



What's New

- SAP BusinessObjects Data Services XI 4.0 (14.0.0)

2010-12-02

Copyright

© 2010 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.

2010-12-02

Contents

Chapter 1	Introduction.....	5
1.1	SAP BusinessObjects information resources.....	5
1.2	Introduction to SAP BusinessObjects Data Services XI 4.0.....	6
Chapter 2	SAP BusinessObjects Data Services XI 4.0 features.....	9
2.1	Documentation.....	9
2.2	SAP integration.....	9
2.3	Security.....	11
2.4	Text Data Processing.....	12
2.5	Architecture.....	12
2.6	Transforms.....	13
2.7	Operational excellence.....	14
2.8	Functions.....	14
2.9	Source and target support.....	15
2.10	Data Quality.....	15
2.10.1	Data Cleanse transform.....	17
2.10.2	Geocoder transform.....	18
2.10.3	Global Address Cleanse transform.....	19
2.10.4	Match transform.....	21
2.10.5	USA Regulatory Address Cleanse transform.....	24
Index		27

Introduction

1.1 SAP BusinessObjects information resources

A global network of SAP BusinessObjects technology experts provides customer support, education, and consulting to ensure maximum information management benefit to your business.

Useful addresses at a glance:

Address	Content
Customer Support, Consulting, and Education services http://service.sap.com/	Information about SAP Business User Support programs, as well as links to technical articles, downloads, and online forums. Consulting services can provide you with information about how SAP BusinessObjects can help maximize your information management investment. Education services can provide information about training options and modules. From traditional classroom learning to targeted e-learning seminars, SAP BusinessObjects can offer a training package to suit your learning needs and preferred learning style.
SAP BusinessObjects Data Services Community http://www.sdn.sap.com/irj/sdn/ds	Get online and timely information about SAP BusinessObjects Data Services, including tips and tricks, additional downloads, samples, and much more. All content is to and from the community, so feel free to join in and contact us if you have a submission.
Forums on SCN (SAP Community Network) http://forums.sdn.sap.com/forum.jspa?forumID=305	Search the SAP BusinessObjects forums on the SAP Community Network to learn from other SAP BusinessObjects Data Services users and start posting questions or share your knowledge with the community.
Blueprints http://www.sdn.sap.com/irj/boc/blueprints	Blueprints for you to download and modify to fit your needs. Each blueprint contains the necessary SAP BusinessObjects Data Services project, jobs, data flows, file formats, sample data, template tables, and custom functions to run the data flows in your environment with only a few modifications.
Product documentation http://help.sap.com/businessobjects/	SAP BusinessObjects product documentation.
Supported Platforms (Product Availability Matrix) https://service.sap.com/PAM	Get information about supported platforms for SAP BusinessObjects Data Services. Use the search function to search for Data Services. Click the link for the version of Data Services you are searching for.

1.2 Introduction to SAP BusinessObjects Data Services XI 4.0

Welcome to SAP BusinessObjects Data Services XI 4.0 version 14.0.0. This What's New document highlights the new features available with this release. Note: The What's New document describes new functionality in major and minor releases only; support packages and patches do not contain new functionality.

For important information about this product release including installation notes and known issues, see the *SAP BusinessObjects Data Services Release Notes*. For resolved issues, see the *SAP BusinessObjects Data Services Customer Issues Fixed*.

SAP offers other products that complement SAP BusinessObjects Data Services and provide additional Enterprise Information Management solutions. These include:

- SAP BusinessObjects Data Federator
- SAP BusinessObjects Data Quality Management for Enterprise Applications
- SAP BusinessObjects Data Quality Management SDK
- SAP BusinessObjects Event Insight
- SAP BusinessObjects Information Steward (expected in 2011)
- SAP Master Data Governance for Embedded MDM
- SAP NetWeaver Business Warehouse (BW)
- SAP NetWeaver Master Data Management for Enterprise MDM
- SAP BusinessObjects Rapid Marts
- SAP BusinessObjects Text Analysis

For more information, see the SAP website or contact an SAP sales representative.

Data Services overview

SAP BusinessObjects Data Services delivers a single enterprise-class solution for data integration, data quality, data profiling, and text data processing that allows you to integrate, transform, improve, and deliver trusted data to critical business processes. It provides one development UI, metadata repository, data connectivity layer, run-time environment, and management console—enabling IT organizations to lower total cost of ownership and accelerate time to value. With SAP BusinessObjects Data Services, IT organizations can maximize operational efficiency with a single solution to improve data quality and gain access to heterogeneous sources and applications.

SAP BusinessObjects Data Services XI 4.0 features

2.1 Documentation

Several changes have been made to the documentation set that is provided with SAP BusinessObjects Data Services.

- Release documentation:
 - The *Release Summary* has been renamed to *What's New*.
 - Resolved issues in the *Release Notes* have been moved to a separate *Fixed Issues* document.
- The *Getting Started Guide* and *Advanced Development Guide* have been removed and most of their content has been moved to the new *Administrator's Guide*. The *Administrator's Guide* contains information about administrative tasks such as monitoring, lifecycle management, security, and so on. It has also been added to the *Technical Manuals*, which is a compilation of the core Data Services documents.

Multi-user content in the *Advanced Development Guide* has been moved to the *Administrator's Guide* and *Designer Guide*.

- The *Management Console: Metadata Reports Guide* and *Management Console: Administrator Guide* have been combined into the *Management Console Guide*.
- The *Installation Guide for Windows* and *Installation Guide for UNIX* have been removed from the *Technical Manuals*. All post-installation, administrative content in the installation guides has been moved to the *Administrator's Guide*.
- The *Integrator's Guide* has been added to the *Technical Manuals*.
- The *Salesforce.com Adapter Guide* has been renamed to *Supplement for Salesforce.com* and added to the *Technical Manuals*.

2.2 SAP integration

SAP BusinessObjects Data Services integrates with your existing SAP infrastructure with the following SAP tools:

Reading business content extractors for ERP and CRM/SRM

Data Services can now read SAP business content extractors for ERP and CRM/SRM. Using extractor sources lets you get data directly from SAP Business Suite applications in a consistent format without

requiring access to the underlying ABAP tables. Data Services supports both extractors that use the Operational Data Provider data replication API and those available in all ECC 6 versions.

When the Operational Data Provider (ODP) data replication API is installed on the SAP NetWeaver platform, in Data Services you get the added benefits of the ability to browse all available extractors, extract data in both initial and changed-data capture (delta) mode, and stream data from the SAP application to the data flow without using staging files.

Support for SAP NetWeaver BW 7.3 staging BAPI

This release adds support for new BW 7.3 staging BAPI for native loading of BW 7.x datastores (no emulation mode translates to ease of use). Additionally, Data Services now includes an updated version for the SAP NetWeaver RFC SDK for better handling of Unicode data and an updated RFC server for parallel processing (affects both loading via staging BAPI and reading via Open Hub). For more information, refer to the SAP NetWeaver BW documentation.

SAP NetWeaver BW 7.3 integration

SAP NetWeaver BW 7.3 provides native support for Data Services now via a Data Services source system and the ability to create Data Services batch jobs to load BW, all from within the BW workbench. For more information, refer to the SAP NetWeaver BW documentation.

SAP System Landscape Directory (SLD)

The system landscape directory of SAP NetWeaver is the central source of system landscape information relevant for the management of your software life-cycle. With the provided directory comprising information about all installable software available from SAP and automatically updated data about systems already installed in a landscape, you get the foundation for tool support to plan software life-cycle tasks in your system landscape.

The SAP BusinessObjects Data Services installation program registers the vendor and product names and versions with the SLD, as well as server and front-end component names, versions, and location.

Solution Manager Diagnostics (SMD)

The SMD component of SAP Solution Manager provides all functionality to centrally analyze and monitor a complete system landscape. Data Services can be monitored by the SMD server if an SMD Agent is installed. The SMD Agent gathers information for the SMD which can then be used for root cause analysis.

Data Services provides support for this performance monitoring through CA/Wily Introscope in Solution Manager Diagnostics through an integration with the NCS library, which is installed automatically with Data Services.

Related Topics

- [Supplement for SAP: Reading from SAP extractors](#)
- [Administrator's Guide: SAP integration](#)
- [Administrator's Guide: Integration overview](#)

2.3 Security

This version of Data Services includes several security improvements.

User management

SAP BusinessObjects Data Services now relies on a Central Management Server (CMS) provided by SAP BusinessObjects Enterprise or Information platform services for user management functions.

The CMS-based platform provides several benefits over previous versions:

- A common set of users for all Data Services applications, including the Designer, Management Console, and central repositories.
- Controlled repository access via both the Designer and Management Console.
- Support for alternate user authentication methods such as LDAP and Active Directory.
- Password enforcement policies such as password complexity and expiration.

For more information about user management, see the *Administrator's Guide*.

Secure Sockets Layer (SSL)

Communication channels between SAP BusinessObjects Data Services components running on different machines in a distributed installation can now be protected with SSL encryption.

A default set of certificates is included to provide a secure default configuration. You can also choose to use your own certificates, if needed.

For more information about SSL encryption between Data Services components, see the *Administrator's Guide*.

Enhanced password encryption

Improved password security is provided through new 128-bit AES encryption libraries. In addition:

- Each Data Services installation now uses a unique encryption key.
- Passwords stored in repository objects can be encrypted using a passphrase when exported to ATL or XML files to allow importing on a different system.
- A command-line utility can be used to encrypt passwords outside of the Data Services environment, such as in a custom application that you integrate with Data Services.

For more information about password encryption, see the *Administrator's Guide*.

Related Topics

- [Administrator's Guide: User and rights management](#)
- [Administrator's Guide: Configuring SSL for Data Services components](#)
- [Administrator's Guide: Password encryption](#)

2.4 Text Data Processing

Text Data Processing is now available on the Data Services platform to process textual content. Text Data Processing analyzes content and automatically identifies and extracts entities (people, dates, places, organizations, and so on) and facts (relationship between entities) in multiple languages.

To support Text Data Processing, Data Services can now extract from unstructured text sources. The new file format types Unstructured text and Unstructured binary let you read all files in a specific folder as long or BLOB rows respectively.

Related Topics

- [Designer Guide: Unstructured file formats](#)

2.5 Architecture

This version of Data Services includes several architecture improvements.

64-bit on Windows

A 64-bit job server is now available on Windows. Server components on all platforms are now 64-bit only; no 32-bit versions are available. For clients, a native 64-bit Designer is available in addition to the 32-bit Designer.

Adaptive processing server

Several SAP BusinessObjects Data Services processes are now run in the Adaptive Processing Server, which is hosted by SAP BusinessObjects Enterprise or Information platform services.

These processes include:

- RFC Server
- Administrator Service (used for log cleanup)
- View Data Service
- Metadata Browsing Service

Performance and memory utilization improvements

The following improvements have been made to improve performance and memory utilization:

- **Single version local repository.** Previously, the local repository created a new version of an object and also retained the older version. The repository no longer stores older versions, which reduces the size of the repository and eliminates the need to manually compact it.
- **Auto Documentation image generation.** The Auto Documentation feature in the Management Console generates images containing job information. The most commonly used icons in the mapping

tree are now pre-loaded to reduce memory usage. The new Image Display option lets you specify that a partial image is displayed for large images.

- **Export ATL/XML.** Exporting objects from the repository to a file now runs as a separate process, and no longer affects memory usage and processing in the Designer.

Related Topics

- [Administrator's Guide: Configuring Metadata Browsing Service and View Data Service](#)

2.6 Transforms

This version of Data Services includes several transform enhancements.

Enhancements to Data Quality transforms are described in the Data Quality section.

Enhanced Validation transform

The Validation transform has been enhanced as follows:

- The new interface provides a better overview on active rules.
- You can define multiple rules for a column.
- You can bind a rule to multiple columns.
- There's a new output path for Rule Violation value pairs for easier reporting on failed rules to help answer the question, "How do I know which rows failed for a particular rule?"
- Validation functions are a separate category of functions and can accept multiple input parameters. In the Validation transform, validation functions can be bound to input columns.
- Validation rules created in SAP BusinessObjects Information Steward can be shared with Data Services. Note: SAP BusinessObjects Information Steward is expected to release in 2011.

Improved joins in the Query transform

Data Services now provides full support for ANSI/SQL92 inner joins and left outer joins. These improvements provide greater pushdown capabilities and better performance.

The Query transform editor user interface has been improved to provide support for mixed join types and guarantee predictable results for the join set. The OUTER JOIN tab has been removed and the functionality in the FROM tab extended to enable you to specify all join types in one location and more accurately define restrictions. Additionally, for increased visibility, join rank and cache can now be set directly in the FROM tab of the Query editor.

Enhanced Hierarchy Flattening transform

The Hierarchy Flattening transform can now handle circular dependencies, rather than aborting with an error.

Related Topics

- [Reference Guide: Validation](#)
- [Reference Guide: Query Editor, Options](#)

2.7 Operational excellence

This version of Data Services includes several enhancements that improve the user's experience.

Designer keyboard accessibility

Data Services Designer is now fully keyboard accessible. The following features have been added:

- Navigation commands on the menu in the main window.
- Command equivalents of mouse operations in context menus.
- TAB key navigation in dialog windows.

Language packs

Additional languages to localize the Data Services user interface and documentation can be added to Data Services now. For availability of languages, see the *Product Availability Matrix* document located in the SAP Service Marketplace: <https://service.sap.com/PAM>. After you install the language pack, you can specify the language to use in the Locale Selector utility or the options menu in Data Services Designer.

2.8 Functions

This version of Data Services includes several enhancements to function support.

New cryptographic functions

Data Services now has cryptographic functions to encrypt and decrypt data using the AES algorithm.

- The key length used for the encryption can be specified as a parameter (128, 192, or 256)
- Based on the passphrase, a key with the required length will be generated. The passphrase is needed to decrypt the data again.

New gen_UUID function

Data Services now has a gen_UUID function to generate a Universally Unique Identifier.

- The ID is unique across space (host, process, thread) and time.
- The function is based on RFC 4122-Version 1 (Timestamp based).

BLOB data support

The base64_encode and base64_decode functions now support BLOB data.

Copy Query functions

You can now copy and paste from one Query to another function, such as lookup_ext, that is called from the Query's output schema.

Related Topics

- [Reference Guide: Functions and Procedures, encrypt_aes](#)
- [Reference Guide: Functions and Procedures, decrypt_aes](#)
- [Reference Guide: Functions and Procedures, gen_uuid](#)
- [Reference Guide: Functions and Procedures, base64_encode](#)
- [Reference Guide: Functions and Procedures, base64_decode](#)

2.9 Source and target support

This version of Data Services includes several improvements to source and target support.

Database datastores

- Database synonyms are now supported in Oracle and DB2 datastores.
- For Teradata datastores, more functions are now optimized for push-down. For quicker data extraction from Teradata, the Teradata fast export functionality can now be used.
- ODBC configuration has been simplified for Linux and UNIX platforms.
- Data Direct ODBC drivers are now packaged with the Data Services installation.

File reader enhancements

For fixed-width flat files, you can control blanks better with the new **Blank trimming** option.

Data Services can now extract from unstructured text sources. The new types Unstructured text and Unstructured binary let you read all files in a specific folder as long or BLOB rows respectively.

Web service datastores

Native web service datastores now support multiple configurations for easier migration between development, test, and production environments. You can change WSDL URLs for each datastore configuration without re-importing them.

Related Topics

- [Designer Guide: Datastores](#)

2.10 Data Quality

The set of transforms that work together to improve the quality of your data has been enhanced with the following features.

Data type definitions

In addition to two existing data type definitions for the Data Quality transforms (name and length), there are two new data type definitions to specify for Data Quality fields:

- **Type**—Used to specify the data type.
- **Scale**—Used to specify the maximum number of digits to the right of the decimal place.

These definitions are listed in the Input tab of the transform editor in the Transform Input Field Name and the Type columns. The Type column contains information for field type, size, and scale attributes in the format <type>(<length>, <scale>). The same format is used in the Type column of the Output tab. The name definition is listed in the Field_Name column.

Global Suggestion Lists

Global Suggestion Lists will be used only to query data due to the introduction of the Global Address engine's Suggestion List option. To view changes to the available options and components, see the *Data Services Reference Guide*. For additional information on updating your current Global Suggestion List processes, see the *Data Services Upgrade Guide*.

View directory expiration information

You can now view directory expiration information in the trace log. To include this information in the trace log, before executing the job, access the Execution Options tab and select Print all trace messages.

Support for NULL

A field that contains NULL is passed through processing with NULL preserved unless there is data available to populate the field on output. When there is data available, the field is output with the data available instead of NULL. The benefit of this treatment of NULL is that the software treats a NULL as unknown instead of empty.

Support for native data types

The Data Quality transforms generally process incoming data types as character data. Therefore, if a noncharacter data type is mapped as input, the software converts the data to a character string before passing it through the Data Quality transforms.

With this release, the Data Quality transforms support native data types. Therefore, when a transform reads an input field that contains an int data type for example, and it is mapped to a corresponding input field that also is an int data type, there is no conversion to character on input and reversion on output.

For more information about specific changes to support native data types, see specific Data Quality transforms in the *Reference Guide*.

Related Topics

- [Designer Guide: Data Quality data type definitions](#)
- [Reference Guide: Transforms, Data Quality Transforms, Global Suggestion Lists](#)
- [Designer Guide: Data Quality, Address Cleanse, View directory expiration dates in the trace log](#)
- [Reference Guide: Data Quality transforms](#)
- [Reference Guide, NULL values and empty strings](#)

2.10.1 Data Cleanse transform

The Data Cleanse solution has been significantly enhanced to include the following functionality:

- Introduction of Cleansing Package Builder in Information Steward empowers data stewards to work closely with Data Services administrators to easily and quickly create cleansing package solutions with a high degree of precision for any data domain. Cleansing Package Builder displays how data is parsed and standardized so users can evaluate the full impact of their choices.
- Universal Data Cleanse functionality in Cleansing Package Builder allows you to use your own sample data to build custom cleansing packages to efficiently parse and standardize any type of data. Data Cleanse can evaluate uncleansed data found in discrete or multiline fields and can recognize a range of alphanumeric values such as part numbers based on the definition of their pattern.
- SAP-supplied person and firm cleansing packages can be used as-is in Data Services or customized and published in Cleansing Package Builder and then referenced in Data Services.

Note:

Cleansing Package Builder is required in order to modify or customize any type of data. The dictionary menu has been removed from the Data Services Designer menu bar.

- Enhanced regional cleansing packages more accurately parse person and firm data according to the regional cultural standards. Newly available person and firm cleansing packages include:
 - Czech Republic and Slovakia
 - Malaysia and Singapore
 - Norway
- The name of the **Output text width conversion** option has been changed to **Character width style** to more accurately describe its use.
- Manual functionality moved from the Data Services Designer to Cleansing Package Builder now allows for the ability to automatically generate a single file cleansing package that is shared between Cleansing Package Builder and Data Services, thus providing easy management and improved security. The file contains all the information needed for Data Cleanse, including:
 - dictionaries
 - parsing rules
 - reference files including phone numbers, email, Social Security numbers and UDPM data

Note:

Existing Data Services 3.2 files can be migrated to Cleansing Package Builder.

Related Topics

- [Designer Guide: Data Quality, Data Cleanse, About cleansing data](#)
- [Reference Guide: Transforms, Data Quality transforms, Data Cleanse](#)

2.10.2 Geocoder transform

The Geocoder transform has been enhanced with the following features.

Point-of-interest and reverse geocoding

Previously, you used the Geocoder transform to input address data and append geographical coordinates to your data.

We have enhanced the transform so that you can now perform point-of-interest (POI) and reverse geocoding.

Point-of-interest geocoding returns a list of locations that meet your POI search criteria, within a geographical area. You can search with an address or geographical coordinates.

Reverse geocoding identifies the closest address based on an input latitude and longitude. The return address type can be points-of-interest, residential areas, or both.

Cascade geocoding

The Geocoder transform assigns the best coordinate for an address, in the following sequence:

- Primary Number Exact (Parcel or POI): assigns to the exact location of the address; for example, 123 Main St. This is the most precise level of assignment.
- Primary Number Interpolated: assigns to the level of the address range; for example, 100-500 Main St.
- Postcode: assigns to the level of the postcode.
- Locality: assigns to the level of city, town, or suburb.

You can determine the accuracy of the assignment based on the Assignment_Level output field.

Returns single or multiple records

Depending on your mapped input and output fields, the Geocoder transform can now output either single or multiple records. If the transform returns multiple records, the results are output into the Result_List output field as an XML string.

Generates report summary

The new Geocoder Summary report includes statistical information about geocoding.

Provides side-of-street information

The new Offset Coordinates option lets you specify whether the offset values of latitude and longitude are returned when the side of the street is known (left or right).

Supports multiple versions of data

For census data, you can use census data from two census periods to compare data, when available.

Note:

GeoCensus functionality in the USA Regulatory Address Cleanse transform will be deprecated in a future version. It is recommended that you upgrade any data flows that currently use the GeoCensus functionality to use the Geocoder transform. For instructions on upgrading from GeoCensus to the Geocoder transform, see the *Upgrade Guide*.

Related Topics

- [Designer Guide: Data Quality, Geocoding](#)
- [Reference Guide: Transforms, Data Quality transforms, Geocoder](#)

2.10.3 Global Address Cleanse transform

New suggestion list options in Global Address Cleanse transform

The Global Address Cleanse transform now supports the Suggestion List option. When the Suggestion List option is enabled, the Global Address Cleanse transform can return a list of assignment candidates to choose from. If you want to query the global address data, continue to use the Global Suggestion List transform.

Note:

The Global Address engine's Suggestion Lists option does not support processing of Chinese and Japanese address data.

Several new suggestion list options and output components have been added to support the Global Address engine's Suggestion Lists processing. For more information about Global Address Cleanse Suggestion Lists and Suggestion List components, see the *Reference Guide*.

Support for Chinese address processing

The Global Address Cleanse transform now supports processing of Chinese addresses. You must purchase the China data directories in order to process Chinese addresses. For more information about processing Chinese addresses, see the *Data Services Designer Guide* section entitled "Processing Chinese addresses."

Updates to increase usability and deployability

As part of this release, the Global Address Cleanse transform underwent modifications to increase usability and deployability. These modifications involved the consolidation of engines and options within the transform along with the elimination of the out-of-process architecture that existed in the version 3.x releases. In making these changes, some address assignment discrepancies were introduced when compared side-by-side with prior versions. This is especially true for certain EMEA regions. These discrepancies are being tracked and evaluated, and with continual efforts being made to improve and tune address assignment functionality, will be pushed into future version 4.x releases.

Number of processing engines reduced

The number of Global Address Cleanse transform engines and engine options are reduced. The Australia, EMEA, and Japan engines have been removed and the options required for processing were

moved to the Global Address engine options groups. The Global Address engine processes all addresses that were previously processed by the Australia, EMEA, and Japan engines. See the *Reference Guide* for the current list of Global Address Cleanse options.

New Character Width option

The **Character Width** option is now available for Japan and China address processing. This option is used to normalize Asian output characters.

Option values include:

- **FULL_WIDTH**: Converts all characters to full-width.
- **HALF_WIDTH**: Converts all characters to half-width.
- **NORMAL_WIDTH**: Converts half-width Chinese and Japanese characters to full-width and converts full-width Latin characters to half-width.

New Status and Information codes

The release includes three new status codes, five new information codes, and two updated information codes descriptions.

New status codes:

- **B**: Base record assignment for New Zealand. Bordering locality for Australia.
- **C**: An Alias and a Bordering locality. Global Address engine (Australia).
- **D**: Deleted record. Global Address engine (Austria and Germany).

New information codes:

- **3200 - INFO_CODE_BAD_OR_NO_BUILDING_NAME**: The building name is missing on input or not in the directory.
- **3220 - INFO_CODE_BUILDING_NAME_TIE**: Possible building names are too close to choose one.
- **3250 - INFO_CODE_NO_RANGE_OR_BUILDING_NAME_INPUT**: The range or building name is missing on input or both are not in the directory.
- **4520 - INFO_CODE_ORG_TIE**: Possible organization names are too close to choose one.
- **5030 - INFO_CODE_CERT_DISALLOW**: The country's postal authority does not permit assignment due to violation of an assignment rule.

Changed information codes:

- **3030 - INFO_CODE_BAD_OR_NO_PRANGE**: The primary range is missing on input or not in the directory.
- **4000 - INFO_CODE_BAD_OR_NO_SECONDARY**: The secondary information is missing on input or not in the directory.
- **3040** - This information code has been removed.

New output components

The following output components are available in the Global Address Cleanse transform:

- **LOCALITY1_FULL**: Includes Locality1, Locality Code, Locality1 Description, and Locality1 Qualifier.
- **LOCALITY2_FULL**: Includes Locality2 and Locality2 Description.
- **LOCALITY3_FULL**: Includes Locality3 and Locality3 Description.

- **LOCALITY4_FULL**: Includes Locality4 and Locality4 Description.
- **REGION1_FULL**: Includes Region1 and Region1 Description.

Enhanced data type support

There are changes to data types for some input and output fields in the Global Address Cleanse transform.

The input field named Suggestion_Start_Selection now has a field type of Int. This field indicates the starting suggestion list number. If the field is left blank, the default value is 1.

Two existing output fields, Count and Error, now have a field type of Int. Previously, they were varchar.

New SERP certification rule for Canada

The new SERP certification for 2011 requires the use of Point-of-Call (POC) reference data. The POC data has more specific primary ranges for the address compared to the non-POC reference data. The new POC reference data file is named `canadapoc.dir`.

New Disable Certification option for Canada

The Disable certification option has been added to the Canada engine options group. This option enables processing of Canada addresses using non-POC reference data for non-mailing purposes.

When you disable certification, you cannot print the SERP Report. Any list created with certification disabled cannot be used for mailing.

When the Disable certification option is set to "NO" the Canada engine will require the POC reference data.

Australia and New Zealand certified options

The Australia, New Zealand, and Canada processing options meet all requirements for AMAS, SendRight, and SERP 2010-2011 certification.

View directory expiration dates

You can now view directory expiration information for a current job in the trace log after you've processed the job. This information is helpful for you to assure you are processing jobs with the most up-to-date directories.

Related Topics

- [Reference Guide: Transforms, Data Quality transforms, Global Address Cleanse](#)
- [Reference Guide: Suggestion List \(Global Address Cleanse transform\)](#)
- [Reference Guide: Suggestion List Output Fields](#)

2.10.4 Match transform

New Proximity matching options

The Match transform now has the ability to match records based on their proximity. There are three types of proximity matching:

1. Geographic Proximity
2. Numeric Proximity
3. Date Proximity

Geographic Proximity

Finds duplicate records based on geographic proximity, using latitude and longitude information. The following new options have been added to support geographic proximity matching. The new options are located on the Match Editor's Options tab.

- **Geo Proximity.** Select this option to match on geographic proximity.
- **Distance Unit.** Specifies the type of distance unit used to calculate the distance between two latitude, longitude pairs. This option is only available when the Geo Proximity option is selected. Select one of the following:
 - Feet
 - Kilometers
 - Meters
 - Miles

- **Max Distance.** Specifies the maximum distance allowed when calculating the distance between two latitude, longitude pairs.

Any distance greater than the Max Distance receives a score of 0. A distance equal to Max Distance receives a score of Max Distance Score. Any distance less than Max Distance receives a proportional score between Max Distance Score and 100. Valid values for this required attribute range from 0 to the maximum double value.

- **Max Distance Score.** Specifies what score to generate when the distance is the same as Max Distance. Enter a value from 0 to 100.

A distance equal to Max Distance receives a score of Max Distance Score. Any distance less than Max Distance receives a proportional score between Max Distance Score and 100.

Numeric Proximity

Finds duplicates based on numerical closeness of data.

Date Proximity

Finds duplicate based on date ranges.

The following options have been added to support numeric and date proximity matching. The new options are located on the Match Editor's Options tab.

- **Numeric Difference.** Select this to match on numeric or date difference.
- **Max Difference.** Specifies the maximum difference allowed in a numeric range. Type a value from 0 to 100.

- **Max Difference Score.** Specifies what score to generate when the difference is the same as the Max Difference. Valid values for this required attribute range from 0 to 100. Any difference greater than the Max Difference receives a score of 0. A difference equal to Max Difference receives a score of Max Difference Score. Any difference less than Max Difference receives a proportional score between Max Difference Score and 100.
- **Numeric Percent Difference.** Select this to match on the percentage difference.
- **Max Percent Difference.** Specifies the maximum difference allowed as a percent of the absolute value. Enter a value from 0 to 100.
- **Max Percent Difference Score.** Specifies what score to generate when the difference is the same as Max Percent Difference. Valid values for this required attribute range from 0 to 100. Any difference larger than the Max Percent Difference receives a score of 0. A difference equal to Max Percent Difference receives a score of Max Percent Difference Score. Any difference less than Max Percent Difference receives a proportional score between Max Percent Difference Score and 100.

Support compound name matching: Approximate Substring Adjustment Score

A new Approximate Substring Adjustment Score option has been added to support matching of compound names found in countries such as Brazil and Mexico. The new option is located on the Match Editor's Options tab.

The option specifies what score to give to words that were not matched to the other words in the compared string. This option is used when comparing Family_Name1 and Family_Name2.

Enter a value from 0 (default) to 100. Enter a value of 0 to disable the option.

New output field

The Person1-3_Family_Name2 output field has been added for use in matching compound family names along with the new Approximate Substring Adjustment Score option.

Support for conditional matching

The following options have been added to support conditional matching. These options are located in the Match Level Options window:

- **Match on Street and RR, or on Box.** Specifies whether to match on PO Box only or on street, rural route, and PO Box. This option affects business and household records matching on address.
- **Address matches blank if Firms match.** Specifies whether to match on firm data when other address data does not match. This only effects records when one has street information and the other has PO Box information. If both records have Street information that do not match, or if both have PO Box information that do not match, the records will not be found as duplicates.
- **Match all resident names if RR, but no Box.** Specifies whether to match on rural route when an input record contains a resident-type name (Current Resident, Occupant, blank, or name not defined) and a rural route address, with no box number.
- **Ignore Firm if Name and Address match.** Specifies whether to ignore non-matching firm data if the name and address information matches. This option works with odd abbreviations or spellings of firm names. This assumes that you are matching on two real names, firm, and address line, and breaking or matching on city or ZIP Code.

Native data type support

Match accepts all usable data types natively supported by the Data Services platform. (Character, Integer, Double, Decimal, Date, Time, DateTime) and creates the appropriate data type on output.

Match Wizard

The following custom options have been added to let you choose any type of data to compare (in addition to party data such as name, address, email, and so on):

- **Custom option.** Select this option to match on custom criteria.
- **Custom name.** Enter a unique name for the custom criteria field.
- **Custom exactness.** If you choose to match on Custom, specify how close the data must be for that criteria in two records to be considered a match.
 - Exact
 - Tight
 - Medium
 - Loose

The following new Match options are also available for use in the Match Wizard:

- Geo proximity
- Numeric difference
- Numeric percent difference

For details on each new option, see the section on new Match options.

Related Topics

- [Designer Guide: Data Quality, Match](#)
- [Reference Guide: Transforms, Data Quality transforms, Match](#)

2.10.5 USA Regulatory Address Cleanse transform

The USA Regulatory Address Cleanse transform has been enhanced with the following features.

DSF2 Walk Sequencer

DSF2 Walk Sequencer is a new Data Quality transform. Use the new transform for assigning DSF2 walk sequencing information to your lists. Find the new transform in the software listed under "Data Quality" in the transform tab of the "Local Object Library".

DSF2 functionality

The software now supports processing your lists with DSF2 functionality. The software can assign both augment and sequencing information when set up in the new DSF2 Walk Sequencer transform and the USA Regulatory Address Cleanse transform.

Multiple data source statistics reporting

The transform can gather statistics for separate sources contained in one input list. Each list must contain a field that holds a unique identifier. The unique identifier can be a name or a number, but it must reside in the same field across all lists.

The software collects statistics for each list using the unique identifier. When the software generates reports, some of the reports will contain a summary for the entire list, and a separate summary per list identified by the unique identifier.

Field size and scale attributes

The software now has type, size, and scale attributes for all of the transform's input and output fields. These attributes appear in the Input and Output tabs in the transform under the **"Type"** column.

Changed field lengths

There are changes to field lengths for some output fields in the USA Regulatory Address Cleanse transform. Affected fields are Multiline1-12 (80 characters per field), Address_Line (90 characters), Locality2 (61 characters), and Suggestion_List (60,000 characters).

Changes to data types

The Suggestion_Start_Selection input field has changed from character to int data type. The Count and Error output fields have changed from character to int data types.

Support for NULL

The transform now supports NULL. When a field contains NULL on input, the transform outputs NULL. NULL is preserved only if the transform does not assign a value to the field as a part of address processing. In this case, the output is the assigned value and not NULL.

USPS certifications

You can use the transform to assist in applying for NCOALink and DSF2 licenses from the USPS. In addition, when you integrate the transform into your own software, you can use Data Services blueprints to apply for CASS self-certification with the USPS.

New suggestion list components option group

The Suggestion List options group in the transform has changed. Changes include new options, changes to existing options, and enhanced XML root elements.

The changes make SAP BusinessObjects Data Services XI version 14.0.0 non-backward compatible.

NCOALink Contact Details group

The option list for NCOALink in the transform has changed. Previously the Contact Details group did not have any subgroups. In the new release, the Contact Details group is now a subgroup under the Contact Detail List group.

View directory expiration dates in the trace log

You can now view directory expiration information for a current job in the trace log.

Changes to fields and options

- **Use Postcode Order Caching** option removed: The option has been removed because postcode order caching no longer provides a time saving benefit.
- **Address_Type OFFICIAL** changed to **Alias_Type OFFICIAL**: The **Address_Type OFFICIAL** field has been replaced with a new field named **Alias_Type OFFICIAL**. This new field contains the alias-type indicator for the input address that was assigned. The **Address_Type** output field with the **Field_Addrclass** of **Official** no longer exists in the transform.
- **Windows Use Extended Memory** option removed: The software no longer includes the option in the Transform Performance group. This option was only applicable for 32-bit systems and the new release of SAP BusinessObjects Data Services requires a 64-bit processor.

Related Topics

- [Designer Guide: USPS DSF2](#)
- [Designer Guide: DSF2, walk sequencing.](#)
- [Designer Guide: Multiple data source reporting](#)
- [Designer Guide: Data type definitions](#)
- [Designer Guide: Platform support for NULL](#)
- [Designer Guide: Completing USPS certifications](#)
- [Designer Guide: View directory expiration dates in the trace log](#)
- [Upgrade Guide: Changed field lengths](#)
- [Upgrade Guide: New suggestion list component options](#)

Index

64-bit 12

A

- accessibility 14
- adaptive processing server 12
- Administrator Service 12
- API
 - Operational Data Provider 9
- Approximate substring adjustment score option 21
- Australia certification 19
- Australia engine 19

B

- BAPI 9
- base64_decode function 14
- base64_encode function 14
- blank trimming 15
- BLOB support 14
- business content extractors 9

C

- Canada
 - Disable Certification option 19
- census data
 - multiple versions 18
- Central Management Server 11
- certifications
 - Australia and New Zealand certified 19
 - USPS 24
- Character Width option 19
- Character width style option 17
- Chinese address processing 19
- Cleansing Package Builder 17
- CMS 11
- compound name matching 21
- conditional matching 21
- content analysis 12
- CRM/SRM business content extractors 9
- cryptographic functions 14
- custom options 21

D

- Data Cleanse 17

- Data Direct ODBC drivers 15
- Data Proximity option 21
- data source statistics reporting 24
- data type
 - native 15
- data type definitions 15
- data type support 21
- data types changes 24
- database datastores 15
- datastores
 - web service 15
- Designer
 - keyboard accessibility 14
- directory expiration 15
- Disable Certification option 19
- DSF2 license 24
- DSF2 Walk Sequencer transform 24

E

- EMEA engine 19
- encryption 11
- engines 19
- ERP business content extractors 9
- expiration
 - directory 15

F

- field lengths 24
- file format type
 - unstructured binary 12
 - unstructured text 12
- file reader 15
- fixed-width flat files 15
- function
 - base64_decode 14
 - base64_encode 14
 - copying in Query transform 14
 - cryptographic 14
 - gen_UUID 14

G

- gen_UUID function 14
- Geo Proximity option 21
- Geocoder Summary report 18
- Geocoder transforms 18
- geocoding 18
- geographic proximity 21

- Global Address Cleanse
 - Suggestion List option 19
- Global Address Cleanse transform 19
- Global Suggestion Lists transform 15

H

- hierarchy flattening 13

I

- information codes 19
- Information Steward 17
- inner join 13

J

- Japan engine 19
- join
 - inner 13
 - left outer 13

K

- keyboard accessibility 14

L

- language packs 14
- left outer join 13
- local repository
 - ATL
 - exporting 12
 - Auto Documentation image generation 12
 - exporting
 - ATL/XML 12
 - repository
 - single local version 12
 - single version 12
 - XML
 - exporting 12
- log cleanup 12
- log, trace 15

M

- Match transform 21
- Match Wizard 21

memory utilization improvements 12
 Metadata Browsing Service 12
 multiple record output 18

N

name matching 21
 native data types 15
 NCOALink license 24
 New Zealand certification 19
 NULL 15, 24
 Numeric Proximity option 21

O

ODBC configuration 15
 Offset Coordinates option 18
 Operational Data Provider API 9
 Output text width conversion 17
 outputting multiple records 18

P

password 11
 performance improvements 12
 POI 18
 point-of-interest geocoding 18
 Primary Number Exact (PRE) 18
 Primary Number Interpolated (PRI) 18
 processing engines 19
 proximity 21

Q

Query transform 13
 copying functions 14

R

report
 Geocoder Summary report 18
 reporting
 multiple data source statistics 24
 Result_List 18
 reverse geocoding 18
 RFC Server 12

S

SAP NetWeaver BW 9
 scale attribute 24
 Scale data type 15
 Secure Sockets Layer 11
 security 11
 SERP report 19
 side-of-street information 18
 size attribute 24
 SLD 9
 SMD 9
 Solution Landscape Repository (SLD)
 9
 Solution Manager Diagnostics (SMD)
 9
 SRM business content extractors 9
 SSL 11
 staging BAPI 9
 statistics reporting
 multiple data source 24

status codes 19
 suggestion lists 19
 synonyms 15

T

Teradata datastores 15
 text data processing 12
 trace log 15
 trimming blanks 15
 type attribute 24
 Type data type 15

U

Universal Data Cleanse 17
 unstructured binary file format type 12
 unstructured text file format type 12
 USA Regulatory Address Cleanse
 transform 24
 Use Postcode Order Caching option
 24
 user management 11
 USPS certifications 24

V

Validation transform 13
 vintages of data 18

W

web service datastores 15
 Windows Use Extended Memory
 option 24