Endeca Content Acquisition System

JSR-170 Compliant Repositories Connector Guide

Version 3.0.2 • March 2012

ENDECA

Contents

Preface.....7
About this guide.....7
Who should use this guide.....7
Conventions used in this guide.....8
Contacting Oracle Endeca Customer Support.....8

Chapter 1: Configuration steps for JSR-170 compliant repositories.....9
Exposing a remote JCR server to the CAS Server.....9
Setting up the CAS Server for JSR-170.....10
Configuration properties for a JSR-170 connector.....10
About obtaining ACL properties.....11

Copyright and disclaimer

Copyright © 2003, 2012, Oracle and/or its affiliates. All rights reserved.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners. UNIX is a registered trademark of The Open Group.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

This software or hardware and documentation may provide access to or information on content, products and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

Rosette® Linguistics Platform Copyright © 2000-2011 Basis Technology Corp. All rights reserved.

Teragram Language Identification Software Copyright © 1997-2005 Teragram Corporation. All rights reserved.

Preface

Oracle Endeca's Web commerce solution enables your company to deliver a personalized, consistent customer buying experience across all channels — online, in-store, mobile, or social. Whenever and wherever customers engage with your business, the Oracle Endeca Web commerce solution delivers, analyzes, and targets just the right content to just the right customer to encourage clicks and drive business results.

Oracle Endeca Guided Search is the most effective way for your customers to dynamically explore your storefront and find relevant and desired items quickly. An industry-leading faceted search and Guided Navigation solution, Oracle Endeca Guided Search enables businesses to help guide and influence customers in each step of their search experience. At the core of Oracle Endeca Guided Search is the MDEX Engine,™ a hybrid search-analytical database specifically designed for high-performance exploration and discovery. The Endeca Content Acquisition System provides a set of extensible mechanisms to bring both structured data and unstructured content into the MDEX Engine from a variety of source systems. Endeca Assembler dynamically assembles content from any resource and seamlessly combines it with results from the MDEX Engine.

Oracle Endeca Experience Manager is a single, flexible solution that enables you to create, deliver, and manage content-rich, cross-channel customer experiences. It also enables non-technical business users to deliver targeted, user-centric online experiences in a scalable way — creating always-relevant customer interactions that increase conversion rates and accelerate cross-channel sales. Non-technical users can control how, where, when, and what type of content is presented in response to any search, category selection, or facet refinement.

These components — along with additional modules for SEO, Social, and Mobile channel support — make up the core of Oracle Endeca Experience Manager, a customer experience management platform focused on delivering the most relevant, targeted, and optimized experience for every customer, at every step, across all customer touch points.

About this guide

This guide describes the tasks necessary to configure the JSR-170 Compliant Repositories CMS connector.

It assumes familiarity with the concepts of the Endeca Content Acquisition System and the Endeca Information Transformation Layer. For more information, see the *Endeca CAS Developer's Guide* and the *Endeca Forge Guide*.

Who should use this guide

This guide is intended for application developers who are building applications using the Endeca Content Acquisition System, and are responsible for gathering, crawling, joining and feeding the data in different source formats into the Endeca pipeline to transform them into Endeca records.

Conventions used in this guide

This guide uses the following typographical conventions:

Code examples, inline references to code elements, file names, and user input are set in `monospace` font. In the case of long lines of code, or when inline monospace text occurs at the end of a line, the following symbol is used to show that the content continues on to the next line: ¬

When copying and pasting such examples, ensure that any occurrences of the symbol and the corresponding line break are deleted and any remaining space is closed up.

Contacting Oracle Endeca Customer Support

Oracle Endeca Customer Support provides registered users with important information regarding Oracle Endeca software, implementation questions, product and solution help, as well as overall news and updates.

You can contact Oracle Endeca Customer Support through Oracle's Support portal, My Oracle Support at *https://support.oracle.com*.

Chapter 1

Configuration steps for JSR-170 compliant repositories

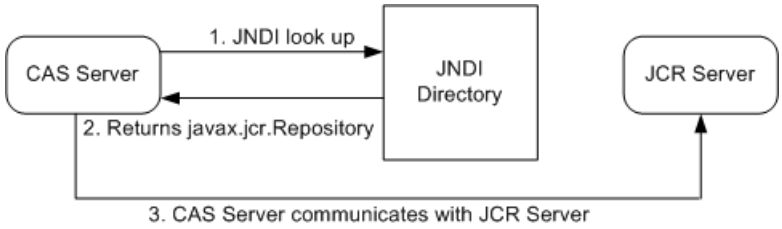
Set up the CAS Server for crawling JSR-170 compliant repositories, and set options specific to the repository in the CAS Console for Oracle Endeca Workbench.

Exposing a remote JCR server to the CAS Server

The instructions below provide an overview of how to expose a Java Content Repository (JCR) server to the CAS Server. This configuration enables communication among the CAS Server, JNDI, and the JCR server components in an implementation.

The CAS Server requires that the `javax.jcr.Repository` instance be exposed in the JNDI directory in order for the CAS Server to communicate with the remote JCR server. If a `javax.jcr.Repository` instance is not available in the JNDI directory, a communication problem occurs when the CAS Server attempts to contact the JCR server.

The following diagram shows the communication process among the components:



1. The CAS Server calls JNDI to look up the name of the remote JCR repository.
2. The JNDI directory returns an instance of a `javax.jcr.Repository`.
3. Using the `javax.jcr.Repository` handle, the CAS Server communicates directly with the remote JCR repository.

The implementation of the steps to expose a remote JCR server are specific to your environment. The steps below provide high-level guidance.

To expose a remote JCR Server to the CAS Server:

1. Configure the JCR server to make the `javax.jcr.Repository` handle available in the JNDI directory.

For more information, see the JNDI documentation at *http://java.sun.com/products/jndi/docs.html* and your JCR-compliant CMS documentation.

2. Configure the Endeca CAS Service to point at the appropriate JNDI directory.

   For more information about how to configure the client side CAS JNDI libraries, see the JNDI InitialContext documentation at *http://java.sun.com/javase/6/docs/api/javax/naming/InitialContext.html*. This configuration usually consists of modifying `cas-service.bat`, `cas-service-wrapper.conf` (Windows) or `cas-service.sh` (UNIX) to pass additional system properties to the JVM running the CAS Service.

   Here are several examples: On UNIX, you modify `cas-service.sh` by adding an additional `JVM_ARGS` setting:

   ```
   JVM_ARGS="$JVM_ARGS <your JNDI setting>"
   ```

   On Windows, if you are running the Endeca CAS Service from the Microsoft Services console, you modify `cas-service-wrapper.conf` by adding an addtional JVM setting:

   ```
   wrapper.java.additional.14=-<your JNDI setting>
   ```

   On Windows, if you are running the Endeca CAS Service from the command prompt, you modify `cas-service.bat` by adding an additional JVM setting:

   ```
   SET JVM_ARGS=%JVM_ARGS% -<your JNDI setting>
   ```

Setting up the CAS Server for JSR-170

To crawl a JSR-170 compliant repository, perform additional configuration steps on the machine where the CAS Server will be running.

It is assumed that your CAS Server is set up by including the libraries for your implementation of the JSR-170 API (Content Repository for Java™ technology API).

To configure the CAS Server for a JSR-170 compliant repository:

1. Add any custom Java libraries that are being used by the repository to the `<install path>\CAS\`*`version`*`\lib\server` directory (on Windows) and to `<install path>/CAS/`*`version`*`/lib/server` (on UNIX).
2. Restart the Endeca CAS Service.

To run a crawl on your repository, you also specify options specific to the repository in the CAS Console.

Configuration properties for a JSR-170 connector

To configure a JSR-170 connector, specify the configuration properties listed below.

Note: In addition to configuring the connector-specific property listed below, you must enter values for the data source username and password.

Create the following configuration properties using either CAS Console or the CAS Server Command-line Utility.

CAS Property Display Name	CAS Property Name	Property Description
JNDI name or RMI address	`jndiName`	(Required). Specify the JNDI name or RMI address that CAS uses to retrieve the repository implementation.
JCR repository retrieval method	`JNDIOrRMI`	(Required). Specify `JNDI` if you specified a JNDI name above or specify `RMI` if you specified an RMI address.

Note: Properties are case sensitive.

About obtaining ACL properties

The CAS Server does not generate ACL properties from JSR-170 compliant repositories, because the JSR-170 specification does not yet provide standard mechanisms for retrieving ACLs (Access Control Lists).

If ACL properties are required in your pipeline, consider the following method:

- If the CAS Server generates ACLs as generic name-value pairs for the repository items that have ACLs, you can use record manipulators to process them in the pipeline.

For additional information on retrieving ACLs, please contact Endeca Customer Support.