Meeting Minutes: Technical requirements workshop for the registry for semantic assets (SEMIC - A07.01)

Project:	SEMIC	Date and Time:	29/01/2024 11:00 - 12:00
Meeting Type:	Webinar	Location:	Virtual
Coordinators:	Claudio Baldassarre Nathan Ghesquière	Issue Date:	04/06/2024

Agenda of the webinar		
10:00 - 10:05	Introduction	Slides 1 - 3
10:05 - 10:15	Summary pre-conference workshops	Slides 4 - 5
10:15 - 10:50	Purpose and objectives	Slides 6 - 12
10:50 - 11:00	Summary and next steps	Slides 13 - 16

Meeting Slides
LINK

Participants		
Name	**Initials**	**Organisation**
Ana Rosa Guzman	ARG	Ministery of Territorial policy and Public Service
Claudio Baldassarre	CB	DIGIT
Claudia Cristina Pollina	CCP	The Department for Digital Transformation

Darius Amilevicius	DA	Committee for the Development of the Information Society
Denis Dechandon	DD	Publications Office of the European Commission
Emidio Stani	ES	PwC EU Services
Florian Barthelemy	FB	PwC EU Services
Javier Andrés Arroyo	JAA	Ministery of Territorial policy and Public Service
Joren Verleyen	JV	PwC EU Services
Mikael af Hällström	MaH	University of Helsinki
Matteo Fortini	MF	The Department for Digital Transformation
Martynas Mockus	MM	Information Society Development Committee
Mantas Zimnickas	MZ	Information Society Development Committee
Norman Calleja	NC	MITA (Malta Information Technology Agency)
Nathan Ghesquière	NG	PwC EU Services
Pascal Derycke	PD	Sciensano
Riitta Alkula	RA	Digital and Population Data Services Agency
Sebastian Sklarß	SS	Init
Sander Van Dooren	SVD	Flemish Government
Torbjörn Ull	TU	Swedish Companies Registration Office

Full Meeting Minutes

Introduction Slides 1 - 3 **Speaker:** Claudio Baldassarre	Claudio Baldassarre (CB) opened the meeting and went over the agenda for the opening meeting: • Recap of the first workshop and the requirements distilled from there • Metadata • Architecture • Summary and next steps

	CB mentioned that the responsibilities from the working group are important, as the EU-wide registry for semantic models will be a community-driven effort with intervention from the SEMIC-team.
Summary pre-conference workshops Slides 4 - 5 **Speaker:** Claudio Baldassarre	CB began by saying that during reviewing the presentations given during the pre-conference last year, SEMIC discovered that there were some interests in this subject. • SEMIC discovered that access to models from other Member States is considered very important. While working on the EU-wide registry, semantic interoperability will be pushed, and everybody will converse to using interoperability. Doing so would increase the reuse and adoption of the same models. • It was determined that the social factor plays an important role in deciding which models to reuse, and it should follow scientifical rules within a framework. • The technical considerations were also considered very important, as Member States should know how to implement a registry and how to potentially combine it with additional features. From these presentations came different considerations and conclusions to keep in mind when designing the EU-wide registry, CB continued. It must serve everybody (so must be in one language) and there should be a governance for Member States to be able to contribute their models. This registry should be a collaborative effort of all the Member States active in the working group, so that the registry will be kept in a healthy way for a long time.
Purpose and objectives Slides 6 - 12 **Speaker:** Nathan Ghesquière	Objectives Nathan Ghesquière (NG) said the most important objective from this working group is to allow Member States to discover what other Member States are modelling. Member States should be able to look for specific models if they have their own registry. The EU-wide registry should be a place where Member States could connect their repositor(y/ies) to. Lastly, the aspect of publishing should be considered. All Member States should be able to publish the models they choose to make accessible for everyone. Working group participants Next up, NG continued by going over the people of whom this working group will consist of. This includes the SEMIC-team, which will drive this working group, Claudio Baldassarre, who is the overall coordination, and participants from the Member States. Activities NG stated that for the activities during this working group, two types of platforms will be used: • *Offline collaboration*: GitHub will be used in between working groups. This platform can be accessed at any time and you can log your feedback offline with the issues-function.

- *Online collaboration*: During a workshop, interactive and collaborative tools will be used, such as Mural for example.

The status of the Member states should be of a mixed level of maturity. It is valuable to have inputs from every Member State and their various levels of maturity, as they could give different insights. The creation of the registry is community-driven, which means it is driven by the users of the portal. Member States willing to participate in this working group must be willing to collaboratively shape this registry and move forward with discussions about making a lightweight pilot in Q3 and Q4 of this year.

Meeting frequency

For the meeting frequency, a general overview is given for when every workshop and meeting will take place.
- The first workshop, which will be held in March, is the first interactive workshop content wise. Here, the adoption requirements and considerations needed for the creation of the registry will be discussed. The difference between a registry and a repository will also be made clear.
- In the second workshop, which will be held in May, the technical requirements will be discussed. These include among others, but not limited to APIs and metadata.
- A closing meeting will be held in June where a brief closing will be held. By then, a roadmap should be able to be put forward so, together with the Member States, the next phase of building a future pilot could be started.

NG then asked if everything was clear and if participants had any thought on the preliminary agenda and timeline.

Sebastian Sklarß (SS) asked what the requirements are for going beyond the possibilities of the Joinup platform. He also mentioned that some people might think that Joinup is a portal for exchanging data models.
- CB answered that Joinup and the registry are two separate products. The registry is an implementation, similar to the LOV-portal, where users search for models. It can be seen as something similar, but with input from the Member States the registry would be customised to the needs of the Member States. Joinup is a platform where SEMIC promotes their content.
- Florian Barthelemy (FB) added that while determining the requirements for the registry is different from the use of the Joinup platform, Joinup could still be used.

Riitta Alkula (RA) asked about the cooperation between the EU-wide registry and the existing repositories from the Member States. She stated that the target of this working group should be clarified, so it would be clear if we are going to start a registry from scratch or look into a solution from the Member States where the registry will be (partially) based on something that already exists.

- CB answered that there is a difference between a registry and a repository. A repository is the place where a Member State stores its models and the EU-wide registry is the place where these repositories would be connected. We recognize that some Member States do not have the maturity for a repository yet, but aiding Member States who do not have a repository yet is not the priority of this working group at this point. SEMIC does not develop its own software but rather uses tools which are open source from the community.

Torbjörn Ull (TU) wanted some clarification if what we would build, would be some kind of portal.
- CB clarified that it would be a portal with one front end. He also added that in the long term, other ways to interact with the interface (e.g. APIs, from the source model, ...) would not be excluded. For now, the preliminary ideas are put forward but together with the working group, the scope the Member States had in mind will be open for discussion. Member States would have the option to choose which models they would make accessible on the registry. This choose would be reflected on the interface we would build.

SS wanted to know how this activity is linked to the 'semantic repository' in EU OOTS that currently is being built for SDG under the responsibility of DIGIT (Pavlina).
- CB stated that this registry would go beyond the scope of OOTS. The semantic repository would be used as input for the EU-wide registry.

Mikael af Hällström (MaH) mentioned domain-specific coordination with governance where models are developed together with other Member States. He said that there should be cross-border coordination so not only national models would be available to be searched on the registry.
- CB recognises the potential and asks Mikael how this cooperation happens and how the registry could support this.
- FB added that it cocreation is indeed important.
- MaH answered that the Finnish team convinced Nordic colleagues to use a common platform. He reiterated the importance of being able to reuse solutions which have been cocreated by Member States. He also stated that it should be possible to search (and find) models from outside of the EU.
- CB said that for the registry, part of the process is a phase of exploration of models from a certain topic. There is a possibility for the registry to onboard Member States to upload models. This example could be used as a use case for the first workshop because it is representative of what we want to achieve with the creation of the registry. If a Member State wants to create a model within a certain domain, it should search for models of other Member States to see how they achieved their models. The EU-wide registry will do this but add one element, namely that the registry would provide a

window for all Member States who have done this exercise already.

- MaH said it could be a good idea to, apart from the creation of the EU-wide registry, create domain-specific communities where the contact information of the creators of public models would be available. This way models could be more easily aligned.

TU added that once a specific domain is identified, the gaps should be filled with concrete semantic work. For this to be done more easily, it should be possible to see what other Member States have achieved in this specific domain.

- CB stated that adding domain-specific communities would be very beneficial to this cause.

SS asked about the social aspects of the registry and how they would be reflected in the workshop planning.

- CB answered that during the pre-conference workshops, a panellist (Ana Rosa Guzman) presented her findings about the obstacles with developing models during collaboration. Some aspects, such as linguistics or cultures would make it that people have trouble understanding each other. The solution was to find a common thread and let it serve as a base from which you could build your model together. This way you could create a mutual understanding of the specific domain. The slides of the pre-conference workshops contain all this information more detailed and will be uploaded to the GitHub repository to fall back on.

Participants introduction

After these questions, NG asked the participants to present themselves, say what their relation to the registry would be and how they would see their involvement in this working group.

- RA explained how they use their platform and how the data models are resolvable. She is searching on how to make the platforms they use more technical interoperable (apart from APIs), so something could also be extracted other than via downloads.
    - NG answered that this application is something that will be discussed in the second workshop, which is about the technical requirements.
- Claudia Cristina Pollina (CCP) said that she is looking forward to collaborating with Italy's national registry. She said she is willing to bring the experience she has to offer and participate in the cocreation design process. She also said she is eager to learn what other member States are doing. She wanted to help with the creation of the registry from the point of view from developers who will also be using these semantic assets. She is, together with Matteo Fortini, part of the statistical team and could invite other people from the team to participate in this working group.

	• Martynas Mockus (MM) said that they are working on a project that will describe the 400 information systems from the public sector in metadata standards since 2006. They are also creating an ontology which involves these information systems. A new law was put in place last December that says these assets must be interoperable, Martynas continued. They do not have an advanced platform yet apart from the use of APIs and looking a bit deeper into data-items and DCAT-AP. Martynas concluded that the semantic assets and everything from SEMIC is important to them because they will invest a lot of money to describe the metadata in a standardised way. • Pascal Derycke (PD) stated that he works for Sciensano, a public health institute. He also works on a similar project for DG Santé, leading a health DCAT-AP for the creation of a health data space. They are supporting health data engine, which aims to build a metadata catalogue for health. He said that he would not be able to make it to the first workshop in March. Nathan replied that it is okay to send someone to the meeting in his place. After the participants introduction, SS asked if the contact data from the people in the working group could be shared. • FB answered that the SEMIC-team is supporting the overall health DCAT-AP modelling process. In addition to us, Bert Van Nuffelen and Jitse De Cock are involved in this and can be contacted as well.
Summary & next steps Slides 13 - 16 **Speaker:** Claudio Baldassarre and Nathan Ghesquière	Closing To close this meeting, SEMIC would like to say that any absences would be facilitated by publishing the scope and the preliminary considerations on GitHub. This is also the platform that will be used during this working group and where you could comment via the 'issue' section. If anyone would have some problems regarding access, do not hesitate to reach out so we can solve the problem. First workshop overview A follow-up e-mail will be sent by the end of February, giving an overview of the first workshop. This way you could see the agenda and raise any comments or questions you might have. CB reiterated to share elements you would want to be taken in consideration. If you could send us a reference document, implementations on the roadmap can be made. To achieve maximum value coming from this working group, staying in touch during the entirety of this working group is important.