

Oracle Commerce

Content Acquisition System Installation Guide

Version 11.2 • October 2015



Contents

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

- Chapter 1: Before you install.....9**
- Overview of the Content Acquisition System.....9
- Prerequisite Oracle Commerce components.....9
- System requirements.....9
- CAS Document Conversion Module.....10
- Recommended reading.....10

- Chapter 2: Installing the Content Acquisition System.....11**
- Installing on Windows.....11
 - Creating a user for the Oracle Commerce services on Windows.....11
 - Installing CAS on Windows.....12
 - Installing CAS silently on Windows.....13
 - Updating the Deployment Template to use the WSDL client stubs and the CAS Deployment Template component.....14
- Installing on UNIX.....14
 - Installing CAS on UNIX.....14
 - Adding the CAS service to inittab on UNIX.....16
 - Installing only the CAS Console on UNIX.....16
 - Installing CAS silently on UNIX.....17
 - Integrating CAS with the Deployment Template on UNIX.....18
- Installing a plug-in into CAS.....18
- Web Crawler installation.....19
- Package contents and directory structure.....19
 - Required installer file permissions.....21
- About changing the role used for the CAS Console extension.....23

- Chapter 3: Uninstalling the Content Acquisition System.....25**
- Uninstalling CAS on Windows.....25
- Uninstalling CAS Console on UNIX.....25
- Uninstalling CAS on UNIX.....26
- Uninstalling a plug-in from CAS.....26
- Uninstalling CAS Console if its extension configuration was changed.....26

Copyright and disclaimer

Copyright © 2003, 2015, Oracle and/or its affiliates. All rights reserved.

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.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

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 unless otherwise set forth in an applicable agreement between you and Oracle. 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, except as set forth in an applicable agreement between you and Oracle.

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>.

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info> or visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> if you are hearing impaired.

Preface

Oracle Commerce 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, Guided Search enables businesses to influence customers in each step of their search experience. At the core of Guided Search is the MDEX Engine™, a hybrid search-analytical database specifically designed for high-performance exploration and discovery. The Oracle Commerce 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. The Oracle Commerce Assembler dynamically assembles content from any resource and seamlessly combines it into results that can be rendered for display.

Oracle Commerce Experience Manager enables non-technical users to create, manage, and deliver targeted, relevant content to customers. With Experience Manager, you can combine unlimited variations of virtual product and customer data into personalized assortments of relevant products, promotions, and other content and display it to buyers in response to any search or facet refinement. Out-of-the-box templates and experience cartridges are provided for the most common use cases; technical teams can also use a software developer's kit to create custom cartridges.

About this guide

This guide describes how to install the Content Acquisition System and set up CAS components after installation on Windows and UNIX.



Note: Unless otherwise indicated, whenever this document specifies UNIX, it applies to Linux and Solaris.

Who should use this guide

This guide is intended for application developers who are building applications using the CAS components and are responsible for installation tasks.

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 Support

Oracle Support provides registered users with answers to implementation questions, product and solution help, and important news and updates about Guided Search software.

You can contact Oracle Support through the My Oracle Support site at <https://support.oracle.com>.

Chapter 1

Before you install

This section provides an overview of the Content Acquisition System, system requirements, and other information you need to know before installing.

Overview of the Content Acquisition System

The Oracle Commerce Content Acquisition System is a set of components that add, configure, and crawl data sources for use in an Oracle Commerce application. Data sources include file systems, Content Management System crawls, Web servers, and custom data sources. The Oracle Commerce Content Acquisition System crawls data sources, converts documents and files to Endeca records and processes them for use in an MDEX Engine.

Prerequisite Oracle Commerce components

CAS Console depends on Oracle Commerce Workbench. Tools and Frameworks, which includes Workbench, must be installed on the machine where you install the CAS Console.

System requirements

See the *Oracle Commerce Supported Environments Matrix* document in the My Oracle Support knowledge base at <https://support.oracle.com/> for information on supported operating systems and Web browsers. Hard disk capacity must be sufficient to store the records written to the Record Store or to record output files. Please contact your Oracle representative if you need more information on hardware sizing.

Minimum hardware requirements:

- x64 processor, minimum 1.8 GHz
- 2 GB of RAM
- At least an 80 GB hard drive, depending on the size of your application data set

CAS Document Conversion Module

In CAS 3.0.1 and later, the CAS Document Conversion Module is installed and enabled as part of the CAS installation.

For a list of supported file formats, see "Appendix B File Formats Supported by the CAS Document Conversion Module" in the *Oracle Commerce CAS Developer's Guide*.

Recommended reading

Before installing, Oracle recommends that you read the following documents for important information about the release.

Getting Started Guide

The *Oracle Commerce Getting Started Guide* gives an overview of Oracle Commerce components and includes information about configuration scenarios. After installing all the components in your deployment, read this guide for information on verifying your installation. You can download the *Oracle Commerce Getting Started Guide* from the Oracle Technology Network.

Release Notes

Refer to the release notes for information about known issues for this release. You can download the *Oracle Commerce Guided Search and Experience Manager Release Notes* from the Oracle Technology Network.

Migration Guide

Refer to the *Oracle Commerce CAS Migration Guide* for information about migrating your implementation from a previous version of Oracle Commerce software. You can download the *Oracle Commerce CAS Migration Guide* from the Oracle Technology Network.

Installing the Content Acquisition System

This section contains installation procedures and describes the contents of the installation directory. Although the CAS Console extension for Oracle Commerce Workbench is distributed in a single package with the CAS Server, it can be installed without installing the CAS Server.

Installing on Windows

This section provides instructions for installing CAS on Windows.

Creating a user for the Oracle Commerce services on Windows

You must run the Oracle Commerce services as a specified user, for which you can control permissions.

Oracle recommends that you create a user account called `endeca` that has the proper file and directory permissions to access all necessary files for your application, and that you set up your Oracle Commerce Workbench Service to run under this account. However, you can use any user that you prefer, as long as it meets these requirements.

To create the user `endeca`:

1. Ensure that you have administrator privileges on the local machine.
2. From the Windows Control Panel, select **Computer Management > Administrative Tools**.
3. In the tree pane of the Computer Management window, select **System Tools > Local Users and Groups**.
4. Right-click **Users** and select **New User**.
5. In the **New User** dialog box, do the following:
 - a) Enter `endeca` for both the User name and the Full name.
 - b) Optionally, enter a description.
 - c) Set a password for user `endeca`.



Note: The user must have a non-blank password, because the installer will not accept a blank password for the services user.

- d) Uncheck **User must change password at next logon**.
6. Click **Create** to create the new user, and then **Close** to exit the dialog box.
 7. Close the Computer Management window, but do not exit Administrative Tools.

8. From Administrative Tools, do the following:
 - a) Open **Local Security Policy**.
 - b) In the Local Security Settings window, select **Local Policies > User Rights Assignments**.
 - c) From the list of policies on the right, open **Log on as a Service**.
 - d) Add user `endeca` to the list of users that can register a process as a service.
 - e) Close the dialog box and exit Administrative Tools.

Installing CAS on Windows

Before installing, ensure that you have met the following pre-requisites:

- Ensure you have administrator privileges on the local machine.
- Close all running programs and, if you previously had CAS installed on this machine, refer to the "Upgrading" section of the *CAS Migration Guide* for information on migration procedures.
- If you are installing CAS Console, the `ENDECA_TOOLS_ROOT` and `ENDECA_TOOLS_CONF` environment variables must be set on the machine running Oracle Commerce Tools and Frameworks before running the CAS installation program. Setting these variables happens automatically when you restart the machine that has Oracle Commerce Tools and Frameworks installed.



Note: Even if you are installing on a drive other than the C drive, you should have approximately 400 MB of space available on the system drive before installing. This space is needed temporarily for the installer files that are unpacked on the C drive during the installation regardless of the drive you specified for your installation. The space is cleared after the installation process completes.

To install CAS on Windows:

1. Download the Content Acquisition System package from the Oracle Software Delivery Cloud.
2. Extract the Content Acquisition System package to a local directory.
The name of the extracted installer file is `OCcas-<version>-win32.exe`.
3. Double-click the installer file `OCcas-<version>-win32.exe` to start the wizard.
4. Select your preferred language, and click **OK**.
5. In the **Introduction** screen, click **Next** to begin the installation process.
6. Accept the terms of the license agreement and click **Next**.
7. In the **Choose Install Folder** screen, accept the default location of `C:\Endeca\CAS` and then click **Next**.
8. In the **Choose Product Features** screen, select required program features and click **Next**.
9. In the **Get User Input** screen, accept the default values for the CAS Server port(8500) and CAS Server shutdown port (8506).
10. In the **Get User Input for Tools and Frameworks Variables** screen, enter the values for `ENDECA_TOOLS_ROOT` and `ENDECA_TOOLS_CONF`.

<code>ENDECA_TOOLS_ROOT</code>	<code>C:\Endeca\ToolsAndFrameworks\11.2.0</code>
<code>ENDECA_TOOLS_CONF</code>	<code>C:\Endeca\ToolsAndFrameworks\11.2.0\server\workspace</code>

11. In the **Pre-Installation Summary** screen, review the settings you have selected and click **Install**.
12. In the **Install Complete** screen, click **Done** to exit the wizard.

The CAS Service starts automatically after installation.

Related Links

[Installing CAS silently on Windows](#) on page 13

The silent installer is useful if you want to add the CAS installation to an install script, or push out the installation on multiple machines. The silent installer is not interactive.

[Updating the Deployment Template to use the WSDL client stubs and the CAS Deployment Template component](#) on page 14

This task is optional. It may be necessary if you did not integrate the CAS Deployment Template component into the Deployment Template during the installation process and then later found that you need the Deployment Template to manage crawling operations.

[Creating a user for the Oracle Commerce services on Windows](#) on page 11

You must run the Oracle Commerce services as a specified user, for which you can control permissions.

Installing CAS silently on Windows

The silent installer is useful if you want to add the CAS installation to an install script, or push out the installation on multiple machines. The silent installer is not interactive.

To launch the silent installer on Windows:

1. From a command prompt, navigate to the directory where you downloaded the installer.
2. Create a response file with the appropriate properties.

Guided Search property Description

NAME

INSTALLDIR Required. Specifies the absolute path to the directory to install CAS (e.g., C:\\Endeca\\CAS). Keep in mind that you cannot install the software in a directory with spaces in its name.



Note: If you do not use the default location, and you are installing more than one Guided Search product on the same machine, ensure that you install each product to a separate location.

CASPORT

Required. The port on which the CAS Service listens. The default is 8500.

CASSHUTDOWNPORT

Required. The shutdown port for the CAS Service. The default is 8506.

ENDECA_TOOLS_ROOT

Required. C:\\Endeca\\ToolsAndFrameworks\\11.2.0

ENDECA_TOOLS_CONF

Required.
C:\\Endeca\\ToolsAndFrameworks\\11.2.0\\server\\workspace

Example:

```
INSTALLDIR=C:\\Endeca\\CAS
CASPORT=8500
CASSHUTDOWNPORT=8506
ENDECA_TOOLS_ROOT=C:\\Endeca\\ToolsAndFrameworks\\11.2.0
ENDECA_TOOLS_CONF=C:\\Endeca\\ToolsAndFrameworks\\11.2.0\\server\\workspace
```

3. Issue a command of the following form:

```
OCcas-<version>-win32.exe -i silent -f cas_response.properties
```

When the CAS installation is complete, the CAS Service starts automatically.

Updating the Deployment Template to use the WSDL client stubs and the CAS Deployment Template component

This task is optional. It may be necessary if you did not integrate the CAS Deployment Template component into the Deployment Template during the installation process and then later found that you need the Deployment Template to manage crawling operations.

To update the Deployment Template:

Copy the new CAS Deployment Template component into the Deployment Template:

- On Windows, copy `<install path>\CAS\<version>\lib\cas-dt\casStubs.jar` to `<install path>\ToolsAndFrameworks\<version>\deployment_template\app-templates\common\config\lib\java` and replace the existing JAR file.
- On UNIX, copy `<install path>/CAS/<version>/lib/cas-dt/casStubs.jar` to `<install path>/ToolsAndFrameworks/<version>/deployment_template/app-templates/common/config/lib/java`.

Both the WSDL client stubs and the CAS Deployment Template component (the `ContentAcquisition-ServerComponent` class) are packaged in `casStubs.jar`.

For details on upgrading applications deployed for a previous version of CAS, see the *Oracle Commerce CAS Migration Guide*.

Installing on UNIX

This section provides instructions for installing CAS on UNIX.

Installing CAS on UNIX

The Oracle Commerce software is distributed as a self-extracting bin file and install script.

CAS unpacks to approximately 400 MB. Because multiple versions may be stored, a destination in a large partition is recommended.

To install CAS on UNIX:

1. Stop the Endeca Tools Service.
2. In your local environment, locate the CAS installation package that you downloaded from the Oracle Software Delivery Cloud. The name of the installation file is as follows:
 - For Intel Linux 64-bit: `OCcas<version>-linux64.bin`
 - For SPARC Solaris: `OCcas<version>-solaris.bin`
3. Determine where you will install CAS. Verify that the target directory where you plan to install has enough available disk space and has write permissions (is not read-only). For example, in this procedure we assume that the target directory is `/usr/local` and that you have write permissions for it. If you do not set these permissions, the install script will not run.
4. From the `/downloads` directory, start the installation with the following command:

```
./OCcas11.2.0-linux64.bin
```

or

```
./OCcas11.2.0-solaris.bin
```

5. Select a locale.
6. In the **Introduction** screen, press **Enter** to continue.
7. Accept the terms of the license agreement.
8. In the **Choose Install Folder** screen, enter `/usr/local` as the installation location.
9. In the **Choose Product Features** screen, enter the number for the program features to be installed. The default is **Typical**.

Product Features	Description
Content Acquisition System	Oracle Commerce Web Crawler and CAS Server, and a rich set of packaged adapters
CAS Samples	Examples of CAS usage. These examples contain code and configuration files.
CAS Console as a Workbench Extension	Web-based application used to crawl various data sources including file systems and content management systems
CAS Deployment Template Integration	Integrates CAS with the Tools and Frameworks Deployment Template. The integration enables you to control CAS operations using the Deployment Template.

10. Enter the CAS Server port.
11. Enter the CAS Server Shutdown port.
12. Enter the fully qualified CAS Server host name.
13. If `ENDECA_TOOLS_ROOT` or `ENDECA_TOOLS_CONF` is not set in the environment, the installer prompts you for these values.
14. In the **Pre-Installation Summary** screen, review the settings you have selected and press **Enter** to continue.
15. In the **Ready to Install** screen, press **Enter** to start the installation.
The Installation Complete message displays at the completion of the installation.
16. Press **Enter** to exit the installation.

Following installation:

- To start the CAS Service, navigate to `/usr/local/endecca/CAS/<version>/bin` and run the following script: `cas-service.sh`
- If you chose to install the CAS Console, you must restart the Oracle Tools Service.

Related Links

[Installing CAS silently on UNIX](#) on page 17

The silent installer is useful if you want to add the CAS installation to your own install script, or push out the installation on multiple machines.

[Updating the Deployment Template to use the WSDL client stubs and the CAS Deployment Template component](#) on page 14

This task is optional. It may be necessary if you did not integrate the CAS Deployment Template component into the Deployment Template during the installation process and then later found that you need the Deployment Template to manage crawling operations.

Adding the CAS service to inittab on UNIX

In a UNIX development environment, the CAS Service can be started from the command line. In a UNIX production environment, however, Oracle recommends starting it from `inittab`.

In UNIX you run the CAS Server using `cas-service.sh`, located in `usr/local/endeca/CAS/<version>/bin`. You can write a script that calls `cas-service.sh` and is referenced in `inittab`.

When writing your script, it is recommended as a best practice that you run the CAS Service as a user other than `root`.

This sample script (named `start_cas_service.sh`) sets the `ENDECA_USER` variable to the “endeca” user and then issues an `su` command to change to the “endeca” user:

```
#!/bin/sh
ENDECA_USER=endeca
CAS_ROOT=/usr/local/endeca/CAS/<version>
CAS_WORKSPACE=/usr/local/endeca/CAS/workspace

# change to user endeca
su $ENDECA_USER -c "cd $CAS_ROOT/bin; \
    $CAS_ROOT/bin/cas-service.sh >> $CAS_WORKSPACE/logs/cas-service-wrapper.log
2>&1"
```

The `start_cas_service.sh` script is then referenced in `inittab` with an entry similar to this example:

```
ecas:2345:respawn:/usr/local/endeca/CAS/<version>/bin/start_cas_service.sh
```

When writing your startup script, keep in mind that it is server-specific, and therefore its details (such as paths and user names) depend on the configuration of your server.

Installing only the CAS Console on UNIX

Use the `configure_tools_and_frameworks_integration.sh` script to install only the CAS Console:

- If you installed CAS but did not install the CAS Console as a Workbench extension (using the `--skip_console_installation` flag).
- If you intend to install the CAS Console with more than one instance of Workbench
- If you intend to run the CAS Service on a separate machine from the one that hosts the CAS Console Workbench Extension

To install only the CAS Console on UNIX:

1. Stop CAS Service.
2. Open a command-line and issue the following:

```
/usr/local/endeca/CAS/<version>/bin/configure_tools_and_frameworks_integration.sh
--skip_dt_integration
[--endeca_tools_root full path to the Tools root directory]
[--endeca_tools_conf full path to the Tools /conf directory]
```

This registers the CAS Console with the specified Workbench installation.



Note: You may omit the flags for `--endeca_tools_root` and `--endeca_tools_conf` if these variables are already set as environment variables.

3. Start CAS Service.

Installing CAS silently on UNIX

The silent installer is useful if you want to add the CAS installation to your own install script, or push out the installation on multiple machines.

The silent installer is not interactive.

To run the silent installer:

1. From a command prompt, navigate to the directory where you downloaded the installer.
2. Create a response file with the appropriate properties. For example, `cas_response.properties`, and specify the response value for each prompt on its own line, in the order in which they appear in the interactive installer.

Guided Search property Name	Description
<code>INSTALLDIR</code>	Required. Specifies the absolute path to the directory to install CAS (e.g., <code>C:\Endeca\CAS</code>). Keep in mind that you cannot install the software in a directory with spaces in its name.
<code>CASPORT</code>	Required. The port on which the CAS Service listens. The default is 8500.
<code>CASHUTDOWNPORT</code>	Required. The shutdown port for the CAS Service. The default is 8506.
<code>CASHOST</code>	Required. Specifies the CAS server host.
<code>ENDECA_TOOLS_ROOT</code>	Required. <code>C:\Endeca\ToolsAndFrameworks\11.2.0</code>
<code>ENDECA_TOOLS_CONF</code>	Required. <code>C:\Endeca\ToolsAndFrameworks\11.2.0\server\workspace</code>



Note: If you do not use the default location, and you are installing more than one Guided Search product on the same machine, ensure that you install each product to a separate location.

Example:

```
INSTALLDIR=C:/Endeca/CAS
CASPORT=8500
CASHUTDOWNPORT=8506
ENDECA_TOOLS_ROOT=C:/Endeca/ToolsAndFrameworks/11.2.0
ENDECA_TOOLS_CONF=C:/Endeca/ToolsAndFrameworks/11.2.0/server/workspace
```



Note: You must specify a value for each prompt. Even if you want to use the default, you must specify that value to the installer.

3. Issue a command of the following form:

```
OCcas<version>-linux64.bin -i silent -f cas_response.properties
```

Following installation:

- Start the CAS Service. Navigate to `CAS/<version>/bin` and run the following command:
`cas-service.sh`
- Restart the Oracle Commerce Workbench Service.

Integrating CAS with the Deployment Template on UNIX

Use the `configure_tools_and_frameworks_integration.sh` script to integrate the CAS component into the Deployment Template. This may be necessary if you installed CAS but did not integrate the CAS component into the Deployment Template and then later found that you need the Deployment Template to manage crawling operations.

To integrate CAS with the Deployment Template on UNIX:

1. Stop CAS Service.
2. Open a command-line and issue the following:

```
/usr/local/endecca/CAS/<version>/bin/configure_tools_and_frameworks_integration.sh
--skip_console_integration
[--endecca_tools_root full path to the Tools root directory]
[--endecca_tools_conf full path to the Tools /conf directory]
```



Note: You may omit the flags for `--endecca_tools_root` and `--endecca_tools_conf` if these variables are already set as environment variables.

3. Start CAS Service.

Installing a plug-in into CAS

After receiving a plug-in (a JAR or set of JAR files) from a plug-in developer, a CAS application developer installs the plug-in into CAS.

The Content Acquisition System detects each plug-in and validates the extensions within it by checking the uniqueness of extension IDs and by checking for the presence of an annotation of either `@CasDataSource` or `@CasManipulator` for each extension.

To install a plug-in into CAS:

1. Stop CAS Service.
2. Navigate to `<install path>\CAS\<version>\lib\cas-server-plugins` and create a `plugin-name` subdirectory for each plug-in.
For example: `CAS\<version>\lib\cas-server-plugins\JDBCDataSourceExt`
3. Copy the plug-in JAR or JARs, and any dependent JAR files, to `<install path>CAS\<version>\lib\cas-server-plugins\plugin-name`.
4. Repeat the steps above as necessary for multiple plug-ins.
5. Start CAS Service.

You can confirm that an extension is installed by running the `listModules` task of the CAS Server Command-line Utility and specifying a `moduleType` of either `SOURCE` or `MANIPULATOR`. The task returns the installed modules. For example, this task shows that a custom data source named `Sample Data Source for testing` is installed:

```
C:\Endeca\CAS\<version>\bin>cas-cmd listModules -t SOURCE
Sample Data Source
 *Id: Sample Data Source
 *Type: SOURCE
 *Description: Sample Data Source for testing

File System
 *Id: File System
```

```
*Type: SOURCE
*Description: No description available for File System
*Capabilities:
  *Binary Content Accessible via FileSystem
  *Data Source Filter
  *Has Binary Content
  *Expand Archives
```

Web Crawler installation

The Oracle Commerce Web Crawler is installed as part of the CAS installation.

The Web Crawler startup script is in `<install path>\CAS\<version>\bin` (and in the equivalent path on UNIX). The Web Crawler configuration files are in `<install path>\CAS\conf\web` (and in the equivalent path on UNIX). See the *Oracle Commerce Web Crawler Guide* for details and usage.

Package contents and directory structure

The CAS directory is the root directory for the Content Acquisition System.

After installing the Content Acquisition System, you see this directory structure:

```
CAS
  <version>
    bin
    ...
  workspace
```

The contents of the CAS directory are described here in detail.

Directory	Contents
<code>version\bin</code>	<ul style="list-style-type: none"> • Scripts to start the CAS Service (<code>cas-service.sh</code> for UNIX and <code>cas-service.bat</code> for Windows). • Scripts to run the CAS Server Command-line Utility (<code>cas-cmd.sh</code> for UNIX and <code>cas-cmd.bat</code> for Windows). • Scripts to run the Component Instance Manager Command-line Utility (<code>component-manager-cmd.sh</code> for UNIX and <code>component-manager-cmd.bat</code> for Windows). • Scripts to run the Record Store Command-line Utility (<code>recordstore-cmd.sh</code> and <code>recordstore-cmd.bat</code>). • Scripts to run the Web Crawler (<code>web-crawler.sh</code> for UNIX systems and <code>web-crawler.bat</code> for Windows). • Scripts to shut down the CAS Service (<code>cas-service-shutdown.sh</code> for UNIX and <code>cas-service-shutdown.bat</code> for Windows). • The executable that runs the CAS Service as a Windows service (<code>cas-service-wrapper.exe</code>). • Script to install CAS Console on UNIX (<code>configure_tools_and_frameworks_integration.sh</code>)

Directory	Contents
<code>version\console</code>	<ul style="list-style-type: none"> The <code>casconsole.war</code> file, which is the CAS Console application. The <code>casconsole.xml</code> file, which is the context file for the CAS Console. The <code>casconsole.properties</code> file, which is the configuration file for the CAS Console. <p>These files are copied to the Oracle Commerce Workbench installation and used from that location.</p>
<code>version\doc</code>	<ul style="list-style-type: none"> The reference documentation (Javadoc) for the CAS Server, the Component Instance Manager, the Record Store, the CAS Extension API, the CAS Deployment Template Component, and the Web Crawler. The <i>Licensing Guide</i>.
<code>version\doc\wsdl</code>	The Web Service (WSDL) files for the CAS Server, the Component Instance Manager, and the Record Store.
<code>version\java</code>	The JDK used to run the CAS components (except CAS Console, which runs in the Oracle Commerce Workbench Service).
<code>version\lib</code>	<ul style="list-style-type: none"> Libraries for the CAS command-line utilities including: the CAS Server utility, the Component Instance Manager utility, the Record Store utility. Libraries for the CAS APIs including: the CAS Server API, the Component Instance Manager API, the Record Store API, and the CAS Extension API.
<code>version\lib\cas-dt</code>	CAS WSDL client stub files for the Deployment Template.
<code>version\lib\cas-server-plugins</code>	Libraries for CAS plug-ins including CMS connectors and custom extensions (if applicable).
<code>version\lib\oit-sx</code>	The CAS Document Conversion Module libraries.
<code>version\lib\recordstore-forge-adapter</code>	The library for the Forge Record Store adapter.
<code>version\lib\web-crawler</code>	The Web Crawler libraries.
<code>version\sample</code>	Directories for the sample applications and sample extensions. The sample applications are documented in the <i>CAS Developer's Guide</i> and also in the <i>CAS Extension API Guide</i> .
<code>version\webapps</code>	The <code>root.war</code> file, which is the CAS Server and Component Instance Manager applications.
<code>version\workspace_template</code>	The template for the workspace directory that contains configuration files.
<code>workspace</code>	The working directory for the CAS Server and the Web Crawler.
<code>workspace\conf</code>	<ul style="list-style-type: none"> The <code>commandline.properties</code> file, which contains the CAS Service settings necessary for the CAS command-line utilities to run.

Directory	Contents
	<ul style="list-style-type: none"> • Three logging configuration files (<code>cas-service.log4j.properties</code> for the CAS Service, <code>recordstore-cmd.log4j.properties</code> for the Record Store, and <code>cas-cmd.log4j.properties</code> for the Command-line Utility). • The Jetty configuration files.
<code>workspace\conf\web-crawler\default</code>	The default configuration files for the Web Crawler, including the <code>log4j.properties</code> logging configuration file.
<code>workspace\conf\web-crawler\non-polite-crawl</code>	Sample crawl configuration files for non-polite crawls. As with the polite version, the settings in these files will override the default settings.
<code>workspace\conf\web-crawler\polite-crawl</code>	Sample crawl configuration files for polite crawls. The settings in <code>site.xml</code> will override the same property settings in the <code>default.xml</code> file, while the <code>crawl-urlfilter.txt</code> file will also override the default version of the file.
<code>workspace\logs</code>	The <code>cas-service.log</code> file, which contains the CAS Service log output, and includes log messages from all crawls managed by the CAS Server.
<code>workspace\output</code>	Default destination directory for the crawl output from the Web Crawler. The <code>output</code> directory is not present upon installation. It is created when the Web Crawler writes to output records for a crawl.
<code>workspace\state</code>	State files for the CAS Service components. State files can include Record Store instances, state directories for data source extension information, and state directories for manipulator extension information.



Note: There is no `logs` directory for the Web Crawler, because by default the Web Crawler sends its standard output to the console. However, you can modify the `log4j.properties` file to send the output to a file.

Required installer file permissions

The CAS installer creates directories and files set to the permissions required by Oracle Security standards.

All directories created by the installer are assigned the 550 permission (`r-x r-x ---`). Shell files are also assigned 550 permission, while all other files are assigned 640 permission (`rw- r-- ---`).



Note: Do not alter these directory and file permissions.

Directory	Permission
<code>CAS/<version>/bin</code> (all <code>.sh</code> files)	550
<code>CAS/<version>/java/bin</code>	550
<code>CAS/<version>/sample/cas-extensions/build.xml</code>	640
<code>CAS/<version>/sample/cas-extensions/data</code>	640
<code>CAS/<version>/sample/cas-extensions/data/document-db</code>	640
<code>CAS/<version>/sample/cas-extensions/data/documents</code>	640

Directory	Permission
CAS/<version>/sample/cas-extensions/src/* (all the files but not directories)	640
CAS/<version>/sample/cas-server-java-client/build.xml	640
CAS/<version>/sample/cas-server-java-client/log4j.properties	640
CAS/<version>/sample/cas-server-java-client/src	640
CAS/<version>/sample/custom-web-crawler-plugin/plugin.xml	640
CAS/<version>/sample/custom-web-crawler-plugin/src	640
CAS/<version>/sample/forge-to-recordstore/*.sh	550
CAS/<version>/sample/forge-to-recordstore/forge-config/*	640
CAS/<version>/sample/forge-to-recordstore/forge-input-data/*	640
CAS/<version>/sample/forge-to-recordstore/forge-partial-input-data/*	640
CAS/<version>/sample/forge-to-recordstore/recordstore-configuration.xml	640
CAS/<version>/sample/multiple-recordstore-to-forge/*.sh	550
CAS/<version>/sample/multiple-recordstore-to-forge/forge-config/*	640
CAS/<version>/sample/multiple-recordstore-to-forge/forge-output-data/*	640
CAS/<version>/sample/multiple-recordstore-to-forge/forge-partial-output-data/*	640
CAS/<version>/sample/recordstore-java-client/*.sh	550
CAS/<version>/sample/recordstore-java-client/build.xml	640
CAS/<version>/sample/recordstore-java-client/conf/*	640
CAS/<version>/sample/recordstore-java-client/src/*	640
CAS/<version>/sample/recordstore-to-forge/*.sh	550
CAS/<version>/sample/recordstore-to-forge/forge-config/*	640
CAS/<version>/sample/recordstore-to-forge/forge-output-data/*	640
CAS/<version>/sample/recordstore-to-forge/forge-partial-output-data/*	640
CAS/<version>/sample/webcrawler-to-recordstore/*.sh	550
CAS/<version>/sample/webcrawler-to-recordstore/*.xml	640
CAS/<version>/sample/webcrawler-to-recordstore/conf/*	640

About changing the role used for the CAS Console extension

By default, only users with administrative rights can view the CAS Console extension in Oracle Workbench. However, you can also make this extension visible to non-administrative users.

See the *Oracle Commerce Guided Search Administrator's Guide* for details on changing the visibility of the CAS Console extension for different user roles.



Note: If you change a user role in Workbench, you must manually remove the CAS Console extension if you choose to unregister it.

Related Links

[Uninstalling CAS Console if its extension configuration was changed](#) on page 26

If you need to uninstall the CAS Console extension for Oracle Oracle Commerce Workbench and you have manually edited its extension configuration (for example to assign the extension to a role other than "admin"), you must manually uninstall the CAS Console as an Oracle Commerce Workbench extension.

Uninstalling the Content Acquisition System

This section contains the procedures for uninstalling the Content Acquisition System.

Uninstalling CAS on Windows

Follow these steps to uninstall the Content Acquisition System from your Windows machine.

The installer creates a timestamped backup of your workspace prior to uninstalling the Content Acquisition System

To uninstall the Content Acquisition System from a Windows machine:

1. Stop Endeca Tools Service and Endeca CAS Service.
2. From the Windows Control Panel, select **Uninstall a program**.
3. Select the **Oracle Commerce Content Acquisition System** from the list of installed software.
4. Click **Uninstall**.

Uninstalling CAS Console on UNIX

Follow this procedure to remove the CAS Console from Oracle Commerce Workbench without uninstalling the Content Acquisition System.

To remove the CAS Console from Oracle Commerce Workbench:

1. Open a command-line prompt and run the following command:

```
/endeca/CAS/<version>/bin/configure_tools_and_frameworks_integration.sh  
--uninstall_console --uninstall_dt_integration
```
2. If you do not have `ENDECA_TOOLS_ROOT` and `ENDECA_TOOLS_CONF` set in the environment, do one of the following:
 - Enter the value for each at the prompt.
 - Run the `/endeca/CAS/<version>/bin/configure_tools_and_frameworks_integration.sh --uninstall_console` script with the flags `--endeca_tools_root full path to the Tools root directory --endeca_tools_conf full path to the Tools /conf directory`
3. Restart the Oracle Commerce Workbench Service

Uninstalling CAS on UNIX

Follow these steps to uninstall the Content Acquisition System from a UNIX machine.

Before you begin the uninstall process, back up any files that you want to retain from the CAS directory.

You must uninstall CAS Console for Oracle Commerce Workbench prior to uninstalling the CAS.

To uninstall the CAS from a UNIX machine:

1. Shutdown CAS Service by navigating to `/usr/local/edeca/CAS/<version>/bin` and running the following command: `cas-service-shutdown.sh`.
2. In a command prompt, issue an `rm` command as in this example:

```
rm -rf edeca/CAS/<version>
```

Uninstalling a plug-in from CAS

Follow these steps to uninstall a plug-in from the Content Acquisition System.

Before you begin the uninstall process, back up any plug-in JAR files that you want to retain from the `CAS\version\lib\cas-server-plugins` directory.

To uninstall a plug-in from the CAS:

1. Stop CAS Service.
2. Navigate to `<install path>\CAS\<version>\lib\cas-server-plugins`.
3. Delete the `plugin-name/` subdirectory for each plug-in.
 For example: `<install path>\CAS\<version>\lib\cas-server-plugins\JDBCDataSourceExt`
4. Repeat the steps above as necessary for multiple plug-ins.
5. Start CAS Service.

Uninstalling CAS Console if its extension configuration was changed

If you need to uninstall the CAS Console extension for Oracle Oracle Commerce Workbench and you have manually edited its extension configuration (for example to assign the extension to a role other than "admin"), you must manually uninstall the CAS Console as an Oracle Commerce Workbench extension.

To manually remove the CAS Console extension to Workbench:

1. Navigate to `%ENDECA_TOOLS_CONF%\conf` (on Windows) or `$ENDECA_TOOLS_CONF/conf` (on UNIX).
2. In `ws-extensions.xml`, remove the `<extension>` element with the id "casconsole".
3. In `ws-mainMenu.xml`, remove the `<menuitem>` element with the id "casconsole".
4. Delete `casconsole.properties`.
5. From `%ENDECA_TOOLS_CONF%\conf\Standalone` (on Windows) or `$ENDECA_TOOLS_CONF/conf/Standalone` (on UNIX), delete `casconsole.xml`.
6. From `%ENDECA_TOOLS_ROOT%\server\webapps` (on Windows) or `$ENDECA_TOOLS_ROOT/server/webapps` (on UNIX), delete `casconsole-version.war`.
7. Restart the Oracle Commerce Workbench Service

Related Links

[About changing the role used for the CAS Console extension](#) on page 23

By default, only users with administrative rights can view the CAS Console extension in Oracle Workbench. However, you can also make this extension visible to non-administrative users.

Index

B

- bin directory
 - contents of 19
 - location of 19

C

- CAS Server API
 - location of 20
- CAS Service
 - starting from inittab 16
- cas-server.log4j.properties file 21
- configuration files
 - for logging 21
 - for non-polite crawls 21
 - for polite crawls 21

D

- Deployment Template
 - upgrading to use WSDL stub files 14

I

- installing
 - CAS Console on UNIX 16
 - on UNIX 14
 - on Windows 12
 - silently on UNIX 17
 - silently on Windows 13

O

- overview of package 9

P

- package
 - contents of 19
 - directory structure of 19
 - overview of 9
- prerequisites 9

R

- recordstore-server.log4j.properties file 21

S

- system requirements 9

U

- uninstalling
 - a plug-in 26
 - CAS Console if extension configuration changed 26
 - CAS Console on UNIX 25
 - on UNIX 26
 - on Windows 25

W

- workspace directory
 - location of 20
- WSDL files
 - location of 20
- WSDL stub files
 - using with Deployment Template 14

