

# Oracle® Enterprise Data Quality

Release Notes

12c (12.2.1)

E56641-01

October 2015

---

This document contains the release information for Oracle Enterprise Data Quality (EDQ) 12c (12.2.1) and includes the following:

- [New Features and Improvements](#)
- [Deprecation Notice](#)
- [Issues Resolved](#)
- [Known Issues and Workarounds](#)
- [Upgrade Considerations](#)

Oracle recommends you review its contents before installing or working with the product.

## New Features and Improvements

The following sections describe the new features introduced in this release as well as improvements.

### WebLogic Clustering and Oracle RAC Support

This release now offers full support for WebLogic clustering, and improved support for Oracle RAC, including the ability for all servers in a cluster to share the same configuration and results schemas in a database, added tolerance for RAC node and DB failures, and automatic job control and load balancing.

### JavaScript Support

Java Script is supported in 12.2.2 by use of the Nashorn engine integrated with Java 8. Scripts written in JavaScript which do not use E4X will not need to be changed (The most common use case for E4X is in custom scripts decoding and creating XML in JMS 'bucket' files.).

### Customer Data Services

The following notes apply to Customer Data Services.

#### Performance

In this release, Key Generation and Matching services are considerably richer than in previous versions. Because of these enhancements, performance may be somewhat slower (higher average response time and lower throughput) than in the previous release.

You can mitigate this by using the new Key Analysis feature which selects the optimum key generation profile for your data to ensure matching is run efficiently.

### **Using Existing Customer Data Services**

You can continue to use your existing Customer Data Services, as well as all existing configuration (especially Watchlist Screening configurations based on Oracle Watchlist Screening 11) in Enterprise Data Quality 12.2.1, without performance degradation.

### **Integrated Installation of CDS and EDQ**

The installation of Oracle Enterprise Data Quality now also includes the installation of CDS widgets, gadgets, and reference data in the landing area. This ensures that an installation of EDQ can access CDS functionality.

### **Improved Control of Key Method Enablement**

This feature allows you to create key profiles which specify the enablement of key methods, and allows the extension of EDQ-CDS to offer a wider menu of key method algorithms. You can assign priority ordering to key methods for efficient candidate selection.

The existing cluster level is deprecated and replaced with a new key profile setting.

### **Key Analysis**

The key analysis feature allows you to analyze the keys generated for the customer data to determine the optimum key profile, and provide a report of analysis results. The feature profiles key values for all key methods, for the full customer data, and scores key methods based on their characteristics, including high frequency key values, and the distribution/diversity of key values. The output of Key Analysis is a recommended Key Profile that can be used to tune the matching performance and efficacy of the CDS matching services.

### **Output New Match Metadata**

This feature provides more granular details about why two records matched, and which logical match identifiers contributed to a match. The feature outputs new EDQ match metadata for each logical match identifier.

### **Enable Matching on Customer-Added Attributes**

This feature improves how you configure matching by allowing you to use up to six custom string and up to three custom date attributes in match, to use any custom attributes in keying, and to adjust the attribute contribution to match scoring. The input interfaces are extended to include new attributes for data and corresponding configuration parameters.

The custom attributes are standardized for keying and matching.

The key method used for each custom attribute is dependent on key profile. The method is fixed for manual profiles. The best option is determined during key analysis for automatic profiles.

## **Inverted Elimination IDs**

Inverted Elimination IDs support cross-segment only matching. These IDs have similar functionality to Elimination Ids (EIDs) but produces a 'No match' result when the identifier values are the same.

This feature allows you to eliminate matches where records share a common value. For example, find duplicate contacts across accounts but NOT [that is, eliminate] matches within account hierarchies

## **Adjusting Match Attribute Weightings**

This feature externalizes match weightings in EDQ to allow external configuration of the weightings. Configuration information is passed on a per-message basis.

Weightings are assigned to logical match identifiers, and contribution values are used in a weighted scoring algorithm.

## **Match Flexibility**

This feature allows matching to be configured in a more efficient and flexible way. You can set up compound comparisons with all match rules on a given logical identifier, with custom scoring on each rule, and optionally configure scoring across logical identifiers. Match rules are more flexible to allow comparison, compound comparison, and scoring results to be used as criteria.

The benefits of this feature include

- Ease of setting up match configuration: many less rules will have to be specified explicitly, therefore meaning much less configuration time required to set up
- Flexibility: rules will be able to take into account the matching or non-matching of all groups within the rule, meaning more accurate information will be able to be returned on a match
- Externalization: allowing weightings across the new groups and allowing these to be externalized will help external configuration of match. For example, giving a higher weighting to a logical group can be used increase the contribution that this group's score can make towards an overall score. It will also be possible to turn off a logical group completely.

## **Update to Address Verification Connector**

The Address Verification processor has been improved to add support for certified outputs (CASS, SERP and AMAS) and to support Traditional Chinese as a specific output script option.

## **Incorporate the DataDirect Hadoop Driver**

This feature broadens the connectivity support in EDQ by implementing the DataDirect Hadoop driver. This allows you to connect to a HIVE data store. For more information, see

<https://www.progress.com/data-sources/apache-hive-drivers-for-odbc-and-jdbc>  
c.

## REST Web Services and Configuration API

A Web services interface allows the display of EDQ jobs in Oracle Enterprise Manager. For example, `http://server:port/edq/config/projects` returns a list of projects (name/id) in JSON format. Adding an: `Accept: text/xml` header to the request returns the results in XML format.

## Deprecation Notice

Rhino (the JavaScript Engine) is deprecated in this release, and will be removed in the next major release of Enterprise Data Quality. Custom scripts that use the E4X functionality of Rhino, which is often used to construct XML, for example to call an external web service, should if possible be rewritten after upgrade to EDQ 12.2.1.

## Issues Resolved

The following sections describe the bugs resolved in this release.

### Convert Number to Date Processor

The convert number to date processor has a milliseconds from base date option that was not working correctly. A fix is included in this release, and the feature is working.

### CED Comparison Failed When Set to Nodata

A problem was reported in Release 11g in which a nodata result was given for an attribute result that had data. The problem has been corrected in 12c.

### Failed to Set Priority from Workflow in Non-English Locale

A problem was reported in which a user could not set priorities in the workflow when using non-English locales. That issue has been resolved in this release.

### JNDI Data Store Configuration Not Offering User Name and Password Protection

This release includes a fix for a problem in which users were not offered the opportunity to password protect their JNDI data stores. Now Username and Password protection is offered.

### Define Purge Rules for Event Data

Purge rules have been extended to cover a situation in which event data was using large amounts of tablespace.

This functionality can be activated by adding an xml file to the housekeeping folder of the config with the following details. The parameter is the time, in days, that the event log entry must be older than to be deleted.

```
.
<housekeeping>
.
  <task name="eventpurger">
    <start>02:00:00</start>
    <interval units="hours">24</interval>
```

```
<parameter>10</parameter>
</task>
```

## EDQ Incorrectly Mapped OPSS Groups with Spaces in Names

An error occurred when trying to map external groups to internal groups when the external group's name included spaces. This is no longer a problem in this release.

## Known Issues and Workarounds

The following sections describe known issues and any workarounds that exist in the 12c (12.2.1).

### Address Cleaning Service

The Address Cleaning service has been changed so that it will always use the country code you provide rather than the default country code in the Run Profile. If you do not provide a country code or it cannot be mapped to the standard country codes in EDQ-CDS, then EDQ-CDS attempts to derive country from the address data provided. When this is not possible, the default country code in the Run Profile is used.

## Upgrade Considerations

The following are some of the major considerations for upgrading to 12c (12.2.1).

### Upgrades in an Oracle WebLogic Server Environment

- Read *Planning an Upgrade of Oracle Fusion Middleware* for guidelines for preparing to upgrade to Oracle Fusion Middleware 12c (12.2.1). This documentation also includes descriptions of terminology changes that you must understand to move forward to a 12c environment.
- If your current EDQ version is 11.1.1.7 or later *and* was installed with Oracle Universal Installer (OUI), you can use the Upgrade Assistant to upgrade your installation directly to 12c (12.2.1).
- All EDQ components must be shut down and remain stopped until you are prompted to start them at the end of these upgrade instructions. The WebLogic Server console must remain running until you are prompted to shut it down during the upgrade procedure.

For all upgrade scenarios, see “Upgrading EDQ On WebLogic Servers” in *Oracle Fusion Middleware Installing and Configuring Enterprise Data Quality*.

### Upgrades in an Apache Tomcat Environment

You can perform a direct upgrade to version 12c of EDQ only from Tomcat version 8. If you are running Tomcat 6, you must upgrade Tomcat to version 7 before proceeding with the EDQ upgrade. See the Apache Tomcat documentation at

<http://tomcat.apache.org>

To upgrade to 12c (12.2.1), see “Upgrading EDQ On Tomcat Servers” in *Oracle Fusion Middleware Installing and Configuring Enterprise Data Quality*.

## Related Documents

For more information about EDQ, see the following documents in the Oracle Enterprise Data Quality documentation set.

### EDQ Documentation Library

The following publications are provided to help you install and use EDQ:

- *Oracle Fusion Middleware Installing and Configuring Enterprise Data Quality*
- *Oracle Fusion Middleware Administering Enterprise Data Quality*
- *Oracle Fusion Middleware Understanding Enterprise Data Quality*
- *Oracle Fusion Middleware Integrating Enterprise Data Quality With External Systems*
- *Oracle Fusion Middleware Securing Oracle Enterprise Data Quality*
- *Oracle Enterprise Data Quality Address Verification Server Installation and Upgrade Guide*
- *Oracle Enterprise Data Quality Address Verification Server Release Notes*

Find the latest version of these guides and all of the Oracle product documentation at

<http://docs.oracle.com>

### Online Help

Online help is provided for all Oracle Enterprise Data Quality user applications. It is accessed in each application by pressing the **F1** key or by clicking the Help icons. The main nodes in the Director project browser have integrated links to help pages. To access them, either select a node and then press **F1**, or right-click on an object in the Project Browser and then select **Help**. The EDQ processors in the Director Tool Palette have integrated help topics, as well. To access them, right-click on a processor on the canvas and then select **Processor Help**, or left-click on a processor on the canvas or tool palette and then press **F1**.

## Documentation Accessibility

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>.

### Access to Oracle Support

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.

---

Oracle Enterprise Data Quality Release Notes, 12c (12.2.1)  
E56641-01

Copyright © 2010, 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, then 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 about 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.

