

# Oracle® Enterprise Data Quality

IBM GNR Integration

Release 11g R1 (11.1.1.7)

E40038-02

October 2013

---

Oracle Enterprise Data Quality (EDQ) can be configured to connect to IBM Global Name Recognition (GNR).

GNR allows EDQ to perform linguistic analysis of names, and linguistically sensitive name searching.

---

---

**Note:** GNR can only be installed on an EDQ instance if you have the required license agreements with both Oracle and IBM.

---

---

This document is intended for systems administrators responsible for installing and configuring EDQ applications.

## 1 Prerequisites

The prerequisites for enabling EDQ GNR connectivity are:

- EDQ version 9.0.4 or later, installed on 64-bit AIX or Linux running 64-bit Java.
- GNR 4.2.1 (4.2 + 4.2.1 fixpack) or later, including the hotfix based on GNR 4.2.2 (4.2 +4.2.2 fixpack).

---

---

**Note:** EDQ does not make use of any of the web services provided by GNR, so you do not need to configure these during GNR installation.

---

---

## 2 Integration Summary

To integrate IBM GNR into EDQ, you perform these tasks:

- [Configuring the EDQ Application Server](#)
- [Building the Search Library](#)
- [Configuring the EDQ GNR Connector](#)

## 3 Configuring the EDQ Application Server

The LD\_LIBRARY\_PATH must be set as required for the installation environment.

The EDQ GNR analytic processors use a shared library (.so) in the lib64 directory of the GNR installation. This directory must be specified in an environment variable passed to the EDQ application server.

In a Linux 64-bit environment, the environment variable is `LD_LIBRARY_PATH`; for example:

```
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:gnr-installation-dir/lib64
```

*gnr-installation-dir* is the GNR installation path; for example, `/opt/GNR/GNM`.

In an AIX environment, the environment variable name is `LIBPATH` instead of `LD_LIBRARY_PATH`.

---

---

**Note:** The environment variable must be available to the application server process.

---

---

## 4 Building the Search Library

The GNR Search Processor uses a native library that must be linked with the GNR libraries.

Oracle supplies these files to create the library:

- Two Makefile templates, one for each platform, that script the building of the search library
- The `namehunter.o` object module file

When building the library on an AIX system, IBM C++ must be available. When building the library on Linux systems, the GCC C++ compiler must be available and it must be of the same version used to create the GNR libraries, as specified in the IBM GNR documentation.

The Makefile template for 64-bit AIX is as follows:

```
# Build library from object file
# -----

GNR=/opt/GNR/GNM
CFLAGS=-qmkshrojb
LIBS=-lNameHunter -lNameTransliterator -lsicui18n -lsicuuc -lsicudata
SDK=aix61_64-xlc9-release
LIBDIRS=-L$(GNR)/sdk/$(SDK)/lib -L$(GNR)/sdk/icu4c/$(SDK)/lib

all: libnimrod.so

libnimrod.so: namehunter.o
        xlc++_r -q64 $(CFLAGS) -o $@ $? $(LIBDIRS) -lNameHunter -lNameTransliterator
        -lsicui18n -lsicuuc -lsicudata
```

The Makefile template for 64-bit Linux is as follows:

```
# Build 64-bit library from object file
# -----

GNR=/opt/GNR/GNM
CFLAGS=-shared -fPIC

all: libnimrod.so

libnimrod.so: namehunter.o
        g++ -m64 $(CFLAGS) -o $@ $? -L$(GNR)/sdk/rhel4_64-gcc34-release/lib
        -L$(GNR)/sdk/icu4c/rhel4_64-gcc34-release/lib -lNameHunter -lNameTransliterator
```

```
-lsicui18n -lsicuuc -lsicudata
```

Before running the Makefile script for your platform, ensure that the value of GNR in the Makefile template is set correctly, according to the GNR installation directory location. On the AIX platform, also ensure that the value of SDK is set correctly, according to the system architecture.

After running the edited Makefile, the newly created libnimrod.so shared library file can be installed anywhere and can be copied to other systems with GNR installs.

## 5 Configuring the EDQ GNR Connector

The EDQ GNR connector requires three types of configuration files to integrate it with GNR:

- The `gnr.properties` properties file in the EDQ installation
- The `nameworks.config` configuration file in the GNR installation
- Search configuration files in the EDQ installation

### 5.1 gnr.properties

The `gnr.properties` file must be manually created and placed in the `gnr` subdirectory of the EDQ configuration directory. It must contain the following properties:

#### **gnr.install**

The GNR installation path. This is the path to the directory containing the following GNR subdirectories:

- `bin`
- `bin64`
- `data` (which contains the GNR data files)
- `lib`
- `lib64`

#### **analytics.config**

The absolute location of the `nameworks.config` configuration file in the GNR installation.

#### **search.jnilib**

The absolute location of the `libnimrod.so` shared library, which was built using the Makefile template.

### 5.2 nameworks.config

During GNR installation, a `nameworks.config` file is created and stored in the GNR data directory.

The critical part of the `nameworks.config` file is the reference files section:

```
[Reference Files]
NameSifter=/opt/GNR/GNM/data/SifterRules.ibm
```

The `NameSifter` value must refer to the `SifterRules.ibm` file in the GNR installation.

## 6 Search Configuration Files

Search configuration files are located in the `gnr/search` subdirectory of the EDQ configuration directory. They are read by the connector and used to set parameters for the Search function.

A sample search configuration file named `search.config` is available in the `support/data/search` subdirectory of the EDQ installation. To create a search configuration file, copy this sample file to the `gnr/search` subdirectory of the EDQ configuration directory and edit the copy to suit your needs.

### 6.1 Support for GNR 3.2 and GNR 4.2 in Search Configuration Files

The search configuration format changed slightly from GNR 3.2 to GNR 4.2, and the EDQ GNR connector supports both versions as far as possible. It also processes data for Organization searches.

The basic differences between the search configuration files in GNR 3.2 and GNR 4.2 are:

- GNR 4.2 specifies the parameter files (tags, variants, etc) in the `[hunter]` section. GNR 3.2 uses the `[search]` section. The EDQ GNR connector looks in the `[hunter]` section first then the `[search]` section.
- The tag and variant files in GNR 4.2 are specified by keys such as `ibmTagFile` and `custTagFile`. In GNR 3.2, just `tagFile` is used. The EDQ GNR connector looks for `tagFile`, `ibmTagFile` and `custTagFile` and loads each if found. The same rules are used for variant and terms files.
- The generic reg file is set by a specific `genericRegFile` setting; in GNR 3.2 this always defaults to the `anglo` reg file.
- Some settings have been added to the `[parms]` sections, and others have been removed.

## 7 Related Documents

For more information, see the following documents in the Oracle Enterprise Data Quality documentation set:

- *Oracle Enterprise Data Quality Release Notes*
- *Oracle Enterprise Data Quality Architecture Guide*
- *Oracle Enterprise Data Quality Installation Guide*

See the latest version of this and all documents in the Oracle Enterprise Data Quality Documentation website at

[http://download.oracle.com/docs/cd/E48549\\_01/index.htm](http://download.oracle.com/docs/cd/E48549_01/index.htm)

## 8 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 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 IBM GNR Integration, Release 11g R1 (11.1.1.7)  
E40038-02

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 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 failsafe, 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.

