



Regional Blueprints User's Guide

- SAP BusinessObjects Data Services XI 4.0 (14.0.0)

2011-01-13

Copyright

© 2011 SAP AG. All rights reserved. SAP, R/3, SAP NetWeaver, Duet, PartnerEdge, ByDesign, SAP Business ByDesign, and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP AG in Germany and other countries. Business Objects and the Business Objects logo, BusinessObjects, Crystal Reports, Crystal Decisions, Web Intelligence, Xcelsius, and other Business Objects products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of Business Objects S.A. in the United States and in other countries. Business Objects is an SAP company. All other product and service names mentioned are the trademarks of their respective companies. Data contained in this document serves informational purposes only. National product specifications may vary. These materials are subject to change without notice. These materials are provided by SAP AG and its affiliated companies ("SAP Group") for informational purposes only, without representation or warranty of any kind, and SAP Group shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP Group products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

2011-01-13

Contents

Chapter 1	Overview.....	5
Chapter 2	Downloading Blueprint Packages.....	7
2.1	Blueprint versions.....	7
2.2	Available regional blueprints.....	7
2.3	Downloading and setting up blueprints.....	8
Chapter 3	Configuring and Running Jobs.....	11
3.1	Editing the datastore	11
3.1.1	Microsoft SQL Server	11
3.1.2	Other database types.....	12
3.2	Running the jobs.....	12
Index		15

Overview

We've identified a number of common data quality scenarios that you are likely to perform with SAP BusinessObjects Data Services. For each scenario, we've included a blueprint that is already set up to solve the business problem in that scenario. Each blueprint contains the necessary project, jobs, data flows, file formats, sample data, template tables, and custom functions to run the jobs in your environment with only a few modifications.

You can download the blueprint packages from the SAP Community Network. On the website, we periodically post new and updated blueprints, custom functions, best practices, whitepapers, and other content. You can refer to this site frequently for updated content and use the forums to provide us with any questions or requests you may have. We've also provided the ability for you to upload and share any content that you've developed with the rest of the SAP BusinessObjects Data Services development community (for instructions on uploading content, see *How to Contribute* at <https://www.sdn.sap.com/irj/scn/submitcontent>).

Instructions for downloading and installing the content objects are also located on the SAP Community Network website.

Downloading Blueprint Packages

2.1 Blueprint versions

The following table shows the version of the Data Quality and Text Data Processing blueprints that can be used for SAP BusinessObjects Data Services. The blueprint version is displayed on the SAP BusinessObjects Data Services Blueprints page of the SAP Community Network website.

Data Services version	Blueprint version	Blueprints available
XI 4.0	XI 4.0	Data Quality Text Data Processing
XI 3.2	XI 3.2	Data Quality
XI 3.1	XI 3.0	Data Quality
XI 3.0	XI 3.0	Data Quality

2.2 Available regional blueprints

Each blueprints package contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in a particular region. The "global" blueprints package illustrates how to configure the same jobs when the data is multinational.

To see the contents of each blueprint, including jobs and custom functions, see the *Content Objects Summary*. To help you compare the available blueprints and decide which to download, see the following table.

Blueprint	Description
Data Quality Blueprints – Brazil	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in Brazil.
Data Quality Blueprints – France	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in France.
Data Quality Blueprints – Germany	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in Germany.
Data Quality Blueprints – Global	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data when the data consists of multiple countries.
Data Quality Blueprints – India	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in India.
Data Quality Blueprints – Japan	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in Japan.
Data Quality Blueprints – Mexico	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in Mexico.
Data Quality Blueprints – USA	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in the United States.
Data Quality Blueprints – USA Regulatory	Contains sample jobs configured to illustrate best practice settings for common Data Quality use cases involving party data in the United States, with regulatory address certification enabled.

2.3 Downloading and setting up blueprints

These are the general steps for downloading and setting up Data Quality blueprint packages for SAP BusinessObjects Data Services.

Note:

If you are using a data flow that includes Data Cleanse, install the cleansing packages before importing the .atl file.

1. To access the SAP Community Network website, go to <https://www.sdn.sap.com/irj/boc/blueprints> in your web browser.
2. Log into your account using your username and password, or create a new account.
3. Review the list of available blueprint packages and other content objects and their descriptions to decide which to download.
4. Select the blueprint package that you want to download. A new page opens.
5. Click the **View this Code Sample** button.
6. In the File Download window, save the .zip file to the Tutorial Files folder in your installed SAP BusinessObjects Data Services path. By default, this folder is installed to \Program Files\SAP BusinessObjects\Data Services\Tutorial Files for 32-bit Windows and \Program Files (x86)\SAP BusinessObjects\Data Services\Tutorial Files for 64-bit Windows. If you are running on UNIX, the Tutorial Files folder exists only on the Windows client workstation, and you should download the .zip file there.
7. In the Tutorial Files folder in Windows Explorer, right-click the .zip file and select to extract the compressed (zipped) folders to the Tutorial Files folder. For example, if you use WinZip for file compression, right-click the .zip file and select **WinZip > Extract to here**.
Extracting creates subfolders and places the files in the appropriate location. The .atl file is saved to the Data Quality Samples folder, and the sample data files are saved to the Data Quality Samples*<subfolder>* folder.
8. In the Designer, import the .atl file. In the Passphrase window, enter the name of the .atl file without the extension (for example, when importing data_quality_blueprints_france.atl, enter the passphrase data_quality_blueprints_france) and click **Import**. Click **OK** to close the warning window.
Importing the file adds a project called DataQualityBlueprintsXxx to your object library, where Xxx represents the name of the region. The project contains jobs whose names begin with DqBlueprintXxx and contain in their name the Data Quality use case that they illustrate. Each real-time job includes "Rt" in its name. Each job contains a data flow with the same name as the job. The import also adds a datastore called DataQualityBlueprintsXxx to your object library, and some file formats called DqXxxIn and DqXxxOut for the sample input and output data.
9. If you are running on UNIX, copy the input files to the job server machine and create the same folder structure that is on the Windows client workstation.
10. If you imported the blueprint .atl files using a Data Services Designer on 32-bit Windows and use a job server on 64-bit Windows, then you must copy the blueprint files to the Data Services installation of the job server machine.

Related Topics

- [Editing the datastore](#)

Configuring and Running Jobs

3.1 Editing the datastore

After you download the blueprint .zip file to the appropriate folder, unzip it, and import the .atl file in the Designer, you must edit the DataQualityBlueprints datastore.

The database that you use for running the blueprints does not need to be the same database that is used for the SAP BusinessObjects Data Services repository. It can be a locally installed database system or any shared database system that you have access to create tables in and read from those tables.

Most of the blueprint jobs write to a flat file on your file system. The following are the only jobs that need access to a database system:

- DqBlueprint $_{xxx}$ _LoadInitial—creates a table called DQ_BLUEPRINTS_ $_{xx}$ _CRM, where $_{xx}$ represents the 2-character country code of the region.
- DqBlueprint $_{xxx}$ _LoadDelta—selects candidate records from DQ_BLUEPRINTS_ $_{xx}$ _CRM for the matching process of a delta load.
- DqBlueprint $_{xxx}$ _RtMatchConsumer—selects candidate records from DQ_BLUEPRINTS_ $_{xx}$ _CRM for the matching process of a real-time match.

Related Topics

- [Microsoft SQL Server](#)
- [Other database types](#)

3.1.1 Microsoft SQL Server

If you have access to write and read data to tables in Microsoft SQL Server 2000, 2005, or 2008, complete the following steps.

1. Select the **Datastores** tab of the Local Object Library, right-click the region-specific DataQualityBlueprints $_{xxx}$ datastore and select **Edit**.
2. In the **Edit Datastore DataQualityBlueprints $_{xxx}$** window, enter your repository connection information in place of the four **CHANGE_THIS** values.
3. Click **OK**. If the window closes without an error message, then the database is successfully connected.

4. Before you run a blueprint job that accesses a database, run `DqBlueprintxxx_LoadInitial` to create the table that the subsequent jobs `DqBlueprintxxx_LoadDelta` and `DqBlueprintxxx_RtMatchConsumer` select rows from.

After you edit the datastore, you should be able to run the blueprint jobs.

3.1.2 Other database types

If you have access to write and read data to tables in another database system (other than Microsoft SQL Server), complete the following steps.

1. Select the **Datastores** tab of the Local Object Library, expand the region-specific DataQuality Blueprints_{xxx} datastore, and expand the **Template Tables** subfolder.
2. Make note of the name of the datastore and the name of the template table. For example, for the Japan blueprints, note that the name of the datastore is `DataQualityBlueprintsJapan` and the name of the template table is `DQ_BLUEPRINTS_JP_CRM`.
3. Delete the template table. Right-click `DQ_BLUEPRINTS_XX_CRM`, select **Delete**, and select **Yes** to confirm your selection.
4. Delete the datastore. Right-click `DataQualityBlueprintsxxx`, select **Delete**, and select **Yes** to confirm your selection.
5. Create a new datastore with the same name as the one you just deleted. In the **Datastores** tab of the Local Object Library, right-click in the white space and select **New**. In the **Datastore** name field, enter the name `DataQualityBlueprintsxxx`, where `xxx` represents the name of the region. In the Database type field, select your database system. Complete the remaining fields with the connection information to the database that you have access to.
6. Click **OK** to close the Create New Datastore window.
7. Open the dataflow `DqBlueprintxxx_LoadInitial` and delete the target. Then add a new template table with the same name by selecting the Template Table icon from the buttons on the right menu and clicking the dataflow canvas. In the Create Template window, enter `DQ_BLUEPRINTS_XX_CRM` for the Template name, where `XX` represents the 2-character country code of the region, and select the `DataQualityBlueprintsxxx` datastore in the In datastore field, where `xxx` represents the region. Click **OK** to close the Create Template window. Connect the last transform to the template table.
8. Before you run a blueprint job that accesses a database, run `DqBlueprintxxx_LoadInitial` to create the table that the subsequent jobs `DqBlueprintxxx_LoadDelta` and `DqBlueprintxxx_RtMatchConsumer` select rows from.

After you edit the datastore, you should be able to run the blueprint jobs.

3.2 Running the jobs

Before you run your jobs, you should have already completed the following tasks:

1. Copy the address cleanse reference files for the region before running jobs that include address cleansing.
2. Copy the geocode reference files for the region before running jobs that include geocoding.
3. Install the cleansing packages for the region before running jobs that include data cleansing.
4. Set accurate values in the substitution parameter configuration Configuration1:
 - `$$RefFilesAddressCleanse`—Enter the path location where you copied the address cleanse reference files.
 - `$$RefFilesGeocoder`—Enter the path location where you copied the geocode reference files.
 - `$$ReportsAddressCleanse/$$ReportsGeocoder/$$ReportsMatch`—Enter YES or NO to specify whether to generate reports when running the blueprints.
 - When running jobs in the USA blueprints package, you must enter values for the following:
 - `$$CompanyName`
 - `$$CompanyAddress`
 - `$$CompanyLocality` (name of city)
 - `$$CompanyRegion` (2-character state code)
 - `$$CompanyPostcode1` (5-digit ZIP Code)
 - `$$CompanyPostcode2` (4-digit ZIP+4 Code)
 - `$$CompanyPhone` (10-digit number without spaces or delimiters)

Index

A

about blueprints 5
available blueprints 7

B

blueprints
 about 5
 available 7
 downloading 8
 versions 7

D

Data Quality regional blueprints
 list of 7
datastore
 editing 11
downloading blueprints 8

J

jobs, running 12

R

running jobs 12

S

SAP BusinessObjects Data Services
 Blueprints
 versions 7
SQL Server
 editing datastore 11

V

versions 7

