

IBM WebSphere Information Integrator  
OmniFind Edition



# 发行说明

版本 8.3



IBM WebSphere Information Integrator  
OmniFind Edition



# 发行说明

版本 8.3

在使用本资料及其支持的产品之前，请务必阅读『声明』中的一般信息。

本文档包含 IBM 的专利信息。它是根据许可协议提供的，并受版权法保护。本出版物包含的信息不包括任何产品保证，且本手册提供的任何声明不应作如此解释。

可以在线方式或通过您当地的 IBM 代表订购 IBM 出版物：

- 要以在线方式订购出版物，可访问“IBM 出版物中心”（IBM Publications Center），网址为 [www.ibm.com/shop/publications/order](http://www.ibm.com/shop/publications/order)。
- 要查找您当地的 IBM 代表，可访问“IBM 全球联系人目录”（IBM Directory of Worldwide Contacts），网址为 [www.ibm.com/planetwide](http://www.ibm.com/planetwide)。

当您发送信息给 IBM 后，即授予 IBM 非专有权，IBM 可以它认为合适的任何方式使用或分发此信息，而无须对您承担任何责任。

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

---

# 目录

第 1 章 关于发行说明 . . . . .	1	第 6 章 对信息中心进行故障诊断 . . . . .	27
第 2 章 V8.3 的新功能 . . . . .	3	企业搜索文档 . . . . .	29
已翻译的企业搜索组件 . . . . .	6	访问有关 <b>WebSphere Information</b>	
第 3 章 限制、局限性和变通方法 . . . . .	7	<b>Integration</b> 的信息 . . . . .	31
第 4 章 对管理控制台的更新 . . . . .	11	提供对文档的意见 . . . . .	33
第 5 章 文档更新 . . . . .	13	与 <b>IBM</b> 联系 . . . . .	35
安装 . . . . .	13	商标 . . . . .	37
管理和配置 . . . . .	14	声明 . . . . .	39
文本分析集成 . . . . .	21		
重建同义词字典 . . . . .	23		
编程和 API . . . . .	23		



---

## 第 1 章 关于发行说明

本文档描述 IBM® WebSphere® Information Integrator OmniFind™ Edition V8.3 的其他功能、局限性、限制、变通方法以及文档更新。在安装和使用此产品之前，请仔细阅读本文档。

要了解更多更新，请参阅 WebSphere Information Integrator OmniFind Edition Support Web 站点提供的技术说明。





---

## 第 2 章 V8.3 的新功能

IBM WebSphere Information Integrator OmniFind Edition V8.3 (企业搜索) 支持新的数据源、提高了搜索质量和性能而且易于使用、显著改进了样本搜索应用程序并扩展了文本分析功能。

### 与桌面搜索产品集成: Google 和 X1 Technologies

可以使用 Google 提供的桌面搜索技术 Google Desktop Search for Enterprise 或 X1 提供的桌面搜索技术 X1 Technologies Desktop Search Solution 来搜索桌面和企业搜索集合。

如果要使用 Google Desktop Search 来搜索企业搜索集合, 则需安装 WebSphere II OmniFind Edition 的 Google Desktop Search 插件。该插件是随 WebSphere II OmniFind Edition 一起交付的。请参阅 CD 映像根目录中 docs/en\_US/pdf 中的插件用户指南 (omnifind\_gds\_userguide\_v83.pdf)。

要了解有关 X1 Technologies Desktop 的更多信息, 请访问 X1 Technologies Web 站点, 网址为 <http://www.x1.com/>。

### 受支持的新数据源

V8.3 支持的新数据源包括:

- WebSphere Portal Document Manager V5.1.0.1
- IBM WebSphere Portal V5.1.0.1 Web 站点
- Lotus® Domino® Document Manager V6.5.1 (以前称为 Domino.Doc®)
- FileNet P8 CM 3.0 和 3.5
- Open Text Livelink Enterprise Server 9.2 和 9.5
- DB2 通用数据库™ z/OS® 版本号 8.x (新版本)
- Lotus QuickPlace® V6.5.1 (以前称为 Team Workplace™ 和 QuickPlace)
- Sybase V11.9.2、V12.0、V12.5 或更高版本
- Workplace Web Content Management V2.5 和 V5.1
- 可通过 IBM WebSphere Information Integrator Classic Federation for z/OS 来访问的大型机数据源。(需要单独购买 WebSphere Information Integrator Classic Federation。)
  - VSAM
  - IMS
  - CA-Datcom
  - Software AG Adabas

下列数据源已被扩展为支持使用本机数据源验证当前用户凭证:

- WebSphere Portal Document Manager
- WebSphere Portal
- Domino Document Manager

- Lotus QuickPlace
- FileNet CS
- Windows® 文件系统
- Documentum

## 搜索质量、相关性和功能

- 可扩展的排名功能
  - 升降字典：影响文档相关性
  - 字段升降：对升降类指定字段以及对升降类指定升降值
  - 静态分数控制：通过使用 URI 模式来提高或降低分数，例如：
 

```
http://servername/directory/*
db2://dbname/schema.tablename/*
domino://servername/dbname/*
```
- 完全支持通配符，并且在短语中也支持通配符
- 无用词字典
- 更好地支持诸如韩国语和德语之类的语言中使用的复合词。例如，查询单词 *Vertrag*（合同）时还将返回包含单词 *Versicherungsvertrag*（保险合同）的文档
- 改进了对多个词单元的同义词支持，例如，不再需要引号
- 增强了对汉语和阿拉伯语的支持
- 提供了用于所有数据源（包括 Web 和 NNTP）的搜寻器插件

## 可伸缩性和性能

- 利用 WebSphere Information Integrator Event Publisher Edition 来将 DB2 通用数据库更改推送到解析器。（必须单独购买 WebSphere Information Integrator Event Publisher Edition。）
- 提供了快速方式重新搜寻功能：通过防止进程检测已被删除的文档来提高性能
- 增加了解析器服务会话数
- 减少了索引和存储内容消耗的磁盘空间量，对于包含亚洲语言文档的集合以及包含许多字段的集合来说尤其如此

## 可使用性和配置

- 支持 Solaris Operating Environment 9
- 提供了用于管理集合的搜索和索引 API：使用搜索和索引 API 来开发定制应用程序并将这些应用程序部署在非企业搜索服务器上
- 可以将 DB2 通用数据库安装在非企业搜索服务器（远程服务器）上
- 仍然需要 WebSphere Application Server，但不再需要 WebSphere Application Server Network Deployment（尽管仍支持它）
- 自动安装 WebSphere Information Integrator Content Edition
- 用于支持自动化的操作 shell 命令
- 能够覆盖缺省的 Microsoft Exchange 字段映射
- 对于数据源搜寻器，支持“作为内容进行处理”：可以设置用于分类和总结等操作的用户定义字段
- 配置和管理 Web 搜寻器更为简便
- 在日志查看器中支持分页

## 功能增强的搜索应用程序

- 可以通过在配置文件中指定选项来进行定制。您不再需要了解 JSP。
- 搜索结果
  - 增强了排序和过滤功能（可以定制过滤器以查看特定文件类型的文档、来自特定来源类型的文档或者与预定义查询匹配的文档）
  - 能够从文档标题中除去冗余文本或无意义的文本
  - 能够配置日期数据的缺省格式
  - 文档来源图标标识了文档来源
  - 访存（单击可查看内容）
  - 在结果中进行搜索
- 外观
  - 可配置的条幅、工具栏、选项卡和链接
  - 增强的类别搜索和浏览功能
  - 增强的首选项配置功能
  - 简化了“搜索”页，提供了更多的空间来处理搜索结果
- 高级搜索
  - 新的查询构建器
  - 更多搜索选项

## 可靠性、可用性和可维护性

- 更好地跟踪有关已删除文档的信息
- 更高效的记录格式
- 管理控制台中的出错帮助：链接至更详细的错误消息信息
- 管理控制台中的数据源监视器显示了已删除文档数
- 提供了选项以使操作员能够在系统工具栏上查看系统信息
- 显示有关已失败的索引构建的消息

## 文本分析和语义搜索功能

- 将文本分析结果写至 JDBC 数据源（DB2 通用数据库和 Oracle）以启用文本挖掘功能：
  - 能够生成 LOAD 脚本
  - 简单示例：可将包含人员的已注释结果映射并添加到包含相关人员属性列表
- 对用户提供的 CAS 处理程序进行基本启用，例如，首先将诸如类别、人员列表或元数据之类的项发送到企业搜索，然后将其放回到内容管理系统中。需要手工将内容添加回到存储库中。请参阅支持 [Web 站点](http://www-306.ibm.com/software/data/integration/db2ii/supportwomnifind.html)（<http://www-306.ibm.com/software/data/integration/db2ii/supportwomnifind.html>）上提供的相应技术说明。
- 对已指定实体进行检测的功能（仅由 IBM Services 提供）
  - 使用产品名、竞争对手名以及雇员等内容来对字典进行扩展
  - 运行注释器以便对包含该信息的文档添加注释
  - 启用语义搜索，允许搜索诸如“查找所有文档，在这些文档中，我的一个雇员与竞争对手的一个产品在同一段中出现”之类的内容。

## 搜索质量工具

企业搜索的搜索质量工具是一些命令行工具，它们通过验证查询排名、修改元数据、消除重复文档和监视 Web 搜寻来帮助您提高查询性能。

搜索质量工具仅在 DeveloperWorks Web 站点上提供。

---

## 已翻译的企业搜索组件

WebSphere II OmniFind Edition 已翻译成许多种语言。但是，并非所有组件都有各种语言的版本。

WebSphere II OmniFind Edition 的某些或全部组件已翻译成下列语言：

阿拉伯语	匈牙利语	斯洛伐克语
巴西葡萄牙语	意大利语	斯拉维尼亚语
捷克语	日语	西班牙语
丹麦语	韩国语	瑞典语
芬兰语	挪威语（博克马尔语）	繁体中文
法语	波兰语	
德语	俄语	
希伯莱语	简体中文	

下列 WebSphere II OmniFind Edition 产品组件已翻译成某些或全部语言：

### 安装程序

除阿拉伯语和希伯莱语以外的所有语言。

### “第一步”程序

除阿拉伯语和希伯莱语以外的所有语言。

### WebSphere Portal 迁移程序

除阿拉伯语和希伯莱语以外的所有语言。

### 管理控制台和管理消息

除阿拉伯语和希伯莱语以外的所有语言。

### 搜索应用程序和搜索应用程序消息

所有语言。

### Google Desktop Search 插件

仅英语。

---

## 第 3 章 限制、局限性和变通方法

WebSphere Information Integrator OmniFind Edition 有若干限制、局限性和变通方法。

### WebSphere Information Integrator Content Edition 技术说明

在使用 Content Edition 搜寻器来创建带有 Documentum 和 FileNet 等数据源的集合前，请阅读相应数据源的技术说明。请访问 WebSphere Information Integrator Content Edition 支持 [Web](http://www-306.ibm.com/software/data/integration/db2ii/supportcontent.html) 站点，网址为 <http://www-306.ibm.com/software/data/integration/db2ii/supportcontent.html>。单击 Self help 下面的 **Technotes**，或者单击 Support 页面上的任何技术说明。

务必阅读下列问题的技术说明：

- FileNet P8 Content Manager 连接器的必需设置：

如果计划搜索 FileNet P8 Content Manager 数据源，则必须使用以下连接器属性：  
Include System Properties = True

- 要搜寻 FileNet CS 数据源，需要配置 RMI 代理连接器：

如果计划搜索 FileNet CS 数据源，则必须在 Windows 服务器上配置 RMI 代理连接器。

- 如果要直接搜寻 Documentum 存储库，则需要执行一个附加的配置步骤：

将本机 Documentum 库（Windows 上的 dmcl40.dll）复制到 *ICE\_HOME*/bin 目录中。

### 搜寻 Content Edition 来源的本机访问控制表信息

当启动 Content Edition 搜寻器以搜寻所有更新内容或者搜寻新文档和已修改文档时，Content Edition 搜寻器无法检测对已搜寻 Content Edition 文档的本机访问控制表信息所作的更改。

对于已建立了索引的文档，如果要对它们的已更改本机访问控制表信息建立索引，则必须在启动 Content Edition 搜寻器时启动完全搜寻。

### Content Edition 搜寻器：端口与企业搜索服务冲突

缺省情况下，WebSphere Information Integrator Content Edition 的数据存储服务器与 WebSphere Information Integrator OmniFind Edition（企业搜索）的公共通信层使用同一个端口，即 6002。

如果计划使用完全安装选项（而不是定制安装选项）来将 WebSphere Information Integrator Content Edition 服务器安装到 WebSphere Information Integrator OmniFind Edition 服务器所在的服务器上，则必须使用 WebSphere Information Integrator OmniFind Edition 安装程序来更改企业搜索公共通信层的端口。

## FileNet 和 Documentum 的必需设置

如果要搜寻 FileNet 或 Documentum 数据源，请遵循下列指示信息：

### FileNet P8 Content Manager 连接器的必需设置

如果计划搜索 FileNet P8 Content Manager 数据源，则必须使用以下连接器属性：

```
Include System Properties = True
```

### 要搜寻 FileNet CS 数据源，需要配置 RMI 代理连接器

如果计划搜索 FileNet CS 数据源，则必须在 Windows 服务器上配置 RMI 代理连接器。

### 如果要直接搜寻 Documentum 存储库，则需要执行一个附加的配置步骤

将本机 Documentum 库（Windows 上的 dmcl40.dll）复制到 IICE\_HOME/bin 目录中。

## 使用搜索 portlet

为了支持搜索受保护的 WebSphere Portal 页面，必须从 WebSphere II OmniFind Edition 搜索 portlet 中提交搜索。从标准 ESSearchApplication 应用程序中提交的搜索将没有正确的凭证，因此无法正确地验证最终用户是否有权访问索引中的文档。

## QuickPlace 搜寻器

只有在下列情况下，QuickPlace 搜寻器才可以搜寻您在其中指定的 **Imported Page**、**Microsoft® Word Page**、**Microsoft Excel Page**、**Microsoft PowerPoint Page** 和 **Multiple Imported Pages** 这些附件：

- 服务器使用 Domino 因特网 ORB 间协议（DIIOP）。
- 启用了搜寻附件功能。
- 已将搜寻器配置为搜寻“\$FILE”字段或所有字段。

## DB2 搜寻器和事件发布队列管理器

使用事件发布功能的 DB2® 搜寻器通过客户机连接来连接到 WebSphere MQ 队列。为了建立客户机连接，队列管理器应该将其 CCSID 设置为 819。

作为 WebSphere MQ 管理员，运行以下命令：

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

## 用于安全套接字层（SSL）Web 搜寻的专用 cacerts 文件副本

Web 搜寻器的两个新配置参数允许使用专用 cacerts 文件来进行 HTTPS 搜寻。

在 crawl.properties 文件中，**cacerts\_path** 参数值是一个字符串，它包含所要使用的 cacerts 文件的绝对路径名。

此外，在 crawl.properties 文件中，**cacerts\_password** 参数值也是一个字符串，它包含读取密钥库时要使用的加密密码。在企业搜索管理控制台中无法编辑这些参数。如果有专用的 cacerts 文件，就可以在没有超级用户许可权的情况下更新 X.509 证书，以及为不同的搜寻器提供不同的证书。

## 在 Windows 和 UNIX 文件系统中执行搜寻所需的许可权

Windows 和 UNIX® 文件系统搜寻器根据企业搜索管理员指定的读许可权来搜寻文档。对于 Windows 来说，管理员是“IBM WebSphere Information Integrator OmniFind Edition”NT 服务的帐户。

在 Windows 上，可以为已搜寻的目录设置用户标识和密码。但是，此用户标识和密码仅用于连接至共享网络文件夹。在这种情况下，Windows 文件系统搜寻器根据此用户标识对共享网络文件夹（而不是本地驱动器）拥有的读许可权来搜寻文件。并且，每次只允许一个用户连接至服务器或网络文件夹。

## 使用 ODBC 包装器来通过 WebSphere II Classic Federation for z/OS 搜寻传统数据源服务器上的 DB2 数据库

DB2 搜寻器可以通过 WebSphere Information Integrator Classic Federation 来搜寻传统数据源服务器上的数据库。要搜寻传统数据源服务器上的 DB2 数据库，必须使用 WebSphere Information Integrator Classic Federation 的 ODBC 包装器来使该数据库与非传统数据源服务器上的 DB2 数据库联合。

要对数据库进行联合：

1. 在传统数据源服务器上安装 WebSphere Information Integrator Classic Federation。
2. 在非传统数据源服务器上安装 WebSphere Information Integrator Classic Federation 客户机模块，该服务器包含要与传统数据源服务器上的数据库进行联合的数据库。
3. 配置 WebSphere Information Integrator Classic Federation 的 ODBC 驱动程序以连接至传统数据源服务器。
4. 连接至联合数据库并创建 ODBC 包装器，以便与传统数据源服务器上的数据库进行联合。
5. 指定 DB2 搜寻器以搜寻非传统数据源服务器上的联合数据库，以便可以通过 WebSphere Information Integrator Classic Federation 来搜寻传统数据源服务器上的数据库。

## HTML 解析器

HTML 解析器不会对文档的所有元数据建立索引，并且，它使用缺省标记列表。对于 V8.3 中的每个新集合都将使用此列表，而不会迁移现有集合。

您可以在管理控制台中控制和更改元数据映射。请转到“解析”页并单击将 **HTML 元数据映射到字段**。如果更改了元数据映射，所作的更改直到重新搜寻文档并再次解析文档后才会生效。

要了解有关 HTML 元数据映射的更多信息，请参阅“使用 HTML 搜索字段”主题或者查阅有关 HTML 解析器的技术说明。

## 搜寻器数据库转换

从旧版本的 WebSphere II OmniFind Edition 升级到 V8.3 后，第一次启动搜寻器时，将把所搜寻文档的数据库转换为新结构。当文档数为 100 万时，将旧数据库转换为新数据库的操作需要 15 分钟。



您必须等待转换完成。转换完成后，搜寻器将开始进行搜寻。您不需要重新启动搜寻器。

## 暂停 Web 搜寻器

在某些情况下，当 Web 搜寻器正在运行但似乎未搜寻任何新文档时，有可能是搜寻器已达到“高级 Web 搜寻器属性”页中定义的临时存储器最大已搜寻文档数限制。要验证搜寻器是否已达到此限制，必须检查 `webrds` 表中的文档数。要检查此表，请运行以下 DB2 命令。对于四服务器安装来说，请在搜寻器服务器上运行此命令：

```
db2 "select count(*) from schema_name.webrds"
```

其中 `schema_name` 是未执行搜寻的搜寻器的模式名。

如果文档数等于或略大于指定的限制，请确保解析器正在运行。如果解析器正在运行，请允许它完成运行。



---

## 第 4 章 对管理控制台的更新

在翻译文档后，已经对管理控制台中的几个字段进行了修订。管理控制台是主要的搜索管理员界面。

### “通配符”页的选项

在不搜索字段的查询中支持通配符复选框的文本已更改为在搜索自由文本的查询中支持通配符。

指定哪些字段支持通配符选项的指示信息文本已更改为在搜索字段的查询中支持通配符。新的单选按钮**无字段**使您能够指定包含通配符的查询不搜索任何字段。

### “外部 JDBC 数据源”页的选项

从此页中除去了下列字段：

- 自由文本搜索
- 字段搜索
- 参数搜索

在配置选项以指示如何搜索外部 JDBC 数据库表时，可以指定能否搜索该字段、能否在搜索结果中返回该字段以及在源数据库表中能否将某个字段用作唯一标识字段。

### Windows 和 UNIX 文件系统搜寻器的已搜寻目录所占的百分比

在企业搜索管理控制台中，对于 Windows 和 UNIX 文件系统搜寻器来说，“搜寻器详细信息”页上的进度条表示已搜寻文件数占所要搜寻的总文件数的百分比。

在 WebSphere II OmniFind Edition V8.2.2 中，进度条表示的是已搜寻目录数占所要搜寻的总目录数的百分比。

### “折叠搜索结果”页

“折叠搜索结果”页上的指示信息以及有关折叠搜索结果的文档内容指出：在搜索结果中只能折叠来自 Web 和 NNTP 来源的文档。此陈述不正确。您可以配置规则，从而将集合的所有源类型中来自同一源的文档进行折叠。

### 集合安全性和站点折叠

如果启用了集合安全性，则除非您指定不想在查询处理期间验证用户凭证，否则不会折叠来自同一个站点的搜索结果。在验证当前凭证时，将检查每个文档的来源，并且有可能路由那些来源以便进行验证。要对来自安全集合的搜索结果进行折叠：

1. 在企业搜索管理控制台的“文档级安全性”页上，选中**在查询处理期间不验证当前凭证**复选框，然后单击**确定**。
2. 监视集合，选择“搜索”页，然后停止并重新启动搜索服务器进程。

当用户对集合执行查询时，对于已配置为要折叠的站点，它们的文档在搜索结果中将折叠起来。在样本搜索应用程序中，用户可以通过单击 **[来自这个来源的更多结果]** 链接来查看折叠的结果。

## 集合安全性和 Windows 文件系统搜寻器

如果启用了集合安全性，并且您配置的是 Windows 文件系统搜寻器，则当下列两个条件都成立时，无法在查询处理期间验证用户凭证：

- 要搜寻的 Windows 服务器不是域的成员
- 要搜寻的目录是远程目录（例如 \\servername\hostname）

---

## 第 5 章 文档更新

在翻译文档后，已经对几个主题进行了修订。信息中心和 PDF 文档都提供了 WebSphere Information Integrator OmniFind Edition 的信息。

---

### 安装

在翻译安装信息后，已经对安装文档的几个部分进行了更改。PDF 文档 (iysi.pdf) 和信息中心都提供了安装信息。

#### Solaris Operating Environment 的必需设置

如果您看到以下错误，则必须检查是否正在使用受支持的 Solaris 内核：

此应用程序需要 Java 运行时环境 (JRE) 才能运行。  
在您的计算机上找不到该环境。  
请使用命令行开关 `-is:javahome` 来指定有效的 JRE。  
要获取更多帮助，请使用 `-is:help` 选项。

受支持的内核是 SunOS 5.9 Generic 112233-12 (2004 年 3 月)。

#### Solaris Operating Environment 的必需补丁

《企业搜索安装指南》指出需要下列级别的 Solaris: Sun Solaris 9 64 位, kernel SunOS 5.9 Generic 112233-12 (2004 年 3 月)。此外，确保安装了下列补丁。如果尚未安装这些补丁，则必须安装它们：

- SunOS 5.9 Patch 111722-04 04/16/2003
- Sun C++ 5.6 Patch 111711-12 2004/10/25

要检查补丁，请运行以下命令：

```
>showrev -a |grep 111722
```

您应该看到类似于以下的输出信息：

```
Patch: 111722-04 Obsoletes: Requires: Incompatibles:  
Packages: SUNWlibms, SUNWlmsx, SUNWlibm
```

补丁级别必须是 04 或更高。

并且，运行以下命令：

```
>showrev -a |grep 111711
```

您应该看到类似于以下的输出信息：

```
Patch: 111711-12 Obsoletes: Requires: Incompatibles:  
Packages: SUNWlibC
```

补丁级别必须是 12 或更高。

如果该命令未返回任何信息，则需要使用 Solaris Management Console 或 `patchadd` 实用程序来安装补丁。

## 硬件和磁盘空间要求

在“大型安装”下面，文本“在每台服务器上需要 8 GB 的 RAM（对于 Windows，需要 4 GB 的 RAM）”不正确。对于 Windows 来说，大型安装需要 8 GB 的 RAM。

并且，所有操作系统的建议调页空间都是 8 GB。

## 启动 WebSphere II OmniFind Edition 安装启动板

在“在单一服务器上安装 WebSphere II OmniFind Edition”、“安装搜寻器组件”、“安装搜索组件”和“安装索引组件”主题中，用于启动安装启动板的代码示例都包含排字错误。请使用以下正确文本：

AIX: 作为 root 用户登录并输入 `./install-aix.bin`

Linux: 作为 root 用户登录并输入 `./install-linux.bin`

Solaris: 作为 root 用户登录并输入 `./install-solaris.bin`

## 从 AIX、Linux 和 Solaris 系统中除去 WebSphere II OmniFind Edition

在第三段中，示例应该是：

```
./uninstall.bin -silent -options response_file_name
```

过程中的步骤 3 应该是：

输入 `./uninstall.bin`

## 在 AIX、Linux 和 Solaris 上启动企业搜索

用于启动 WebSphere Application Server Network Deployment 的步骤 3 和步骤 5 包含排字错误。这些步骤应该是：

3. 输入 `./startManager.sh`

5. 输入

```
./startNode.sh  
./startServer.sh servername
```

## 第一次使用 V8.3 搜寻 Web 数据源

在安装 WebSphere II OmniFind Edition 之前，请务必备份数据。

WebSphere II OmniFind Edition V8.3 使用新的数据库模式来存储搜寻到的文档。第一次启动 Web 搜寻器时，将自动地把数据库转换为新模式。要了解更多信息，请参阅第 9 页的『搜寻器数据库转换』。

---

## 管理和配置

在翻译管理和配置信息后，已经对管理文档中的几个部分作了更改。PDF 文档（`iiysa.pdf`）和信息中心都提供了管理和配置信息。

## 在 WebSphere Application Server 中启动和停止 ESSearchApplication 和 server1 应用程序

管理信息中有关启动和停止 ESSearchApplication 和 server1 应用程序的指示信息不正确。用于 WebSphere Application Server V6.0 和 V5.1 的指示信息是不同的。请使用下列指示信息来启动或停止这些应用程序:

要停止应用程序:

- 对于 WebSphere Application Server V6.0:
  1. 作为 root 用户登录, 然后切换到安装了 WebSphere Application Server 的 bin 目录。
  2. 运行下列命令:

AIX、Linux 和 Solaris:

```
./stopServer.sh server1  
./stopServer.sh ESSearchServer
```

Windows:

```
stopServer server1  
stopServer ESSearchServer
```

- 对于 WebSphere Application Server V5.1, 打开 WebSphere Application Server 管理控制台, 然后停止应用程序。

要启动应用程序:

- 对于 WebSphere Application Server V6.0:
  1. 作为 root 用户登录, 然后切换到安装了 WebSphere Application Server 的 bin 目录。
  2. 运行下列命令:

AIX、Linux 和 Solaris:

```
./startServer.sh ESSearchServer  
./startServer.sh server1
```

Windows:

```
startServer ESSearchServer  
startServer server1
```

- 对于 WebSphere Application Server V5.1, 打开 WebSphere Application Server 管理控制台, 然后启动应用程序。

## 更新企业搜索的安全性设置

如果在安装企业搜索 (WebSphere II OmniFind Edition) 后启用 WebSphere Application Server 安全性, 请执行本过程中的所有步骤。

如果在安装企业搜索前启用 WebSphere Application Server 安全性, 则只需执行步骤 2。企业搜索安装程序会自动使用 WebSphere 用户名和密码来更新 es.cfg 文件 (步骤 3 和步骤 4)。

要更新企业搜索安全性设置:

1. 确保输入的企业搜索管理员标识在 WebSphere Application Server 使用的用户注册表中是有效用户。
2. 对于 ESSearchApplication 应用程序, 转至以下目录:

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

使用有效 WebSphere Application Server 用户标识来更新 ESSearchApplication 应用程序的 config.properties 文件中的用户名和密码条目。

3. 使用有效 WebSphere Application Server 用户标识来更新 ES\_NODE\_ROOT/nodeinfo/es.cfg 文件中的 WASUser 条目。
4. 使用 eschangewaspw.bat 或 eschangewaspw.sh 命令来更改 es.cfg 文件中的 WASPassword 条目。请参阅下一节以获取有关如何更改此密码的指示信息。

要在单服务器安装中更改密码:

1. 作为企业搜索管理员登录。
2. 通过输入此命令以停止服务器: `esadmin stop`
3. 停止 server1 和 ESSearchServer 企业应用程序。
  - 对于 WebSphere Application Server V6.0:
    - a. 作为 root 用户登录, 然后切换到安装了 WebSphere Application Server 的 bin 目录。
    - b. 运行下列命令:

```
AIX、Linux 和 Solaris:  
./stopServer.sh server1  
./stopServer.sh ESSearchServer
```

```
Windows:  
stopServer server1  
stopServer ESSearchServer
```

- 对于 WebSphere Application Server V5.1, 打开 WebSphere Application Server 管理控制台, 然后停止应用程序。
4. 使用 AIX、Linux 或 Solaris 操作系统命令或 Microsoft Windows 更改密码工具来更改企业搜索管理员用户标识的系统密码。
  5. 运行以下脚本, 其中 *newValue* 是步骤 4 中指定的密码:
    - AIX、Linux 和 Solaris: `eschangewaspw.sh newValue`
    - Windows: `eschangewaspw newValue`
  6. 启动 server1 和 ESSearchServer 应用程序:
    - 对于 WebSphere Application Server V6.0:
      - a. 作为 root 用户登录, 然后切换到安装了 WebSphere Application Server 的 bin 目录。
      - b. 运行下列命令:

```
AIX、Linux 和 Solaris:  
./startServer.sh ESSearchServer  
./startServer.sh server1
```

Windows:

```
startServer ESSearchServer
startServer server1
```

- 对于 WebSphere Application Server V5.1, 打开 WebSphere Application Server 管理控制台, 然后停止应用程序。

7. 通过输入下列命令, 重新启动 WebSphere II OmniFind Edition 公共通信层 (CCL):

- AIX、Linux 和 Solaris:

```
stopccl.sh
startccl.sh -bg
```

- Windows:

```
stopccl
startccl
```

8. 通过输入此命令来重新启动企业搜索: `esadmin start`

要在多服务器安装中更改密码:

1. 在安装了索引服务器的计算机上, 作为企业搜索管理员登录。
2. 通过输入此命令来进入系统维护方式: `esadmin system maintenance`
3. 使用 AIX、Linux 或 Solaris 操作系统命令或 Microsoft Windows 更改密码工具来更改企业搜索管理员用户标识的系统密码。
4. 运行以下脚本, 其中 *newValue* 是步骤 3 中指定的密码:

- AIX、Linux 或 Solaris: `eschangewaspw.sh newValue`

- Windows: `eschangewaspw newValue`

5. 通过在索引服务器上输入下列命令, 重新启动 WebSphere II OmniFind Edition 公共通信层 (CCL):

- AIX、Linux 或 Solaris:

```
stopccl.sh
startccl.sh -bg
```

- Windows:

```
stopccl
startccl
```

6. 在安装了搜寻器服务器的计算机上, 作为企业搜索管理员登录。
7. 通过使用 UNIX 操作系统命令或 Windows 更改密码工具来更改企业搜索管理员用户标识的系统密码。此密码必须与步骤 3 中指定的密码相匹配。
8. 运行以下脚本, 其中 *newValue* 是步骤 4 中指定的密码:

- AIX、Linux 或 Solaris:

```
eschangepw.sh newValue
```

- Windows:

```
eschangepw newValue
```

9. 通过输入以下命令在搜寻器服务器上重新启动搜寻器:

- AIX、Linux 或 Solaris:

```
stopccl.sh
startccl.sh -bg
```

- Windows:

```
stopccl  
startccl
```

10. 在其中一台安装了搜索服务器的计算机上，作为企业搜索管理员登录。
11. 停止 IBM HTTP Web 服务器。（此步骤可防止 Network Dispatcher 将查询路由到此服务器。）
12. 通过输入下列命令在搜索服务器上停止 CCL:
  - AIX、Linux 或 Solaris:

```
stopccl.sh
```
  - Windows:

```
stopccl
```
13. 打开 WebSphere Application Server 管理控制台，然后停止 server1 和 ESSearchServer 应用程序。
14. 通过使用 UNIX 操作系统命令或 Windows 更改密码工具来更改企业搜索管理员用户标识的系统密码。此密码必须与步骤 3 中指定的密码相匹配。
15. 运行以下脚本，其中 *newValue* 是步骤 3 中指定的密码:
  - AIX、Linux 或 Solaris:

```
eschangepw.sh newValue
```
  - Windows:

```
eschangepw newValue
```
16. 通过输入下列命令在搜索服务器上重新启动 CCL:
  - AIX、Linux 或 Solaris:

```
startccl.sh -bg
```
  - Windows:

```
startccl
```
17. 在第二台搜索服务器上，重复步骤 10 到步骤 16。
18. 启动 IBM HTTP Web 服务器。
19. 在 WebSphere Application Server 管理控制台中，启动 server1 和 ESSearchServer 企业应用程序。
20. 在安装了索引服务器的计算机上，作为企业搜索管理员登录，然后输入此命令以启动企业搜索：`esadmin start`。
21. 登录到企业搜索管理控制台，监视集合，单击搜索页，然后单击停止和启动以重新启动搜索服务器。对企业搜索系统中的每个集合重复此步骤。

## 在安装 WebSphere II OmniFind Edition V8.3 后进行重新搜寻

在安装 WebSphere II OmniFind Edition V8.3 后，将重新搜寻所有文档（那些已由 Web 和 NNTP 搜寻器搜寻过的文档除外）。

在安装 V8.3 后首次启动非 Web 和非 NNTP 搜寻器时，将搜寻所有数据源中的所有文档并再次对它们建立索引，而不只是针对上次搜寻数据源后的新文档或更改过的文档。

在搜索集合中每个表、目录或文件夹中的所有文档后，搜寻器将转换为执行正常更新处理并开始只搜寻新文档或更改过的文档。通常，仅当对搜寻器配置所作的更改要求进行完全重新搜寻时，才有必要执行该操作。例如，如果数据源中字段或列的显示名



发生了更改，则搜寻器必须重新搜寻所有文档，才能对文档中的新信息建立索引。

## 恢复排名属性的缺省值

调整排名属性后，可以通过编辑

`ES_NODE_ROOT/master_config/colid.runtime.nodeid/runtime.properties` 文件来将属性重新设置为缺省值。

在多服务器系统中，需编辑每台搜索服务器的 `runtime.properties` 文件。这些文件位于索引服务器上的 `ES_NODE_ROOT/master_config/colid.runtime.nodeid` 目录中，例如：

```
ES_NODE_ROOT/master_config/mycolid.runtime.node3/runtime.properties
ES_NODE_ROOT/master_config/mycolid.runtime.node4/runtime.properties
```

`node3` 和 `node4` 是各搜索服务器的服务器标识。

要将集合的排名属性值复原为缺省值：

1. 标识要将缺省排名值复原至的集合的标识。集合标识包含在 `ES_NODE_ROOT/master_config/collections.ini` 文件中。为了便于查看，可以对此文件进行排序，如以下示例所示：

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

在此示例中，`collection1` 的标识为 `coll`。

2. 在文本编辑器中，打开 `ES_NODE_ROOT/master_config/colid.runtime.nodeid/runtime.properties` 文件，其中 `colid` 是要更改的集合的标识，`nodeid` 是运行此搜索会话的搜索服务器的标识。对 `colid.runtime.nodeid/runtime.properties` 文件进行下列更改：

- 删除下列属性：

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- 在您编辑的 `runtime.properties` 文件中，如果 `trevi.sourcetype=1`（链接排名集合），则打开 `ES_INSTALL_ROOT/default_config/runtime.1/runtime.properties` 文件并将下列缺省属性复制并粘贴到 `colid.runtime.nodeid/runtime.properties` 文件中：

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- 在 `colid.runtime.nodeid/runtime.properties` 文件中，如果 `trevi.sourcetype=2`（日期排名集合），则打开 `ES_INSTALL_ROOT/default_config/runtime.2/runtime.properties` 文件并将下列属性复制并粘贴到 `colid.runtime.nodeid/runtime.properties` 文件中：

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

- 在 `colid.runtime.nodeid/runtime.properties` 文件中，如果 `trevi.sourcetype=3`（没有排名集合），则打开 `ES_INSTALL_ROOT/default_config/runtime.0/runtime.properties` 文件并将下列属性复制并粘贴到您正在编辑的 `colid.runtime.nodeid/runtime.properties` 文件中：

```
trevi.autorank.dfthreshold1
trevi.autorank.dfthreshold2
trevi.autorank.dfthreshold3
trevi.autorank.rc0.*
trevi.autorank.rc1.*
```

3. 在管理控制台中，停止并重新启动此集合的搜索进程以使新值生效。

对要将缺省排名值复原至的每个集合重复这些步骤。

## Web 搜寻器 HTTP 返回码

添加了专用的 HTTP 返回码 618 来标识最初被重定向（3xx）但没有重定向 URL 的文档。由于这些返回码未包含有用的内容，所以不会在企业搜索系统中传递它们。

## 仅搜寻新文档和已修改文档

可以启动除 Web 和 NNTP 搜寻器以外的搜寻器以便仅搜寻新文档和已修改文档。在这种情况下，对于比上次搜寻期间搜寻的所有文档更旧的文档，该搜寻器不会搜寻它们。

例如，对于 Windows 和 UNIX 文件系统搜寻器来说，如果您将一个文件复制到已搜寻的目录中，并且该文件的上次修改日期早于所有已搜寻文档的上次修改日期，则搜寻器不会搜寻该文件。尽管该文件是刚刚复制的，搜寻器也不会搜寻该文件。

## 索引构建状态示例

一些索引构建状态示例不正确。以下文本替换“企业搜索命令、返回码和会话标识”主题中的示例：

```
<?xml version="1.0"?>
<Monitor Type="MainIndexHistory" Count="1">
  <IndexStatus Id="1">
    <StartTime>1131987633901</StartTime>
    <Progress>0</Progress>
    <CurrentPhase>0</CurrentPhase>
    <TotalPhase>0</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>
<ScheduleStatu>
</Monitor>
```

## 索引构建状态命令的元素

“企业搜索命令、返回码和会话标识”主题中的索引构建状态命令元素表包含不正确的属性和描述:

属性 `progress` 指定了两次。将第二个 `progress` 属性替换为 `Status` 属性并提供以下描述: “0 表示索引构建和复制, -1 表示索引构建请求失败, 1 表示索引构建和 / 或复制正在进行”。

`IndexCopyTime` 的描述不正确。正确的描述是“将索引复制到所有搜索服务器耗用的总时间”。

缺少 `IndexBuildTime` 属性的描述。正确的描述是“所有索引构建阶段的总时间”。

有两个新属性, 它们的描述如下所示:

**JobID** 与每个索引构建请求相关联的唯一标识

### **MessagesAvailable**

这个布尔值指示发生故障时是否有错误消息

## 配置并发索引构建

在单个集合中启用并发索引构建(通过选中各个集合的索引重组和刷新请求可以同时运行框)之后, 您可能会看到与正在停止的索引刷新相关的错误消息。在特定索引重组阶段, 系统会强制停止运行刷新构建并且不允许启动任何新的刷新构建, 以使重组的索引构建能够完成。

您可能会看到下列消息:

FFQC001E: 强制停止集合 *collection ID* 的索引复制。

FFQC0008E: 为集合 *collection ID* 构建 *index build type* 索引时出错。

您可以忽略这些消息。

并且, 如果在单个集合中启用并发索引构建, 则解析器也将与索引构建并行运行。请确保有足够的系统资源来供索引构建进程和解析进程使用。

---

## 文本分析集成

在翻译文本分析信息后, 已经对文本分析集成文档的几个部分进行了更改。PDF 文档 (*iiyst.pdf*) 和信息中心都提供了文本分析信息。

### 在数据库中存储分析结果

下列信息对文本分析集成信息进行补充。

如果要使用 DB2 数据库或装入文件来存储 `Cas2Jdbc` 处理程序返回的结果, 请使用 JDBC 3 类驱动程序, 而不是 4 类驱动程序。

按原始文档建议的那样上载引用了 4 类驱动程序的 Cas2Jdbc 映射文件会引起错误，从而无法成功地上载文件。在某些系统配置上，引用 4 类驱动程序会导致企业搜索系统锁定。

请使用 3 类驱动程序（名为 COM.ibm.db2.jdbc.net.DB2Driver），此驱动程序也被称为“网络驱动程序”。使用此驱动程序时，要求 DB2 运行一个特殊的 applet 服务器。要运行该 applet 服务器，请参阅 WebSphere II OmniFind Edition 软件包中提供的 DB2 通用数据库文档。

确保驱动程序库的维护级别始终与 DB2 applet 服务器的维护级别相同。请参考 Cas2Jdbc 映射文件中 DB2 通用数据库安装的 Java 目录下的下列驱动程序库：db2java.zip 和 db2jcc.jar。

以下 Cas2Jdbc 映射文件示例使用 3 类驱动程序以及正确的 ZIP 和 JAR 文件：

```
. . .  
<databaseConnection>  
  <connectionUrl>db2://hostname:port/databasename  
  </connectionUrl>  
  <driver type="jdbc">COM.ibm.db2.jdbc.net.DB2Driver</driver>  
  
  <authentication>  
    <username>username</username>  
    <password>password</password>  
  </authentication>  
  
  <driverLibraries>  
    <driverLibrary>/opt/IBM/db2/V8.1/java/db2java.zip  
    </driverLibrary>  
    <driverLibrary>/opt/IBM/db2/V8.1/java/db2jcc.jar  
    </driverLibrary>  
  </driverLibraries>  
</databaseConnection>  
  
. . .
```

## 删除无用词

添加了下列文本以阐明当查询只包含无用词时发生的情况。*Text Analysis Integration* 文档 (iyst.pdf) 中的“Custom stop word dictionaries”以及 *Programming Guide and API Reference* (iysp.pdf) 中的“Controlling query behavior”对无用词字典作了描述。

如果查询只包含无用词，则不会在查询处理期间除去任何无用词。

如果在无用词处理期间除去了所有查询项，则结果集将是空的。为了确保返回搜索结果，当所有查询项都是无用词时，将禁用“删除无用词”功能。

例如，如果 *car* 一词是无用词并且您搜索 *car*，则搜索结果将包含与 *car* 一词匹配的文档。如果搜索 *car buick*，则搜索结果只包含与 *buick* 一词匹配的文档。

## 内置功能

在“内置功能”主题中接近末尾的地方，以下文本不正确：

通过将相关功能结构映射至索引或数据库并在语义搜索查询中指定范围，可以只检索与查询精确匹配的文档部分。

正确的文本是:

通过将相关功能结构映射至索引和数据库并在语义搜索查询中指定范围, 可以只检索与查询精确匹配的文档部分。

## 检索与语义搜索查询匹配的文档部分

在“检索与语义搜索查询匹配的文档部分”主题中, 以下文本不正确: “解析 `Result.getProperty (“TargetElement”)` 返回的字符串以查找出现标识”。

正确的文本是“解析 `Result.getProperty (“TargetElement”)` 返回的字符串以查找所有出现标识”。

## 安装和运行企业搜索基本注释器

应该在“安装和运行企业搜索基本注释器”主题中添加以下文本:

“要为 `OmniFind` 注释器设置类路径, 必须在 `setUIMAClasspath` 脚本末尾调用 `setOfAnnotEnv` 脚本。”

## 重建同义词字典

要在 `WebSphere Information Integrator OmniFind Edition V8.3` 系统中使用您为 `V8.2.2` 构建的同义词字典, 必须迁移字典 (`.dic`) 文件。

在安装 `WebSphere Information Integrator OmniFind Edition V8.3` 之后, 请执行下列步骤以迁移定制同义词字典:

1. 在“集合”视图中, 在配置搜索服务器选项页上, 从集合中除去旧的同义词字典。
2. 在“系统”视图中, 在配置同义词字典页上, 从企业搜索系统中删除旧的同义词字典。
3. 将同义词字典的源 XML 文件转换为新的字典 (`.dic`) 文件。

请参阅 *Text Analysis Integration* 一书中的 *Creating a synonym dictionary* 以了解有关如何使用 `essyndictbuilder` 工具的指示信息。

4. 在“系统”视图中, 在配置同义词字典页上, 将新的同义词字典添加到企业搜索系统中。如果 `.dic` 文件小于 8 MB, 则该文件可以在您的系统上。如果 `.dic` 文件大于 8 MB, 它就必须放在索引服务器上。
5. 在“集合”视图中, 在配置搜索服务器选项页上, 使新的同义词字典与要使用它的集合相关联。

对于企业搜索系统中的每个定制同义词字典, 重复这些步骤。

---

## 编程和 API

在翻译编程和 API 信息后, 已经对应用程序编程和 API 参考文档的几个部分进行了更改。PDF 文档 (`iiysp.pdf`) 和信息中心都提供了编程和 API 参考信息。

### 安装客户机管理 API 工具箱

在已进行编号的列表中, 进行下列更改:

- 步骤 1: 对于多服务器安装, 在两台搜索服务器上都有 `es.siapi.toolkit.jar` 归档文件。

- 步骤 3: 不再需要 `es.client.cfg` 文件。与服务器进行通信时所需的配置文件打包在 `es.siapi.toolkit.jar` 文件中, 该文件是 `es.cfg`。此文件将被解压缩到工具箱解压缩后所在目录的根目录中。但是, 您必须编辑此文件并在文件末尾添加下列条目:

- `es_server_hostname=fully qualified host name of the enterprise search index server`

这是必需的字段, 它指定索引服务器主机名, 例如 `omnifind.server.ibm.com`。

- `.logFileName=absolute path of a log file`

这是可选字段, 它在特定的日志文件 (例如 `c:\temp\siapi.log`) 中注册跟踪信息。

## 将文档添加至集合

在“将文档添加至集合”主题的步骤 4 中, 不支持下列 API:

- `Field createField (java.lang.String fieldName, boolean[] value)`
- `Field createField (java.lang.String fieldName, double[] value)`
- `Field createField (java.lang.String fieldName, int[] value)`
- `Field createField (java.lang.String fieldName, java.lang.Sting[] value)`

## 重组和刷新索引

可以设置分段计数属性, 此属性向搜索和索引 API 代码指示应该何时刷新或重组索引。

以下属性设置中的值“2”表示刷新索引两次后将重组索引:

```
indexes.index.setProperty  
(IAdminConstants.BUILD_FRAGMENTATION_COUNT, 2)
```

请参阅索引构建样本以获取更多示例。

## 管理样本应用程序

在创建应用程序标识前, 必须注册该应用程序标识。请参阅“注册应用程序标识”主题以获取样本应用程序。

## 编译样本管理应用程序

在“编译样本管理应用程序”主题的步骤 3 中, 对于所有操作系统, 将文件 `es.client.cfg` 替换为 `es.cfg`。

## getInternalFieldName 方法

此方法的语法不正确。正确的语法是:

```
getInternalFieldName
```

## 编译样本搜索和数据侦听器应用程序

在“ESSearchApplication 搜索应用程序”一节中, Windows 后面的文本目录不正确。正确的文本是“Windows: C:\Program Files\IBM\es\samples\ESSearchApplication”。

并且, 在“数据侦听器客户机应用程序”一节中, Windows 后面的文本也不正确。正确的文本是“Windows: C:\Program Files\IBM\es\samples\datalistener”。

## 用于非 Web 来源的搜寻器插件

受支持的搜寻器列表不完整。您可以为 V8.3 中所有受支持的搜寻器编写应用程序（插件）。

### 检索目标 XML 元素

“在指定语义搜索的语义不明的项中，可以在 xmlf2 查询项中的一个 XML 元素（或注释）前面添加 # 符，可以在 xmlf2 查询项中的一个 XML 元素前面添加 # 符”一句包含重复的文本。

正确的句子是“在指定语义搜索的语义不明的项中，可以在 xmlf2 查询项中的一个 XML 元素（或注释）前面添加 # 符”。

### 查询语法

描述在查询中使用双引号（" "）的一节中，以下示例不正确：“短语中单独出现的通配符（\*）将被忽略，查询 "app\* pea\*" 的结果应该与 *apples pears* 相同。”

正确的句子是“短语中单独出现的通配符（\*）将被忽略，查询 "apples \* pears" 的结果应该与 *apples pears* 相同。”

### 样本搜索应用程序

在“样本搜索应用程序”主题中，指向两个主题链接不正确。“简单搜索”链接应该是“简单搜索和高级搜索”。此链接指向描述简单样本搜索应用程序和高级样本搜索应用程序的主题。

“高级搜索”链接不正确地指向了关于编译搜索应用程序的主题。





---

## 第 6 章 对信息中心进行故障诊断

如果您单击“帮助”按钮时主题未显示，或者您尝试启动信息中心但它未启动，则表明信息中心的配置可能不正确。

要解决此问题，请停止并重新启动信息中心。对于多服务器安装，因为在两台搜索服务器上安装了信息中心，所以必须在这两台服务器上启动和停止信息中心。

要停止并重新启动信息中心：

1. 通过运行以下命令来停止信息中心。对于多服务器安装来说，请在搜索服务器上运行以下命令：

**AIX**     /usr/IBM/WebSphereIIInformationCenter/ibm\_help/IC\_end.sh

**Linux 和 Solaris**

        /opt/IBM/WebSphereIIInformationCenter/ibm\_help/IC\_end.sh

**Windows**

        cd C:\WebSphereIIInformationCenter\ibm\_help

        IC\_end.bat

2. 从 /ibm\_help/eclipse/configuration 目录中删除 config.ini 以外的所有文件。
3. 启动信息中心：

**AIX**     /usr/IBM/WebSphereIIInformationCenter/ibm\_help/IC\_start.sh &

**Linux 和 Solaris**

        /opt/IBM/WebSphereIIInformationCenter/ibm\_help/IC\_start.sh &

**Windows**

        cd C:\WebSphereIIInformationCenter\ibm\_help

        IC\_start.bat

输入命令以启动信息中心之后，您不会看到指示信息中心已启动的确认信息。

4. 通过在浏览器中输入以下 URL 来测试信息中心：<http://localhost:8888>



---

## 企业搜索文档

您可以阅读 PDF 或 HTML 格式的 WebSphere Information Integrator OmniFind Edition 文档。

WebSphere Information Integrator OmniFind Edition 安装程序可以自动安装信息中心。安装程序在搜索服务器上安装信息中心。对于多服务器安装，在两台搜索服务器上都要安装信息中心。如果未安装信息中心，当您单击“帮助”时，将打开 IBM Web 站点上的信息中心。要查看 HTML 格式的企业搜索主题，请启动信息中心。

要查看 PDF 文档，请转到 docs/locale/pdf。例如，要查找英文文档，请转到 docs/en\_US/pdf。您也可以查看 WebSphere Information Integrator OmniFind Edition 支持站点上提供的 PDF 文档、下载、修订、技术说明和信息中心。

下表提供了可用的文档、文件名和位置。

表 1. 企业搜索的 PDF 文档和信息中心文档

文档	文件名	位置
《企业搜索安装指南》（信息中心也提供了此文档的主题）	iiysic82.pdf	docs/locale/pdf/
<i>Quick Start Guide</i>	iiyss.pdf	docs/locale/pdf/
《管理企业搜索》（信息中心也提供了此文档的主题）	iiysac82.pdf	docs/locale/pdf/
<i>Programming Guide and API Reference for Enterprise Search</i> （信息中心也提供了此文档的主题）	iiysp.pdf	docs/locale/pdf/
<i>Messages Reference</i> （信息中心也提供了此文档的主题）	iiysm.pdf	docs/locale/pdf/
<i>Installation Requirements for Enterprise Search</i> （信息中心也提供了此文档的主题）	iiysr.txt 或 iiysr.htm	docs/locale/（也可以从“第一步”程序中打开此文件。）
<i>Text Analysis Integration</i> （信息中心也提供了此文档的主题）	iiyst.pdf	docs/locale/pdf/
<i>Plug-in for Google Desktop Search</i>	omnifind_gds_userguide_v83.pdf	docs/en_US/pdf/（此文件只有英文版）
《发行说明》	iiysnc82.pdf	仅在 IBM WebSphere Information Integrator OmniFind Edition 文档 Web 站点上提供。



---

## 访问有关 WebSphere Information Integration 的信息

您可以通过打电话或浏览 Web 获得有关 WebSphere Information Integration 产品的信息。

此处提供的电话号码仅在美国有效:

- 要订购产品或了解一般信息: 1-800-IBM-CALL (1-800-426-2255)
- 要订购出版物: 1-800-879-2755

您也可以在 Web 上找到有关 WebSphere Information Integration 的信息, 网址为 [www.ibm.com/software/data/integration/db2ii/](http://www.ibm.com/software/data/integration/db2ii/)。此站点提供了下列方面的最新信息:

- 产品文档
- 产品下载
- 修订包
- 发行说明和其他支持文档
- 有关 WebSphere Information Integration 的新闻
- 链接到 Web 资源, 例如白皮书和 IBM Redbooks™
- 链接到新闻组和用户组
- 链接到 WebSphere Information Integration 产品的在线信息中心
- 订购书籍

要访问产品文档:

1. 访问 Web 地址 [www.ibm.com/software/data/integration/db2ii/](http://www.ibm.com/software/data/integration/db2ii/)。
2. 从下拉列表中选择产品, 例如 WebSphere Information Integrator OmniFind Edition。
3. 单击页面左边的 Support 链接。
4. 在 Learn 部分中, 选择要使用的链接。如果您选择的产品有信息中心, 则可选择信息中心的链接。请参阅第 32 页的图 1 以获取示例。

## Learn

- **Product documentation and manuals** (2 items)
- **Redbooks** (1 item)
- **V8.2 Documentation and release notes**

## Information Center

Provides fast, online centralized access to product information.

- [1.0](#)

图 1. WebSphere Information Integration Support Web 站点上的产品文档链接示例

---

## 提供对文档的意见

请发送您对本资料或其他 IBM WebSphere Information Integration 文档的任何意见。

您的反馈有助于 IBM 提供高质量的信息。请发送您对本资料或其他 WebSphere Information Integration 文档的任何意见。您可以使用下列任何方法来提供意见：

1. 使用 [www.ibm.com/software/awdtools/rcf/](http://www.ibm.com/software/awdtools/rcf/) 上的在线读者意见表发送您的意见。
2. 通过电子邮件将您的意见发送至 [ctscrcf@cn.ibm.com](mailto:ctscrcf@cn.ibm.com)。确保包括产品的名称、产品的版本号和资料的名称及部件号（如果适用的话）。如果您对特定文本有意见，请包括此文本的位置（例如，标题、表号或页码）。





---

## 与 IBM 联系

在中国，请致电下列其中一个号码以与 IBM 联系：

- 800-810-1818 或 (010) 84981188 分机 5151，可获得售前客户服务；
- 800-810-1818 或 (010) 84981188 分机 5200，可获得售后客户服务；
- 800-810-1818 或 (010) 84981188 分机 5017，可获得市场营销与销售的信息；

要查找您所在国家或地区的 IBM 营业处，可在网上查看 IBM 全球联系人目录 (Directory of Worldwide Contacts)，网址为：[www.ibm.com/planetwide](http://www.ibm.com/planetwide)。



---

## 商标

本主题列示 IBM 商标和某些非 IBM 商标。

请参阅 <http://www.ibm.com/legal/copytrade.shtml> 以了解有关 IBM 商标的信息。

下列各项是其他公司的商标或注册商标：

Java 和所有基于 Java 的商标和徽标是 Sun Microsystems, Inc. 在美国和 / 或其他国家或地区的商标或注册商标。

Microsoft、Windows、Windows NT 和 Windows 徽标是 Microsoft Corporation 在美国和 / 或其他国家或地区的商标。

Intel、Intel Inside（徽标）、MMX 和 Pentium 是 Intel Corporation 在美国和 / 或其他国家或地区的商标。

UNIX 是 The Open Group 在美国和其他国家或地区的注册商标。

Linux 是 Linus Torvalds 在美国和 / 或其他国家或地区的商标。

其他公司、产品或服务名称可能是其他公司的商标或服务标记。



---

## 声明

本信息是为在美国提供的产品和服务编写的。IBM 可能未在所有国家或地区提供本文中讨论的产品、服务或功能特性。有关您当前所在区域的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。但是，评估和验证任何非 IBM 产品、程序或服务，则由用户自行负责。

IBM 公司可能已拥有或正在申请与本文档内容有关的各项专利。提供本文档并未授予用户使用这些专利的任何许可。您可以用书面方式将许可查询寄往：IBM Director of Licensing IBM Corporation North Castle Drive Armonk, NY 10504-1785 U.S.A.

有关双字节（DBCS）信息的许可查询，请与您所在国家或地区的 IBM 知识产权部门联系，或用书面方式将查询寄往：IBM World Trade Asia Corporation Licensing 2-31 Roppongi 3-chome, Minato-ku Tokyo 106-0032, Japan

本条款不适用英国或任何这样的条款与当地法律不一致的国家或地区：International Business Machines Corporation“按现状”提供本出版物，不附有任何种类的（无论是明示的还是暗含的）保证，包括但不限于暗含的有关非侵权、适销和适用于某种特定用途的保证。某些国家或地区在某些交易中不允许免除明示或暗含的保证。因此本条款可能不适用于您。

本信息中可能包含技术方面不够准确的地方或印刷错误。此处的信息将定期更改；这些更改将编入本资料的新版本中；IBM 可以随时对本资料中描述的产品和 / 或程序进行改进和 / 或更改，而不另行通知。

本信息中对非 IBM Web 站点的任何引用都只是为了方便起见才提供的，不以任何方式充当对那些 Web 站点的保证。那些 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无须对您承担任何责任。

本程序的被许可方如果要了解有关程序的信息以达到如下目的：（i）允许在独立创建的程序和其他程序（包括本程序）之间进行信息交换，以及（ii）允许对已经交换的信息进行相互使用，请与下列地址联系：

IBM Corporation  
J46A/G4  
555 Bailey Avenue  
San Jose, CA 95141-1003  
U.S.A.

只要遵守适当的条件和条款，包括某些情形下的一定数量的付费，都可获得这方面的信息。

本资料中描述的许可程序及其所有可用的许可资料均由 IBM 依据 IBM 客户协议、IBM 国际软件许可协议或任何同等协议中的条款提供。

此处包含的任何性能数据都是在受控环境中测得的。因此，在其他操作环境中获得的数据可能会有明显的不同。有些测量可能是在开发级的系统上进行的，因此不保证与一般可用系统上进行的测量结果相同。此外，有些测量是通过推算而估计的，实际结果可能会有差异。本文档的用户应当验证其特定环境的适用数据。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

所有关于 IBM 未来方向或意向的声明都可随时更改或收回，而不另行通知，它们仅仅表示了目标和意愿而已。

本信息包含日常商业运作所使用的数据和报表的示例。为了尽可能全面地说明这些数据和报表，这些示例包括个人、公司、品牌和产品的名称。所有这些名称都是虚构的，如与实际商业企业所使用的名称和地址有任何雷同，纯属巧合。

版权许可：

本信息包括源语言形式的样本应用程序，这些样本说明不同操作平台上的编程方法。如果是为按照在编写样本程序的操作平台上的应用程序编程接口（API）进行应用程序的开发、使用、经销或分发为目的，您可以任何形式对这些样本程序进行复制、修改、分发，而无须向 IBM 付费。这些示例并未在所有条件下作全面测试。因此，IBM 不能担保或暗示这些程序的可靠性、可维护性或功能。用户如果是为了按照 IBM 应用程序编程接口开发、使用、经销或分发应用程序，则可以任何形式复制、修改和分发这些样本程序，而无须向 IBM 付费。

凡这些样本程序的每份拷贝或其任何部分或任何衍生产品，都必须包括如下版权声明：

Outside In (®) Viewer Technology, ©1992-2005 Stellent, Chicago, IL., Inc. All Rights Reserved.

IBM XSLT Processor Licensed Materials - Property of IBM ©Copyright IBM Corp., 1999-2005. All Rights Reserved.





中国印刷