



FUSION REGISTRY

NEW FEATURES

FUSION REGISTRY
VERSION 10

New Features

CONTENTS

| | | |
|-----|---|---|
| 1 | Version 10.0 | 4 |
| 1.1 | Data Browser..... | 4 |
| 1.2 | SDMX-JSON | 4 |
| 1.3 | Kafka Producer Interface for Structures | 4 |
| 1.4 | Codelist Inheritance | 4 |
| 1.5 | Security Improvement: CORS Policy Management | 4 |
| 1.6 | Security Improvement: Stronger AES-256 bit Encryption for Sensitive Configuration Information | 4 |
| 1.7 | Active Directory Role Mappings CSV Import / Export | 4 |
| 1.8 | SDMX Data Availability API | 4 |
| 1.9 | Removed: Draft Structures | 5 |
| 2 | Version 9.8 | 5 |
| 2.1 | Data Validation Settings..... | 5 |
| 2.2 | Data Portal (beta)..... | 5 |
| 3 | Version 9.7 | 5 |
| 3.1 | SDMX REST API Compliance | 5 |
| 4 | Version 9.6 | 5 |
| 4.1 | Advanced Content Security: Fine-Grained Access Permissions | 5 |
| 4.2 | Support for Custom Database Connection Strings..... | 5 |
| 4.3 | SDMX Structure Mapping: N-to-N Mapping | 5 |
| 4.4 | SDMX Structure Mapping: Mapped and Unmapped Data Output Streams | 6 |
| 4.5 | SDMX Structure Mapping: Include Metrics..... | 6 |
| 5 | Version 9.5.1 | 6 |
| 5.1 | Annotations in Data | 6 |
| 5.2 | Fusion Edge Server Environments..... | 6 |
| 5.3 | Root User Account Independent from Authentication Service | 7 |
| 6 | Version 9.5.0 | 7 |
| 6.1 | Quick Convert..... | 7 |
| 6.2 | Added Fail-On-Error to Data Transformation | 7 |
| 6.3 | Excel Conversion – Report Converted Time Periods | 7 |
| 6.4 | Support First N and Last N Observations in same Data Query..... | 7 |
| 6.5 | Support Custom Database connection string | 7 |
| 6.6 | Excel Report Template – Variable Dimension | 7 |
| 7 | Version 9.4.4 | 7 |
| 7.1 | Support single Dataflow mapping to multiple Dataflows | 8 |
| 8 | Version 9.4.3 | 8 |
| 8.1 | Mapping: Support Group Attribute mapping to a Dimension | 8 |

| | | |
|-----|---------------------------------|---|
| 9 | Version 9.4.0 | 8 |
| 9.1 | Structure Synchronization | 8 |
| 9.2 | Validity of Items | 8 |
| 9.3 | Embedding Charts | 8 |
| 9.4 | Partial Code Mapping | 8 |
| 9.5 | Series Mapping | 8 |

Version History

| Version # | Implemented By | Revision Date | Reason |
|-----------|----------------|-------------------|----------------------------|
| 20190326 | Matt Nelson | 26 March 2019 | Initial Version |
| 20190503 | Phil Lazarou | 3 April 2019 | Release 9.6 New Features |
| 20190809 | Glenn Tice | 8 August 2019 | Release 9.7 New Features |
| 20190815 | Phil Lazarou | 15 August 2019 | Release 9.8 New Features |
| 20190927 | Phil Lazarou | 27 September 2019 | Release 9.8.3 New Features |
| 20200113 | Glenn Tice | 13 January 2020 | Release 10.0 New Features |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

1 Version 10.0

1.1 Data Browser

New stand-alone web tool for data search, visualisation and analysis.

Fusion Data Browser is a web application suitable for both internal and external users to search and explore the catalogue of datasets, select series and cubes of interest, build visualisations using tables and charts, and retrieve selected data in a variety of formats including SDMX, Excel and CSV.

1.2 SDMX-JSON

Support for input and output of SDMX data and structures using the SDMX-JSON v1.0 specification.

1.3 Kafka Producer Interface for Structures

Integration with Apache Kafka allowing Fusion Registry to act as a 'producer' where changes to structural metadata content are published on a specified Kafka topic.

Kerberos authentication is supported.

Apache Kafka is a high-performance stream-processing platform that in some ways acts like traditional message-queue middleware with publish and subscribe, but is instead designed for handling real-time data feeds. The main use case for the Fusion Registry Kafka Producer is to allow other processes attached to a Kafka platform to keep a copy of relevant SDMX structures, and to be notified when structures are updated in the master Fusion Registry.

1.4 Codelist Inheritance

Codelists can now extend other Codelists using an inheritance model. This is useful where an existing Codelist (perhaps published by a maintenance agency like the ECB) has most of the codes required, but where additional codes need to be added to satisfy a specific use case.

1.5 Security Improvement: CORS Policy Management

Cross Origin Resource Sharing (CORS) is a security mechanism which restricts whether resources on a web page such as dynamic calls to web services can be obtained from a server in another domain.

In Version 10, the CORS policy can be set explicitly on the Fusion Registry server to be either permissive (i.e. allow requests from any origin), or to only allow requests from web clients on specific domains.

1.6 Security Improvement: Stronger AES-256 bit Encryption for Sensitive Configuration Information

Sensitive information such as database access passwords that are stored by Fusion Registry as part of an installation's configuration are now encrypted using AES-256 bit.

1.7 Active Directory Role Mappings CSV Import / Export

Maintenance of large Active Directory mapping tables has been simplified by allowing the rule sets to be exported to, and imported from CSV. Previously, the only way to maintain the rules was using the Fusion Registry administration GUI which was time consuming for large tables.

1.8 SDMX Data Availability API

Fusion Registry 10 supports the SDMX standard Data Availability API. The API can be used by client data applications to discover what data is available for a particular dataset given specific dimension values, without having to query for the actual data.

A complete of the API can be found here: <https://github.com/sdmx-twg/sdmx-rest/wiki/Data-Availability>

1.9 Removed: Draft Structures

The concept of 'draft' structures has been removed. In Fusion Registry 9, structures could be marked as draft during development allowing them to be kept private to the owning Agency until published.

In Fusion Registry 10, access to structures can be controlled using Content Security rules, allowing them to be kept private until release by assigning them to an appropriate security group.

2 Version 9.8

2.1 Data Validation Settings

A new section has been added to the Registry settings to permit fine-grained control for data validation, allowing the specification over whether specific errors should be ignored, display an error, or block conversion or publication.

2.2 Data Portal (beta)

Allows structures and data to be obtained from an external SDMX Web Service and pulled into the current Registry. It is possible to specify that this process is performed at regular time periods.

3 Version 9.7

3.1 SDMX REST API Compliance

Updates have been made to the SDMX REST API to make it fully compliant with the latest version of the standard: <https://github.com/sdmx-twg/sdmx-rest/releases/tag/v1.4.0>.

4 Version 9.6

4.1 Advanced Content Security: Fine-Grained Access Permissions

By default, all Registry content is public. From Version 9.6, the Advanced Content Security sub-system allows access to specific structural metadata and data content to be restricted to particular groups of users.

Fine-grained rules can be defined, for instance controlling which user groups can see specific codes in a Codelist. A practical example is restricting access to confidential series with rules that limit visibility of the 'CONFIDENTIAL' code in the 'CL_CONF_STATUS' Codelist to authorised people. People without the correct authorities will be unable to see any series referencing the CONFIDENTIAL code.

Note that Advanced Content Security controls who can see data and structure content. Control over who can create and maintain structures, and load and modify data is still managed by granting Agency and Data Provider roles respectively through Active Directory or LDAP.

A video demonstrating the Advanced Content Security key concepts and features is available here:

<https://www.youtube.com/watch?v=yYkubDDR-Wo>

4.2 Support for Custom Database Connection Strings

Custom database strings can now be specified when creating connections to SQL databases. The feature allows administrators to take advantage of advanced database features like load-balancing and fail-over by including the necessary parameters in the connection string.

4.3 SDMX Structure Mapping: N-to-N Mapping

The SDMX Structure Mapping services have been enhanced to allow multiple dimensions on the source data structure to be mapped to multiple dimensions on the target.

Consider the following mapping:

A:B = C

This means that a source series with the values A and B in two specific dimensions should map to the value C in the target dimension.

N-to-N mapping rules support any number of dimensions on both the source and target so the following rule patterns are also valid:

A:B:C = D

A = B:C

A:B:C:D = F:G:H:I:J

N-to-N Mapping generalises and replaces the Series Mapping functionality introduced in Fusion Registry release 9.4 which allowed arbitrary series key to series key mapping. Series mapping can now be performed using the N-to-N mapping function by specifying complete series keys on both the input and output sides of an N-to-N rule.

4.4 SDMX Structure Mapping: Mapped and Unmapped Data Output Streams

Series not matching any mapping rules were previously discarded. In release 9.6, the SDMX Structure Mapping services have been enhanced to optionally output unmapped series on a separate stream by setting the 'Inc-Unmapped' header on the Data Transformation REST web service to 'true'. Explicitly outputting unmapped series allows further processing or corrective action to be taken.

The web service response format may be either multipart/mixed with mapped and unmapped data in separate parts, or Zip. If Zip is chosen, the result will be a single Zip archive containing two files: 'out' containing successfully mapped series, and 'unmapped' for series that could not be mapped.

4.5 SDMX Structure Mapping: Include Metrics

The Data Transformation REST web service now supports the 'Inc-Metrics' header. When set to 'true', the response will be a multipart/mixed message with a JSON section containing metrics on the processing time, the source data, the mapped output (series that have been successfully mapped) and any unmapped series.

5 Version 9.5.1

5.1 Annotations in Data

The Fusion Data Store now supports Annotations in Series and Observations.

To delete an Annotation from a Series, submit the Dataset with action DELETE, the dataset should contain the series with the Annotation(s) to delete, ensuring there are no Observations or series attributes present.

To delete an Annotation from an Observation, submit the Dataset with action DELETE, the dataset should contain the observation with the Annotation(s) to delete, ensuring there are no Observation attributes or series attributes present.

Stored Annotations can be queried from the Fusion Data Store by including the query parameter **includeAnnotations=true** as part of the query URL. This is documented in the Web Service Guide. Annotations are supported in both Structure Specific, Generic, and JSON formats.

5.2 Fusion Edge Server Environments

This feature is a step towards a larger feature to enable the user to select what data and metadata gets published to the Fusion Edge server.

The user is able to define multiple Fusion Edge Server environments, and select what content gets published to each environment (data/structures/reference metadata/provision agreements). The user is also able to add

an embargo time to the Edge Server publication file, this is read by the Fusion Edge server, and will not be made available until the embargo time has passed.

This feature is available under Admin->Edge Server

5.3 Root User Account Independent from Authentication Service

Prior to v9.5.1 the Fusion Registry had 2 concepts for security; local security, or an external authentication service such as Active Directory. Now the Fusion Registry has a concept of a single Root User account, which the user can log in with regardless of whether an external service is used. When a user logs in with the root account, the external service is not used for authentication. For additional security, the root user account can be set to lock itself after a specified number of failed login attempts – a locked account can only be reset from the database.

6 Version 9.5.0

6.1 Quick Convert

The user is able to drag and drop a dataset in the Data Load page in order to perform a conversion to SDMX-ML format.

6.2 Added Fail-On-Error to Data Transformation

An optional parameter available to the calling client to tell the data transformation process to fail if it encounters an error in the source dataset. This is documented in the Web Service Guide.

6.3 Excel Conversion – Report Converted Time Periods

For Excel Datasets report which time periods were converted (from period/to period). This is reported when using the Data Transformation User Interface, the information is also available from the Data Validation web service; the Web Service guide has more details on the Validation Service.

6.4 Support First N and Last N Observations in same Data Query

This is now supported in the Fusion Store, enabling queries to determine the earliest and latest observation date for any series. Previously the firstNObservations and lastNObservations was mutually exclusive.

Note: this is only supported in the Fusion Data Store

6.5 Support Custom Database connection string

Prior to this feature the Fusion Registry automatically created the database connection string using the information about Server, Port, and Credentials. Now the user is able to specify that they will provide their own connection string. This enables user to take advantage of advanced database features such as load balanced servers, where both servers are required in the connection string.

6.6 Excel Report Template – Variable Dimension

Previous to this feature all Dimensions with multiple possible values for the user to report were placed in the Excel table as a table row or column. Now it is possible to assign any number of Dimensions as being variable. A variable Dimension turns the Dimension into a drop down select in the Report, allowing the user to select one value for the Dimension that is applied to the entire report.

7 Version 9.4.4

Data Transformation: Map-Structure header to provide Structure Map URN

Previously this HTTP Header was expecting the URN of the Dataflow to map the data to. From 9.4.4 onwards it also accepts the URN of a DataflowMap or Data Structure Map to use in order to map the dataset.

7.1 Support single Dataflow mapping to multiple Dataflows

Prior to 9.4.4 the Dataflow mapping was supported as 1:N in the User Interface but the dataset transformation service did not support multiple Dataflows mapping to a single Dataflow. Now it is possible to load data according to one Dataflow and have the option of which Dataflow to transform the data into.

8 Version 9.4.3

8.1 Mapping: Support Group Attribute mapping to a Dimension

It is now possible to create a Dataflow Map or Data Structure Map which maps a Group Attribute in the Source Data Structure, to a Dimension Id in the target Data Structure. The Data Transformation service uses this information when transforming a dataset using the mapping rules

9 Version 9.4.0

9.1 Structure Synchronization

Allow the user to create define multiple Fusion Registry environments in their Fusion Registry server. They can then synchronize they environment with each target environment, to be able to PUSH or PULL structure to or from the target environment.

9.2 Validity of Items

Support the ability to add valid from and to dates to Items including Codes, Concepts, Mapped Codes (codelist mapping), and Hierarchical Codes. Each Item Scheme can be viewed in the User Interface with respects to a specific time period, where invalid items are not shown. Data validation takes into account the Observation time to ensure Code Ids for the Observation Attributes, or Series Dimensions/Attributes are valid for that specified period.

Note: The validity information is stored as an Annotation in cases where this is not natively supported by SDMX.

9.3 Embedding Charts

The updated data browser User Interface now supports AMCharts version 4, with the ability to embed charts in an Iframe or from a URL.

9.4 Partial Code Mapping

This enables a user to create a mapping between to Codelists, where a substring of the source code Id is used to map to one or more target codes. For example if the Code Id starts with 'ABC' then the target Code Id is either aaa, bbb, or xyz.

9.5 Series Mapping

This feature enables a user to create a Mapping between two Data Structures at the Series Key level instead of at the individual Code level. For example input series A:UK:EMP maps to A:044:GB:TW, a subsequent version then introduced the concept of wildcarding these maps. So for example if the previous example the code 'A' represents the Frequency Dimension's code of Annual, the user could specify *:UK:EMP -> *:044:GB:TW to represent that any input frequency maps to the same frequency on the output.