

**Oracle® Enterprise Data Quality**

Installation Notes

Version 9.0

January 2013

**ORACLE®**

**Copyright © 2006, 2013, Oracle and/or its affiliates. All rights reserved.**

Oracle ® Enterprise Data Quality, version 9.0

Copyright © 2006, 2013, 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 RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation shall be subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License (December 2007). Oracle America, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

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. 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

---

<b>Table of Contents</b> .....	<b>3</b>
<b>1 Introduction</b> .....	<b>3</b>
<b>2 Requirements</b> .....	<b>4</b>
2.1 Minimum hardware requirements.....	4
2.2 Virtual hardware.....	5
2.3 VPN clients, anti-virus and anti-spyware software.....	5
2.4 Enabling features.....	5
<b>3 Microsoft Windows installation</b> .....	<b>5</b>
3.1 The repository database.....	6
3.2 Changing the password for the default user.....	6
<b>4 Getting started</b> .....	<b>8</b>
4.1 The OEDQ Launchpad.....	8
4.2 The Director client.....	10
4.3 Dashboard.....	11
4.4 The Match Review application.....	11
4.5 OEDQ documentation.....	12
<b>5 Upgrading OEDQ</b> .....	<b>12</b>
5.1 Enabling secure external tasks.....	13
5.2 Upgrading OEDQ on Tomcat.....	13
<b>6 Troubleshooting OEDQ</b> .....	<b>13</b>
6.1 Director Client works but unable to access online help or OEDQ Launchpad.....	13

## 1 Introduction

This document is a guide to installing OEDQ onto a standard platform, using the supplied installation program. It covers:

- System and licensing requirements;
- The installation process;
- Getting started with OEDQ;

- A guide to the OEDQ documentation, and
- A guide to troubleshooting OEDQ.

The OEDQ installation program will install PostgreSQL, Tomcat and OEDQ onto a server running an appropriate Windows operating system. Together, Windows, PostgreSQL and Tomcat form the standard OEDQ platform. If you wish to use an alternative operating system, application server or RDBMS, please refer to **Advanced Installation Notes.pdf**, supplied with the distribution.

## 2 Requirements

OEDQ is a Java Web Application using a Java Servlet Engine, a Java Web Start GUI and a SQL RDBMS system for data storage.

It is certified to work on various platforms that support Java version 1.6 or later. For full information on the supported platforms for OEDQ, please check the Oracle Technology Network (OTN) website.

The OEDQ installation process includes the installation of Java 1.6, Apache Tomcat 6 and Postgres 8.2.

If required, OEDQ can be simply configured to use a separately installed Oracle RDBMS rather than Postgres.

### 2.1 Minimum hardware requirements

Depending on the tasks that OEDQ is required to perform, it can place heavy demands on the hardware used to run it.

For installation on a standalone desktop or laptop, a recommended **minimum** hardware configuration to begin using OEDQ is:

- 2 GB of memory, although 4 GB is preferred.
- 2 GHz processor
- 50 GB of hard disk space

For working with large datasets, complex processes or multiple users a recommended **minimum** configuration is:

- 8 GB of memory
- 4-processor architecture
- 250 GB of available hard disk space. A reasonable rule of thumb is that OEDQ will need disk space amounting to 10 times the size of the data it is working with.

**Note:** The above notes do not represent sizing advice for any specific deployment, where it may be appropriate to deploy considerably larger machines, or many machines, depending on

the processing needs placed on the OEDQ system. Please contact your account representative if you require detailed advice about the optimal hardware set and deployment options.

## **2.2 Virtual hardware**

It is possible to install OEDQ on virtualized machines, using a virtualization tool such as VMware. However both the virtual machine and the physical machine it is deployed on must fulfil the minimum hardware requirements.

If load balancing software is used to deploy multiple virtual machines onto a single physical server, care must be taken to ensure that the load balancing software is carefully tuned. OEDQ will generally impose a load similar to an ETL tool or data warehousing software. Between batches, very little load is imposed on the system. When processing a batch of work, OEDQ will rapidly drive hardware to be CPU or IO bound. Unless the virtualized load balancing is correctly configured it is likely to throttle this behavior, resulting in suboptimal performance. Your virtualization software documentation and experts should be consulted to ensure that any load balancing software is correctly configured.

## **2.3 VPN clients, anti-virus and anti-spyware software**

The OEDQ installer configures a number of Windows services that use network interfaces. It has been noted that a number of VPN clients, anti-virus and anti-spyware applications can cause problems with registering these services. If you have a VPN client, anti-virus or anti-spyware product installed, it should be disabled during the installation process. In addition, any services associated with the VPN client should be stopped during the installation process.

## **2.4 Enabling features**

From the Launchpad it is possible to enable and disable OEDQ functionality according to your license agreement with Oracle.

Oracle allows its customers to evaluate any features on a trial basis, but reserves the right to audit any user to ensure that production systems are in compliance with their license agreement.

The feature configuration specifies which sets of processors are activated in an installation, whether or not real-time processing is enabled, and whether or not the Dashboard application is enabled.

If a family of processors is not selected, processors in this family will not be available for you to use. If real-time processing is not enabled, it will not be possible to use OEDQ's Web Service or JMS interfaces. If Dashboard is not enabled, it will not be possible to publish data quality metrics to the Dashboard.

# **3 Microsoft Windows installation**

The supplied installer is for use on the following platforms:

- Windows XP,
- Windows 2003 Server,
- Windows Vista,
- Windows 2008 Server, and
- Windows 7

To install OEDQ you will need local administrator privileges.

OEDQ consists of a server component and a client application. This installation will install both components onto the target machine, which will become an OEDQ server. Instructions for installing just the OEDQ Director client on a machine are contained in the [Director client](#) section.

To install the OEDQ server and client applications on Windows Vista, Windows 2008 Server or Windows 7 (which use Microsoft's User Account Control (UAC) functionality):

- Right click on **dnDirectorSetup.exe** and select 'Run as Administrator'.

To install OEDQ server and client applications on earlier versions of Windows:

- Execute **dnDirectorSetup.exe**

### Non-Windows installations

The supplied installer will only run on the versions of Windows listed above. The artifacts required to install OEDQ on a non-Windows platform, and/or onto an existing application server, are supplied in the zip file.

## 3.1 The repository database

This installation will install two instances of PostgreSQL on the OEDQ server. These instances are referred to as the Director Schema and the Results Schema. The database engines for each instance are optimized for their specific use. The JDBC connection strings are exposed in the file **director.properties** which is located in the **config** directory beneath the OEDQ installation directory. At the time of writing, it is possible to use either PostgreSQL or Oracle as the RDBMS. Should you wish to switch to using Oracle, use the 'Configuration' utility, launched from the 'OEDQ Launchpad' (see below).

## 3.2 Changing the password for the default user

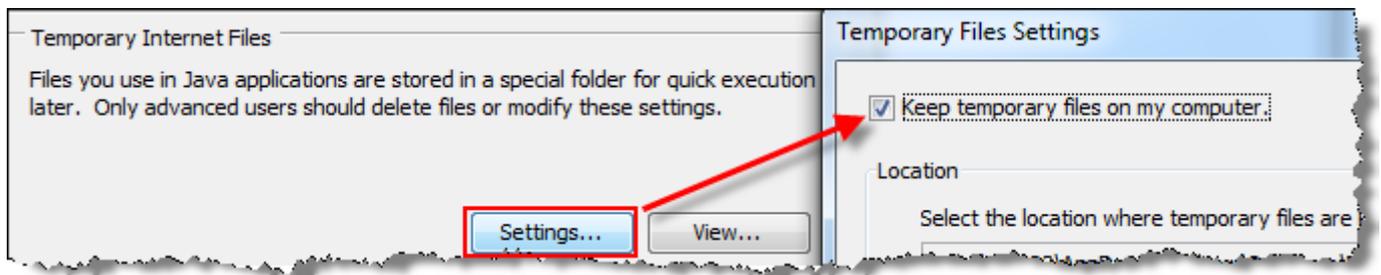
Once the OEDQ installation is complete, you must change the password of the default dnadmin user to ensure the security of the system.

You will be prompted to do this when logging into any client application for the first time, using the user name `dnadmin` and the initial password `dnadmin`.

Please ensure that you do not forget or lose the new password as without it you may lose access to the OEDQ system. You may wish to add further users with rights to add user accounts to ensure that this does not happen.

The OEDQ client applications can now be launched using the default user name `dnadmin` and the password you have specified.

**Note:** Java caching must be enabled in order to start the OEDQ Java WebStart applications on client machines. This setting is controlled in the Java Control Panel as follows:



## 4 Getting started

The following sections guide you through the steps needed to become familiar with the main components of OEDQ.

### 4.1 The OEDQ Launchpad

To open the OEDQ Launchpad, point your web browser at:

`http://<server name>:<port number>/dndirector`

where *<server name>* is the name of the server onto which you installed OEDQ and *<port number>* is the HTTP or HTTPS port that your application server is running against. For installations using the Windows Installer, the ports will normally be 9002 for HTTP and 9004 for HTTPS, unless these ports were already in use at the time of installation.

For custom installations (not using the Windows installer), the default port number for most installations is 8080, but be aware that other application servers may vary. Also, if you have deployed the application server to run against a different port, you should use your port number here.

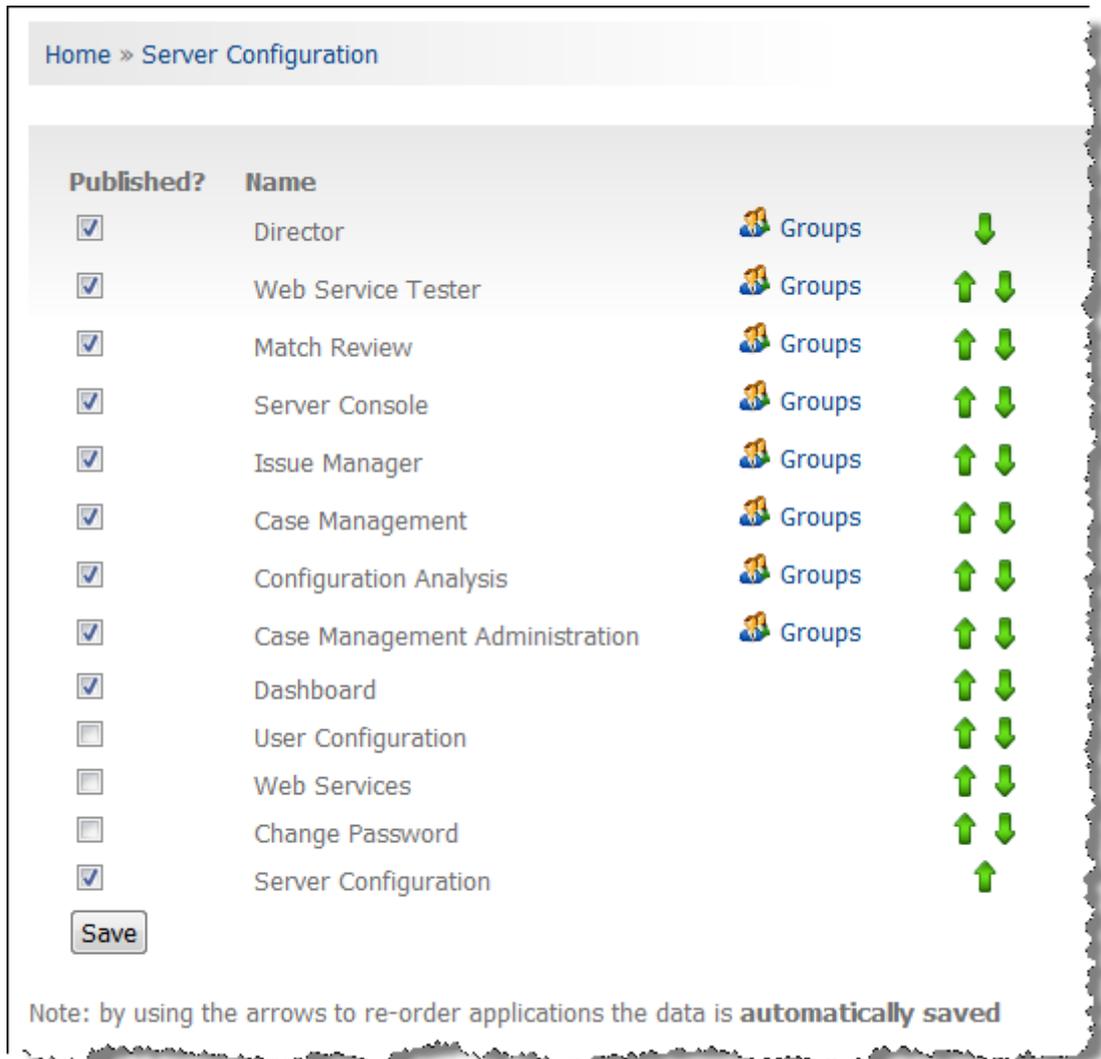
The Launchpad provides access to a number of services related to the OEDQ system as well as an alternative way of starting the client applications. The following services are displayed by default on the Launchpad when OEDQ has just been installed:

 Director	Launch the Director client application. See the Director Client section for details.
 Dashboard	Launch the Dashboard web application. See the Dashboard section for details.
 Match Review	Launch the Match Review application. See the Match Review section for details.
 Web Services	Displays details of the Web Services configured in the OEDQ server.
 User Configuration	<p>Provides a list of OEDQ users and their groups. This web application also allows a sufficiently privileged user to configure OEDQ users, permission groups and password/security rules.</p> <p>An initial installation of the OEDQ server includes one user, called <code>dnadmin</code>, with complete rights to the system. It is recommended that a user be</p>

	created for each person authorized to use a particular OEDQ system and that their role (and hence access permissions) be reflected in the permissions granted to their user account.
 Server Configuration	Allows the following configuration tasks to be achieved after installation: <ul style="list-style-type: none"> <li>• Configuration directory setup</li> <li>• Feature, interface and dashboard enablement</li> <li>• Repository database configuration</li> </ul>
 Change Password	Allows a user to change their password. The user must log in using their own password, then provide and confirm their new password.
 Help Contents	Provides information about how to use OEDQ and a list of the 'processors' that are available (according to the enabled modules) for constructing processes.

#### **Utilities available from the Launchpad**

The Launchpad may be customized to remove any of the above applications (for example, where they are not used), to change the order in which the application links appear, and to add any of the other OEDQ applications:



The Applications screen can be accessed by selecting Server Configuration from the Launchpad, then Applications. You will have to log in to an account with the correct privileges to access this screen.

## 4.2 The Director client

It is not necessary to install OEDQ onto a machine in order to run any of its client applications, including the main configuration application, Director. The Director client is designed to be installed and run remotely via the Launchpad on any machine which has Java Web Start, and Java 6 or later, installed.

### Java Web Start

Java Web Start must be installed on the client machine before it can be used to launch the Director client. Java Web Start integrates with the web browser on the client machine and can download, install, run and automatically update the Director client. The OEDQ distribution includes a Java runtime installer, which includes Java Web Start, in the Third Party Software directory. Alternatively, Java installers can be downloaded from the Oracle Java web site (<http://java.sun.com/javase/downloads/index.jsp>). The Director client

needs Java 6 or later. The Java installation will register Java Web Start with your web browser. Once this is done, the Director client can be installed or started using Java Web Start as outlined below.

### Starting the client application

To launch the Director client:

- Open the Launchpad by pointing your web browser at:

`http://<server name>:port number/dndirector`

where *<server name>* is the name of the server onto which you installed OEDQ, and the port number is the HTTP or HTTPS port. HTTP requests are automatically redirected to the HTTPS port.

The default port numbers for installations using the OEDQ Windows Installer are 9002 for HTTP and 9004 for HTTPS, though different port numbers may be allocated if either of these ports were in use at the time of installation.

For non-Windows installations, the default HTTP port is normally 8080, but be aware that if you have deployed the application server to run against a different port, you should use your custom port number here.

- Click on the Director icon. You may be prompted to either open or save the JNLP application; choose the Open option.
- After the Director user application has been downloaded onto your machine, you will receive a warning about the security certificate. To run the Director user application you must choose to accept this certificate.
- The installation process will ask you if you wish to create shortcuts for the client application. If you choose Yes, the installer will create a Start Menu shortcut that can be used to start the Director user application independently of the Launchpad.

## 4.3 Dashboard

The Dashboard user application is installed as part of OEDQ and is a separately enabled module. Dashboard allows ready publication of data quality metrics derived from OEDQ processes designed using Director. Dashboard can be accessed by clicking on the Dashboard button on the Launchpad, as described on page [8](#).

## 4.4 The Match Review application

The Match Review application is installed as part of OEDQ. Match Review allows a user to view an overview of the reviews assigned to them and to launch the review application.

Match Review can be accessed by clicking on the Match Review button on the Launchpad, as described on page [8](#).

## 4.5 OEDQ documentation

OEDQ is supplied with an extensive set of online documentation, which can be accessed as follows:

### Online help

Online help is provided for the following user applications:

- Director
- Server Console
- Case Management
- Configuration Analysis

The online help for each application can be accessed by pressing the F1 key or by clicking on the Help icons within each application.

The Director online help can also be accessed via the Launchpad as described on page [8](#). The Director online help is the largest set and contains a great deal of information on the system in general.

### Project browser context sensitive help

All of the main nodes in the Director project browser, such as projects, snapshots, issues and so on, have integrated links to help pages which explain the terms used and the purposes of the object. This context sensitive help can be accessed by right-clicking on an object in the Project Browser and choosing 'Help' from the context menu, or by left-clicking on the node and pressing F1.

### Processor-specific context sensitive help

Each of the supplied data quality processors has associated help documentation, explaining what the processor does, which data types it can handle, its inputs and outputs and examples of how you might use the processor. This help can be accessed by right-clicking on a processor on the canvas and choosing 'Processor Help' from the context menu, or by left-clicking on a processor, either on the canvas or in the tool palette, and pressing F1.

## 5 Upgrading OEDQ

The Windows installer will perform most of the steps required to upgrade your OEDQ installation without intervention. Any specific upgrade notes will be detailed in the release notes distributed with your OEDQ distributon.

**SECURITY NOTE:** When upgrading from a version of OEDQ prior to 8.1, it is **strongly recommended** that you enable secure external tasks and migrate all external tasks to use the command area, as described in [section 5.1 "Enabling secure external tasks"](#).

## 5.1 Enabling secure external tasks

There are a number of steps that must be taken to enable secure external tasks in OEDQ:

- Set the `externaltasks.restricted` property in **director.properties** to `true`;
- The default command area is **config/commandarea**. If you wish to use an alternative directory, you will need to create it.
- If you are not using the default command area, update the `commandarea` property in **director.properties** to point to the directory you wish to use;
- Move all scripts that are called by your external tasks into the command area. Note that if an external processor, such as perl or wscript, is required to execute any of your scripts, the processor must be invoked from within a script or batch file that resides in the command area.
- Update all the external tasks that use those scripts to point to the new location of the scripts.

## 5.2 Upgrading OEDQ on Tomcat

When upgrading an instance of OEDQ running on Tomcat, Oracle strongly recommends that all cached JSP pages be deleted beforehand.

To do this, delete the contents of the Tomcat **work** directory. An example of the path to this directory is: **C:\Program Files\Datanomic\dnDirector\tomcat\6.0\work**

# 6 Troubleshooting OEDQ

The following list includes some common problems encountered when attempting to install or run OEDQ.

## 6.1 Director Client works but unable to access online help or OEDQ Launchpad

If you can access the Director Client application but are unable to get to the OEDQ Launchpad or access the online help, please check your browser settings. If your browser is configured to use a proxy server it may be attempting to access the OEDQ Application Server via the proxy. In Internet Explorer, check the Tools > Internet Options menu item, go to the Connections tab and press the LAN Setting button to review the settings.