

**IBM DB2 Information Integrator
OmniFind Edition**



リリース・ノート

バージョン 8.2.2

**IBM DB2 Information Integrator
OmniFind Edition**



リリース・ノート

バージョン 8.2.2

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

本書には、IBM の専有情報が含まれています。その情報は、使用許諾条件に基づき提供され、著作権により保護されています。本書に記載される情報には、いかなる製品の保証も含まれていません。また、本書で提供されるいかなる記述も、製品保証として解釈すべきではありません。

IBM の資料は、オンライン、または IBM 担当員を通してご注文いただけます。

- 資料をオンラインで注文する方法については、www.ibm.com/shop/publications/order にある「IBM Publications Center」にアクセスしてください。
- お客様の IBM 担当員に連絡する方法については、www.ibm.com/planetwide にある「IBM Directory of Worldwide Contacts」にアクセスしてください。

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。

本マニュアルに関するご意見やご感想は、次の URL からお送りください。今後の参考にさせていただきます。

<http://www.ibm.com/jp/manuals/main/mail.html>

なお、日本 IBM 発行のマニュアルはインターネット経由でもご購入いただけます。詳しくは

<http://www.ibm.com/jp/manuals/> の「ご注文について」をご覧ください。

(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： IBM DB2 Information Integrator
OmniFind Edition
Release Notes
Version 8.2.2

発 行： 日本アイ・ピー・エム株式会社

担 当： ナショナル・ランゲージ・サポート

第1刷 2005.7

この文書では、平成明朝体™W3、平成明朝体™W7、平成明朝体™W9、平成角ゴシック体™W3、平成角ゴシック体™W5、および平成角ゴシック体™W7を使用しています。この(書体*)は、(財)日本規格協会と使用契約を締結し使用しているものです。フォントとして無断複製することは禁止されています。

注* 平成明朝体™W3、平成明朝体™W7、平成明朝体™W9、平成角ゴシック体™W3、平成角ゴシック体™W5、平成角ゴシック体™W7

© Copyright International Business Machines Corporation 2004, 2005. All rights reserved.

© Copyright IBM Japan 2005

目次

リリース・ノートについて	1	AIX での共用メモリー・セグメント	13
Web 上の技術サポート情報	1	文書の更新情報	15
このリリースでの新規事項	3	エラー・メッセージの変更	15
エンタープライズ・サーチ基本アノテーター・パッケージ	4	照会構文の更新	15
既知の制限、問題、および対処方法.	7	URI の縮小	15
文書サイズの制限.	7	FileNet Content Services の場合の直接文書アクセスの使用可能化.	16
フィックスパックのインストールによって再フル・クローラーが起こる	7	汎用クローラー・プラグイン	16
Java パーサー・アノテーターのヒープ・サイズの制限	7	「個人用プロファイル」ページに関するヘルプがユーザー・インターフェースと一致しない.	17
PEAR (Processing Engine ARchive) ファイルの最大サイズ	9	Web クローラーのヘルプ・ページに誤った値がある	18
AIX 上の Notes クローラーには I/O Completion Port モジュールが必要である	9	indexBuildItem の説明	18
セキュア Notes コレクションの照会で結果文書数が不正確な場合がある	11	索引への共通分析構造のマッピング	19
ユーザー名にピリオドを含ませることができる	11	WebSphere Portal からのモデル・ベースの分類のマッピング	19
フィールド検索でのワイルドカード文字.	11	特記事項	21
サーバー上の索引の場所	12	商標.	23
Linux および AIX の場合の孤立 URLFetcher プロセス	12	IBM と連絡を取る	25
		製品情報	25

リリース・ノートについて

リリース・ノート文書には、IBM® DB2® Information Integrator OmniFind™ Edition バージョン 8.2.2 (バージョン 8.2 にフィックスパック 2 を適用したもの) について、このリリースでの新規事項、既知の制限事項、問題とその解決方法、および資料に対する更新内容が記述されています。

DB2 II OmniFind Edition バージョン 8.2.2 は、アップグレード専用です。バージョン 8.2.2 をインストールするには、あらかじめバージョン 8.2 またはバージョン 8.2.1 がインストールされている必要があります。

SuSE Linux® Enterprise Server 9 にインストールする前に、DB2 II OmniFind Edition バージョン 8.2.2 サポート Web サイトで、問題の回避策とアップデート情報を確認してください。

Web 上の技術サポート情報

DB2 II OmniFind Edition 製品は、WebSphere® Information Integrator OmniFind Edition と名前が変わりました。製品関連 Web ページでは WebSphere Information Integrator OmniFind Edition と記述されていますが、製品インターフェース、資料、およびトラブルシューティング情報 (technotes) では、まだ DB2 ブランドのままになっていることがあります。

技術サポート情報は、以下の Web サイトにあります。

IBM DB2 Information Integrator OmniFind Edition

www.ibm.com/software/data/integration/db2ii/supportomni.html

IBM WebSphere Information Integrator

<http://www.ibm.com/software/data/integration/db2ii/support.html>

IBM WebSphere Application Server

<http://www.ibm.com/software/webservers/appserv/was/support/>

IBM DB2 Universal Database™ for Linux、UNIX®、および Windows® バージョン 8 製品

www.ibm.com/software/data/db2/udb/winos2unix/support

このリリースでの新規事項

IBM DB2 Information Integrator OmniFind Edition バージョン 8.2.2 には以下の機能があります。

オペレーティング・システム

バージョン 8.2.2 は、追加で以下のオペレーティング・システムをサポートします。

- AIX® 5.3 (Maintenance Level 1 および August 2004 C++ Runtime for AIX PTF が必要)
- SuSE Linux Enterprise Server 9 with Service Pack 1 (UnitedLinux SP1) (libstdc++-3.3.3-43.28 が必要)

SuSE Linux Enterprise Server 9 にインストールする前に、DB2 II OmniFind Edition バージョン 8.2.2 サポート Web サイトで、問題の回避策とアップデート情報を確認してください。

大規模文書

コレクションには、32 MB までのサイズの文書を入れることができます。

Lotus Domino® 文書に対する文書レベルのセキュリティー

- ネイティブ・リポジトリによって管理されるアクセス制御によるユーザー信用証明情報の (照会処理中の) リアルタイム検証
- ユーザー・プロファイルおよび LDAP ユーザー・レジストリーを介したユーザー検証。信用証明情報を暗号化されたフォーマットで保管することによって、ユーザーは Domino ソースを照会するたびに信用証明情報を指定する必要はありません。

クローラー

- Lotus Notes® クローラー: データベース、ビュー、フォルダーに加えて、ディレクトリーをクロールする機能
- Web クローラー:
 - パスワード保護された Web サイト (HTTP 基本認証またはフォーム・ベースの認証を使用するサイト) をクロールするためのユーザー・インターフェースのサポート
 - 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機能
 - ソフト・エラー・ページのカスタム・メッセージを戻す機能

検索の品質

- フェデレーテッド検索 (一度に複数のコレクションを検索する機能)
- 照会における末尾ワイルドカード文字のサポート
- 検索結果内でのサマリー表示長のカスタマイズのサポート
- 同じ URI からの検索結果を縮小するためのユーザー・インターフェースのサポート

言語分析

- コレクション内の文書のカスタム・テキスト分析を行う機能を提供する、非構造化情報管理アーキテクチャー (unstructured information management architecture (UIMA)) のサポート
- ネイティブ XML 検索のサポート
- カスタム同義語辞書のサポート

API および検索アプリケーション

- カスタム・クライアント・データ・リスナー・アプリケーションの構成およびモニターのためのユーザー・インターフェースのサポート
- カスタムのセキュリティーおよびビジネス・ルールの指定を可能にする汎用クローラー・プラグイン
- 照会機能、検索結果オプション、フェデレーテッド・サーチ、文書レベルのセキュリティーをデモンストレーションするサンプル検索アプリケーションに対する多くの改善

エンタープライズ・サーチ基本アノテーター・パッケージ

エンタープライズ・サーチ基本アノテーター・パッケージを使用することにより、エンタープライズ・サーチ・アノテーターの出力を基にして新しいアノテーターを開発し、エンタープライズ・サーチ・アノテーターとカスタム・アノテーターの両方を UIMA Software Development Kit (SDK) に おいてテストすることができます。

エンタープライズ・サーチ基本アノテーター・パッケージは、エンタープライズ・サーチで使用されるテキスト分析手順が含まれた ZIP ファイルです。これらのアノテーターは、文書を解析する際、常にカスタム分析の前に実行します。

UIMA Software Development Kit (SDK) は、IBM developerWorks Web サイト (<http://www.ibm.com/developerworks/db2/zones/db2ii/>) から入手できます。

アノテーター・パッケージをインストールするには、次のようにします。

1. `ES_INSTALL_ROOT/packages/uima` ディレクトリーにインストールされたエンタープライズ・サーチ (DB2 II OmniFind Edition) 内で、アノテーター・パッケージ `OF_base_annotators.zip` を見つけます。
2. ご使用の環境での UIMA SDK インストールのルート・ディレクトリーに、この ZIP ファイルをコピーします。
3. この ZIP ファイルを解凍して、エンタープライズ・サーチ基本アノテーターのファイルを、指定された UIMA SDK インストールのディレクトリー構造に追加します。

エンタープライズ・サーチ基本アノテーター・パッケージに組み込まれている辞書は、英語文書の処理に必要な辞書のみです。ご使用の開発環境で他の言語を処理したい場合は、以下の手順を実行してください。

1. `ES_INSTALL_ROOT/configurations/parserservice/jediidata/frost4` にインストールされたエンタープライズ・サーチ内でエンタープライズ・サーチ辞書を見つけてみます。
2. `UIMA_SDK_INSTALL/data/frost4` にインストールされた ローカル UIMA SDK に、辞書の内容をコピーします。

次のようにして、UIMA SDK が正常にインストールされたことを検証します。

1. UIMA_SDK_INSTALL/bin/cvd[.bat/.sh] ディレクトリーの CAS Visual Debugger (CVD) をオープンします。
2. 「実行」 → 「TAE のロード」をクリックします。
3. UIMA_SDK_INSTALL/docs/examples/descriptors/analysis_engine ディレクトリー内で、TAE 指定子ファイル of_tokenization.xml を見つけます。
4. サンプル文書をロードし、TAE を実行します。CVD にタイプ uima.tt.TokenAnnotation の注釈が表示されるようになります。

エンタープライズ・サーチ・アノテーターを使用して処理を行うには、次のようにします。

1. エンタープライズ・サーチ・アノテーターによって定義されたタイプをカスタム・アノテーターが使用する場合、カスタム・アノテーター指定子の typeSystem セクションに、ファイル of_typesystem.xml への参照を組み込みます。
of_typesystem.xml ファイルは、
UIMA_SDK_INSTALL/docs/examples/descriptors/analysis_engine ディレクトリーにあります。ファイルへの参照を組み込む方法のサンプルについては、
analysis_engine ディレクトリーにある jtok.xml ファイル を参照してください。
2. ファイル of_tokenization.xml およびカスタム・アノテーター指定子を含んだ集約 TAE 指定子を作成します。
3. その集約 TAE 指定子を、UIMA SDK 内でテストを行う際に使用します。

アノテーターをエンタープライズ・サーチにアップロードするために PEAR ファイルを作成するときには、この PEAR ファイルがカスタム・アノテーター指定子のみを含むようにしてください。ファイル of_tokenization.xml、of_typesystem.xml や、of_base_annotators.zip の内容、ステップ 2 で作成した集約指定子を含まないようにします。

エンタープライズ・サーチにおける分析機能についての詳細は、インフォメーション・センターまたは「エンタープライズ・サーチプログラミング・ガイドおよび API リファレンス」にある、「カスタム・テキスト分析の組み込み」および「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」のトピックを参照してください。

既知の制限、問題、および対処方法

ここでは、DB2 Information Integrator OmniFind Edition バージョン 8.2.2 について現在知られている制限、問題、および対処方法を示します。

制限および制約事項は、他のリリースには該当しない場合があります。

文書サイズの制限

32 MB までの文書は、クロールおよび索引付けを行うことができます。LOB データが入った DB2 Universal Database 表をクロールおよび索引付けできるのは、文書全体の 32 MB サイズ制限をデータが順守している場合に限られます。

クローラー・プロパティのヘルプ・ファイル内の説明文で、最大ページ・サイズは 32 MB を超えることができるとの記述は無視してください。

フィックスパックのインストールによって再フル・クロールが起こる

DB2 II OmniFind Edition バージョン 8.2.2 をインストールした後、Web クローラーおよび NNTP クローラーによってクロールされた文書を除くすべての文書が再クロールされます。

バージョン 8.2.2 インストール後に Web クローラーおよび NNTP クローラー以外のクローラーを初めて開始すると、すべてのデータ・ソース内のすべての文書が再びクロールおよび索引付けされます。対象となる文書は、データ・ソースが最後のクロール以降に変更された文書または新規文書ではありません。

クローラーは、コレクション内の各表、ディレクトリー、またはフォルダーに入っているすべての文書をクロールした後、通常の更新処理に切り替え、新規または変更文書のクロールを開始します。

通常、再フル・クロールが必要なのは、クローラーの構成に対する変更がそれを必要とする場合に限られます。例えば、あるデータ・ソース内のフィールドまたは列の表示名が変更された場合、クローラーは、新しい情報を文書と索引付けできるように、すべての文書を再クロールする必要があります。

バージョン 8.2.2 のインストール後に自動的に発生する再フル・クロールには時間がかかりますが、エンタープライズ・サーチ索引の品質には影響しません。

Java パーサー・アノテーターのヒープ・サイズの制限

Java パーサー・アノテーターを実行している `collection_ID.casprocessor` セッションのコレクション・エラー・ログに `OutOfMemory` (メモリー不足) メッセージが表示された場合、ヒープ・サイズを 200 MB に増やしてください。

エンタープライズ・サーチにアップロードされたカスタム・テキスト分析アノテーターは、Java セッションで実行されます。例えば、エンタープライズ・サーチにお

いて、これらのセッションは *collection_ID.casprocessor* セッションです。ここで、*collection_ID* は、このアノテーターが属している コレクション ID です。これらのセッションには、200 MB までのヒープ・サイズの制限があります。

ヒープ・サイズを増やすには、次のようにします。

1. エンタープライズ・サーチ・システムを停止します。すべてのオペレーティング・システムで、コマンド `esadmin stop` を入力します。
2. 次のようにして CCL を停止します。
 - Linux および AIX の場合、コマンド `stopccl.sh` を入力します。
 - Windows の場合、「スタート」 → 「コントロールパネル」 → 「サービス」を選択して、DB2 Information Integrator OmniFind Edition Windows サービスを停止します。
3. アノテーターが属しているコレクション ID を見つけます。コレクション ID を見つけるには、コマンド `esadmin check` を実行するか、または、*ES_NODE_ROOT/master_config* ディレクトリーに移動し、作成した各コレクションの *collection_ID_config.ini* ファイルを見つけてみます。
4. *collection_ID_config.ini* ファイル内で、*collection_ID.casprocessor* セッションに対する `max_heap` 値を増やします。
5. *collection_ID_config.ini* ファイルをソートし、CAS プロセッサ項目を探します。Linux での col7 の場合の例を以下に示します。

```
% cd $ES_NODE_ROOT
% sort col7_config.ini
...

session3.collectionid=col7
session3.configDir=col7.casprocessor
session3.description=Fenced CAS Processor Session
session3.displayname=Fenced CAS Processor Server
session3.domain=.
session3.flags=0
session3.id=col7.casprocessor
session3.max_heap=200

...
```

6. `session_number.max_heap=200` の行を編集して、新しい最大ヒープ・サイズ (200) を含むようにします。例えば col7 について、`session3.max_heap=200` とします。
7. 次のようにして CCL を再始動します。
 - Linux および AIX の場合、コマンド `startccl.sh -bg` を入力します。
 - Windows の場合、「スタート」 → 「コントロールパネル」 → 「サービス」を選択して、DB2 Information Integrator OmniFind Edition Windows サービスを開始します。
8. エンタープライズ・サーチを再始動します。すべてのオペレーティング・システムで、コマンド `esadmin start` を入力します。
9. エンタープライズ・サーチ管理コンソールから、このコレクションのパースーを開始します。アノテーターは、新しい最大ヒープ・サイズを使用して JVM 上で実行します。

PEAR (Processing Engine ARchive) ファイルの最大サイズ

カスタム・テキスト分析の場合、Processing Engine ARchive ファイル (PEAR ファイル) の最大サイズは 8 MB です。

エンタープライズ・サーチ管理コンソールを使用してアップロードできるのは、8 MB よりも小さい PEAR ファイルのみです。8 MB よりも大きい PEAR ファイルには、辞書などのリソースが含まれている場合がよくあります。そのような PEAR ファイルをアップロードするには、次のようにします。

1. ユーザーのアノテーター・コードを変更して、リソースのロードを UIMA リソース・マネージャーを介して行う (AnnotatorContext の getResourceObject メソッドを使用する) ようにするか、リソースをポイントするアノテーター構成パラメーターの値を使用するようにします。
2. エンタープライズ・サーチ・システムのインデクサー・ノードの専用ディレクトリーにリソース・ファイルをコピーします。
3. アノテーター構成 (<externalResource> 指定、またはご使用の構成パラメーターの <namevaluePair> のいずれか) を変更し、ステップ 2 で作成したディレクトリーを使用して、リソースへの絶対パスを組み込みます。
4. 更新した構成で PEAR ファイルを再作成します。このファイルにリソースが含まれていないことを確認し、エンタープライズ・サーチにアップロードします。

AIX 上の Notes クローラーには I/O Completion Port モジュールが必要である

IBM AIX プラットフォームで Notes クローラーおよびディスクバリー・プロセスを使用するには、AIX 製品 CD-ROM から I/O Completion Port (IOCP) モジュール (bos.iocp.rte) をインストールし、その状況を **Available** に設定する必要があります。

bos.iocp.rte モジュールをインストールした後、適切なフィックスを適用する必要があります。その手順については、次のリンクを参照してください。

<http://www-1.ibm.com/support/docview.wss?uid=swg21086556>

この IOCP モジュールがないと、Notes ディスカバリー・プロセスによる新規クローラーの作成中に障害が起こり、次のエラー・メッセージが表示されます。

FFQM0105E サーバーからエラーを受信しました -
メッセージ: FFQG0024E 予期しない例外がありました: ディスカバー

さらに、ENOEXEC エラーを含む次のようなメッセージが、

ES_NODE_ROOT/logs/system_yyyymmdd.log ファイルに書き込まれます。(一部のメッセージ・テキストは、読みやすくするため複数行に分割されています。)

```
5/20/05 18:08:52.423 JST [Error] [ES_ERR_EXCEPTION_DEFAULT_MESSAGE] [] [discovery]
dominoserver.ibm.com:0:2108088751:control:ComponentDiscoveryW.java:
com.ibm.es.control.discovery.server.ComponentDiscoveryW.discover:86
FFQ00277E An exception was caught with the detail 'java.lang.UnsatisfiedLinkError:
/opt/lotus/notes/65010/ibmpow/liblsxbe_r.a:
load ENOEXEC on shared library(s) /opt/lotus/notes/latest/ibmpow/libnotes_r.a'
and a stack trace of 'java.lang.UnsatisfiedLinkError:
/opt/lotus/notes/65010/ibmpow/liblsxbe_r.a:
load ENOEXEC on shared library(s) /opt/lotus/notes/latest/ibmpow/libnotes_r.a
at java.lang.ClassLoader$NativeLibrary.load(Native Method)
```

```

at java.lang.ClassLoader.loadLibrary0(ClassLoader.java:2120)
at java.lang.ClassLoader.loadLibrary(ClassLoader.java:1998)
at java.lang.Runtime.loadLibrary0(Runtime.java:824)
at java.lang.System.loadLibrary(System.java:908)
at lotus.domino.NotesThread.load(NotesThread.java:306)
at lotus.domino.NotesThread.checkLoaded(NotesThread.java:327)
at lotus.domino.NotesThread.sinitThread(NotesThread.java:181)
at com.ibm.es.crawler.discovery.notes.NotesLibrary$NotesOperation.discover
(Unknown Source)
at com.ibm.es.crawler.discovery.api.DiscoveryAPI.discover(Unknown Source)
at com.ibm.es.control.discovery.server.ComponentDiscoveryW.discover
(ComponentDiscoveryW.java:72)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:85)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:58)
at sun.reflect.DelegatingMethodAccessorImpl.invoke
(DelegatingMethodAccessorImpl.java:60)
at java.lang.reflect.Method.invoke(Method.java:391)
at com.ibm.es.ccl.sessionwrapper.CallThread.run(CallThread.java:77)

```

Notes クローラーを作成する前に、次のようにして IOCP モジュールが正しくインストールされていることを確認してください。

1. 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ていることを確認するため、次の `lspp` コマンドを入力します。

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

この `lspp` コマンドの出力は、以下の例のようになります。

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

2. IOCP ポートが **Available** であることを確認するため、次の `lsdev` コマンドを入力します。

```
$ lsdev -Cc iocp
```

この `lsdev` コマンドの出力は、以下の例のようになります。

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

3. IOCP ポート状況が **Defined** である場合、以下のステップを実行して状況を **Available** に変更します。
 - a. root ユーザーとしてクローラー・サーバーにログインし、次のコマンドを発行します。

```
# smit iocp
```
 - b. 「**Change / Show Characteristics of I/O Completion Port**」を選択し、「**STATE to be configured at system restart**」を「**Defined**」から「**Available**」に変更します。
 - c. クローラー・サーバーを再始動します。
 - d. もう一度 `lsdev` コマンドを入力し、IOCP ポートの状況が **Available** に変更されていることを確認します。

セキュア Notes コレクションの照会で結果文書数が不正確な場合がある

照会処理中に現在の信用証明情報を認証するというオプションが選択されている場合、セキュア・コレクションを検索する照会では結果セット内の検索された文書数が不正確になる場合があります。

エンタープライズ・サーチ・コレクションを検索すると、検索結果に、結果文書の総数が示されます。以下の条件がすべて満たされている場合、この数は実際の結果文書数よりも大きくなる場合があります。

- コレクションの作成時に、コレクションに対してセキュリティーが使用可能になっていた。
- コレクションに Notes クローラーが含まれている。
- 文書レベルのセキュリティーのオプションを Notes クローラーについて構成するときに、照会処理中に現在の信用証明情報を認証するというオプションが指定されていた。

この問題は、SIAPI ResultSet.getAvailableNumberOfResults 値 および ESSearchApplication サンプル検索アプリケーションに影響します。

ユーザー名にピリオドを含ませることができる

エンタープライズ・サーチ管理者名の追加に関するオンライン説明が 誤っています。

エンタープライズ・サーチ管理コンソールを使用して、エンタープライズ・サーチ管理者 ID をシステムに追加するとき、そのページに関する説明には、ユーザー名はピリオドを含むことができないと明記されています。その説明は誤りです。ピリオドを含むユーザー名 (j_user@this.company.com など) が許可されます。

フィールド検索でのワイルドカード文字

照会用語として *field:pref** を使用することは、一般的に許容されています。DB2 II OmniFind Edition は、*field* という名前のフィールドのメンバーとしてマークされた *pref* で始まる語が含まれている文書を検索します。このルールには例外があり、ワイルドカード照会にあまり関係のない少数の特殊なフィールドの場合には、照会用語 *field:pref** は、*field:pref* であるかのように扱われます。それらの特殊なフィールドは、次のとおりです。

- site
- link
- docid
- samegroupas
- url
- group (ニュースグループのこと。site に似ていますが、NNTP で使用されます。)

例えば、照会 *url:porta** では何も結果が戻されず、*url:portal* では結果が戻されるということがあります。

サーバー上の索引の場所

コレクションを作成するときに、索引の場所として、サーバー・データのルート・ディレクトリー上のデフォルトの場所以外を指定した場合、このデフォルトではない索引の場所のディスク・スペース使用量はモニターされません。

マルチサーバー・インストールの場合、コレクションを作成するときに、索引の場所としてデフォルトの場所以外を指定したい場合は、索引の場所が、まず 4 つのサーバーすべての上に存在することが必要です。

Linux および AIX の場合の孤立 URLFetcher プロセス

問題: Web クローラーが 2 つのプロセスを使用します。1 つのプロセスが終了したときに、そのプロセスは 2 番目のプロセスを終了するためのシグナルを出します。たまたま、1 つのプロセスが突然終了し、もう一方のプロセスが取り残されて、いつまでも実行を続ける場合があります。

次に示す UNIX ps コマンドの出力に、次のプロセスのような孤立プロセスがあると、これが発生していることが分かります。

```
F S UID          PID PPID  C PRI  NI ADDR      SZ WCHAN  STIME TTY          TIME CMD
0 S user1      26222     1  0  75   0   - 152641 schedu 10:34 pts/8      00:00:00
```

```
/opt/IBMJava2-141/jre/bin/java -classpath ../../URLFetcher.jar:...
```

次の点に注目してください。

- PPID (親プロセス ID) が 1 となっているのは、そのプロセスが孤立していたことを示します。
- classpath の中に URLFetcher.jar があれば、これが Web クローラー・プロセスであることを示します。

解決策

孤立プロセスのプロセス ID (PID) を見つけ、それを強制終了します。上の例では、26222 が孤立プロセスです。Linux および AIX では、kill コマンドを引数を付けずに使用します。プロセス表に、「子」項目があるかもしれません。子項目は、孤立プロセスによって作成されたスレッドを示しています。これらの子項目は、上位のプロセスを強制終了するときに、削除する必要があります。

孤立 URLFetcher プロセスがあれば、そのプロセスを強制終了してください。これは、そのようなプロセスがプロセス表にあると、オペレーティング・システムがそのためにリソースを予約するためです。これらの孤立プロセスがないかどうか、定期的に検査する必要があります。特に、Web クローラーが異常終了したと考えられる場合は、検査が必要です。そのようなプロセスが存在した場合は、強制終了してください。

重要: DB2 II OmniFind Edition バージョン 8.2.1 では、C++ コンポーネントがなくなると URL フェッチャーは自ら終了するため、まれにしか孤立プロセスは発生しません。それでもこの問題が発生した場合は、セッションを編集して別のユーザーの httpauth.ini を強制終了してしまわないように注意してください。HTTPEdit ユーティリティーは、URLFetcher JAR ファイルから実行する JVM ですが、プロセス表の中でよく似ているように見えます。

AIX での共用メモリー・セグメント

AIX 上でメッセージ SQL1224N が表示される場合、共用メモリー・セグメントの数が AIX 上で許される数を超えている可能性があります。

この問題を解決するには、IBM Support Web サイトの『SQL1224N Shared Memory Segments』のトピックを参照してください。

文書の更新情報

ここでは、DB2 Information Integrator OmniFind Edition バージョン 8.2.2 の資料に対する修正および若干の追加情報について説明します。

エラー・メッセージの変更

翻訳のためにマニュアルが発行された後で、エラー・メッセージが更新されました。

FFQD1034E

エラー: 環境変数 CMBROOT/IBMCMROOT が定義されていません。DB2 Information Integrator for Content が正しくインストールされていることを確認してください。

説明: 環境変数 CMBROOT/IBMCMROOT が定義されていません。DB2 Information Integrator for Content がインストールされていない可能性があります。

ユーザー応答:

DB2 Information Integrator for Content が正しくインストールされていることを確認してください。

FFQD1035E

エラー: cmbicmsrvs.ini ファイルが見つかりませんでした。DB2 Information Integrator for Content が正しくインストールされていることを確認してください。

説明: 指定されたパスに cmbicmsrvs.ini ファイルが見つかりませんでした。DB2 Information Integrator for Content がインストールされていない可能性があります。

ユーザー応答:

DB2 Information Integrator for Content が正しくインストールされていることを確認してください。

照会構文の更新

単一の - (マイナス) 条件からなる照会は無効です。マイナス条件は、事後フィルターとしてのみ使用でき、結果を生成するためには使用できません。

検索照会に適用可能な構文のタイプが説明されている「照会構文」トピックは、「プログラミング・ガイドおよび API リファレンス」(iiysp.pdf) とインフォメーション・センターにあります。

URI の縮小

検索結果内で URI を縮小する方法についてのトピックに、誤解を招くような説明や正しくない情報があります。

- 資料にある説明では、すべてのデータ・ソース・タイプについて、検索結果内で URI を縮小できるかのように読み取れます。しかし、URI を縮小できるのは、URL を持つ文書 (つまり、Web クローラーまたは NNTP クローラーによってクロールされた文書) に限られます。
- また、資料に記述されている `samegroupas` 照会構文の使用法は正しくありません。照会でこのキーワードを使用するには、完全修飾の URL を指定する 必要があります。複数の URI をまとめて縮小するために作成されたグループの名前を指定することはできません。

例えば、同じグループ (例: News という名前のグループ) に属している文書を照会するには、グループ名ではなく、そのグループ内のいずれか 1 つの文書の URI を指定します。

正: `samegroupas:http://mycompany.server3.com/news/`

誤: `samegroupas:news`

FileNet Content Services の場合の直接文書アクセスの使用可能化

DB2 Information Integrator Content Edition (旧名: VeniceBridge) サーバーによって管理される FileNet Content Services リポジトリを、VeniceBridge クローラーを使用してクロールする場合、ターゲットの FileNet Content Services ライブラリーおよび FileNet Web Services は、FileNet Content Services コネクタと共に動作し、直接文書アクセスを使用するように構成されていなければなりません。直接文書アクセスのサポートをセットアップする方法については、FileNet Web Services の「*Administrator's Guide*」および FileNet Content Services 管理資料を参照してください。

汎用クローラー・プラグイン

クローラー・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施行、およびエンタープライズ・サーチ索引内の文書とのメタデータの関連付けに使用する、Java™ クラスの名前を指定できます。Web クローラーまたは NNTP クローラーと共に 汎用クローラー・プラグインを使用することはできません。

クローラー・プラグインは、製品にインプリメントされていますが、エンタープライズ・サーチの資料には説明がありません。

汎用クローラー・プラグインは、バージョン 8.2 の DB2 II OmniFind Edition で使用可能だったセキュリティー・トークン・プラグインのすべての機能をカバーします。セキュリティー・トークン・プラグインは引き続きサポートされていますが、代わりに汎用クローラー・プラグインを使用することをお勧めします。

プラグインを使用するには、新規コレクションを作成する必要があります。以前に作成したコレクションと共に新しいプラグインを使用することはできません。

プラグインは、クローラーがクロールするそれぞれの文書に対して呼び出される Java クラスから構成されます。この Java クラスには、文書 ID (URI)、割り当てられたセキュリティー・トークン、およびメタデータが渡されます。クラスは、新規または変更されたセキュリティー・トークンのセットおよびメタデータを戻すか、あるいは、文書を無視またはスキップすることを示すことができます。

クロール・スペース内のすべての文書がクロールされた後は、プラグインは新規または変更された文書に対してのみ呼び出されます。更新されなかった文書に対するセキュリティ・トークンおよびメタデータを変更したい場合は、クロール・スペース内のすべての文書を再クロールし、索引を再編成してください。

セキュリティを施行するためのプラグインの使用

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ストリング) を関連付けることによって実施されます。セキュリティ・トークンとしてグループ ID がよく使用されます。

デフォルトでは、各文書にパブリック・トークンが割り当てられ、誰でもその文書を利用できるようになっています。パブリック・トークンは、管理者が指定した値、またはクロールされる文書内のフィールドから取り出された値で置き換えることができます。

プラグインを使用すると、クロールされる文書のセキュリティ・トークンの値を、ユーザー独自のビジネス・ルールにより決定できます。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。それらのセキュリティ・トークンは、文書をフィルターに掛けてセキュリティ・トークンと一致する文書のみを取り出すのに使用され、これによって、ユーザーが参照を許可された文書のみが検索結果に戻されることとなります。

メタデータを追加するためのプラグインの使用

最終変更日付データや、Lotus Notes® 文書のフィールド・データ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対して作成されます。クローラー・プラグインによって、各文書について索引付けが必要なメタデータの値を決定する際に、ユーザー独自のビジネス・ルールを適用できます。

メタデータは、名前と値のペアとして定義されます。フィールド名を指定する照会またはフリー・テキスト照会で、メタデータを検索できます。

DB2 II OmniFind Edition Support Web サイトに、このトピックに関する詳細、およびサンプルのプラグイン・コードに関する情報があり、このサンプルはユーザー特有のロジックに合わせて変更できます。

「個人用プロファイル」ページに関するヘルプがユーザー・インターフェースと一致しない

サンプル検索アプリケーションのユーザー・インターフェースが、資料が翻訳された後に更新されました。

ESSearchApplication 検索アプリケーション中の「個人用プロファイル」ページで、「データ・ソース信用証明情報」ラベルと「データ・ソース名」ラベルは、それぞれ「ドメイン信用証明情報」と「ドメイン名」に変更されました。このページのヘルプ・ファイルでは、まだ古いラベルが使われています。

Web クローラーのヘルプ・ページに誤った値がある

Web クローラーの構成に関するオンライン・ヘルプに、誤ったデフォルト値が指定されています。

Web クローラーを構成するには、以下の訂正を参照してください。

「Web クローラーのプロパティ」ページ

- 「**文書検索の並行スレッド数**」のオンライン・ヘルプでは、デフォルト値および最大値は 200 であると説明されています。正しいデフォルト値は 50 です。最大値は 200 です。
- 「**最大ページ・サイズ**」のオンライン・ヘルプでは、許容値は 250 KB から 10000 KB であると説明されています。最大値は 32 MB です。

「Web クローラーの拡張プロパティ」ページ

- 「**URL の最大数**」のオンライン・ヘルプでは、デフォルト値は 2 000 000 URL であると説明されています。正しいデフォルト値は 3 000 000 です。
- 「**アクティブ・ホストの最大数**」のオンライン・ヘルプでは、デフォルト値は 200 であると説明されています。正しいデフォルト値は 500 です。
- 「**固有文書の最大数**」のオンライン・ヘルプでは、デフォルト値は 7 000 000 であると説明されています。正しいデフォルト値は 1 000 000 です。
- 「**クロールする新規文書の比率 (%)**」のオンライン・ヘルプでは、デフォルト値は 99 % であると説明されています。正しいデフォルト値は 75 % です。
- 「**一時ストレージ内にクロールした文書を保管する最大数**」のオンライン・ヘルプでは、デフォルト値は 5 000 000 文書であると説明されています。正しいデフォルト値は 500 000 です。

「Web クロール・スペース」ページ

- 「**URL パスの深さ**」のオンライン・ヘルプでは、デフォルト値は 25 であると説明されています。正しいデフォルト値は 16 です。

indexBuildItem の説明

「プログラミング・ガイドおよび API リファレンス」には、「索引作成構成ファイルの作成」というトピックがあります。サブセクション「<indexBuildItem> エlement」の説明は正しくありません。

正しい説明文は次のとおりです。

タイプ A がタイプ B のサブタイプである (例えば、?Person? が ?Entity? のサブタイプである) 場合、かつ、両方のタイプに対して <indexBuildItem> Element Ia と Ib が定義されている場合、処理は次のようになります。

- Ib に定義された各索引ルールは、タイプ B のフィーチャー構造と、タイプ A のフィーチャー構造に適用されます。
- Ia に定義された各索引ルールは、タイプ A のフィーチャー構造に適用されません。

この処理により、スパン "entity" の下のすべての "Entity" 注釈を保管するようなルールを定義することができ、それは "Person" 注釈にも適用されます。また、"Person" と名づけられた追加スパンを作成するような "Person" 注釈のルールも別に定義することができます。それぞれのルールは互いに独立して適用されるので、同じフィーチャーに対して異なる <attributemapping> エレメントを含むことができます。

索引への共通分析構造のマッピング

索引への共通分析構造のマッピングのための XML ファイル は、xmlns:namespace ではなく、xmlns を使用しなければなりません。

XML ファイルは、次のステートメントで始まっていないければなりません。

```
<?xmlversion="1.0" encoding="UTF-8"?>  
<indexBuildSpecification xmlns="http://www.ibm.com/of/822/consumer/index/xml">
```

xmlns 属性に :namespace を付加しないでください。そうしなければ、パーサーの開始時にエラー・メッセージ FFQT0138E が発生します。

<indexBuildItem> エレメント内では、<typeName> ではなく <name> を使用してください。サンプル構成には、誤ったタグが含まれています。正しくは、

```
<indexBuildItem><name>com.ibm.uima.tt.PersonAnnotation</name>
```

であり、

```
<indexBuildItem><typeName>com.ibm.uima.tt.PersonAnnotation</typeName>
```

は誤りです。

WebSphere Portal からのモデル・ベースの分類のマイグレーション

「エンタープライズ・サーチの管理」の「WebSphere Portal からのモデル・ベース分類構造のマイグレーション」というトピックには、説明を追加する必要があります。

マイグレーション・ユーティリティは、初めて実行するときには、WebSphere Portal インストールへのパスを必要とします。このユーティリティは、エンタープライズ・サーチ構成ファイルを更新します。それらの構成更新が正常に行われるためには、エンタープライズ・サーチ・システムが停止されている必要があります。そうでない場合、モデル・ベースの分類のファイルをマイグレーションするオプションは、ぼかし表示されます。マイグレーション・ユーティリティの 2 回目以降の実行では、エンタープライズ・サーチ・システムがアクティブであってもかまいません。

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あり、米国以外の国においては本書で述べる製品、サービス、またはプログラムを提供し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、製造元によって明示的に指定されたものを除き、他社の製品、プログラムまたはサービスを使用した場合の評価と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。

〒106-0032
東京都港区六本木 3-2-31
IBM World Trade Asia Corporation
Licensing

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）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すが、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。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。

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。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めかしたり、保証することはできません。お客様は、IBM 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。

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した創作物には、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。

Outside In (®) Viewer Technology, ©1992-2004 Stellent, Chicago, IL., Inc. All Rights Reserved.

IBM XSLT Processor Licensed Materials - Property of IBM ©Copyright IBM Corp., 1999-2004. All Rights Reserved.

商標

ここでは、IBM の商標と、特定の IBM 以外の商標をリストします。

以下は、IBM Corporation の商標です。

IBM
AIX
AIX 5L
DB2
DB2 Universal Database
Domino
Domino.doc
Hummingbird
Informix
Lotus
Lotus Notes
Notes
OmniFind
POWER4
POWER5
RISC System/6000
Tivoli
WebSphere
Workplace
xSeries
z/OS

以下は、それぞれ各社の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または登録商標です。

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。

Intel、Intel Inside (ロゴ)、MMX および Pentium は、Intel 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。

Linux は、Linus Torvalds の米国およびその他の国における商標です。

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。

IBM と連絡を取る

お客様の国または地域で IBM に連絡する方法については、Web の www.ibm.com/planetwide にある「IBM Directory of Worldwide Contacts」にアクセスしてください。

製品情報

DB2 Information Integrator 製品に関する情報は、Web により入手できます。

Web の www.ibm.com/software/data/integration/db2ii/support.html にアクセスしてください。

このサイトには、次の最新情報が入っています。

- 技術ライブラリー
- 資料の注文方法
- 製品のダウンロード
- ニュースグループ
- フィックスパック
- ニュース
- Web リソースへのリンク



Printed in Japan

