

Oracle® Universal Content Management
OracleTextSearch Component Installation and Administration
Guide
Release 10gR3

May 2010

Copyright © 2008, 2010, Oracle. All rights reserved.

Primary Author: Karen Johnson

Contributor: Hui Ye

The Programs (which include both the software and documentation) contain proprietary information; they are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright, patent, and other intellectual and industrial property laws. Reverse engineering, disassembly, or decompilation of the Programs, except to the extent required to obtain interoperability with other independently created software or as specified by law, is prohibited.

The information contained in this document is subject to change without notice. If you find any problems in the documentation, please report them to us in writing. This document is not warranted to be error-free. Except as may be expressly permitted in your license agreement for these Programs, no part of these Programs may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose.

If the Programs are delivered to the United States Government or anyone licensing or using the Programs on behalf of the United States Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the Programs, including documentation and technical data, shall be subject to the licensing restrictions set forth in the applicable Oracle license agreement, and, to the extent applicable, the additional rights set forth in FAR 52.227-19, Commercial Computer Software--Restricted Rights (June 1987). Oracle USA, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

The Programs are not intended for use in any nuclear, aviation, mass transit, medical, or other inherently dangerous applications. It shall be the licensee's responsibility to take all appropriate fail-safe, backup, redundancy and other measures to ensure the safe use of such applications if the Programs are used for such purposes, and we disclaim liability for any damages caused by such use of the Programs.

Oracle, JD Edwards, PeopleSoft, and Siebel are registered trademarks of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

The Programs may provide links to Web sites and access to content, products, and services from third parties. Oracle is not responsible for the availability of, or any content provided on, third-party Web sites. You bear all risks associated with the use of such content. If you choose to purchase any products or services from a third party, the relationship is directly between you and the third party. Oracle is not responsible for: (a) the quality of third-party products or services; or (b) fulfilling any of the terms of the agreement with the third party, including delivery of products or services and warranty obligations related to purchased products or services. Oracle is not responsible for any loss or damage of any sort that you may incur from dealing with any third party.

Contents

Preface	v
1 Introduction to the OracleTextSearch Component	
1.1 Purpose	1-1
1.2 Considerations	1-1
1.2.1 Hardware and Software Considerations	1-1
1.2.2 Search Function Considerations	1-2
1.2.3 Use of Stopwords	1-2
1.3 Benefits and Features of Using Oracle Text 11g	1-3
1.3.1 Indexing and Query Speeds and Techniques	1-3
1.3.2 Fast Rebuild	1-4
1.3.3 Query Syntax	1-4
1.3.4 Search Operators	1-4
1.3.4.1 Search Thesaurus	1-6
1.3.5 Case Sensitivity and Stemming Rules	1-7
1.3.6 Search Results Data Clustering	1-7
1.3.7 Snippets	1-7
1.3.8 Additional Changes	1-7
2 Installation and Configuration	
2.1 Requirements and Considerations	2-1
2.2 Pre-Installation Configuration	2-1
2.2.1 Running Oracle Database Configuration Scripts	2-2
2.2.2 Assigning an Oracle Database Role	2-2
2.3 Component Installation	2-2
2.3.1 Component Wizard Installation	2-3
2.3.2 Component Manager Installation	2-4
2.4 Post-Installation Configuration for Oracle Text 11g	2-4
2.5 Post-Installation Configuration for Oracle Secure Enterprise Search 11g	2-6
2.5.1 Configuring Oracle SES 11g and Oracle Database 11g	2-6
3 Managing the OracleTextSearch Component	
3.1 Determining Fields to Optimize	3-1
3.2 Assigning/Editing Optimized Fields	3-1
3.3 Performing a Fast Rebuild	3-2

3.4	Optimizing the Search Collection.....	3-2
3.5	Modifying the OracleTextSearch Fields Displayed on Search Results	3-3
3.6	Finding Status Information	3-3
3.7	Text Search Admin Page.....	3-3

4 Searching with the OracleTextSearch Component

4.1	Performing a Search.....	4-1
4.2	Search Results With OracleTextSearch	4-1

A Third Party Licenses

A.1	Apache Software License	A-1
A.2	W3C Software Notice and License	A-1
A.3	Zlib License	A-2
A.4	General BSD License.....	A-3
A.5	General MIT License.....	A-3
A.6	Unicode License	A-3
A.7	Miscellaneous Attributions.....	A-4

Index

Preface

This guide provides information for installing, configuring, and administering the OracleTextSearch component on Oracle Content Server. This component works as the primary search engine with Oracle Database 11g and is included because Autonomy VDK is not sold with Oracle Universal Content Management as of June 2008.

The information contained in this document is subject to change as the product technology evolves and as hardware, operating systems, and third-party software are created and modified.

Audience

This guide is intended for developers and administrators who want to implement OracleTextSearch as the primary search engine for Oracle Content Server.

Documentation Accessibility

Our goal is to make Oracle products, services, and supporting documentation accessible, with good usability, to the disabled community. To that end, our documentation includes features that make information available to users of assistive technology. This documentation is available in HTML format, and contains markup to facilitate access by the disabled community. Accessibility standards will continue to evolve over time, and Oracle is actively engaged with other market-leading technology vendors to address technical obstacles so that our documentation can be accessible to all of our customers. For more information, visit the Oracle Accessibility Program Web site at <http://www.oracle.com/accessibility/>.

Accessibility of Code Examples in Documentation

Screen readers may not always correctly read the code examples in this document. The conventions for writing code require that closing braces should appear on an otherwise empty line; however, some screen readers may not always read a line of text that consists solely of a bracket or brace.

Accessibility of Links to External Web Sites in Documentation

This documentation may contain links to Web sites of other companies or organizations that Oracle does not own or control. Oracle neither evaluates nor makes any representations regarding the accessibility of these Web sites.

Access to Oracle Support

Oracle customers have access to electronic support through My Oracle Support. For information, visit <http://www.oracle.com/support/contact.html> or visit

<http://www.oracle.com/accessibility/support.html> if you are hearing impaired.

Related Documents

For more information, see these Oracle resources available from Oracle Technology Network (OTN) at <http://www.oracle.com/technology/index.html>.

For information about Oracle Text 11g:

- *Oracle Text Application Developer's Guide*
- *Oracle Text Reference*

For information about Oracle Database 11g:

- *Oracle Database Concepts*
- *Oracle Database Administrator's Guide*
- *Oracle Database Utilities*
- *Oracle Database SQL Reference*
- *Oracle Database Reference*
- *Oracle Database Application Developer's Guide - Fundamentals*

For information about PL/SQL:

- *Oracle Database PL/SQL User's Guide and Reference*

For information about Oracle Universal Content Management Content Server 10gR3:

- *Oracle Content Server Installation Guide* (for Microsoft Windows or UNIX)
- *Oracle Content Server Managing Repository Content*
- *Oracle Content Server Managing Security and User Access*
- *Oracle Content Server Managing System Settings and Processes*
- *Oracle Content Server Release Notes*
- *Oracle Content Server Working with Content Components*

Conventions

The following text conventions are used in this document:

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
<i>italic</i>	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
<code>monospace</code>	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.
<i>Install_Dir</i>	This notation is used to refer to the location on your system where the content server instance is installed.

Introduction to the OracleTextSearch Component

This chapter covers the following topics:

- ["Purpose"](#) on page 1-1
- ["Considerations"](#) on page 1-1
- ["Benefits and Features of Using Oracle Text 11g"](#) on page 1-3

1.1 Purpose

The OracleTextSearch component enables the use of Oracle Text 11g as the primary full-text search engine for Oracle Content Server. Oracle Text 11g offers state-of-the-art indexing capabilities that match or exceed the search capabilities offered by Oracle Content Server with Autonomy VDK.

The component enables administrators to specify certain metadata fields to be optimized for the search index and to customize additional fields. It also enables a fast index rebuild and index optimization.

When the OracleTextSearch component is installed it adds a new menu bar and options to the Search Results page, which enables users to manage the view of content items found by the search.

1.2 Considerations

This section provides information concerning issues that are important to consider when using the OracleTextSearch component. These include:

- ["Hardware and Software Considerations"](#) on page 1-1
- ["Search Function Considerations"](#) on page 1-2
- ["Use of Stopwords"](#) on page 1-2

1.2.1 Hardware and Software Considerations

The following are items that concern hardware and software issues:

- This component runs on Oracle Content Server 10gR3 under the supported operating systems.
- Oracle Content Server 10gR3 supports all languages defined in the Certification Matrix.

- Oracle Text 11g runs on Oracle Database 11g. The Oracle Content Server system database can be Oracle Database 11g, Microsoft SQL Server, or other databases as listed in the Oracle Content Server 10gR3 Certification Matrix. However, if the system database is not Oracle Database 11g, then a provider for the OracleTextSearch component must be configured.

Note: To function properly, the OracleTextSearch component requires release 11.1.0.7.0 or higher of the Oracle Database 11g. OracleTextSearch component supports Oracle Database release 11.2. Additionally, the JDBC version must be 10.2.0.4.0 or higher.

- The OracleTextSearch component enables Oracle Text 11g as a separate search collection instance on Oracle Database 11g for Oracle Content Server, which allows the search collection to reside on a separate machine and not compete with Oracle Content Server for processors and memory. This can improve indexing and search response time.
- The OracleTextSearch collection instance can be installed on a different platform than the Oracle Content Server installation.
- Because Oracle Content Server uses the SDATA section feature in Oracle Text 11g, up to 32 SDATA fields can be defined with the OracleTextSearch component interface. For more detailed information about the SDATA section feature, see ["Indexing and Query Speeds and Techniques"](#) on page 1-3.
- Besides the OracleTextSearch component, no other components or customizations are required for Oracle Content Server to work with Oracle Text 11g.
- Oracle Content Server 10gR3 with the OracleTextSearch component can be configured to support Oracle Secure Enterprise Search 11g (Oracle SES release 11.2) as its back-end search engine. Oracle SES 11g enables a secure, high quality, easy-to-use search across all enterprise information assets. Additional configuration is required to support Oracle SES 11g with Oracle Database 11g, Oracle Content Server, and the OracleTextSearch component. See ["Post-Installation Configuration for Oracle Secure Enterprise Search 11g"](#) on page 2-6. See also *Oracle Secure Enterprise Search Administrator's Guide*.

1.2.2 Search Function Considerations

The following are items that concern search functions:

- While Oracle Content Server 10gR3 provides numerous search options using a variety of databases (Oracle, Microsoft SQL Server, IBM DB2). By default, the database that serves as the search index is the same system database used by Oracle Content Server to manage metadata and other configuration information (users, security groups, etc.).
- The OracleTextSearch component can be used with large deployments, but it only affects queries run against the search collection instance. OracleTextSearch does not affect system queries run against an Oracle database.

1.2.3 Use of Stopwords

Stopwords are words that a search engine filters out or ignores in a search query when they are combined with other keywords. These words are typically not meaningful and omitting them increases the speed and efficiency of searches. However, with the

OracleTextSearch component, this function can adversely affect returned search results if the query contains stopwords.

A typical example of adverse search results when using the OracleTextSearch component would occur with a list of accounts as follows:

```
any
any/sd/any/admin
int
int/sd
int/sd/admin
dept
dept/sd
dept/sd/admin
```

If you perform a Contains search on the word *any*, the query will return all of the accounts whether or not they contain the word *any*. because 'any' is an Oracle Text stopword. A list of default stopwords includes: a, he, out, up, be, more, their, at, had, one, will, from, it, than, and, is, only, when, corp, not, she, also, in, says, was, by, ms, to, about, her, over, because, most, there, has, or, with, its, that, are, of, which, could, some, an, inc, we, can, mz, after, his, s, been, mr, they, have, other, would, last, the, as, on, who, for, such, any, into, were, co, no, all, if, so, but, mrs, this.

For more detailed information about using stopwords or for instructions on how to remove them, refer to the *Oracle Text Reference 11g* document.

1.3 Benefits and Features of Using Oracle Text 11g

This section covers the following topics:

- ["Indexing and Query Speeds and Techniques"](#) on page 1-3
- ["Fast Rebuild"](#) on page 1-4
- ["Query Syntax"](#) on page 1-4
- ["Search Operators"](#) on page 1-4
- ["Case Sensitivity and Stemming Rules"](#) on page 1-7
- ["Search Results Data Clustering"](#) on page 1-7
- ["Snippets"](#) on page 1-7
- ["Additional Changes"](#) on page 1-7

1.3.1 Indexing and Query Speeds and Techniques

Using Oracle Text 11g, Oracle Content Server offers a significant increase in index speeds. Oracle Text indexing is transactional. Oracle Content Server sends a batch of document to Oracle Text, commits the batch, then starts the Oracle Text indexer. Oracle Content Server is notified of which documents failed to index and only those documents are resubmitted to be indexed. Oracle Content Server also supports the use of parallel indexing with the database, which can leverage multiple CPUs on the database server. This parallel indexing option can be enabled via the following Oracle Content Server configuration variable in the config.cfg file:

```
OracleTextIndexingParallelDegree=1
```

Oracle Content Server uses some of the newest Oracle Text 11g features. For example, Oracle Content Server automatically creates a new search index zone for each text information field in order to provide better search speed. Using information zones

enables Oracle Content Server to query data as if it were full-text data. All text-based information fields (text, long text, and memo) are automatically added to as separate zones. In addition to the zones created for text information fields, Oracle Content Server provides an extra zone named `IdcContent`, which enables custom components, Inbound Refinery components, applications, or users to create XML content with tags that will be indexed as full-text metadata fields.

Oracle Content Server also uses the `SDATA` section feature in Oracle Text 11g to index important text, date, and integer fields and define them as Optimized Fields. The `SDATA` section is a separate XML structure managed by the Oracle Text engine which allows the engine to respond rapidly to requests involving data and integer ranges. Oracle Content Server uses an `SDATA` field in the database for each Optimized Field. Oracle Text 11g allows up to 32 `SDATA` fields, so Oracle Content Server can use up to 32 Optimized Fields. The Content ID and Document Title fields are automatically defined as Optimized Fields for `SDATA`. You can define Optimized Fields using the OracleTextSearch component Text Search Admin page.

Note: If you want to change the set of Optimized Fields defined for Oracle Text 11g, remember that the maximum allowed number of Optimized Fields is 32. This maximum number includes the Content ID and Document Title fields.

1.3.2 Fast Rebuild

OracleTextSearch provides a Fast Rebuild option on the Text Search Admin page. This option allows the search engine to add new information to the search collection without requiring a full collection rebuild. A Fast Rebuild is required in the following cases:

- Adding or removing information fields
- Changing any Optimized Field
- Changing an information field to be an Optimized Field

A Fast Rebuild does not cause all the information (metadata and full-text) to be re-indexed. It adds the changes throughout the collection and updates it. Oracle Content Server search functionality is not affected during a Fast Rebuild cycle.

1.3.3 Query Syntax

Queries defined in Universal Query Syntax, introduced in Content Server release 7.5, are supported and generally do not need any modification. This includes queries saved by users, queries defined in custom components, and queries defined in Site Studio pages.

1.3.4 Search Operators

Oracle Text supports the same search operator capabilities as with Autonomy VDK, including the following defaults:

- CONTAINS
- MATCHES
- Has Prefix
- Not Contains

- Range searches for dates and integers

The Oracle Text 11g engine supports additional search operators and functions which are not exposed in the user interface by default, but can be exposed via customization that adds to the operator definition HDA table. These include the operators shown in [Table 1–1](#). For details and examples of these operators see *Oracle Text Reference*.

Table 1–1 Oracle Text search operators and functions

Query Operator	Description
ABOUT	Performs a theme search where available, and increases the number of relevant documents returned from the query.
ACCUMulate (,)	Searches for documents that contain at least one occurrence of any of the query terms. Increases relevance as more terms are found.
Broader Term (BT, BTG, BTP, BTI)	Expands a query to include the term that has been defined in a thesaurus as a broader or higher level term.
DEFINEMERGE	Defines how the score of child nodes of the AND and OR should be merged.
DEFINESCORE	Defines how a term or phrase, or a set of term equivalences will be scored.
EQUIValence (=)	Specifies alternate substitution terms in a query.
FUZZY	Expands queries to include words which are spelled similarly, or sound similar to the specified term.
HASPATH	Finds all XML documents which contain a specified section path.
INPATH	Searches within a particular path in an XML document.
MDATA	Queries MDATA (MetaDATA) sections.
MINUS (-)	Lowers the relevance of documents that contain a particular term, but does not necessarily exclude them.
Narrower Term (NT, NTG, NTP, NTI)	Expands a query to include all the terms which have been defined in a thesaurus as the narrower or lower level terms for a specified term.
NEAR (;)	Returns a score based on the proximity of two or more query terms.
Preferred Term (PT)	Replaces a term in a query with the preferred term that has been defined in a thesaurus for the term.
Related Term (RT)	Replaces a term in a query with the related term that has been defined in a thesaurus for the term.
SDATA	Performs tests on SDATA sections and columns, which contain structured data values.
soundex (!)	Expands queries to include words that have similar sounds.
stem (\$)	Searches for terms that have the same linguistic root as the query term.
Stored Query Expression (SQE)	Calls a stored query expression created with the CTX_QUERY.STORE_SQE procedure.
SYNonym (SYN)	Expands a query to include all the terms that have been defined in a thesaurus as synonyms for the specified term.

Table 1–1 (Cont.) Oracle Text search operators and functions

Query Operator	Description
threshold (>)	This operator at the expression level eliminates documents in the result set that score below a threshold number. This operator at the query term level selects a document based on how a term scores in the document.
Translation Term (TR)	Expands a query to include all foreign language equivalent terms defined in a thesaurus.
Translation Term Synonym (TRSYN)	Expands a query to include all the defined foreign equivalents of the query term, the synonyms of the query term, and the foreign equivalents of the synonyms.
Top Term (TT)	Replaces a term in a query with the top term that has been defined for the term in the standard hierarchy in a thesaurus.
weight (*)	Multiplies the score by the given factor, topping out at 100 when the score exceeds 100.
wildcards (% _)	Expands word searches into pattern searches.
WITHIN	Narrows a query into a document sections.

1.3.4.1 Search Thesaurus

Certain queries, such as stem and Related Term, may be more effective if you use an Oracle Text thesaurus. Oracle Text enables you to create case-sensitive or case-insensitive thesauri which define synonym and hierarchical relationships between words and phrases. You can then search and retrieve documents that contains relevant text by expanding queries to include similar or related terms as defined in the thesaurus. For example, you can populate a thesaurus with specific product names, associated models, associated features, and so forth.

- Default thesaurus:** If you do not specify a thesaurus by name in a query, by default, the thesaurus operators use a thesaurus named DEFAULT. However, Oracle Text does not provide a DEFAULT thesaurus.

As a result, if you want to use a default thesaurus for the thesaurus operators, you must create a thesaurus named DEFAULT. You can create the thesaurus through any of the thesaurus creation methods supported by Oracle Text:

- CTX_THES.CREATE_THESAURUS (PL/SQL)
- ctxload utility

- Supplied thesaurus:** Oracle Text does not provide a default thesaurus, but Oracle Text does supply a thesaurus, in the form of a file that you load with ctxload, that can be used to create a general-purpose, English-language thesaurus.

The thesaurus load file can be used to create a default thesaurus for Oracle Text, or it can be used as the basis for creating thesauri tailored to a specific subject or range of subjects.

Note: See the *Oracle Text Reference* to learn more about using ctxload and the CTX_THES package, and see chapter 9, "Working With a Thesaurus in Oracle Text," in the *Oracle Text Application Developer's Guide*.

1.3.5 Case Sensitivity and Stemming Rules

Oracle Content Server automatically ensures that queries are executed as case-insensitive. By default, all full-text and text field search queries are case-insensitive. Oracle Content Server also handles case-insensitive search queries for information stored as Optimized Fields.

Oracle Content Server does not apply any stemming rules by default for Oracle Text 11g, but stemming rules can be applied by using the `stem()` function. Stemming rules may be used to have searches account for plurals, verbs, and so forth. Other methods for implementing stemming rules include modifying the standard query definition in the `searchindexerrules` configuration file, and by making configuration changes in the Oracle Text engine (Oracle Database).

Oracle Content Server handles content in non-English languages by using the `WORLD_LEXER` feature in the Oracle Text engine. This enables Oracle Text to automatically identify the language and apply the proper tokenization rules.

1.3.6 Search Results Data Clustering

With the `OracleTextSearch` component, Oracle Content Server retrieves additional information about a search result list and displays it in a new menu bar on the Search Results page. This information summarizes how many documents are attached to specific values in specific information fields. Oracle Content Server supports data clustering for up to four information fields (the default fields are Security Group and Document Type).

This can be useful if you have a query that returns many items. For example, a result set could include 200 content items, including 100 documents that belong to the Public security group, 75 that belong to the Sales group, and 25 that belong to the Marketing group. The menu option for Security Group will show you the list of values and how many documents belong to each value. You can select one of the values (Public, Sales, Marketing) from the menu and it will list only those documents in the result set that belong to that value.

1.3.7 Snippets

Oracle Content Server can retrieve document snippets as part of search results to show the occurrence of search terms in context of their usage. This feature is enabled by default. To disable this feature as it can improve search query performance, set the following configuration entry in the `config.cfg` file:

```
OracleTextDisableSearchSnippet=true
```

1.3.8 Additional Changes

Additional changes because of the use of Oracle Text 11g include:

- XML content is automatically indexed.
- There are no visible changes in the Search user interface other than removal of Substring as a search operator option. The default search operators are CONTAINS, MATCHES, Has Prefix, and Not Contains. Substring-based queries will still work.
- Queries using the MATCHES operator on a non-optimized field will behave like a CONTAINS query. For example, if `xDepartment` is not optimized, then the query `xDepartment MATCHES 'Marketing'` will behave like `xDepartment`

CONTAINS 'Marketing' and return hits on documents that have an xDepartment value of 'Marketing Services' or 'Product Marketing'.

- Relevancy ranking can be changed in Oracle Text 11g through use of an operator called DEFINESCORE. This operator can be added via a component to the WhereClause value of OracleTextSearch in the SearchQueryDefinition table (in the searchindexerrules configuration file). More information about this operator is available in the *Oracle Text Reference* document.
- The PDF Highlighting feature has been disabled.
- The Spell Checking feature can be enabled, but it requires a custom component just as it did with Autonomy VDK.

Installation and Configuration

This section covers the following topics:

- ["Requirements and Considerations"](#) on page 2-1
- ["Pre-Installation Configuration"](#) on page 2-1
- ["Component Installation"](#) on page 2-2
- ["Post-Installation Configuration for Oracle Text 11g"](#) on page 2-4
- ["Post-Installation Configuration for Oracle Secure Enterprise Search 11g"](#) on page 2-6

2.1 Requirements and Considerations

The following items are important when considering use of the OracleTextSearch component:

- The OracleTextSearch component enables Oracle Text 11g as a separate search collection instance. This instance can be installed on a different platform than the Oracle Content Server installation.
- Oracle Text 11g works with Oracle Database 11g, therefore if you choose to use Oracle Text 11g as the search index engine, then the OracleTextSearch component must connect to an Oracle 11g database.

The Oracle Content Server system database can be Oracle Database 11g, Microsoft SQL Server, or other databases as listed in the Oracle Content Server Release 10gR3 certification matrix. However, if the system database is not Oracle Database 11g, then Oracle Database 11g and a provider for the database must be installed. See ["Post-Installation Configuration for Oracle Text 11g"](#) on page 2-4.

Note: To function properly, the OracleTextSearch component requires version 11.1.0.7.0 or higher of Oracle Database 11g. Additionally, the JDBC version must be 10.2.0.4.0 or higher.

- If the OracleTextSearch component is enabled, but it is not configured and is not being used, then Oracle Content Server continues to use the previous search engine until the OracleTextSearch component is fully configured.

2.2 Pre-Installation Configuration

Before an administrator can manage the OracleTextSearch component, Oracle Database 11g must be configured to work with the component.

- If you perform a new Oracle Content Server installation and plan to install the OracleTextSearch component during the installation process, an Oracle Database 11g administrator must run two configuration scripts *before* the component is installed. See "[Running Oracle Database Configuration Scripts](#)" on page 2-2.
- If you update Oracle Content Server from an earlier release of 10gR3, an Oracle Database 11g administrator must run two configuration scripts *before* installing the component. See "[Running Oracle Database Configuration Scripts](#)" on page 2-2.
- After Oracle Database 11g has been configured for the OracleTextSearch component, an Oracle Database role must be assigned to the Oracle Content Server user who will administer the component. This role assignment is required whether you are updating an Oracle Content Server Release 10gR3 installation or installing a new Oracle Content Server Release 10gR3. See "[Assigning an Oracle Database Role](#)" on page 2-2.

2.2.1 Running Oracle Database Configuration Scripts

Whether you are updating an earlier version of Oracle Content Server Release 10gR3 or installing a new Oracle Content Server Release 10gR3, the following steps *must* be completed before you install the OracleTextSearch component:

1. Log on to Oracle Database 11g as an administrator.
2. Run the script `contentserverrole`. This script creates a new role.
3. Log on to Oracle Database 11g as an Oracle Content Server privileged user.
4. Run the PL/SQL script `contentprocedures`.

Note: If you run the `contentprocedures` script multiple times, you might see the following message:

```
ERROR at line 1:  
ORA-02303: cannot drop or replace a type with type or table  
dependents
```

The error message can be ignored.

Note: After the OracleTextSearch component is installed, the two scripts are available in the directory `<Install_Dir>/custom/OracleTextSearch/scripts`.

2.2.2 Assigning an Oracle Database Role

After Oracle Database 11g has been configured for the OracleTextSearch component, the Oracle Database 11g administrator must assign the role `contentserver_role` to the database user that Oracle Content Server uses to manage the component and its connection to Oracle Database 11g. This task also can be done by running an SQL command.

2.3 Component Installation

This section contains instructions for installing the OracleTextSearch component for Oracle Content Server Release 10gR3.

Note: Whether you install a new Oracle Content Server or upgrade an existing Release 10gR3 of Oracle Content Server, an Oracle Database administrator role must be assigned to the Oracle Content Server user who will manage the OracleTextSearch component (see step 3 of "[Assigning an Oracle Database Role](#)" on page 2-2).

If the Oracle Content Server system database is *not* Oracle Database 11g, or if you choose to use a separate Oracle Database instance for the Oracle Text search engine, you must complete the "[Post-Installation Configuration for Oracle Text 11g](#)" on page 2-4 to set up a provider for Oracle Database 11g.

When you upgrade Oracle Content Server Release 10gR3, you must install and enable the OracleTextSearch component on the content server using either the [Component Wizard Installation](#) or the [Component Manager Installation](#).

Before you can install the component, if you are running Oracle Content Server Release 10gR3, you must first download the "Oracle Universal Content Management Patchset for Content Server 10.1.3.x.x Update Bundle Component" (CS10gR3UpdateBundle) from the My Oracle Support site (<http://metalink.oracle.com>). The 10.1.3.x.x Update Bundle Component contains the OracleTextSearch component file (typically named OracleTextSearch.zip). Extract the file and place it in a temporary location, then proceed with the installation instructions.

When the installation is complete, you must perform the [Post-Installation Configuration for Oracle Text 11g](#).

2.3.1 Component Wizard Installation

Use this procedure to install the component using the Component Wizard:

1. Start the Component Wizard by selecting **Programs** from the **Start** menu. Then select **Oracle Content Server**, then *<instance>*, then **Utilities**, and then select **Component Wizard**.

The Component Wizard main screen and the Component List screen are displayed.

2. On the Component List screen, click **Install**.

The Install screen is displayed.

3. Click **Select. Navigate** to the location where you downloaded the component zip file and select it.

4. Click **Open**.

The zip file contents are added to the Install screen list.

5. Click **OK**.

The Component Wizard prompts whether you want to enable the component.

6. Click **Yes**.

The component is listed as enabled on the Component List screen.

7. Restart the Oracle Content Server.

2.3.2 Component Manager Installation

Use this procedure to install the component using the Component Manager:

1. Open a new browser window and log in to the Oracle Content Server as the system administrator.
2. Select the Administration tray (or link).
3. Click **Admin Applets** to open the Administration page.
4. Click **Admin Server**.
5. Click the Oracle Content Server instance.
6. In the sidebar click **Component Manager**.

The Component Manager screen is displayed.

7. Click **Browse** next to the Install New Component field.
8. Navigate to the component zip file and select it.
9. Click **Install**.

A page is displayed listing the component items that will be installed.

10. Click **Continue** to continue with the installation.

A Content Server message is displayed that indicates the installation is successful.

11. On the message page click to return to the Component Manager.
12. Click the component name in the Disabled Components box.
13. Click **Enable** to enable the component.

The component is listed in the Enabled Components box.

14. In the side menu click **Start/Stop Content Server**.
15. Restart the Oracle Content Server.

2.4 Post-Installation Configuration for Oracle Text 11g

If you choose to use Oracle Text 11g as the search index engine with the OracleTextSearch component, then after the OracleTextSearch component is installed, complete the following steps to configure the Oracle Content Server instance to use Oracle Text 11g. Some of the post-installation configuration is necessary even if you are using Oracle Database 11g as your system database and want to use the OracleTextSearch component with a separate database instead of the system database.

Caution: If you are updating Oracle Content Server from an earlier version of release10gR3 instead of installing a new Oracle Content Server Release 10gR3, before performing post-configuration tasks the Oracle Database administrator must run a script to configure settings in Oracle Database 11g and assign a new role to the Oracle Content Server administrator. See "[Pre-Installation Configuration](#)" on page 2-1.

If the Oracle Content Server system database is *not* Oracle Database 11g, complete the following steps:

1. Point the JAVA_CLASSPATH_defaulttjdbc setting to an Oracle database driver (such as ojdbc14.jar) in the <Install_Dir>/bin/intradoc.cfg file. For example, if

you installed Oracle Content Server Release 10gR3 with an SQL database, the value would be:

```
JAVA_CLASSPATH_defaultjdbc=$SHAREDDIR/classes/jtlds.jar
```

To use OracleTextSearch, append this value to also point to the Oracle jar:

```
JAVA_CLASSPATH_
defaultjdbc=$SHAREDDIR/classes/jtlds.jar;$SHAREDDIR/classes/ojdbc14.jar
```

2. Define a database provider:

- a.** Select **Administration**, then click **Providers**.
- b.** On the Providers page, click **Add** to define a database provider.
- c.** On the Database Provider Information page, use the following settings.

For more information on creating a database provider, see *Managing System Settings and Processes*.

Element	Setting
Provider Name	Oracle11g_provider
Provider Description	Oracle 11g Database Provider
Provider Class	intradoc.jdbc.JdbcWorkspace
Connection Class	intradoc.jdbc.JdbcConnection
Configuration Class	oracletextsearch.server.OracleTextProviderConfig
Test Query	select 1 from dual
Database Type	(check the box for) JDBC
Database Type	<i>database_type</i>
Database Name	<i>database_name</i>
JDBC Driver	oracle.jdbc.OracleDriver
JDBC Connection String	jdbc:oracle:thin:@sta00894.us.oracle.com:1511:ade
JDBC User	<i>instance_name</i>
JDBC Password	<i>password</i>
Number of Connections	5

d. Click Update.

For any system database to be used with the OracleTextSearch component, complete the following:

- 1.** Set the following configuration variables in the config.cfg file:

```
SearchIndexerEngineName=OracleTextSearch
IndexerDatabaseProviderName=provider_name
```

Note: If the system database is Oracle Database 11g and it also will be used for OracleTextSearch, then the second variable should be: `IndexerDatabaseProviderName=SystemDatabase`.

The `SearchIndexerEngineName` variable tells Oracle Content Server what search engine to use; in this case the engine is the `OracleTextSearch` search engine.

The `IndexerDatabaseProviderName` value tells Oracle Content Server what database provider to use to store the search index; in this case the provider is `provider_name`.

2. (Optional.) Oracle Text treats any non-alphanumeric characters as word breaks, which may cause difficulties when using certain search operators. You can choose to use different search operators against different fields, or you can change non-alphanumeric character specifications for `OracleTextSearch` using the following variable in the `config.cfg` file:

```
(ORACLETEXTSEARCH)AdditionalEscapeChars=character_to_be_replaced:character_to_be_used
```

For example, when indexing, Oracle Text would treat `"oly_12345"` the same as `"oly 12345"`. A `CONTAINS` search for Content ID `"oly_12345"` would match `"oly[A]12345"` and `"oly[B]12345"` and so forth, because the underscore is treated as a wildcard in a search. The solution is to specify the character, as in the following example:

```
(ORACLETEXTSEARCH)AdditionalEscapeChars=_:#
```

3. Restart the Oracle Content Server and rebuild the index.

Note: In this case, the index rebuild must be a full collection rebuild rather than the [Fast Rebuild](#). The Fast Rebuild option would not work unless at least one full rebuild is successfully completed.

Caution: Depending on the size of your search index and available system resources, the search index rebuild process can take up to a couple of days. Therefore, the rebuild should be done at times of non-peak system usage.

2.5 Post-Installation Configuration for Oracle Secure Enterprise Search 11g

If you choose to support Oracle Secure Enterprise Search (SES) 11g with Oracle Database 11g and the `OracleTextSearch` component, after the component is installed and configured with Oracle Content Server Release 10gR3, and after Oracle SES 11g is installed, complete the following procedures to configure Oracle SES 11g and Oracle Content Server Release 10gR3 to support Oracle SES 11g as the back-end search index engine.

- ["Configuring Oracle SES 11g and Oracle Database 11g"](#) on page 2-6

2.5.1 Configuring Oracle SES 11g and Oracle Database 11g

1. After installing Oracle SES 11g, edit the file `$ORACLE_HOME/network/admin/sqlnet.ora` to comment out the following two lines:

```
tcp.invited_nodes  
tcp.validate_checking
```

2. If Oracle SES is running, shut it down (mid-tier and database):

```
$ORACLE_HOME/bin/searchctl stopall
```

3. Start the Oracle Database:

```
$ORACLE_HOME/bin/searchctl start_backend
```

4. Find the database connection information in the following file:

```
$ORACLE_HOME/search/webapp/config/search.properties
```

5. Use the database connection information to connect your favorite client (SQLPlus, SQL Developer, Toad, and so forth) to the database.

6. Follow the instructions in ["Running Oracle Database Configuration Scripts"](#) on page 2-2 and ["Assigning an Oracle Database Role"](#) on page 2-2 to create a database user assigned the database role `contentserver_role`, which is used to connect to the Oracle Content Server provider.

Managing the OracleTextSearch Component

This chapter covers the following topics:

- ["Determining Fields to Optimize"](#) on page 3-1
- ["Assigning/Editing Optimized Fields"](#) on page 3-1
- ["Performing a Fast Rebuild"](#) on page 3-2
- ["Optimizing the Search Collection"](#) on page 3-2
- ["Text Search Admin Page"](#) on page 3-3

3.1 Determining Fields to Optimize

Consider the following questions when determining what fields to optimize. Optimization improves search behavior for these actions.

- Do you want an exact match in a query?
- Do you want that match to work faster in a search?
- Do you want to sort search results by field?

By default the OracleTextSearch component optimizes the Content ID and Document Title metadata fields.

Oracle Content Server uses an SDATA field in the database for each optimized field, and there is a limit of 32 SDATA fields available in Oracle Text 11g. This limit is why Oracle Text Search can define a maximum number of 32 Optimized Fields. The display of integer fields in OracleTextSearch is dynamic and depends on the Oracle Content Server system configuration.

3.2 Assigning/Editing Optimized Fields

To select metadata Non-Optimized Fields and assign them to be Optimized Fields for search purposes, or to edit Optimized Fields and make them Non-Optimized, complete these steps:

1. Log on to Oracle Content Server as system administrator.
2. Click **Administration** in the navigation panel.
3. Click **Oracle Text Search Admin** in the navigation panel.

The [Text Search Admin Page](#) is displayed.

4. To make a metadata field optimized, click the desired field name in the Non-Optimized Fields column, then click the appropriate arrow to move the field to the Optimized Fields column.
5. To edit an Optimized Field and make it Non-Optimized, click the desired field name in the Optimized Fields column, then click the appropriate arrow to move the field to the Non-Optimized column.
6. When you have completed moving fields, click **Index Fast Rebuild** to update the search collection to use the new and modified fields.

Note: The Fast Rebuild will not function if a search collection rebuild is already in progress.

3.3 Performing a Fast Rebuild

The Fast Rebuild option on the Text Search Admin page allows the search engine to add new information to the search collection without requiring a full collection rebuild. A Fast Rebuild does not re-index metadata or full-text items.

A Fast Rebuild is required in the following cases:

- Adding or removing information fields
- Changing any Optimized Field
- Changing an information field to be an Optimized Field

The Repository Manager can be used to perform a full search collection rebuild, if necessary.

To perform a Fast Rebuild, complete these steps:

1. Log on to Oracle Content Server as system administrator.
2. Click **Administration** in the navigation panel.
3. Click **Oracle Text Search Admin** in the navigation panel.

The [Text Search Admin Page](#) is displayed.

4. Click **Index Fast Rebuild**.

A Fast Rebuild of the search collection is performed.

Note: A Fast Rebuild will not be performed if a rebuild of the search collection is already in progress.

3.4 Optimizing the Search Collection

After numerous search collection rebuilds, it may be useful to run the Index Optimization option on the Text Search Admin page to optimized search performance. This function scans the branches and stubs of the search collection index and adjusts links to speed up the time required to find metadata and full-text items.

To optimize the search collection using the OracleTextSearch component, complete these steps:

1. Log on to Oracle Content Server as system administrator.
2. Click **Administration** in the navigation panel.

3. Click **Oracle Text Search Admin** in the navigation panel.
The [Text Search Admin Page](#) is displayed.
4. Click **Index Optimization**.
Optimization of the search collection index is performed.

3.5 Modifying the OracleTextSearch Fields Displayed on Search Results

The OracleTextSearch component provides three default menu options on the Search Results page (set by the Oracle Database configuration script):

```
DrillDownFields=dDocType, dSecurityGroup, dDocAccount
```

Administrators can add one more option from the list of Optimized Fields to further customize the search results. Edit the configuration to add the option to the list of DrillDownFields.

Note: To ensure optimal performance, no more than four fields should be included. If one of the fields added to the list is not an Optimized Field, then the index will need to be rebuilt.

3.6 Finding Status Information

Several methods exist to find the status of what is being indexed from a document or to view query status.

Find What is Indexed From a Document

To find out exactly what is being indexed from a document, turn on logging in the Oracle Content Server Repository Manager:

1. Select **Administration** in the navigation portal bar, then select **Admin Applets**.
2. On the Admin Applets screen, select **Repository Manager**.
3. Select the **Indexer** tab.
4. Select **Automatic Update Cycle**, then click **Configure**.
5. For the Index Debug Level, select **Trace**. (When rebuilding the search collection, use the same options in the Collection Rebuild section.)

A text file containing the exact text extracted from the document will be created in the `<Install_Dir>/Search/<active_index>/bulkload/~export` directory.

6. Open the file with a text editor to view the content.

3.7 Text Search Admin Page

The Text Search Admin page is used to select which metadata fields are optimized for searching via Oracle Text and to rebuild the OracleTextSearch index. To access this page, select **Administration** from the main Oracle Content Server interface, then select **Text Search Admin**.

Text Search Admin

Administration --> Text Search Admin

Optimal Field Configuration

Non Optimized Fields	Optimized Fields
<p>Standard</p> <p>Author</p> <p>Content Type</p> <p>Security Group</p> <p>Custom</p> <p>Customer Code</p> <p>Customer Name</p> <p>Discussion Type</p> <p>Partition ID</p> <p>Stellent Language</p> <p>Stellent Long Text</p> <p>Stellent Text</p> <p>Stellent Text 1</p> <p>Stellent Text 2</p> <p>Storage Rule</p>	<p>Content Type</p> <p>Security Group</p> <p>Customer Code</p> <p>Customer Name</p> <p>Discussion Type</p> <p>Partition ID</p> <p>Stellent Language</p> <p>Stellent Text</p> <p>Stellent Text 1</p> <p>Stellent Text 2</p> <p>Storage Rule</p> <p>Web Flag</p> <p>Account</p> <p>AlternateFormat</p> <p>File Extension</p>
<p>Update Reset</p>	

Rebuild And Optimization

Fast Rebuild Status: Idle
 Index completed at 6/16/08 9:50 PM.

Optimization Status: Idle
 Index optimization completed at 6/16/08 3:19 PM

Element	Description
Non-Optimized Fields column	Lists available non-optimized metadata fields that can be optimized for searching by the Oracle Text search engine. See "Assigning/Editing Optimized Fields" on page 3-1.
Optimized Fields column	Lists the metadata fields that are optimized for searching by Oracle Text search engine. See "Assigning/Editing Optimized Fields" on page 3-1.
Update button	Updates the list of field assignments.
Reset button	Clears the list of fields selected to be optimized.
Index Fast Rebuild button	Performs a fast rebuild of the Oracle Text search index, implementing the fields selected to be optimized. See "Performing a Fast Rebuild" on page 3-2
Index Optimization button	Optimizes the Oracle Text search index to improve performance time for retrieving information for a search. This may be useful in certain situations, such as when Oracle Content Server has run for a long time and many new content items have been checked in. Optimizing the search index condenses the content item entries to make searching more efficient. See "Optimizing the Search Collection" on page 3-2.

Searching with the OracleTextSearch Component

This chapter covers the following topics:

- "Performing a Search" on page 4-1
- "Search Results With OracleTextSearch" on page 4-1

4.1 Performing a Search

Performing a search with the OracleTextSearch component enabled is generally the same except for the following:

- There are no visible changes in the Search:Expanded Form page other than removal of Substring as a search operator option. The default search operator is CONTAINS. Substring-based queries will still work.
- Queries using the MATCHES operator on a non-optimized field will behave like a CONTAINS query. For example, if `xDepartment` is not optimized, then the query `xDepartment MATCHES 'Marketing'` will behave like `xDepartment CONTAINS 'Marketing'` and return hits on documents that have an `xDepartment` value of `'Marketing Services'` or `'Product Marketing'`.

4.2 Search Results With OracleTextSearch

When users run a search using the Search:Expanded Form, the Search Results page displays an additional menu bar with options that enable users to selectively view search results. The options represent categories used to filter the search results. The options can be context-sensitive, so if only one content item is returned for an option, then it shows only the one result in the menu itself, as shown in [Figure 4-1](#). The default set of options include content type, security group, and account.

Note: Two default menu options on the OracleTextSearch menu bar can be replaced by customized menu options: Security Group and Document Type.

If more than one content item is found for an option, an arrow is displayed next to the option name. When you move your cursor over the option name, a popup displays the list of the categories found in the search results for that option and the number of content items for each of the categories. You can click any category name in the popup to change the search results page to list only those items that match the category. For example, [Figure 4-2](#) shows a Security Group list with following categories and

number of items found: Administration- (3), Marketing- (1), Public- (14), Secure- (5), Production- (1).

Figure 4–1 Search results with OracleTextSearch default menu

Search Results Found 11 items

Change View Query Actions

Filter by Category: Content Type:ADACCT Security Group Account

Actions					
<input type="checkbox"/>	ID	Title	Date	Author	Actions
<input type="checkbox"/>	PPT_TEST1	PPTTestDoc1	6/6/08	sysadmin	
<input type="checkbox"/>	SEARCHINDEX_000021	TestDoc19	9/4/02	sysadmin	
<input type="checkbox"/>	SEARCHINDEX_000010	TestDoc8	9/4/02	sysadmin	
<input type="checkbox"/>	SEARCHINDEX_000013	TestDoc11	9/4/02	user 1	
<input type="checkbox"/>	SEARCHINDEX_000011	TestDoc9	9/4/02	sysadmin	

Figure 4–2 Search results with snippets display and expanded OracleTextSearch menu

Search Results Found 24 items

Change View Query Actions

Filter by Category: Content Type Security Group Account Author:sysadmin

Administration- (3) Marketing- (1) Public- (14) Secure- (5)
Production- (1)

<input type="checkbox"/>	Description	Content ID	Actions
<input type="checkbox"/>	<p>Title: Chinese Test Author: sysadmin Content Type: ADACCT - Acme Accounting Department Account: MSP Security Group: Administration Native File Extension: doc Release Date: 6/19/08 4:20 PM Snippet: idcnll SEARCHINDEX_000026 idcnll sysadmin idcnll N/A idcnll idcnll Score: 3</p>	SEARCHINDEX_000026	
<input type="checkbox"/>	<p>Title: Contemporary Resume RTF Author: sysadmin Content Type: ADACCT - Acme Accounting Department Account: MSP Security Group: Administration Native File Extension: rtf Release Date: 6/16/08 11:10 AM Snippet: idcnll SEARCHINDEX_000025 idcnll sysadmin idcnll N/A idcnll idcnll Score: 3</p>	SEARCHINDEX_000025	
<input type="checkbox"/>	<p>Title: errortext Author: sysadmin Content Type: ADACCT - Acme Accounting Department Account: MSP Security Group: Public Native File Extension: txt</p>	SEARCHINDEX_000024	

Element	Description
Filter by Category	Displays the categories used to filter the search results; for example, Content Type, Security Group, Account.
Content Type	(Default) Lists the types and the number of each type of content items in the search results. Clicking one of the content type names will change the search results list to show only those items that match the content type.

Element	Description
Security Group	<p>(Default) Lists the security groups and number of content items assigned to each group in the search results. Default security groups include Public and Secure.</p> <p>Clicking one of the security group names will change the search results list to show only those items that match the security group.</p>
Account	<p>(Default) Lists the account types and number of items assigned to each account in the search results.</p> <p>Clicking one of the account types will change the search results list to show only those content items that match the account.</p>

Third Party Licenses

This appendix includes a description of the Third Party Licenses for all the third party products included with this product.

- ["Apache Software License"](#) on page A-1
- ["W3C Software Notice and License"](#) on page A-1
- ["Zlib License"](#) on page A-2
- ["General BSD License"](#) on page A-3
- ["General MIT License"](#) on page A-3
- ["Unicode License"](#) on page A-3
- ["Miscellaneous Attributions"](#) on page A-4

A.1 Apache Software License

```
* Copyright 1999-2004 The Apache Software Foundation.
* Licensed under the Apache License, Version 2.0 (the
* "License"); you may not use this file except in compliance
* with the License.
* You may obtain a copy of the License at
*   http://www.apache.org/licenses/LICENSE-2.0
*
* Unless required by applicable law or agreed to in writing,
* software distributed under the License is distributed on an
* "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND,
* either express or implied.
* See the License for the specific language governing
* permissions and limitations under the License.
```

A.2 W3C Software Notice and License

```
* Copyright 1994-2000 World Wide Web Consortium,
* (Massachusetts Institute of Technology, Institut National de
* Recherche en Informatique et en Automatique, Keio University).
* All Rights Reserved.  http://www.w3.org/Consortium/Legal/
*
* This W3C work (including software, documents, or other related
* items) is being provided by the copyright holders under the
* following license. By obtaining, using and/or copying this
* work, you (the licensee) agree that you have read, understood,
* and will comply with the following terms and conditions:
*
* Permission to use, copy, modify, and distribute this software
```

```
* and its documentation, with or without modification, for any
* purpose and without fee or royalty is hereby granted, provided
* that you include the following on ALL copies of the software
* and documentation or portions thereof, including
* modifications, that you make:
*
* 1. The full text of this NOTICE in a location viewable to
* users of the redistributed or derivative work.
*
* 2. Any pre-existing intellectual property disclaimers,
* notices, or terms and conditions. If none exist, a short
* notice of the following form (hypertext is preferred, text is
* permitted) should be used within the body of any redistributed
* or derivative code: "Copyright [$date-of-software] World
* Wide Web Consortium, (Massachusetts Institute of Technology,
* Institut National de Recherche en Informatique et en
* Automatique, Keio University). All Rights Reserved.
* http://www.w3.org/Consortium/Legal/"
*
* 3. Notice of any changes or modifications to the W3C files,
* including the date changes were made. (We recommend you
* provide URIs to the location from which the code is derived.)
*
* THIS SOFTWARE AND DOCUMENTATION IS PROVIDED "AS IS," AND
* COPYRIGHT HOLDERS MAKE NO REPRESENTATIONS OR WARRANTIES,
* EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, WARRANTIES
* OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR
* THAT THE USE OF THE SOFTWARE OR DOCUMENTATION WILL NOT
* INFRINGE ANY THIRD PARTY PATENTS, COPYRIGHTS, TRADEMARKS OR
* OTHER RIGHTS.
*
* COPYRIGHT HOLDERS WILL NOT BE LIABLE FOR ANY DIRECT, INDIRECT,
* SPECIAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF ANY USE OF THE
* SOFTWARE OR DOCUMENTATION.
*
* The name and trademarks of copyright holders may NOT be used
* in advertising or publicity pertaining to the software without
* specific, written prior permission. Title to copyright in this
* software and any associated documentation will at all times
* remain with copyright holders.
*
```

A.3 Zlib License

```
* zlib.h -- interface of the 'zlib' general purpose compression library version
1.2.3, July 18th, 2005
Copyright (C) 1995-2005 Jean-loup Gailly and Mark Adler
This software is provided 'as-is', without any express or implied warranty. In no
event will the authors be held liable for any damages arising from the use of this
software.
Permission is granted to anyone to use this software for any purpose, including
commercial applications, and to alter it and redistribute it freely, subject to
the following restrictions:
1. The origin of this software must not be misrepresented; you must not claim
that you wrote the original software. If you use this software in a product, an
acknowledgment in the product documentation would be appreciated but is not
required.
2. Altered source versions must be plainly marked as such, and must not be
misrepresented as being the original software.
3. This notice may not be removed or altered from any source distribution.
```


Jean-loup Gailly jloup@gzip.org
 Mark Adler madler@alumni.caltech.edu

A.4 General BSD License

Copyright (c) 1998, Regents of the University of California
 All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

"Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

"Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

"Neither the name of the <ORGANIZATION> nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

A.5 General MIT License

Copyright (c) 1998, Regents of the Massachusetts Institute of Technology

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

A.6 Unicode License

UNICODE, INC. LICENSE AGREEMENT - DATA FILES AND SOFTWARE

Unicode Data Files include all data files under the directories

<http://www.unicode.org/Public/>, <http://www.unicode.org/reports/>, and

<http://www.unicode.org/cldr/data/> . Unicode Software includes any source code published in the Unicode Standard or under the directories

<http://www.unicode.org/Public/>, <http://www.unicode.org/reports/>, and

<http://www.unicode.org/cldr/data/>.

NOTICE TO USER: Carefully read the following legal agreement. BY DOWNLOADING, INSTALLING, COPYING OR OTHERWISE USING UNICODE INC.'S DATA FILES ("DATA FILES"), AND/OR SOFTWARE ("SOFTWARE"), YOU UNEQUIVOCALLY ACCEPT, AND AGREE TO BE BOUND BY,

ALL OF THE TERMS AND CONDITIONS OF THIS AGREEMENT. IF YOU DO NOT AGREE, DO NOT DOWNLOAD, INSTALL, COPY, DISTRIBUTE OR USE THE DATA FILES OR SOFTWARE.

COPYRIGHT AND PERMISSION NOTICE

Copyright 1991-2006 Unicode, Inc. All rights reserved. Distributed under the Terms of Use in <http://www.unicode.org/copyright.html>.

Permission is hereby granted, free of charge, to any person obtaining a copy of the Unicode data files and any associated documentation (the "Data Files") or Unicode software and any associated documentation (the "Software") to deal in the Data Files or Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, and/or sell copies of the Data Files or Software, and to permit persons to whom the Data Files or Software are furnished to do so, provided that (a) the above copyright notice(s) and this permission notice appear with all copies of the Data Files or Software, (b) both the above copyright notice(s) and this permission notice appear in associated documentation, and (c) there is clear notice in each modified Data File or in the Software as well as in the documentation associated with the Data File(s) or Software that the data or software has been modified.

THE DATA FILES AND SOFTWARE ARE PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OF THIRD PARTY RIGHTS. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR HOLDERS INCLUDED IN THIS NOTICE BE LIABLE FOR ANY CLAIM, OR ANY SPECIAL INDIRECT OR CONSEQUENTIAL DAMAGES, OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER IN AN ACTION OF CONTRACT, NEGLIGENCE OR OTHER TORTIOUS ACTION, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THE DATA FILES OR SOFTWARE.

Except as contained in this notice, the name of a copyright holder shall not be used in advertising or otherwise to promote the sale, use or other dealings in these Data Files or Software without prior written authorization of the copyright holder.

Unicode and the Unicode logo are trademarks of Unicode, Inc., and may be registered in some jurisdictions. All other trademarks and registered trademarks mentioned herein are the property of their respective owners

A.7 Miscellaneous Attributions

Adobe, Acrobat, and the Acrobat Logo are registered trademarks of Adobe Systems Incorporated.

FAST Instream is a trademark of Fast Search and Transfer ASA.

HP-UX is a registered trademark of Hewlett-Packard Company.

IBM, Informix, and DB2 are registered trademarks of IBM Corporation.

Jaws PDF Library is a registered trademark of Global Graphics Software Ltd.

Kofax is a registered trademark, and Ascent and Ascent Capture are trademarks of Kofax Image Products.

Linux is a registered trademark of Linus Torvalds.

Mac is a registered trademark, and Safari is a trademark of Apple Computer, Inc.

Microsoft, Windows, and Internet Explorer are registered trademarks of Microsoft Corporation.

MrSID is property of LizardTech, Inc. It is protected by U.S. Patent No. 5,710,835. Foreign Patents Pending.

Oracle is a registered trademark of Oracle Corporation.

Portions Copyright 1994-1997 LEAD Technologies, Inc. All rights reserved.

Portions Copyright 1990-1998 Handmade Software, Inc. All rights reserved.

Portions Copyright 1988, 1997 Aladdin Enterprises. All rights reserved.

Portions Copyright 1997 Soft Horizons. All rights reserved.

Portions Copyright 1995-1999 LizardTech, Inc. All rights reserved.

Red Hat is a registered trademark of Red Hat, Inc.

Sun is a registered trademark, and Sun ONE, Solaris, iPlanet and Java are trademarks of Sun Microsystems, Inc.

Sybase is a registered trademark of Sybase, Inc.

UNIX is a registered trademark of The Open Group.

Verity is a registered trademark of Autonomy Corporation plc

Index

A

assigning Optimized Fields, 3-1

D

database provider, 2-5

databases

used with Oracle Text, 1-2

default optimized fields, 3-1

defining a provider, 2-5

displaying categories in search results, 4-1

E

editing Optimized Fields, 3-1

F

fast rebuild

overview, 3-2

when to perform, 3-2

I

Index Fast Rebuild

performing, 3-2

index fast rebuild

when to perform, 3-2

Index Optimization option, 3-2

installing the Oracle Text Search component, 2-3

M

maximum number of fields to optimize, 3-1

O

Optimized Fields

assigning, 3-1

editing, 3-1

optimizing the search collection, 3-2

Oracle Text Search component

considerations, 1-1

installing, 2-3

search results menu options, 4-1

S

search collection

optimizing, 3-2

search collection instance, 1-2

platform, 1-2

search results

displaying categories, 4-1

Search Results page

Oracle Text Search fields, 3-3

with Oracle Text Search, 4-1

searching with Oracle Text Search component, 4-1

stopwords

in search queries, 1-2

T

Text Search Admin page, 3-3

