

Installation Guide Document

Product Name: CDE Curation Tool

Version 3.2

Submitted to: NCI Center for Bioinformatics

Prepared by:



Larry Hebel
101 West Renner Road, Suite 130
Richardson, TX 75082
September, 2, 2005

Contact Information:

E-mail lhebel@scenpro.com

Telephone 972/437-5001

FAX 972/437-3611

Revision History

Author(s)/Responsible Party	Rev No.	Date	Page(s)
Sumana Hegde	1	10/23/06	All

Table of Contents

1	INTRODUCTION.....	4
2	INSTALLATION STEPS.....	5
3	SITE CONFIGURATION.....	6
4	REQUIRED TECHNOLOGY	11
5	TROUBLESHOOTING	12

1 Introduction

The purpose of this document is to detail the Installation Steps and Site Configuration Options for the CDE Curation Tool Version 3.2.

The Curation Tool source and all related documentation can be accessed at <http://gforge.nci.nih.gov/projects/cdecurate>. All caCORE companion projects can be referenced from <http://gforge.nci.nih.gov>.

Assumptions

1. These steps rely on the installation of the standard caCORE 3.2 Technology Stack for runtime and builds. **If the environment has not been previously used for one of the other caDSR Tools (e.g. caDSR Sentinel, CDEBrowser, etc.), please refer to Section 5, Required Technology, for special notes and technology URL references.**
2. The caDSR database is installed and verified using Oracle SQL Plus or other appropriate database query tool.
3. The Oracle JAR should be copied from the Oracle Client installation to the JBoss 4.0.2 installation. The Curation Tool accesses the caDSR database from the web server using a 'thin' connection which does not require the installation of the Oracle OCI.

Document Definitions and Syntax

1. Download Package
The zip file available from the NCICB GForge Curation Tool project site noted above.
2. Folder Path Separator ('/')
The folder path separator on Windows is the character '\ ' and on Unix is the character '/ '. For simplicity all folder (directory) paths in this document use the '/' character. Change as needed for your environment.
3. Folder name 'cdecurate'
All examples represent the root for the extracted package as 'cdecurate'.
4. Folder name 'JBoss'
All examples represent the root for the installed JBoss 4.0.2 web server as 'JBoss'.
5. Use of quotes (" ") and apostrophes (' ')
The appearance of quotes is restricted to examples in which the quote should be used as a typed key. Apostrophes are restricted to clarification within this document and do not appear in examples or any data used for the installation of the software.

2 Installation Steps

Should any errors occur during the completion of these steps, refer to Section 6 Troubleshooting, for assistance.

1. Download package -
Extract the complete content of the download package zip or tar.gz file, preserving the folder names.
2. Make **Site Configuration** is executed as per the environment.
3. Execute '**ant build-osi**' to build cdecurate.war file.
4. Move the **cdecurate.war** to the 'jboss/server/default/deploy' folder.
5. Restart **JBoss** if necessary.

Depending on the installation JBoss may be configured for new or existing deployments. Check the JBoss server.log to verify the deployment of the cdecurate.war file. If you cannot verify the deployment try restarting JBoss.

6. Open a browser window and enter **http://<server name>/cdecurate** in the address.

Replace **<server name>** with the appropriate reference to the server and the port on which JBoss is listening. This should open the Curation Tool Login window which confirms JBoss has deployed the application.

3 Site Configuration

As noted in section 2, step 2 above, any configuration changes to the Curation Tool must be done before building a war file.

- **Database Specific Properties**

The database specific properties are stored in 'cdecurate/conf/web/web.xml' extracted from the download package. Some of the values of these xml tags can be changed as needed but do not remove or comment-out any of the tags.

1. Database Username

The user name / account with write access to the database connection is stored in the **username** tag as shown below.

```
<init-param>
  <param-name>username</param-name>
  <param-value>cdecurate</param-value>
</init-param>
```

It is strongly recommended this be an account which is used exclusively by the Curation Tool, hence the default value of 'cdecurate'. By creating this dedicated account it becomes possible to identify specific database changes made by the logic independent of user interaction. In other words, when the Curation Tool code needs access to or makes changes in the database due to an Automated or Unattended execution, it will use this username account.

2. Database Password

The password to match the user name account for database access is stored in the **password** tag as shown below.

```
<init-param>
  <param-name>password</param-name>
  <param-value>cdecurate</param-value>
</init-param>
```

3. Tool session timeout

Tool level session timeout can be configured in the session-config tag as shown below.

```
<session-config>
  <session-timeout>45</session-timeout>
</session-config>
```

- **Site Specific Properties**

Configure the datasource for Curation tool connection in **oracle-ds.xml** file at 'jboss/server/default/deploy' folder as follows.

```
<datasources>
  <local-tx-datasource>
    <jndi-name>jdbc/CDECurateDS</jndi-name>
    <connection-url>jdbc:oracle:thin:@database description</connection-url>
    <driver-class>oracle.jdbc.driver.OracleDriver</driver-class>
    <exception-sorter-class-name>
      org.jboss.resource.adapter.jdbc.vendor.OracleExceptionSorter
    </exception-sorter-class-name>
    <check-valid-connection-sql>select 1 from dual</check-valid-connection-sql>
    <blocking-timeout-millis>10000</blocking-timeout-millis>
    <idle-timeout-minutes>1</idle-timeout-minutes>
    <max-pool-size>85</max-pool-size>
    <min-pool-size>1</min-pool-size>
    <application-managed-security/>
  </local-tx-datasource>
</datasources>
```

The **database description**, it can take either of two (2) forms, (1) the **<server>:<port>:<sid>** or (2) **<alias>**.

The first form, **<server>:<port>:<sid>**, indicates a thin client connection and does not require the installation of the Oracle OCI on the application server machine. This is the same information that appears in the tnsnames.ora file for the alias description. This first form is recommended. The second form is only provided for backward compatibility.

The second form, **<alias>**, indicates a thick client connection and requires the installation of the Oracle OCI on the application server machine. The database alias referencing the caDSR database must appear in the Oracle tnsnames.ora file. Find the alias for the caDSR database then copy and paste it as the value for this attribute. **The alias name must appear exactly as entered in the tnsnames.ora file.**

The tnsnames.ora file is normally located in %oracle_home%/network/admin/. The database alias appears to the left of the equal sign. For example, the alias 'cadsr.nci.nih.gov' may appear in tnsnames.ora as:

```
cadsr.nci.nih.gov = (DESCRIPTION =
  (ADDRESS_LIST = ...
  )
  (CONNECT_DATA = (SID = ...))
  )
)
```

Often multiple database aliases are listed in the tnsnames.ora file, one for each Oracle Database for which access is granted. One of these must be configured for access to the caDSR database. If unsure which, please consult your Oracle DBA.

- **Tool Configuration Options**

Some of the Curation Tool logics are controlled by configuration options stored in a table within the caDSR database. These options are set via the script **load_tool_options.sql** located at 'cdecurate/conf/<tier_name>', where <tier_name> can be stage, prod etc. The options contained in this script allow for dynamic changes without the need to build or deploy a cdecurate.war file.

The options are identified by the tool name and a property. For the Curation Tool the tool name in the script is always 'CURATION'. Details of the property and values of the configuration options follow. Each is identified by the property name followed by a description of its use. In addition to the information provided here there are extensive comments in the load_tool_options.sql file concerning valid values for each property.

1. URL

The URL to access the Curation Tool Login page, should be stored in the format – "http://<Curation Tool Alias>". E.g. NCICB production is <http://cdecurate.nci.nih.gov>. If <Curation Tool Alias> does not exist it should follow this format – "http://<server_name>:<port>/cdecurate"

2. EVS.URL

The EVS URL to access the EVS data from Curation Tool, should be stored in this format – "http://"<EVS RMI Alias>/cacore32/http/remoteService". E.g. NCICB production is <http://cabio.nci.nih.gov/cacore32/http/remoteService>. If <EVS RMI Alias> does not exist it should follow this format – "http://<server_name>:<port>/cacore32/http/remoteService".

3. REFDOC_FILECACHE

The directory for file cache to create url for uploading blobs from the caDSR database to the web server. E.g /local/content/cdecurate/filecache/. This directory must exist in the application server.

4. REFDOC_FILEURL

The prefix URL for building the file anchor tag for files uploaded to the file cache, should be stored in the format – "<URL>/filecache" where <URL> is as defined above (1). E.g. NCICB production is <http://cdecurate.nci.nih.gov/filecache/>.

Curation tool also requires the URLs of other NCICB applications stored in the tools_options table. If not already exists in the table, include the following in the load_tool_options.sql.

5. URL (CDE Browser)

The URL to access the CDE Browser Details page from Curation Tool should be stored in the format – "http://<CDE Browser Alias>". E.g. NCICB production is <http://cdebrower.nci.nih.gov>. If < CDE Browser Alias> does not exist it should follow this format – "http://<server_name>:<port>/CDEBrowser"

6. URL (EVS)

The URL to access the EVS data from Curation Tool, should be stored in this format – “http//”<EVS RMI Alias>/cacore32/http/remoteService”. E.g. NCICB production is <http://cabio.nci.nih.gov/cacore32/http/remoteService>. If <EVS RMI Alias> does not exist it should follow this format – “http//<server_name>:<port>/cacore32/http/remoteService”.

Execute the cdecurate/conf/osi/**load_tool_options.sql** in the SQL IDE.

This script can be run from Eclipse, SQL Plus, Toad or any SQL IDE connected to the desired caDSR database instance.

- **Logging Specific Property**

Curation Tool uses log4j to capture the logged information. Add the following information to **log4j.xml** for JBoss located at JBoss’s server/default/conf/ directory to make a separate log file for Curation Tool. This text must be placed inside the **<log4j:configuration>** xml tag.

```
<!-- ===== -->
<!-- Setup the cdecurate Appender -->
<!-- ===== -->
<appender name="CDECURATE_FILE"
class="org.jboss.logging.appender.RollingFileAppender">
  <errorHandler
class="org.jboss.logging.util.OnlyOnceErrorHandler"/>
  <param name="File"
value="${jboss.server.home.dir}/log/cdecurate.log"/>
  <param name="Append" value="false"/>
  <param name="MaxFileSize" value="1MB"/>
  <param name="MaxBackupIndex" value="3"/>
  <layout class="org.apache.log4j.PatternLayout">
    <param name="ConversionPattern" value="%d %-5p [%c] %m%n"/>
  </layout>
</appender>

<!-- ===== -->
<!-- Setup the cdecurate category -->
<!-- ===== -->
<category name="gov.nih.nci.cadsr.cdecurate.tool" additivity="false">
  <priority value="WARN"/>
  <appender-ref ref="CDECURATE_FILE"/>
</category>
```

- **Java SDK Specific Property**

In order for Curation Tool to access EVS Server, certain java policy permissions must be granted. If this is not done then EVS search results will not be available within the tool. The

server should have the System environmental variable JAVA_HOME pointing to the directory of the JDK.

The java.policy file in the jdk's jre/lib/security folder should have this entry:

```
grant {  
    permission java.security.AllPermission;  
};
```

- **JBoss Security Property**

If JBoss is configured to use security manager at startup, ensure that the policy file referenced has appropriate entries to allow necessary access for CDE Curation Tool to function properly.

Policy file used in NCICB environment -
\$JBASS_HOME/server/default/conf/server.policy

```
grant {  
    permission java.security.AllPermission;  
};
```

4 Required Technology

The product versions which follow were used during the development and system testing of the CDE Curation Tool. Although new versions of the products have been released they have not yet been certified with the Curation Tool version 3.2.

Note: deviating in any way from the products and versions listed below may cause errors. When attempting to deploy the Curation Tool to a different technology stack, please first deploy as documented in this Installation Guide and then begin the migration.

Product	Version	Web Site & Notes
JBoss	4.0.4	http://www.jboss.org/downloads/index
JDK	1.5.0_06	http://java.sun.com/j2se/1.5.0/ Recommend downloading the J2EE 1.5 SDK.
Struts	1.2.2	http://struts.apache.org/
ANT	1.6.2	http://ant.apache.org/
Eclipse	3.0.1	http://www.eclipse.org/platform/ This is a development environment only and provides no runtime or build components to the Sentinel Tool.
IE	6.0+	http://www.microsoft.com/windows/ie/downloads This is the latest Internet Explorer from Microsoft.

5 Troubleshooting

Following is a list of known errors and resolutions that may occur during the build and deployment process. If you have a question or situation not covered in this document please contact the NCICB Help Desk via email at ncicb@pop.nci.nih.gov.

1. **The Curation Tool Login screen appears and a valid user name and password is not accepted.**

If the error “**java.sql.exception: TNS: listener could not resolve SERVICE_NAME**” is found in server console or log file, an entry is needed in the TNSNAMES.ORA file. See Database Specific Property in Section 3 above. Also make sure the database description under oracle-ds.xml, as specified in site configuration, is valid.

Verify the username and password tag values entered in web.xml are valid and successfully connects to the database by other query tools like SQLPlus or Toad.

Verify the Oracle OCI client is installed on the same machine as JBoss. Please check with the System Administrator or DBA for this.

2. **Using Firefox, Mozilla and other web browsers.**

The Curation Tool has only been certified with Microsoft Internet Explorer version 6.0 and higher.

3. **java.net.bind.exception JVM_8080 port not available.**

In `<jboss_home>/server/all/deploy/jbossweb.sar/META-INF/jboss-service.xml`, edit the following:

- Under the "Configure the Request Listeners" section, look for the add listener call of "org.mortbay.http.SocketListener" and change the port setting line to (8080 is the default):

```
<Set name="Port"><SystemProperty name="jetty.port"
default="18080"/></Set>
```
- Under the "Configure the Request Listeners" section, look for the add listener call of "org.mortbay.http.ajp.AJP13Listener" and change the port setting line to (8009 is the default):

```
<Set name="Port">18009</Set>
```

Also the call "org.mortbay.http.ajp.AJP13Listener", change the port setting line to (443 is the default):

```
<Set name="confidentialPort">1443</Set>
```

4. **Error in EVS Connection.**

Make sure EVS connection string stored in the value of EVS.URL property in tool_options_ext table in caDSR is valid by accessing the URL. If the URL is valid and the server is up the page with the following text will be opened.

“Welcome to the HTTP Proxy. The Server has been deployed. call doPost to send request!”

5. **Goes to error page with NullPointerException at the opening of create or edit pages.**

If you cannot open create or edit pages of any AC, verify the ‘cdecurate.log’ file located at ‘<jboss_home>/server/default/log’ folder. If the error message ‘getCSCSIList for other : java.sql.SQLException: Invalid column name’ is found in the log file, there may be an API error. Modify the stored procedure SBREXT_CDE_CURATOR_PKG.GET_CSCSI_LIST to remove the comment text next to ‘LEVEL’ field of the query.

6. **API_CON_106 error while creating Data Element Concept**

Make sure that ‘caBIG’ context data exists in the ‘Contexts’ table of database either by querying the database or locating the ‘caBIG’ under Contexts dropdown list of the contexts on the main search page.

If you have a question or situation not covered in this document please contact the NCICB Help Desk via email at ncicb@pop.nci.nih.gov.