



European Commission
Directorate - General for Informatics
Directorate B – Digital Services
DIGIT B2 - Interoperability
Digital Europe Programme

A07.01 – Development of a SEMIC registry of data models

**Action ‘Promoting semantic interoperability amongst
the European Union Member States (SEMIC)’**

Specific Contract 098 under Framework Contract BEACON

Date: 16/02/2024

Doc. Version: v0.1

DOCUMENT CONTROL INFORMATION

Settings	Value
Document Title:	Development of a SEMIC registry of data models
Project Title:	Action 'Promoting semantic interoperability amongst the European Union Member States (SEMIC)'
Document Author(s):	Nathan Ghesquière Emidio Stani Joren Verleyen
Project Owner (European Commission):	Alexandra Balahur
Project Manager (European Commission):	Claudio Baldassarre
Contractor's Project Manager (CPM):	Florian Barthélemy
Doc. Version:	0.1
Sensitivity:	Limited
Date:	16/02/2024

REVISION HISTORY

Date	Version	Description	Author(s)
16/02/2024	0.1	First version	Nathan Ghesquière Emidio Stani Joren Verleyen
19/02/2024	0.1	Review and approval	Claudio Baldassarre
1/03/2024	0.2	Revisited	Claudio Baldassarre

DOCUMENT LOCATION

The latest version of this controlled document is stored in
<https://citnet.tech.ec.europa.eu/CITnet/confluence/display/SEMIC/A07.01+Development+of+a+SEMIC+Registry+of+Data+Models>

Table of Contents

1	Introduction.....	4
1.1	Context and scope of this document.....	4
1.2	Objectives of the registry	4
1.3	Outcome and benefits of the registry	4
2	Users categories for the registry	4
2.1	Categories.....	5
2.2	Potential roles in the registry	5
2.2.1.	Model owners.....	5
2.2.2.	Model publishers	5
2.2.3.	Registry adopters	Error! Bookmark not defined.
3	Use cases.....	6
3.1	Context	6
3.1.1	Use case 1: Publication.....	6
3.1.2	Use case 2: Discovery	6
3.1.3	Use case 3: Adoption.....	6
3.1.4	Use case 4: Collaboration	6
3.1.5	Use case 5: Integration.....	7
4	Registry vs repository	8
4.1	Introduction.....	8
4.2	Connecting national repositories to the EU-wide Registry of semantic models.....	8
5	Next steps.....	8

1 Introduction

1.1 Context and scope of this document

In the rapidly evolving landscape of digital governance and interoperability initiatives within the European Union (EU), many Member States are actively pursuing the objective to become interoperable within national border and cross-boarder. Part of the effort relies in the adoption of existing data models of shared semantics. This pre-requisite gives rise to the need, for any Member State, to discover what are the existing data models currently in use by others. This information become useful at the time of designing a new model that captures a domain in the public sector. Where Interoperability is our objective, discoverability, as well as Findability, Accessibility, and Resuability of the data models become issues to tackle. Recognizing this challenges, SEMIC endeavours to address the rising requirements, and interest from the community, for a registry of semantic models to become the centre of semantic models discoverability across the Member States.

1.2 Objectives of the registry

In a first phase, The key objectives for the registry, are Findability and Accessibility of semantic models across Member States in the EU. In a second phase the registry will grow to offer advanced features to combine existing models and design new data models aided by AI. It will also adopt the principles for knowledge elicitation that will improve semantic models designed across-boarder language and culture. The registry will consider enhancing user experience for several identified categories of data practitioners in different fields of public sectors where data are key (e.g. statistics).

1.3 Outcome and benefits of the registry

The outcomes of the registry are envisioned as follows:

- Standardization and Harmonization: Establishing a standardized repository for semantic models to promote consistency, harmonization, and alignment of vocabularies, ontologies, and data schemas.
- Findability and Accessibility: Facilitating the discovery and retrieval of existing semantic models by providing a platform that register the content of national repositories.
- Reusability: Promoting reusability of semantic models through consistently selecting the models that shows optimal design.
- Collaboration: Support knowledge-sharing and co-design among Member States as by product of discovering each other's work on modelling the public sector.

2 Users categories for the registry

This section highlights the main users' categories that classify the knowledge workers who can use the features made available by the registry. This information is also used to design the registry features with best fit in mind for the identified categories.

2.1 Categories

While the registry would be accessible for everyone, it should fulfil the needs of different user categories:

- Public administration workers: are responsible for overseeing the activities in a specific office. They can formulate model requirements in the form of short documents, or competency questions. They can perform model search, selection and pass this information to the Model Engineers.
- Model engineers: possess technical expertise and experience, to process the requirement generated by the administrator worker, iteratively transform the requirements into conceptual models, find models that contains that information and carry the same intended semantics, can combine models, can extend, or specialize models, can reiterate their result with the administrator work to see if the result cover the information need initially expressed.
- Software developers: can streamline the publication of a new model from and to the registry, with the use of dedicated API or web services made available.
- Registry casual browser: are able to complete a search, analyse the result, make a selection of models, follow link to their sources, share the models with a co-worker.

2.2 Potential roles in the registry

Thinking of the lifecycle of a semantic models exposed via the registry, we can advance users' roles associated to governance aspects of the registry.

2.2.1. Model owners

The owner of the original model intended as the authoritative source of the model semantic concepts and relations definition. Model owners are the ones who manage the lifecycle of the models in their repository, and they exercise the ultimate right on the model presence via the registry.

2.2.2. Model publishers

The model publishers consist of users that expose their semantic models to the registry. The publishers of these semantic models have a knowledge of the domain represented by the model and can compile metadata that will make the model findable and accessible on the EU-wide registry.

Model publishers can belong to:

- Government agencies: On either the regional, national, and even the EU-wide level, governmental bodies that engage in semantic standardization activities should have the possibility to make models available.
- Research institutions: Institutions and researchers that have created semantic models as part of their research could make their models available on the registry.

3 Use cases

3.1 Context

The following use cases can be seen as part of one storyline. It concerns 3 agents, namely Member State X, Member State Y and Member State Z. To make these use cases more concrete, the model of which will be spoken in the use cases is from a certain topic which will remain the same throughout the different use cases (e.g. public healthcare).

3.1.1 Use case 1: Publication

Member State X wants to make a certain model available on the EU-wide registry. For the model to be made available on the registry, a model engineer and a model publisher must properly decorate the model with publishing metadata in a format that is processable by the registry. The model will be published to the registry with a connection via API or web service or web interface. Member State X now made their semantic model available for all registry users.

3.1.2 Use case 2: Discovery

Member State Y wants to customise a semantic model for their public healthcare domain but wants to align with other customisation efforts done by other Member States, in particular its close neighbours. The model uploaded on the EU-wide registry by Member State X would be a good foundation for Member State Y to start its modelling activities from. To use the model by Member State X, a public administrator from Member State Y must search for the specific model it wants to use by searching for the models' characteristics, domain, used classes and/or attributes, etc.

3.1.3 Use case 3: Adoption

After filtering on certain criteria, the public administrator from Member State Y finds the model it wants to use. It can access the URL that directs the user to the storage location of the model. It can then download, or refer to it, using the URL provided by the registry, after which the model can be used in a local environment. After the model engineer from Member State Y is done with adapting the model to its own requirements, the public administrator can make their model available to the registry.

3.1.4 Use case 4: Collaboration

Member State Z sees the models made available by Member State X and Member State Y. Member State Z proposes the co-design of a model for another, related topic (e.g. environmental health). They collectively gather information and related models already made

available on the registry by other Member States. Member State Z would, after the model is finished, make the model public on the registry in the designated community.

3.1.5 Use case 5: Integration

A governmental body within Member State Z wants to create systematic access to the registry to get easy access to certain models. For this, a software developer from Member State Z creates an integration to the registry via API or web services. Consequently, Member State Z can obtain up to date information on the latest semantic modelling efforts from other actors in their own environment.

4 Registry vs repository

4.1 Introduction

It is important to highlight the difference between a registry and a repository. While both serve as central hubs for organizing and accessing data models, the way they function and the objectives they fulfil are fundamentally different.

A registry primarily serves as a centralized directory or index that catalogues metadata about data models, ontologies, vocabularies, without physically storing said semantic models.

In contrast, a repository functions as a centralized storage facility for storing and managing actual data artifacts, including data models, ontologies, vocabularies, and related documentation. Unlike a registry, which focuses on metadata management, a repository hosts the actual content of semantic assets, making them accessible for download, reuse, and integration into various applications and systems.

Public Administrations (such as Italy, France, Norway, Finland, etc.) together with existing public repositories (such as LOV, TNO, etc.) have created their own repository of semantic data models allowing end users to be able to find data models. However, the implementation of each repository is different, as well as the way end users can find a data model.

4.2 Connecting national repositories to the EU-wide Registry of semantic models

The registry would be inherently built to make the connection with National repositories of semantic models. This implies that the model publishers (as described in section 2.2.1) should have a repository in place where they can store their semantic models, ontologies, or vocabularies, and that can be connected to the EU-wide Registry.

To this point, there can be different ways of exposing the decentralised information to the registry:

- Upstream, where data flows from Public Administrations to the European Registry. Here, the model publisher pushes their models to the registry through a standardised API. This is the same [method used by data.europa.eu](#).
- Downstream, where data flows from the European Registry to the Public Administrations. In this case, the registry would crawl the models from different repositories.

Employing the upstream flow to streamline data about the models to the Registry is favoured. Consequently, by providing an interface to the European Registry, Public Administrations will be able to connect to the registry by providing their models metadata.

As an example, the same [approach](#) has been adopted by data.europa.eu which allows Member States to retrieve and upload automatically metadata regarding their dataset via API.

5 Next steps

This document will be verified with the active working group in March 2024. Once the scoping and conceptualisation laid out in this document are verified, the working group will

focus on the features and technical considerations for the registry in May 2024. Any feedback is welcomed in the ['issues'](#) tab of this Github page.