Oracle® Endeca Information Discovery Integrator

Integrator ETL Migration Guide

Version 3.2.0 • January 2016

ORACLE®

Copyright and disclaimer

Copyright © 2003, 2016, 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

Preface

Oracle® Endeca Information Discovery Integrator is a powerful visual data integration environment that includes:

The Integrator Acquisition System (IAS) for gathering content from delimited files, file systems, JDBC databases, and Web sites.

Integrator ETL, an out-of-the-box ETL purpose-built for incorporating data from a wide array of sources, including Oracle BI Server.

In addition, Oracle Endeca Web Acquisition Toolkit is a Web-based graphical ETL tool, sold as an add-on module. Text Enrichment and Text Enrichment with Sentiment Analysis are also sold as add-on modules. Connectivity to data is also available through Oracle Data Integrator (ODI).

About this guide

This guide helps you upgrade your Oracle Endeca Information Discovery Integrator ETL implementation by describing the major changes between versions 3.1.x and 3.2.x.

Who should use this guide

This guide is intended for system administrators and developers who are upgrading Oracle Endeca Information Discovery Integrator ETL on Windows or Linux.

Conventions used in this guide

The following conventions are used in this document.

Typographic conventions

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

Typographic conventions

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 Name>`	This formatting is used for variable values, such as <install path>.
`File Path`	This formatting is used for file names and paths.

Symbol conventions

The following table describes symbol conventions used in this document.

Symbol conventions

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.

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*.

Chapter 1

Upgrading to Oracle Endeca Information Discovery Integrator ETL 3.2.x

This section provides instructions for upgrading to Oracle Endeca Information Discovery Integrator ETL 3.2.x. Before you start the upgrade processes, check the remaining sections of this guide to learn about the changes that will affect you during or after an upgrade.

Required reading

Upgrading Integrator ETL

Required reading

In addition to reading this document, Oracle recommends that you read the following documents for important information about the release.

Release Announcement

The Release Announcement outlines the new features that were added in version 3.2.x.

Release Notes

The Release Notes for Integrator ETL provide information about known issues and bug fixes for this release.

Installation Guide

The *Oracle Endeca Information Discovery Integrator ETL Installation Guide* provides an overview of Integrator ETL and includes information about configuration scenarios. It contains installation instructions for Integrator ETL, and information on how to verify your installation.

Upgrading Integrator ETL

To upgrade Integrator ETL, uninstall the existing version and install the new one. You can still use your existing workspace after you install Integrator ETL version 3.2.x, but the 3.1.x graphs must be converted to the 3.2.x format.

If you have custom plugins deployed to Integrator ETL 3.2.x, you should back up the plugins before beginning the Integrator ETL upgrade. You can reload them after the upgrade.

To upgrade Integrator ETL:

1. Uninstall version 3.1.x. For information, see the *Oracle Endeca Information Discovery Integrator Installation Guide*, version 3.1.x.

    After you uninstall, the workspace you used for creating your Integrator ETL graphs should remain in place on your machine.

2. Install Integrator ETL version 3.2.x. For information, see the *Oracle Endeca Information Discovery Integrator ETL Installation Guide*, version 3.2.x.

3. Start Integrator ETL and point it to use the workspace you used for your version 3.1.x graphs.

Now you have upgraded to the new version of Integrator ETL.

Because the graph format has changed from Integrator 3.1.x to Integrator ETL 3.2.x, you need to convert your Integrator 3.1.x graphs. For details, see *Converting 3.1.x graphs to the 3.2.x format on page 8*.

If you are reloading a custom plugin, keep in mind that the installation path has changed. For example, on Windows, the 3.1.x default installation path has changed from:

```
C:\Oracle\Endeca\Discovery\3.1.0\IntegratorETL
```

to this 3.2.x default path:

```
C:\Oracle\Endeca\Discovery\3.2.0\IntegratorETL
```

After the upgrade, Integrator ETL still refers to the `customcomponents.xml` at the previous location. Therefore, you must update the reference to the new 3.2.x location.

Chapter 2

Required Changes in Integrator ETL

The following changes are required for Integrator ETL.

Platform support

Converting 3.1.x graphs to the 3.2.x format

Changes to OBI Server graphs

Requirement for trans folder

Platform support

This release adds support for the following platforms:

- Windows Server 2012 R2
- Java 7 and Java 8
- Oracle WebLogic Server 12c (12.1.3)

This release removes support for the following platforms:

- Oracle Linux Release 5 for x64
- Red Hat Enterprise Linux Server version 5 for x64
- Red Hat Enterprise Linux Advanced Platform version 5 for x64
- Windows Server 2008 R2 Enterprise
- Oracle WebLogic Server 11gR1 (10.3.6)

Converting 3.1.x graphs to the 3.2.x format

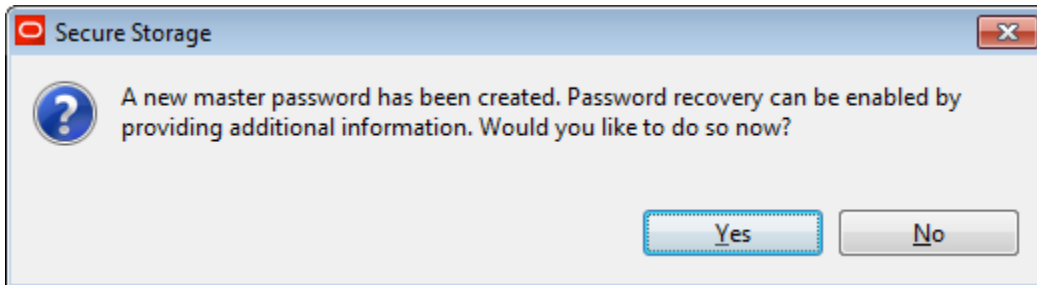Before using your 3.1.x graphs, they must be converted to the Integrator ETL 3.2.x format.

The formats of the graphs and of the graph parameters have changed from Integrator 3.1.x to Integrator ETL 3.2.x. Before you can re-use your 3.1 graphs, they must be fully converted to the 3.2 formats. Note that you have the option of re-implementing your graphs rather than attempting to migrate them.

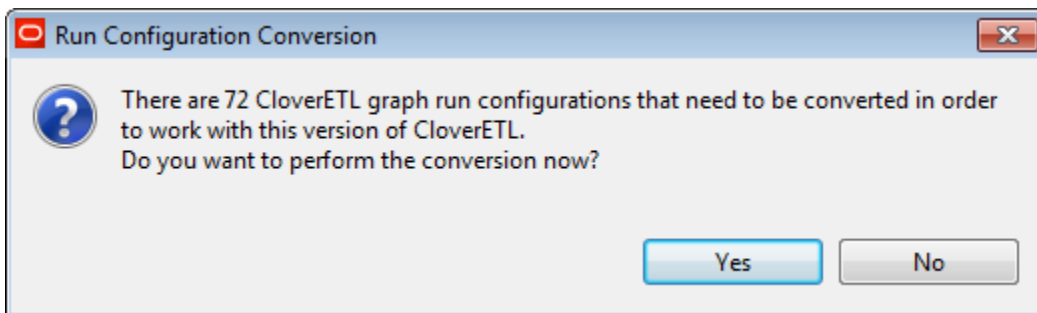To convert Integrator ETL 3.1 graphs:

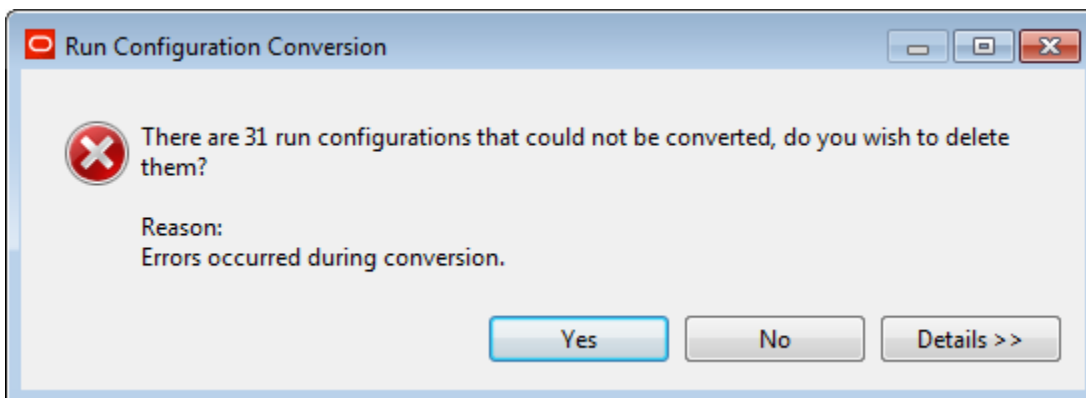1. Install Integrator ETL version 3.2.

2. Start Integrator ETL 3.2 on a workspace that has 3.1 projects. You should see the following "Secure Storage" message. You can click No and set up a master password later, as described in the *Oracle Endeca Information Discovery Integrator ETL User's Guide*.

3. You should also see a "Run Configuration Conversion Version" message, similar to this example:
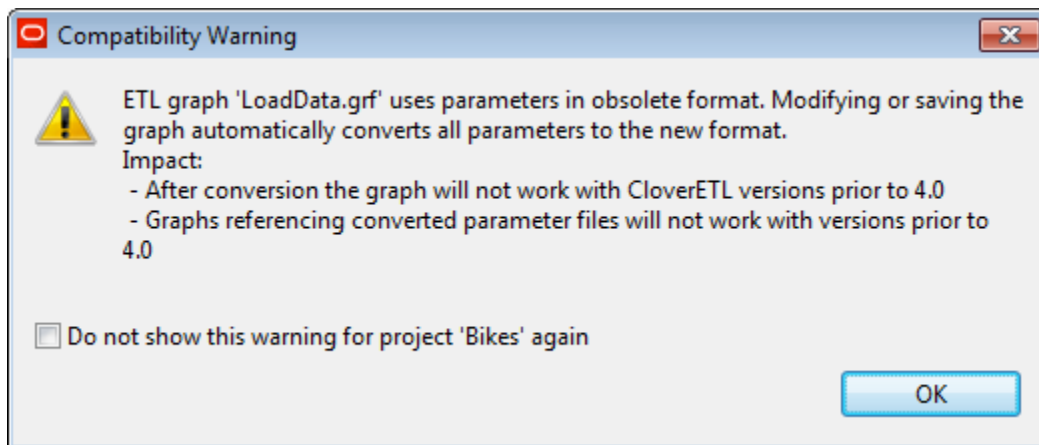

4. Click **Yes** to convert the graphs.

5. If one or more graphs could not be converted, a message similar to this example is displayed:

You can:
- Click **Details** to see which graphs could not be converted. After seeing the details, you can use the **Yes** or **No** buttons.
- Click **Yes** to delete the unconverted graphs.
- Click **No** to keep the unconverted graphs.

6. If you then open a graph that uses obsolete-format parameters, you will see a "Compatibility Warning" message similar to the following example. In this case, saving the graph will convert the parameters.

Changes to OBI Server graphs

You should re-implement your OBI Server graphs.

The graphs used to load data from Oracle Business Intelligence Server should be recreated using the Load Data from OBI Server wizard rather than trying to migrate existing graphs.

In particular, the implementation of encrypted passwords (for OBI Server connections) now uses the Secure Graph Parameters feature, which requires you to set up a Master Password. For details, see the "Loading Data from Oracle Business Intelligence Server" topic in the *Integrator ETL User's Guide*.

Requirement for trans folder

Integrator ETL requires the `trans` folder to be present in the project folder.

Be sure to check that the `trans` folder is present in the project folder, because otherwise the graph may fail.

Chapter 3
Behavioral changes in Integrator ETL

The changes listed here may not require any action on your part, but will affect how your application behaves after you upgrade.

Secure Graph Parameters feature

Social Media property

Text Enrichment component now validates field type

Secure Graph Parameters feature

The Secure Graph Parameters feature can secure graph parameters.

The feature lets you create a Master Password in Integrator ETL, which can then be used to encrypt graph parameters, such as a parameter that stores the password for a database connection. Although use of this feature is optional, Oracle recommends its use in order to provide security for graph parameters that may reference sensitive data.

For details, see the *Integrator ETL User's Guide* and the *Integrator ETL Designer Guide*.

Social Media property

The Social Media data library is no longer packaged separately.

In the EID 3.1.x release, the data library called `TE_SocialData.zip` was packaged as separate library like the language libraries. In EID 3.2.0, the text enrichment features provided by `TE_SocialData.zip` are now part of the Salience engine itself, and the separate library for `TE_SocialData.zip` has been removed.

To process Social Media data, set the new `te.short-content.enabled` property to `true` in the Text Enrichment properties file.

Text Enrichment component now validates field type

The Text Enrichment component has added validation for field types. This will validate the field type and will report an error if a wrong field type is detected.