Oracle® Endeca Information Discovery

Getting Started Guide

Version 3.2.0 • Revision A • January 2017



Copyright and disclaimer

Copyright © 2003, 2017, 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.

Table of Contents

Copyright and disclaimer	
Preface	4
About this guide	
Conventions used in this document	4
Contacting Oracle Customer Support	5
Chapter 1: Overview	6
Chapter 2: System requirements	
Chapter 3: Installing Oracle Endeca Information Discovery	9
Installation overview	
Downloading the required installation packages	
Setting up the installation directories	
Configuring the orchestration script	
Running the orchestration script	
Verifying your installation	
Connecting Studio to the Provisioning Service	
Additional resources	
Chapter 4: Uninstalling Oracle Endeca Information Discovery	17
Appendix A: Installing Integrator ETL	18
Downloading Integrator ETL	
Installing Integrator ETL	
Appendix B: Installing Oracle Endeca Information Discovery on Linux	20

Preface

Endeca Information Discovery Studio is an industry-leading application composition environment and discovery experience that allows business users to easily upload and mash up multiple diverse data sources, and then quickly configure discovery applications - all within the context of an enterprise framework that maintains existing governance and enterprise definitions.

Studio includes world-class search, guided navigation, and filtering, as well as offering an array of powerful interactive visualizations, for rapid intuitive analysis that requires zero training.

About this guide

This guide provides instructions for installing Oracle Endeca Information Discovery on a single instance of WebLogic Server. It also provides instructions for installing Integrator ETL.

Conventions used in this document

The following conventions are used in this document.

Typographic conventions

The following table describes the typographic conventions used in this document.

Typeface	Meaning
User Interface Elements	This formatting is used for graphical user interface elements such as pages, dialog boxes, buttons, and fields.
Code Sample	This formatting is used for sample code phrases within a paragraph.
Variable	This formatting is used for variable values. For variables within a code sample, the formatting is Variable.
File Path	This formatting is used for file names and paths.

Symbol conventions

The following table describes symbol conventions used in this document.

Symbol	Description	Example	Meaning
>	The right angle bracket, or greater-than sign, indicates menu item selections in a graphic user interface.	File > New > Project	From the File menu, choose New, then from the New submenu, choose Project.

Preface 5

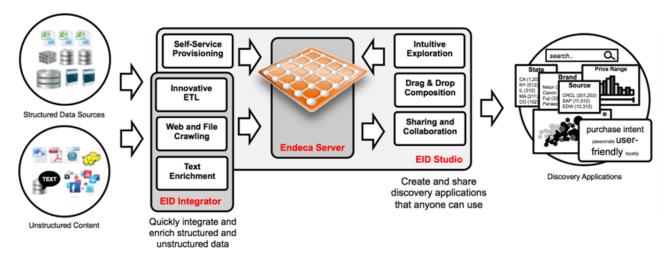
Contacting Oracle Customer Support

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

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

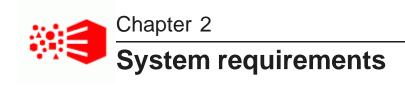


Oracle Endeca Information Discovery is a data discovery platform that guides people to better decisions based on diverse and changing data. It leverages the Endeca Server, a hybrid search-analytical database that provides a centralized platform to rapidly deploy interactive analytic applications and keep pace with changing business requirements while maintaining information governance.



Oracle Endeca Information Discovery consists of the following modules:

- **Studio**, where you can create and configure applications, views, and dashboards. Studio includes the Provisioning Service, which allows you to upload data directly from spreadsheets or JSON files, or import data from configured Oracle BI Server or JDBC connections.
- Oracle Endeca Server, which powers Search, Guided Navigation™, and analytics capabilities.
- Integrator ETL and the Integrator Acquisition System, which provide comprehensive data integration options. Integrator ETL includes connectors for standard structured and unstructured data sources, an extensive data enrichment library, and direct connectors to the Oracle Endeca Server.



Before installing, verify that your machine meets the following requirements.

Requirement	Details		
Hardware	Minimum hardware requirements:		
	x64 processor, minimum 1.8 GHz		
	At least 2 GB of RAM, depending on the size of the application data set		
	80 GB hard drive, depending on the size of the application data set		
	Recommended hardware:		
	x64 3.0+ GHz processors; we recommend Intel Xeon (including Nehalem) or AMD Opteron processors		
	8 GB of RAM or more, depending on the size of the application data set		
	 High performance network-attached storage (for example, attached via a dedicated iSCSI or fibre channel network) or high performance locally-attached RAID storage (for example, a RAID 6 or RAID 0+1 array with battery-backed write caching, operating on 72GB or 146 GB spindles at 10k or 15k RPM spindle speed) 		
	Gigabit Ethernet		
	For improved performance, Oracle recommends turning off hyper-threading.		
	Note: In this guide, the term "x64" refers to any processor compatible with the AMD64/EM64T architecture. You might need to upgrade your hardware, depending on the data you are processing.		
	All run-time code must fit entirely in RAM. Likewise, hard disk capacity must be sufficient based on the size of your data set.		
Operating system	Windows Server 2012 R2 Enterprise		
	Windows 7		
	Note that Windows 7 isn't supported for production environments.		

System requirements 8

Requirement	Details	
Disk space	Oracle EID requires approximately 1.6GB of disk space. At the component level, this breaks down to:	
	WebLogic Server: 725 MB	
	Endeca Server: 800 MB	
	Additionally, the temporary directory used by each component's installer must contain about 2.5 times the amount of space the component requires after installation. By default, this directory is defined by the TMP system variable.	
JDK	Oracle EID requires the Java 7 or 8 JDK, which you can download from http://www.oracle.com/technetwork/java/javase/downloads/index.html .	
	When installing the JDK, change its location to C:\Java\jdk1.X.X_XX\. You can't use the default location because its filepath includes a space.	
Browser (for Studio)	Firefox ESR (recommended)	
	Internet Explorer 11	
	Safari Mobile 9.x (iPad only)	
	Note: You can only use the iPad to log in to Studio and view applications; you can't create or configure them on the iPad.	
User privileges	You must have administrator privileges on your machine to install EID.	



These installation instructions create a development instance of Oracle Endeca Information Discovery (Endeca Server, Studio, and the Provisioning Service) on one machine. In this instance, all components are installed into a single WebLogic Server.

Installation overview

Downloading the required installation packages

Setting up the installation directories

Configuring the orchestration script

Running the orchestration script

Verifying your installation

Connecting Studio to the Provisioning Service

Additional resources

Installation overview

If you want to demo Oracle Endeca Information Discovery without committing to a full-scale installation, you can install it on a single machine with default configuration suitable for a development environment.



Note: Single-machine installations can only be used in development environments for development, testing or training.

This installs the following components:

- WebLogic Server 12.1.3
- Endeca Server 7.7.0
- Studio 3.2.0
- Provisioning Service 3.2.0

Downloading the required installation packages

Before you can install Oracle EID, you must download the installation packages for each of the modules from the Oracle Software Delivery Cloud. The Endeca Server download page includes links to the required WebLogic Server installation package.

To download the installation packages for Oracle Endeca Information Discovery and WebLogic Server:

- 1. Go to the Oracle Software Delivery Cloud and sign in.
- 2. Accept the Export Restrictions.
- 3. Check **Programs** if it isn't already.
- 4. In the **Product** text box, type Oracle Endeca Server.
- 5. Click Select Platform, check the appropriate platform for your environment, and click Select.
 - Oracle Endeca Server displays in the Selected Products table.
- 6. Repeat steps 4 and 5 for Oracle Endeca Information Discovery Studio.
 - Oracle Endeca Information Discovery Studio is added to the Selected Products table.
- 7. Click Continue.
- 8. Verify that Available Release and both products are checked, then click Continue.
- 9. Review and accept the Oracle Standard Terms and Restrictions and click Continue.
 - The list of available media packs displays.
- 10. In the File Download popup, click the download links for the following packages:
 - Oracle Endeca Server (7.7.0) for Microsoft Windows x64 (64-bit)
 - Oracle WebLogic Server 12c (12.1.3) Generic and Coherence
 - Oracle Endeca Information Discovery Studio (3.2) Install Scripts for Microsoft Windows x64 (64-bit)

You should also make a note of each package's part number: you will need this information to identify it.

The installation packages are downloaded to your machine.

Setting up the installation directories

After you download the installation packages, you must set up the installation directories.

This procedure assumes you've downloaded the following packages:

- Oracle Endeca Server (7.7.0) for Microsoft Windows x64 (64-bit)
- Oracle WebLogic Server 12.1.3 Generic installer
- Oracle Endeca Information Discovery Studio (3.2) Install Scripts for Microsoft Windows x64 (64-bit)

To set up the installation directories:

Unzip the install scripts package to a convenient location.

This creates a file called EID_3.2_Studio_InstallWin.zip.

2. Unzip EID_3.2_Studio_InstallWin.zip to a convenient location.

Unzipping this file creates the eidorch directory, which contains the following subdirectories:

Directory	Description	
installers	Contains the installers for Studio and the Provisioning Service.	
orchScripts	Contains all of the orchestration scripts and related files for Studio and the Provisioning Service.	

3. Unzip the Endeca Server installation package to a convenient location.

This creates two zip files:

- endecaserver.zip
- windows_OC.zip
- 4. Copy the file endecaserver.zip from the location where you unzipped the Endeca Server package to the directory eidOrch\installers.

Don't unzip this file.

5. Copy windows_OC.zip from the location where you unzipped the Endeca Server package to the directory eidOrch\orchScripts and unzip it.

This creates a directory called windows in eidOrch\orchScripts.

- 6. Rename the windows directory to ES_windows_OC.
- 7. Unzip the WebLogic Server installation package and copy fmw_12.1.3.0.0_wls.jar (WebLogic 12.1.3 Generic installer) to the directory eidOrch\installers.

Configuring the orchestration script

Before you run the orchestration script, you must update its configuration file with information specific to your installation.

The file is named eidOrch\orchScripts\config_EID_win.prop. You can edit it in any text editor; for example, Notepad.

If you want, you can install with most of the default settings defined in the file; however, you must set USE_SSL to FALSE, as SSL is only supported in production environments.

The EID orchestration script includes the following configuration parameters.

Configuration property	Description
START_MODE	Configures the start mode of the WebLogic Server. Set this to DEV.
USE_SSL	Determines whether the EID components will use SSL for secure communication. Set this to FALSE.

Configuration property	Description	
JAVA_HOME	The absolute path to the installed JDK. This should be C:\Java\jdk1.X.X_XX\.	
	Note that this path must not include spaces.	
ORACLE_HOME	The absolute path to the Oracle Middleware directory where WebLogic Server and EID will be installed.	
	This directory must not exist, or the installation will fail.	
ORACLE_INV_PTR	Linux only.	
	The location of the Oracle Installer Pointer file.	
INSTALLER_LOCATION	Absolute path to the installation directory. Set this to the location you extracted the installer files in.	
STOP_EID_PROCESSES	Linux only.	
	Determines whether all EID processes will be stopped if the installation fails, or the orchestration script is rerun after a successful installation.	
ENDECA_SERVER_DOMAIN_NAME	The name of the Endeca Server domain.	
ENDECA_SERVER_PORT	The HTTP port Endeca Server listens on for requests when using non-secure communication (not using SSL).	
ENDECA_SERVER_SECURE_PORT	The HTTP port Endeca Server listens on for requests when using secure communication (using SSL).	
ENDECA_STUDIO_DOMAIN_NAME	The name of the Studio domain.	
ENDECA_STUDIO_PORT	The HTTP port Studio listens on for requests when using non-secure communication (not using SSL).	
ENDECA_STUDIO_SECURE_PORT	The HTTP port Studio listens on for requests when using secure communication (using SSL).	
ENDECA_PS_DOMAIN_NAME	The name of the Provisioning Service domain.	
ENDECA_PS_PORT	The HTTP port the Provisioning Service listens on for requests when using non-secure communication (not using SSL).	
ENDECA_PS_SECURE_PORT	The HTTP port the Provisioning Service listens on for requests when using secure communication (using SSL).	

Running the orchestration script

After you configure the orchestration script, you can run it to install EID.

Before running the orchestration script, be sure you have downloaded all required packages and have set up the orchestration script directories. For details, see Setting up the installation directories on page 10.

To run the orchestration script:

- 1. Open a command prompt in administrator mode and go to eidOrch\orchScripts.
- 2. Run the orchestration script:

```
run_EID_install.bat config_EID_windows.prop [ --temp-directory <tempFolderName> ]
```

The temp-directory parameter is optional. If specified, the directory must exist.

3. When prompted, enter the following data:

Module	Data to Enter	
Endeca Server	 WebLogic Server Admin Username WebLogic Server Admin Password (and confirmation) Endeca Server Admin Username Endeca Server Admin Password (and confirmation) 	
Endeca Information Discovery Studio	 WebLogic Server Admin Username WebLogic Server Admin Password (and confirmation) 	
Provisioning Service	 WebLogic Server Admin Username WebLogic Server Admin Password (and confirmation) Endeca Server hostname 	

WebLogic Server is installed to the location you specified in your configuration. WebLogic domains are created for Endeca Server, Studio, and the Provisioning Service. All domains are started and Studio is deployed to its domain.

Next, you must verify your installation. See Verifying your installation on page 13.

Verifying your installation

After running the orchestration script, you should verify that all components were installed successfully.

To verify your installation:

- 1. To verify that Endeca Server and the Provisioning Service have been installed successfully, start a browser, then in the address bar, enter http://<hostname>:<port>/eid-ps/status, where:
 - <hostname> is the name or IP address of the machine where you installed EID

<port> is the Provisioning Service port

If the Provisioning Service and Endeca Server are functioning correctly, a page similar to the following is returned:

Oracle Endeca Information Discovery Provisioning Service 3.2.0.0 Revision 2689

Endeca Server Communication SUCCESS

Backend Database Access SUCCESS

If either check fails, the message returned is Error: check logs.

- 2. To verify that Studio has deployed successfully, start a browser, then in the address bar enter http://<hostname>:<port>/eid, where:
 - <hostname> is the name or IP address of the machine where you installed EID
 - <port> is the Studio port

The Studio login page should open.

Connecting Studio to the Provisioning Service

You need to configure the connection between Studio and the Provisioning Service so that you can create Studio applications via file upload.

Before you configure the Provisioning Service connection, make sure that Endeca Server is running.

To connect Studio to the Provisioning Service:

- Log in to Studio using the default credentials:
 - Login: admin@oracle.com
 - Password: Welcome123
- When prompted, change the Studio password.

This must be at least 6 characters long and contain at least 1 alpha-numeric character.

3. From the Studio menu, select Control Panel.



- 4. In the Control Panel menu, under Information Discovery, click Provisioning Service.
- 5. On the **Provisioning Service** page, make the following changes to the connection string:
 - Change the value of port to the port used for installation. For a non-SSL installation, the default is 8201.
 - Change the value of server to be the name of the machine you installed EID on.
 - Remove the sslConfig section.

When you finish, the connection string should look something like:

```
{
  "port": "8201",
  "server": "myhostname"
}
```

6. Click Save.

Additional resources

This guide walked you thought a basic installation scenario, but doesn't account for differences you might encounter in a real staging environment, such as alternative platforms or specific use cases for loading data. If you need more information on these or any other topics, you can refer to the following resources.

Documentation set

You can access the full EID documentation set on OTN (Oracle Technology Network).

Depending on your role in the deployment process, you should consult the following guides to find more specific in-depth information:

Topic	Related guide(s)	
EID terminology	Oracle Endeca Information Discovery Glossary	
Installing EID and its components	 Oracle Endeca Server Installation Guide Integrator ETL Installation Guide Studio Installation Guide Integrator Acquisition System Installation Guide IKM SQL to Endeca Server Installation and Usage Guide 	
Administering EID	 Oracle Endeca Server Administrator's Guide Studio Administration and Customization Guide Provisioning Service Administration Guide 	
EID security	 Oracle Endeca Server Security Guide Security Guide for Integrator Studio Security Guide Important: While security is not required by this sample installation, in real-world scenarios, Oracle recommends that you install and run Oracle Endeca Information Discovery securely. 	
Loading data into EID	Integrator ETL User's Guide Oracle Endeca Server Data Loading Guide	

Topic	Related guide(s)	
Endeca Server development	Oracle Endeca Server Developer's Guide	
Managing data sources for applications	Studio Administration and Customization Guide	
Creating and configuring Studio applications	Studio User's Guide	
Endeca Server APIs	 Oracle Endeca Server Developer's Guide API Reference (this is located in EndecaServer EndecaServer 	
Extending Studio	 Studio Administration and Customization Guide The Studio Javadoc, which is packaged with the rest of the Studio documentation 	

OTN product page

You can find other information useful for getting up to speed, including videos and best practices, on the OTN page for Oracle Endeca Information Discovery:

http://www.oracle.com/technetwork/middleware/endeca/overview

User forums

The Oracle forums include an *Information Discovery forum*, where you can post questions and view information from other users and experts.

Education and training

Oracle offers training courses on configuring and using Oracle Endeca Information Discovery. For information on available courses, see the *Oracle University site*.



After you have finished exploring the Oracle Endeca Information Discovery suite of modules, you may want to uninstall the product from your local machine.

For information about uninstalling the Oracle Endeca components, see the corresponding following installation guides:

- · Oracle Endeca Server Installation Guide
- · Oracle Endeca Information Discovery Integrator ETL Installation Guide
- Oracle Endeca Information Discovery Studio Installation Guide



Integrator ETL is used with enterprise systems that need to ingest and combine data from a variety of sources. In enterprise environments, you may prefer to use Integrator ETL to extract, transform, and load data for Oracle Endeca Information Discovery.

For these types of deployments, the data is managed centrally by a data architect. The data architect creates the Endeca Server domains, then uses Studio to create available Endeca Server connections for users to create applications from.

Downloading Integrator ETL
Installing Integrator ETL

Downloading Integrator ETL

Integrator ETL is downloaded from the Oracle Software Delivery Cloud.

To download Integrator ETL:

- 1. Go to the Oracle Software Delivery Cloud and sign in.
- 2. Accept the Export Restrictions.
- 3. Check **Programs** if it isn't already.
- 4. In the **Product** text box, type Oracle Endeca Information Discovery Integrator.
- Click Select Platform, check the appropriate platform for your environment, and click Select.
 Oracle Endeca Information Discovery Integrator displays in the Selected Products table.
- 6. Click Continue.
- 7. Verify that **Available Release** and **Oracle Endeca Information Discovery Integrator 3.2.0...** are checked, then click **Continue**.
- 8. Review and accept the Oracle Standard Terms and Restrictions and click **Continue**.
 - The list of available media packs displays.
- 9. In the File Download popup, click the download link for the Oracle Endeca Information Discovery Integrator ETL (3.2.0) for Microsoft Windows x64 (64-bit) package.
 - You should also make a note of the package's part number: you will need this information to identify it. The Integrator ETL installation package is downloaded to your machine.

Installing Integrator ETL 19

Installing Integrator ETL

You install Integrator ETL by running a script provided in its installation package, which you downloaded with the other Oracle EID packages.

Integrator ETL is installed into an Eclipse IDE, which you must download as part of the installation procedure. You must also download the Eclipse Remote System Explorer (RSE) plugin. When the installer runs, it automatically installs both Eclipse and the RSE plugin along with Integrator ETL.

To install Integrator ETL:

- Download the Eclipse IDE for Java Developers version Indigo (3.7) from eclipse.org.
 - Be sure to download the correct version. DON'T download the Eclipse IDE for Java EE Developers. Don't extract the file, as the Integrator ETL installer does this for you.
- Download the latest version of the RSE plugin from the Target Management Project Downloads page.
 Don't extract the file, as the installer does this for you.
- Go to the location you downloaded the Integrator ETL installation package to and extract it to a convenient location.
 - This adds a file called EID_3.2.0_IntegratorETL.zip to the directory.
- 4. Extract EID_3.2.0_IntegratorETL.zip to a convenient location.
 - This creates a directory called EID_3.2.0_IntegratorETL.
- 5. Go to EID_3.2.0_IntegratorETL and run:

install.bat

- 6. Select the installation directory when prompted.
 - You can either accept the default (C:\Oracle\Endeca\Discovery\3.2.0\IntegratorETL), or choose a new one. If you specify a directory that doesn't exist, the installer will create it automatically.
 - Oracle recommends that you specify a directory that doesn't exist, as the installer will overwrite an existing one.
- 7. Enter the absolute paths to the Eclipse and RSE .zip files when prompted.

When the installer finishes, Integrator ETL is installed in the specified location.

Appendix B Installing Oracle Endeca Information Discovery on Linux

An orchestration script is provided to install Oracle Endeca Information Discovery on Linux operating systems. Installing Oracle Endeca Information Discovery on Linux differs in the following ways from the documented procedure:

- Use the appropriate procedures for your Linux operating system to install the JDK 7 or 8.
- Download the Linux versions of the Oracle Endeca Information Discovery Packages.



Note: The recommended installation package for WebLogic Server is compatible with both Windows and Linux.

- The Endeca Server installation package includes a directory named linux. When setting up the orchestration script directories, rename this directory to ES_linux_OC.
- · To run the orchestration script, use the command

sh run_EID_install.sh config_EID_window.prop [--temp-directory <tempFolderName>]' should change into 'sh run_EID_install.sh config_EID_linux.prop [--temp-directory <tempFolderName>]

Index

E	Endeca Server additional resources 15 downloading 10	Р	system requirements 7 uninstalling 17 orchestration script, configuration properties 11
I			Provisioning Service, connecting from Studio 14
0	Integrator ETL additional resources 15 downloading 18 installing 19	S	Studio additional resources 15 downloading 10 Provisioning Service connection,
	Oracle Endeca Information Discovery about 6 additional resources 15 downloading 10 installation, verifying 13 orchestration script, running 13 orchestration script, setting up directories 10	W	configuring 14 WebLogic Server, downloading 10