Federation of Semantic Assets Repositories
The SEMIC.EU approach
v.1
May 2011
Federation of Semantic Assets Repositories

1. Background

Currently, SEMIC.EU provides a repository infrastructure storing over 500 reusable semantic assets. Semantic assets are considered key resources for achieving semantic interoperability and can be in the form of data models, taxonomies, ontologies, code lists and semantic data exchange formats. When the project started four years ago, SEMIC.EU has been a one-of-a-kind infrastructure. However recently, in the European Union there is a growing number of national web resources (repositories) storing, indexing and making available semantic artefacts and metadata with the goal to facilitate and enable seamless data exchange. The artefacts they store come in various formats (e.g. XML schemas, csv, txt, RDF). They are covering a wide range of domains and are implemented with different technologies, languages, and modelling formalisms. These repositories are implemented, populated and maintained as separate entities at the national level. They are usually central parts of National Interoperability Frameworks, Metadata Standards Catalogues, eGovernment strategies etc. Nevertheless, from a European perspective they can be described as unlinked, non-interoperable metadata containers as it is not possible for someone to have an overview on all possible assets available from this diverged set of repositories.

2. Semantically interlinking Assets Repositories

This fragmentation should be addressed to allow seamless retrieval of information on reusable semantic artefacts in the MSs. The solution should ideally be technology and platform-independent to allow users all over Europe search for relevant resources across various repositories. For this to happen, a federation of repositories shall be implemented to make decentralised resources available through a single point of access.

This federation has to respect the autonomy of each repository but at the same time allow the cross-querying and discovery of relevant assets stored in different places. A common schema for assets description that allows precise and rich description of repositories content is a key enabler of seamless data exchange. The approach is to semantically interlink the federated repositories, preserving their autonomy and decouple while at the same time avoiding imposing specific technologies.

To implement such type of federation the following steps should be followed:

- Draft and agree on an Asset Description Metadata Schema (ADMS). ADMS describes in a generic way a "semantic asset" in a similar sense that Dublin Core provides a generic description of a resource.

- Annotate the nationally stored semantic assets with ADMS and make them machine-readable. As soon as ADMS is been adopted, existing national repositories could:
  
  a) Map their own representation of a "semantic asset" to ADMS.
b) When this mapping becomes available, resources (semantic assets) stored locally can be exported as ADMS-annotated data using a machine-readable format (e.g. RDF).

c) This exported data can be browsed and queried while users could in a transparent way identify the assets that are relevant to their interests.

3. **Work so far and next steps**

We try to bring together representatives from our MSs, semantic assets repository owners, metadata and vocabulary experts and standardization bodies to draft a simple but highly reusable schema, by building on top and reusing existing work in the area (e.g. DCAT\(^1\), DCMI\(^2\)).

The federation of assets repository through the use of ADMS has been already discussed with the MSs representatives during the last ISA Trusted Information Exchange WG, 3\(^{rd}\) March 2011. The feedback has been positive and supportive.

Work has already started to define *Asset Description Metadata Schema*\(^3\). The figure below gives the first draft specification that is currently under discussion.

![Draft ADMS model expressed in RDF](image)

**Figure 1:** Draft ADMS model expressed in RDF

A small prototype is being developed with the participation of a small number of repositories to test the feasibility and benefits of the approach. Based on the lessons-learnt, we intend to invite all repository owners in MSs to describe and export their assets as ADMS-annotated metadata. SEMIC.EU will provide a browsing and querying mechanism over the aggregated ADMS metadata.

4. **Benefits**

- Definition and adoption of ADMS as a common schema to describe semantic assets. ADMS can be used as a common language to describe semantic artefacts stored in different systems. It can also provide directions and could be used as a template for new initiatives to create similar repositories at the national level.

- Semantically interlinking SEMIC.EU with national asset repositories across the EU, creating a flexible federation which facilitates cross-repository search and consequently boosts reusability of existing semantic assets.

- SEMIC.EU becomes not only a semantic asset repository and clearinghouse but also a semantic asset portal and single point of access for assets stored in national repositories.

---

\(^1\) [http://www.w3.org/egov/wiki/Data_Catalog_Vocabulary](http://www.w3.org/egov/wiki/Data_Catalog_Vocabulary)

\(^2\) [http://dublincore.org/specifications/](http://dublincore.org/specifications/)

5. **Draft Development plan**

   a) ADMS first stable draft is already available as ADMS v0.6
   b) Small prototype with 3 repositories to be implemented - Q2 2011
   c) All national repository owners to be contacted to participate in the implementation - Q2 2011
   d) The storing, browsing and querying infrastructure for the ADMS-annotated semantic assets descriptions to be deployed and become available – Q2 2012

**ISA Contact:**

Vassilios Peristeras, DG DIGIT, ISA Unit, vassilios.peristeras@ec.europa.eu
Montoyer (MO) 34, Floor: 3, Office: 103; B-1049 Brussels, Belgium
Tel: +32 (0) 2 29 81014