 223
  - 動的ランキング 229
- 検索結果 (続き)
  - ランキング 236
  - ランキング調整クラスの構成 237, 239, 240, 241
  - ワイルドカード文字 205
  - ワイルドカード文字展開 208
  - URI パターン構成 236
- 検索結果のランキング
  - 静的 230, 231
  - 説明 229
  - テキスト・ベースのスコアリング 229
  - デフォルト値の復元 231
  - 動的 229
  - ランキング調整クラス 237, 239, 240, 241
  - ランキング調整ワード辞書 235
  - runtime.properties ファイル 231
  - URI パターン 236
- 検索結果フィールド、説明 200
- 検索サーバー
  - 応答時間履歴 351
  - 開始 327, 332, 351
  - 検索キャッシュ 216
  - 高頻度の照会 351, 353
  - 最新の照会 351, 353
  - 辞書の再デプロイ 222
  - システム状況 351
  - 照会カウンターの計算 353
  - ストップワード辞書 219
  - ストップワード辞書の関連付け 221
  - 説明 9, 215
  - 停止 330, 332, 351
  - 同義語辞書 217, 218
  - 同義語辞書の関連付け 219
  - プロキシー・サーバー構成 276
  - モニター 351, 353
  - ランキング調整ワード辞書 233
  - ランキング調整ワード辞書の関連付け 234, 235
  - HTTPS の構成 274
  - SSL の構成 274
- 検索センター、WebSphere Portal
  - 説明 381
  - バージョン 6、エンタープライズ・サーチとの統合 392
- 検索バー、WebSphere Portal
  - バージョン 5.1、エンタープライズ・サーチへのリダイレクト 387
  - バージョン 6、エンタープライズ・サーチへのリダイレクト 394
- 検索ポータルレットのデプロイメント
  - es.wp5.install.jar ファイル 384
  - es.wp6.install.jar ファイル 389, 399
  - wp5\_install スクリプト 384
  - wp5\_uninstall スクリプト 388
  - wp6\_cluster\_install スクリプト 399
- 検索ポータルレットのデプロイメント (続き)
  - wp6\_cluster\_uninstall 403
  - wp6\_install スクリプト 389
  - wp6\_uninstall スクリプト 397
- 検出 5
- 検証、現行のクレデンシャルの 295, 314, 315, 318
- コード・ページ
  - サポートされる 189
  - 自動検出 189
- 高頻度の照会
  - 計算 353
- 高頻度の照会のモニター 351
- 「構文解析」ページ、説明 17
- コマンド、エンタープライズ・サーチ 409
- コレクション
  - アプリケーション ID セキュリティー 291
  - アンカー・テキスト・セキュリティ 302
  - 検索 200
  - 検索アプリケーションとの関連付け 244
  - 検索サーバー 215
  - 構文解析 147
  - 「コレクション」ビューを使用した作成 37
  - コレクション・ウィザードを使用した作成 35
  - サイズの見積もり 37
  - 削除 40
  - 作成方法 35
  - システム状況 336
  - セキュリティ 290
  - 説明 3
  - 重複文書セキュリティ 290
  - 重複文書の検出 205
  - デフォルトの移行設定 407
  - ドラフト 35
  - フェデレーション 35
  - 文書レベルのアクセス制御の迂回 324
  - 編集 38
  - モニター 336
  - リソースの見積もり 335
  - ID の判別 41
  - MigrationWizard.log ファイル 405
  - WebSphere Portal からの移行 405
- コレクション ID 41
- コレクション ID、構文規則 37
- 「コレクション」ビュー
  - コレクションの作成 37
  - 説明 17
- コレクション管理者
  - 説明 288

コレクション管理者 (続き)  
 役割の構成 289  
コレクション・ウィザード 35  
コレクション・レベルのセキュリティー  
 アプリケーション ID 291  
 アンカー・テキスト分析 302  
 説明 285, 290  
 重複文書の検出 290  
 有効化 37

## [サ行]

サーバー・モード、Content Edition リポ  
 ジトリー 52  
最近クローラされた URL のモニター  
 339  
最小再クローラ間隔 111  
最新の照会  
 計算 353  
最新の照会のモニター 351  
最大再クローラ間隔 111  
サイトの詳細レポート  
 作成 343  
 説明 339  
索引  
 アラート 360  
 アンカー・テキスト (anchor text) 303  
 キューからの削除 351  
 縮小表示した URI 200, 211, 212  
 スケジューリング 194  
 スケジュールの変更 195  
 スケジュールを使用可能にする 195,  
 350  
 スケジュールを使用不可にする 195,  
 350  
 説明 7, 193  
 同時作成 196  
 並行作成 196  
 変更の検出 198  
 モニター 350, 351  
 有効範囲 200, 209  
 ワイルドカード文字 200, 205, 208  
 ワイルドカード文字の影響 208  
 startIndexBuild コマンド 198  
 URI の除去 200, 213  
 URI フォーマット 135  
 「索引」ページ、説明 17  
索引キュー 351  
索引サーバー  
 開始 327  
 停止 330  
索引作成 193  
 開始 350  
 システム状況 351  
 スケジューリング 195  
 説明 193

索引作成 (続き)  
 停止 199, 350, 351  
 同時 196  
 並行 196  
 変更の検出 198  
 startIndexBuild コマンド 198  
索引スケジュールを使用可能にする 195  
索引スケジュールを使用不可にする 195  
索引付けしないディレクティブ  
 構成 122  
 説明 121  
索引展開  
 索引サイズへの影響 208  
 索引作成時間への影響 208  
 説明 205  
索引の並行作成 196  
削除  
 キューから索引を 351  
 クローラー 47  
 コレクション 40  
作成  
 クイック・リンク 226  
 クローラー 45  
 コレクション 35, 37  
 有効範囲 210  
 ルール・ベースのカテゴリ 152  
 HTML 検索フィールド 159  
 Web クローラー・レポート 343  
 XML 検索フィールド 155  
差分索引  
 スケジューリング 194  
 説明 7, 193  
 並行作成 196  
 変更の検出 198  
サマリー  
 カスタマイズ 224  
 動的 223  
サンプル検索アプリケーション  
 検索機能 243, 245  
 セキュリティーの使用不可化 322  
 説明 12, 245  
 デフォルトのデプロイメント 264  
 複製 268  
 config.properties ファイル 246, 264  
 HTTPS の施行 274  
 SSL の施行 274  
シード・リスト・クローラー  
 構成 98  
 URI フォーマット 135  
 WebSphere Portal との統合 381, 396  
 「システム」ビュー、説明 17  
システム状況  
 クローラー 338  
 検索サーバー 351  
 コレクション 336  
 索引作成 351

システム状況 (続き)  
 パーサー 348  
 Web クローラー 339  
システムのリストア 371, 373  
システム・バックアップ 371, 372  
システム・リソース  
 見積もり 335  
システム・リソースの見積もり 335  
自動検出  
 言語 188  
 コード・ページ 189  
縮小表示した URI  
 構成 212  
 セキュリティーの制約事項 325  
 説明 211  
縮小表示の検索結果  
 構成 212  
 セキュリティーの制約事項 325  
 説明 211  
主索引  
 スケジューリング 194  
 説明 7, 193  
 並行作成 196  
 変更の検出 198  
上位結果  
 バー・チャート 270  
照会展開  
 索引サイズへの影響 208  
 索引作成時間への影響 208  
 説明 205  
照会の検証 295  
照会ログ構成 367  
除去、索引から URI を 213  
資料  
 検索 445  
 HTML 445  
 PDF 445  
シングル・サインオン・セキュリティー  
 構成 301  
 ID 管理 300  
スクリプト  
 esbackup.bat 372  
 esbackup.sh 372  
 escrcm.sh 66  
 escrcm.vbs 68  
 escrdb2.sh 57  
 escrdb2.vbs 58  
 escrnote.sh 87  
 escrnote.vbs 89  
 escrvbr.sh 52  
 escrvbr.vbs 53  
 esrestore.bat 373  
 esrestore.sh 373  
 startectl 373  
スケジューリング  
 クローラー 43, 48



スケジューリング (続き)  
索引作成 194, 195  
ストップワード辞書  
コレクションとの関連付け 221  
再デプロイ 222  
システムへの追加 221  
説明 219  
スレッド  
パーサー 168  
Web クローラー 341  
スレッドの詳細のモニター 339  
静的ランキング  
移行されたコレクションにおける 405  
コレクションの有効化 37  
説明 230  
セキュリティー  
アクセス制御 288  
アンカー・テキスト分析 302  
エンタープライズ・アプリケーション  
に対する使用不可化 322  
エンタープライズ・サーチの使用可能  
化 304  
管理役割 289  
グローバル、WebSphere Application  
Server 304, 305  
クローラー・プラグイン 132  
検索アプリケーション ID 291  
検索のための HTTPS 構成 274  
検索のための SSL 構成 274  
コレクションの有効化 37, 285  
コレクション・レベル 290, 324  
縮小表示の検索結果 325  
シングル・サーバー・セットアップ  
307  
シングル・サインオン・サポート 300  
説明 285  
重複文書の検出 290  
認証 288, 322  
文書レベル 292, 294, 295, 301, 324  
文書レベルのアクセス制御の迂回 324  
マルチサーバー・セットアップ 308  
ユーザー・プロファイル 299  
ID 管理 296, 301  
LDAP ユーザー・レジストリー 305  
Lotus Domino 文書 314, 315  
WebSphere グローバル・セキュリティー  
 322  
Windows ドメイン 318  
「セキュリティー」ビュー、説明 17  
セキュリティー・トークン  
クローラー構成 294  
コレクションの無効化 324  
文書レベルのセキュリティー 294  
セッション ID、エンタープライズ・サー  
チ 409

セットアップ・スクリプト  
WebSphere Portal 383  
セマンティック検索 161, 164, 169  
ソート可能フィールド  
数値ソート 200  
ストリング・ソート 200  
ソフト・エラー・ページ、Web クローラ  
ー 112  
**[夕行]**  
高い再呼び出し照会  
説明 237  
デフォルトのランキング調整要因 241  
タスクの概要、管理コンソール 17  
単純テキスト言語 187  
中国語  
改行文字の除去 190  
N-gram セグメンテーション 190  
重複文書の検出  
グローバル分析 205, 290  
セキュリティーの使用可能化 290  
説明 205, 285  
ランキング調整クラスの構成 237  
直接モード、Content Edition リポジトリ  
 50  
データ・ソース・タイプ  
エンタープライズ・サーチでサポート  
されている 2  
外部サポート 2, 11  
リレーショナル・データベース 55  
CA-Datacom データベース 55  
Content Edition リポジトリ 49, 50,  
52  
DB2 Content Manager 項目タイプ 65  
DB2 for iSeries データベース 55  
DB2 for z/OS 55  
DB2 データベース 55, 74, 279  
Domino Document Manager データベー  
ス 70  
Exchange Server のパブリック・フォ  
ルダ 73  
IMS データベース 55  
Informix データベース 55  
JDBC データベース 74, 76, 78, 279  
Lotus Notes データベース 83, 86  
Lotus Quickr for Domino 94  
Lotus Quickr for WebSphere Portal 98  
NNTTP ニュースグループ 82  
Oracle データベース 55, 74, 279  
QuickPlace データベース 94  
Software AG Adabas データベース  
55  
SQL Server データベース 55, 74  
Sybase データベース 55  
UNIX ファイル・システム 100

データ・ソース・タイプ (続き)  
VSAM データベース 55  
Web Content Management サイト 124  
Web サイト 101  
WebSphere Portal サイト 126  
Windows ファイル・システム 129  
データ・フロー、エンタープライズ・サー  
チ・システム 13  
データ・リスナー  
構成 131  
再始動 131, 354  
モニター 354  
停止  
エンタープライズ・サーチ・サーバー  
327, 330  
クローラー・サーバー 338  
検索サーバー 332, 351  
索引作成 350, 351  
パーサー 348  
テキスト処理  
アノテーター 161  
共通分析構造 161  
テキスト分析エンジン 161  
テキスト分析 (text analysis)  
共通分析構造 165, 167  
テキスト分析エンジン 162, 163  
XML エLEMENTのマッピング 164  
テキスト分析エンジン  
共通分析構造のリレーショナル・デー  
タベースへのマッピング 167  
コレクションとの関連付け 163  
索引への分析結果のマッピング 165  
システムへの追加 162  
説明 161  
分析結果のリレーショナル・データベ  
ースへのマッピング 167  
XML エLEMENTのマッピング 164  
テキスト・ベースのスコアリング 229  
デフォルトの検索アプリケーション 264  
同義語辞書  
コレクションとの関連付け 219  
再デプロイ 222  
システムへの追加 218  
説明 217  
同時索引作成 196  
動的要約 223  
動的ランキング 229  
トラステッド・サーバー構成 315  
ドロップされた文書  
説明 355  
に関するレポート 356  
のログ・ファイル 357  
ログ・ファイルの構成 355

## [ナ行]

### 日本語

改行文字の除去 190

N-gram セグメンテーション 190

### 認証

エンタープライズ・アプリケーション  
対する使用不可化 322

説明 288

ネイティブ XML 検索 169

## [ハ行]

### パーサー

改行文字の除去 190

開始 348

拡張子のないファイル 176

空白文字の除去 190

言語検出 188

言語処理 187

コード・ページ検出 189

構文解析、文書タイプの 173

サポートされる Stellent 文書タイプ  
182

サポートされる言語 187

システム状況 348

スレッド 168

説明 5, 147

データ分析タスク 5

停止 348

ネイティブ XML 検索 169

パーサー・サービスの文書タイプ  
174, 176

パーサー・タイプの選択 171

複合語分析 168

不明な文書タイプ 176

文書フォーマットの検出 171

モニター 348

ASCII パーサー 176

HTML 置換規則 177, 178

N-gram セグメンテーション 190

Stellent パーサー用の文書タイプ 179

### パーサー・サーバー

スレッド構成 168

### バー・チャート

上位結果 270

上位結果の Java クラス 270

### パスワード、エンタープライズ・サーチ管

理者 22, 24

### パスワード保護された Web サイト

フォーム・ベースの認証 115

HTTP 基本認証 114

### バックアップ・スクリプト

実行中 372

説明 371

### パラメトリック・フィールド

数値ソート 200

説明 200

### 低い再呼び出し照会

説明 237

デフォルトのランキング調整要因 241

### 表意文字言語

表示

ドロップされた文書のログ・ファイル  
357

ログ・ファイル 368

URI の詳細 336

ファイアウォール、Exchange Server 文書  
のクローラ 314

### ファイル拡張子

コレクション・パーサーによってサポ  
ートされている 174, 176

サポートされる Stellent パーサー  
182

除外、Web クロール・スペースから  
106

Stellent パーサーでサポートされてい  
る 179

フィールド、ランキング調整クラスへのマ  
ッピング 239

### フィールド検索

ストリング・ソート 200

説明 200

フェデレーテッド・コレクション 35

フォーム・ベースの認証 114, 115

複合語、構文解析 168

複合テキスト言語 187

複数の構造化されたテーブルのプラグイン  
76, 78

### 複製

クローラー 43

検索アプリケーション 268

不明な文書タイプ 176

### プラグイン

複数の構造化されたテーブルのクロー  
ラ 76, 78

JDBC データベース・クローラー 76,  
78

プラグイン、クローラーの 132

フリー・スペースのアラート 362

フリー・テキスト検索、説明 200

プロキシー・サーバー 116

検索サーバー構成 276

文書コンテンツ、説明 200

### 文書サマリー

カスタマイズ 224

プロパティの編集 224

### 文書タイプ

検出 171

構文解析 173

パーサー・サービスの 174, 176

### 文書タイプ (続き)

Stellent パーサーでサポートされてい  
る 182

Stellent パーサー用 179

### 文書トラッキング

使用可能化 355

使用不可化 355

説明 355

レポート 356

ログ・ファイル 357

ログ・ファイルの構成 355

### 文書の重要度

移行されたコレクションにおける 405

コレクションの有効化 37

静的 230

デフォルト値の復元 231

ランキング調整クラス 237, 241

ランキング調整ワード辞書 235

URI パターン 236

### 文書ランキング

デフォルト値の復元 231

### 文書レベルのセキュリティ

クローラー構成 43

クローラー・プラグイン 132

現行のクレデンシャルの検証 295

索引付きアクセス制御 294

事後フィルターの結果 292

事前フィルターの結果 292

シングル・サインオン・サポート 300

セキュリティ・トークン 294

説明 285, 292

ユーザー・プロファイル 299

リアルタイムの検証 295

ID 管理 296, 301

Lotus Domino 文書 315

Lotus Domino 文書の 314

Windows ファイル・システムでの  
318

分類構造、WebSphere Portal からの移行  
405

### 編集

クローラーのプロパティ 46

クロール・スペース 47

検索アプリケーションのプロパティ  
246, 264

コレクション 38

データ・リスナー・アプリケーション  
131

ポート番号、エンタープライズ・サーチ  
27

### ポートレット

エンタープライズ・サーチ 381

説明 381

Lotus Quickr ソース用のセットアップ  
396

WebSphere Portal 5.1 からの除去 388

ポートレット (続き)

- WebSphere Portal 5.1 との統合 384
- WebSphere Portal 6 との統合 389, 396
- WebSphere Portal クラスタからの除去 403
- WebSphere Portal クラスタとの統合 399
- WebSphere Portal バージョン 6 からの除去 397
- 本製品のアクセシビリティ機能 447

## [マ行]

マッピング

- 共通分析構造のリレーショナル・データベースへの 167
- 共通分析構造を 索引へ 165
- 分析結果のリレーショナル・データベースへの 167
- ランキング調整クラスに対するフィールドの 239
- HTML 検索フィールド 159
- XML エレメントの共通分析構造への 164
- XML 検索フィールド 155
- マルチバイト・エンコード 189
- メタデータ・フィールド、データ収集のバー・チャート 270
- 戻りコード、エンタープライズ・サーチ 409
- モニター
  - エンタープライズ・サーチ 335
  - 応答時間履歴 351
  - クローラー 338
  - 検索サーバー 351, 353
  - 高頻度の照会 351, 353
  - コレクション 336
  - 最新の照会 351, 353
  - データ・リスナー 354
  - ドロップされた文書 356, 357
  - パーサー 348
  - ログ・ファイル 368
  - URI の詳細 336
  - Web クローラー 339
  - Web クローラーのクロール速度 342
  - Web クローラーのスレッドの詳細 341
  - Web クローラー・アクティブ・サイト 341
- 「モニター」ビュー、説明 17
- モニター担当者
  - 説明 288
  - 役割の構成 289

## [ヤ行]

- ユーザー・エージェント 102
- ユーザー・セキュリティ・コンテキスト・ストリング 296
- ユーザー・プロファイル
  - 構成 301
  - 説明 299
- 有効範囲
  - 検索 209
  - 作成 210
  - 説明 209
  - URI パターン 209, 210
  - URI フォーマット 135

## [ラ行]

- ランキング調整クラス
  - 構成 239, 240
  - 説明 237
  - 高い再呼び出し照会 237, 241
  - 重複文書の検出 237
  - デフォルト値 241
  - 低い再呼び出し照会 237, 241
  - フィールドのマップ 239
- ランキング調整要因
  - ランキング調整クラスの 239, 241
  - ランキング調整クラスの構成 237, 240
  - ランキング調整ワード辞書の 233
  - URI パターンの 236
- ランキング調整ワード辞書
  - コレクションとの関連付け 235
  - 再デプロイ 222
  - システムへの追加 234
  - 説明 233
- リストア・スクリプト
  - 実行中 373
  - 説明 371
- リンクをたどらないディレクティブ
  - 構成 122
  - 説明 121
- ルール・ベースのカテゴリー
  - カテゴリー化タイプの選択 151
  - 作成 152
  - 説明 149
- ルール・ベースの分類構造、WebSphere Portal からの移行 405
- ローカル・ユーザー・セキュリティ、QuickPlace クローラー 317
- 「ログ」ページ、説明 17
- ログイン、管理コンソールへの 21
- ログ・ファイル
  - 移行ウィザード 408
  - サイズ構成 367
  - 最大サイズ 363

ログ・ファイル (続き)

- 重大度レベル 363
- 循環 363
- 照会ログ 367
- 説明 359
- デフォルトの場所 359
- ドロップされた文書の表示 357
- 表示 368
- フィルター 368
- 文書トラッキングに対する 355
- モニター 357, 368
- E メール・オプション 365
- SMTP サーバー構成 364
- ロケール
  - 検索 187
  - 構文解析 187
- ロボット排他プロトコル
  - ユーザー・エージェント ID 102
  - Web クローラーによる準拠 103

## [ワ行]

- ワイルドカード文字
  - 索引展開 205, 208
  - 照会展開 205, 208
  - 照会内の 205

## [数字]

- 1 バイト・エンコード 189

## A

- AdminLinkBarInclude.jsp ファイル 387
- AIX オペレーティング・システム
  - イベント・パブリッシング構成 57
  - Content Edition クローラーの構成 52
  - DB2 Content Manager クローラー構成 66
  - DB2 クローラー構成 57
  - Lotus Domino Document Manager クローラー構成 87
  - Lotus Notes クローラー構成 87
  - QuickPlace クローラー構成 87
- API
  - 検索および索引 11, 243
  - 説明 11
- ASCII パーサー 176

## B

- banner.jspf ファイル 394
- banner\_searchControl.jspf ファイル 394
- bos.iocp.rte モジュール 92

## C

CCLServer\_date.log ファイル 27  
ccl.properties ファイル 29  
config.properties ファイル 307, 308  
    カスタマイズ 266  
    複製 268  
    プロパティの説明 246  
    編集 264  
Content Edition クローラー  
    構成 49  
    サーバー・モード 52  
    直接モード 50  
    AIX オペレーティング・システムでの  
        セットアップ 52  
    Linux オペレーティング・システムで  
        のセットアップ 52  
    Solaris オペレーティング環境でのセッ  
        トアップ 52  
    URI フォーマット 135  
    Windows でのセットアップ 53  
cookies.ini ファイル  
    構成 118  
    説明 117  
    フォーマット 118  
Cookie、Web クロールの  
    構成 118  
    説明 117  
    フォーマット 118  
crawler\_rdb\_plugin.xml ファイル 78  
crawl.rules ファイル 119

## D

DB2 Content Manager クローラー  
    構成 65  
    AIX オペレーティング・システムでの  
        セットアップ 66  
    Linux オペレーティング・システムで  
        のセットアップ 66  
    Solaris オペレーティング環境でのセッ  
        トアップ 66  
    URI フォーマット 135  
    Windows でのセットアップ 68  
DB2 Content Manager の Java コネクター  
    66, 68  
DB2 クローラー  
    イベント・パブリッシング 55  
    イベント・パブリッシング構成 57,  
        58  
    構成 55  
    AIX での WebSphere MQ インストー  
        ル 57  
    AIX でのセットアップ 57  
    Linux での WebSphere MQ インスト  
        ール 57

DB2 クローラー (続き)  
    Linux でのセットアップ 57  
    Solaris での WebSphere MQ インスト  
        ール 57  
    Solaris でのセットアップ 57  
    URI フォーマット 135  
    WebSphere II Classic Federation 64  
    WebSphere II Event Publisher Edition  
        構成 59  
    WebSphere MQ 構成 62  
    Windows での WebSphere MQ インス  
        トール 58  
    Windows でのセットアップ 58  
DB2 データベース  
    外部ソースとしてアクセス 279  
    DB2 クローラーでのアクセス 55  
    JDBC データベース・クローラーでア  
        クセス 74  
Default.jsp ファイル 387, 394  
DIIOP プロトコル、クローラー構成 91  
Directory Assistance 構成 318  
distinctRecentQueryCheck パラメーター  
    353  
Dublin Core エレメント 159

## E

E メール通知  
    アラート 365  
    メッセージ 365  
    SMTP サーバー構成 364  
EAR ファイル  
    ESAdmin アプリケーション 322  
    ESSearchApplication アプリケーション  
        322  
    ESSearchServer アプリケーション 322  
esadmin startSearch command 332  
esadmin stopIndex コマンド 199  
esadmin stopSearch command 332  
esadmin system startall コマンド 409  
esadmin system stopall コマンド 409  
ESAdmin アプリケーション  
    セキュリティの使用不可化 322  
    ログイン 21  
esadmin コマンド 409  
esapi.jar ファイル 384, 389, 399  
esbackup.bat スクリプト 372  
esbackup.sh スクリプト 372  
exchangeproxypw コマンド 276  
exchangepw スクリプト 22, 24  
eschangetrustpw コマンド 274  
eschangewaspw コマンド  
    シングル・サーバー構成 307  
    マルチサーバー構成 308  
escrcm.sh スクリプト 66  
escrcm.vbs スクリプト 68

escrdb2.sh スクリプト 57  
escrdb2.vbs スクリプト 58  
escrnote.sh スクリプト 87  
escrnote.vbs スクリプト 89  
escrvbr.sh スクリプト 52  
escrvbr.vbs スクリプト 53  
ESPACServer.ear ファイル 384, 389, 399  
esrestore.bat スクリプト 373  
esrestore.sh スクリプト 373  
ESSearchAdapter.ear ファイル 384  
ESSearchApplication アプリケーション  
    開始 273  
    セキュリティの使用不可化 322  
    config.properties ファイル 246, 264  
ESSearchPortlet.war ファイル 384, 389,  
    399  
ESSearchRegistrationPortlet.war ファイル  
    384  
ESSearchServer アプリケーション  
    再始動 264, 266  
    セキュリティの使用不可化 322  
es.cfg ファイル 22, 24, 29, 30, 274, 307,  
    308  
es.search.provider.jar ファイル 389, 399  
es.security.jar ファイル 384, 389, 399  
es.wp5.install.jar ファイル 384  
es.wp6.install.jar ファイル 389, 399  
ES\_INSTALL\_ROOT、説明 22, 24  
ES\_NODE\_ROOT、説明 22, 24  
es\_special\_field.default\_field 予約フィー  
    ルド 241  
es\_special\_field.default\_metadata\_field 予約  
    フィールド 241  
es\_special\_field.regular\_text 予約フィー  
    ルド 241  
Exchange Server クローラー  
    構成 73  
    セキュア文書 314  
    URI フォーマット 135

## F

followindex.rules ファイル  
    構成 122  
    説明 121

## G

global.rules ファイル 119

## H

HTML 検索フィールド  
    エレメントのマッピング 158, 159  
    作成 159

HTML 検索フィールド (続き)  
説明 158  
Dublin Core エlement 159

HTML 置換規則 177, 178

HTML 文書  
検索 158, 159  
構文解析 177, 178  
置換規則 177, 178

HTTP 基本認証 114

HTTP 状況コード  
Web クローラーが受け取った 343  
Web クローラー・レポート 343

HTTP プロキシ・サーバー 116

HTTPS、検索サーバー構成 274

## I

ID 管理  
グループ抽出 296  
構成 301  
使用不可化 296  
シングル・サインオン・サポート 300  
説明 296  
ユーザー・セキュリティ・コンテキ  
スト 296  
ユーザー・プロファイル 299  
XML 照会ストリング 296

IOCP、クローラー構成 92

IP アドレス  
二重サポート 30  
IPv6 サポート 31  
Loopback Adapter 30

IP アドレス、変更 29

IPv6 プロトコル 31

I/O 完了ポート・モジュール、クローラー  
構成 92

## J

JavaScript サポート、Web クローラーの  
105

JDBC 外部ソース

構成 279  
削除 279  
編集 279  
JDBC ドライバー 279

JDBC データベース・クローラー

構成 74  
サポートされるドライバー 74  
複数の表のクローリング 76, 78  
複数の表をクローリングするプラグイン  
76, 78

URI フォーマット 135

JDBC ドライバー

JDBC 外部ソースの 279

JDBC ドライバー (続き)

JDBC データベース・クローラーの  
74

## L

LDAP 外部ソース

構成 279  
削除 279  
編集 279

LDAP ユーザー・レジストリー 305

Linux オペレーティング・システム  
イベント・パブリッシング構成 57  
Content Edition クローラーの構成 52  
DB2 Content Manager クローラー構成  
66  
DB2 クローラー構成 57  
Lotus Domino Document Manager クロ  
ーラー構成 87  
Lotus Notes クローラー構成 87  
QuickPlace クローラー構成 87  
Solaris オペレーティング環境  
イベント・パブリッシング構成 57

Loopback Adapter の構成 30

Lotus Domino Document Manager クロー  
ラー  
構成 70  
AIX オペレーティング・システムでの  
セットアップ 87

DIIOOP プロトコル構成 91  
IOCP 構成 92

Linux オペレーティング・システムで  
のセットアップ 87

NRPC プロトコル 87, 89  
Solaris オペレーティング環境でのセッ  
トアップ 87  
URI フォーマット 135  
Windows でのセットアップ 89

Lotus Domino ドメイン 314, 315

Lotus Domino トラストド・サーバー  
315

Lotus Domino ユーザー構成、QuickPlace  
クローラー 317

Lotus Notes  
エンタープライズ・サーチとの統合  
379  
検索バー 379  
プラグイン更新サイト 379  
プラグインのインストール 379

Lotus Notes クローラー  
検証、現行のクレデンシャルの 315  
構成 83  
使用のヒント 86  
フィールド・マッピング規則 86

Lotus Notes クローラー (続き)

文書レベルのセキュリティ構成 314  
AIX オペレーティング・システムでの  
セットアップ 87

DIIOOP プロトコル構成 91

IOCP 構成 92

Linux オペレーティング・システムで  
のセットアップ 87

Lotus Domino トラストド・サーバ  
ー 315

NRPC プロトコル 87, 89

Solaris オペレーティング環境でのセッ  
トアップ 87

URI フォーマット 135

Windows でのセットアップ 89

Lotus Notes 用のプラグイン

インストール 379  
更新サイト 379

Lotus Quickr

シード・リスト・クローラーの構成  
98

QuickPlace クローラー構成 94

WebSphere Portal との統合 381, 396

## M

MIME タイプ、Web クローリング・スペース  
に入れる 106

## N

newHtmlTagReplacement パラメーター  
177

NNTP クローラー、構成 82

nodes.ini ファイル 29, 274

NRPC プロトコル、クローラー構成 87,  
89

N-gram セグメンテーション 190

## O

OmniFind Enterprise Edition

概要 1  
管理コンソール 10  
クローラー・サーバー 5  
検索アプリケーション 12  
検索サーバー 9  
コマンド 409  
コンポーネント 3  
索引サーバー 7  
シングル・サーバーでのパスワードの  
変更 22  
セッション ID 409  
データ・フロー・ダイアグラム 13  
二重 IP サポート 30

OmniFind Enterprise Edition (続き)  
パーサー 5  
ポート番号構成 27  
マルチサーバーでのパスワードの変更  
24  
戻りコード 409  
API 11  
IP アドレスの変更 29  
IPv6 プロトコル・サポート 31  
Lotus Notes の統合 379  
WebSphere Portal との統合 381  
Oracle データベース  
外部ソースとしてアクセス 279  
DB2 クローラーでのアクセス 55  
JDBC データベース・クローラーでア  
クセス 74

## P

parserdriver.collection.properties ファイル  
177  
parserTypes.cfg ファイル 171, 174, 176

## Q

QuickPlace クローラー  
構成 94  
ローカル・ユーザー・セキュリティー  
317  
AIX オペレーティング・システムでの  
セットアップ 87  
DIIOOP プロトコル構成 91  
Directory Assistance 構成 318  
IOCP 構成 92  
Linux オペレーティング・システムで  
のセットアップ 87  
Lotus Domino ユーザー構成 317  
NRPC プロトコル 87, 89  
Solaris オペレーティング環境でのセッ  
トアップ 87  
URI フォーマット 135  
Windows でのセットアップ 89

## R

removeCjNewlineChars オプション 190  
removeCjNewlineCharsMode オプション  
190  
robots.txt ファイル  
ユーザー・エージェント ID 102  
Web クローラーによる準拠 103  
runtime-generic.properties ファイル 224,  
353

## S

SearchBarInclude.jsp ファイル 387  
SIAPI (検索および索引 API) 11, 243  
siapi.jar ファイル 384  
SMTP サーバー構成 364  
Solaris オペレーティング環境  
Content Edition クローラーの構成 52  
DB2 Content Manager クローラー構成  
66  
Lotus Domino Document Manager クロ  
ーラー構成 87  
Lotus Notes クローラー構成 87  
QuickPlace クローラー構成 87  
Solaris オペレーティング・システム  
DB2 クローラー構成 57  
SQL Server データベース  
DB2 クローラーでのアクセス 55  
JDBC データベース・クローラーでア  
クセス 74  
SSL、検索サーバー構成 274  
startccl スクリプト 373  
startIndexBuild コマンド 198  
Stellent パーサー  
構文解析、文書タイプの 173  
説明 171  
デフォルトの文書タイプ 182  
文書タイプの関連付け 179  
stellentTypes.cfg ファイル 171  
stellenttypes.cfg ファイル 179  
stellent.properties ファイル 179  
T  
tar ファイル  
クローラー 134  
URI フォーマット 135

## U

UIMA  
共通分析構造 165, 167  
共通分析構造の索引へのマッピング  
165  
共通分析構造のリレーショナル・デー  
タベースへのマッピング 167  
コレクションとの関連付け 163  
索引への分析結果のマッピング 165  
システムへのテキスト分析エンジンの  
追加 162  
説明 161  
分析結果のリレーショナル・データバ  
ースへのマッピング 167  
XML エLEMENTのマッピング 164  
Unicode エンコード 189

UNIX ファイル・システム・クローラー  
構成 100  
URI フォーマット 135

## URI

エンタープライズ・サーチにおけるフ  
ォーマット 135  
カテゴリ規則 149, 152  
クイック・リンク 226  
検索結果の縮小表示 211, 212  
索引からの除去 213  
詳細情報の表示 336  
静的スコアの操作 236  
有効範囲 209, 210

## URI の詳細

ドロップされた文書 356  
モニター 336

URL にできる限り早くアクセスする 112

URL にできる限り早く再アクセスする  
112

URL パスの深さ 106

USC スtring 296

## V

vbr\_access\_services.jar ファイル 52, 53

## W

Web Content Management

WebSphere Portal との統合 381

Web Content Management クローラー

構成 124

サイト URL のコピー 127

URI フォーマット 135

Web クローラー

アクティブ・サイト 339, 341

開始 URL 106, 112

グローバル・クローラー・スペース 119

クローラー履歴 339

クローラー規則 106

クローラーされた文書の日付 122

クローラー速度 342

クローラー・スペース の制限 106

構成 101

最近クローラーされた URL 339

再クローラー間隔 111

サイトの詳細 339

索引付けしないディレクティブ 121,  
122

システム状況 339

スレッドの詳細 339, 341

ソフト・エラー・ページ 112

パスワード保護された Web サイト  
114, 115

プロキシ・サーバー 116

Web クローラー (続き)  
 モニター 339  
 ユーザー・エージェント 102  
 リンクをたどらないディレクティブ  
 121, 122  
 レポートの作成 343  
 Cookie 117  
 Cookie 構成 118  
 Cookie フォーマット 118  
 followindex.rules ファイル 121, 122  
 JavaScript サポート 105  
 robots.txt ファイル 102, 103  
 URL 状況 339  
 URL にできる限り早くアクセスする  
 112  
 Web クローラーの IP アドレス規則 106  
 Web クローラーのアドレス規則 106  
 Web クローラーの開始 URL 106, 112  
 Web クローラーの再クロール間隔 111  
 Web クローラーの接頭部規則 106  
 Web クローラーのドメイン規則 106  
 Web クロール・スペースの制限 106  
 WebSphere Application Server ユーザー  
 シングル・サーバー構成のパスワード  
 307  
 マルチサーバー構成のパスワード 308  
 WebSphere II Classic Federation 64  
 WebSphere II Event Publisher  
 Edition、DB2 クローラー構成 59  
 WebSphere MQ、クローラー・サーバー構  
 成 57, 58  
 WebSphere MQ、DB2 クローラー構成  
 62  
 WebSphere Portal  
 エンタープライズ・サーチとの統合  
 381  
 エンタープライズ・サーチのセットア  
 ップ・スクリプト 383  
 カテゴリ・ツリーの移行 405  
 クラスター化システム 398  
 検索センター、説明 381  
 検索バー、説明 381  
 コレクションの移行 405  
 デフォルトの移行設定 407  
 バージョン 5.1、エンタープライズ・  
 サーチの除去 388  
 バージョン 5.1、検索バーの構成 387  
 バージョン 5.1、統合スクリプト 384  
 バージョン 6、エンタープライズ・サ  
 ーチの除去 397  
 バージョン 6、検索センターの構成  
 392  
 バージョン 6、検索バーの構成 394  
 バージョン 6、統合スクリプト 389  
 分類構造の移行 405

WebSphere Portal クラスター  
 エンタープライズ・サーチからの除去  
 403  
 統合スクリプト 399  
 統合のガイドライン 398  
 WebSphere Portal クローラー  
 構成 126  
 URI フォーマット 135  
 WebSphere Portal との統合  
 クラスター化システム 398  
 セットアップ・スクリプト 383  
 説明 381  
 es.wp5.install.jar ファイル 384  
 es.wp6.install.jar ファイル 389, 399  
 Lotus Quickr 381  
 Web Content Management 381  
 wp5\_install スクリプト 384  
 wp6\_cluster\_install スクリプト 399  
 wp6\_install スクリプト 389  
 WebSphere グローバル・セキュリティ  
 検索アプリケーションのプロパティ  
 264  
 検索アプリケーション・カスタマイザ  
 ー 266  
 使用不可化 322  
 Windows  
 IPv6 プロトコル・サポート 31  
 Windows オペレーティング・システム  
 イベント・パブリッシング構成 58  
 クローラー構成 89  
 Content Edition クローラーの構成 53  
 DB2 Content Manager クローラー構成  
 68  
 DB2 クローラー構成 58  
 Windows ドメイン 318  
 Windows ファイル・システム・クローラ  
 ー  
 構成 129  
 文書レベルのセキュリティ構成 318  
 URI フォーマット 135  
 wp5\_install スクリプト 384  
 wp5\_uninstall スクリプト 388  
 wp6\_cluster\_install スクリプト 399  
 wp6\_cluster\_uninstall スクリプト 403  
 wp6\_install スクリプト 389  
 wp6\_uninstall スクリプト 397  
 WpsMigratorLog.log ファイル 408

## X

XML エlement  
 共通分析構造へのマッピング 164  
 検索 154, 164  
 検索フィールドへのマッピング 155

XML 検索フィールド  
 エlementのマッピング 154, 155,  
 164  
 作成 155  
 説明 154, 164  
 XML 照会構文、ネイティブ 169  
 XML フラグメント、ネイティブ XML 検  
 索 169  
 XML 文書  
 検索 155  
 ネイティブ XML 検索 169  
 XPath、ネイティブ XML 検索 169

## Z

zip ファイル  
 クロール 134  
 URI フォーマット 135







**IBM**



SD88-6374-04



**日本アイ・ビー・エム株式会社**  
〒106-8711 東京都港区六本木3-2-12

Spine information:

OmniFind Enterprise Edition バージョン 8.5

エンタープライズ・サーチの管理

