Building Semantic Interoperability in Europe

Federation of Semantic Asset Repositories

23 April 2012

Joao.Frade@pwc.be

@fradester
Agenda

Promoting Open Government Metadata

The need for ADMS

ADMS in practice

Get involved
ISA is undertaking a number of initiatives around Interoperability

What is interoperability?
Ability of disparate organisations to interact towards mutually beneficial and agreed goals, involving the sharing of information and knowledge
### Political context

Cooperating partners with compatible visions, aligned priorities, and focused objectives

<table>
<thead>
<tr>
<th>Interoperability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal Interoperability</td>
<td>• Aligned legislation so that exchanged data is accorded proper legal weight</td>
</tr>
<tr>
<td>Organisational Interoperability</td>
<td>• Coordinated processes in which different organisations achieve a previously agreed and mutually beneficial goal</td>
</tr>
<tr>
<td>Semantic Interoperability</td>
<td>• Precise meaning of exchanged information which is preserved and understood by all parties</td>
</tr>
<tr>
<td>Technical Interoperability</td>
<td>• Planning of technical issues involved in linking computer systems and services</td>
</tr>
</tbody>
</table>
What is a semantic asset?

A semantic asset is

- highly reusable metadata (e.g. xml schemata, generic data models)
- & reference data (e.g. code lists, taxonomies, dictionaries, vocabularies)

Semantic agreement(s)

- Organisations use semantic assets to share information and knowledge (within themselves and with others)
- Interoperable interfaces depend on semantic assets
Different metadata and reference data cause **semantic conflicts**

This is caused by unnecessary redundancy, variants, lack of awareness and management of semantic assets
What is semantic asset management?

• adopting policies, processes, and tools (repositories)
• to plan, evaluate, and improve the use and reuse of semantic assets to share information and knowledge

This means that semantic assets should be:

• properly documented and stored in a repository (shared drive, wiki, content management system, etc)
• easily discoverable and retrievable,
• accessible in a reusable format (open license)
**GCloud in the UK**

Departments must reuse and share ICT solutions, rather than purchasing new or bespoke solutions. The mandation of specific open standards will make ICT solutions fully interoperable to allow for reuse, sharing and scalability across organisational boundaries.

**SharedFirst in the US**

*Shared First*” foundational paradigm which provides Federal agencies with guidance on the identification, implementation, and operation of shared services for commodity, support, and mission IT functions.

Why are semantic assets important?
What is a repository?

Implemented as:
- Content Management System (CMS)
- Wiki
- Webpage / web-application

Repository

Reference Library
- Guidelines
- Templates
- Patterns

Assets (standards)
- Semantic Assets
- Technical Assets
- ...

External Reference Models

External Assets
Public administrations should identify important metadata with reuse potential for developing eGovernment applications and systems.

Public administrations should systematically document this metadata.

Public administrations should check their metadata to identify inconsistencies, overlaps, conflicts and opportunities of harmonization.

Public administrations should make their metadata open for reuse.

**Main benefits**

- Cost savings
- Coherent Architecture
- Interoperability
Ignorance
Reusable metadata + reference data are not documented, mainly because administrations don’t consider this exercise important. This results in serious semantic IOP problems within each country as developers use ad hoc data models, metadata, code lists, taxonomies, etc for developing eGov systems.

Scattered and/or Closed semantic assets
Reusable metadata + reference data may be documented but a) not in a centralised and organized way and/or b) they are not available and accessible as "open metadata" for developers, etc.

Open for Humans
Reusable metadata + reference data are documented, and are made available as "open semantic assets" but are not systematically published in a reusable format (e.g. only available as pdf documents).

Open and machine readable
Reusable metadata + reference data are centrally documented, they are published as "open semantic assets", in a machine readable format and/or provide an API for computers to access, query and reuse them. Electronic Metadata Management Systems (MMSs) are introduced (e.g. the SEMIC platform, Digitalisér.dk) to support the established metadata architecture and policies.

Linked open semantic assets
Semantic Assets are documented using linked data principles and are managed by advanced MMSs.
There are several semantic asset repositories in Europe.

There is a growing number of repositories storing, indexing and making available semantic assets.

LEGEND
- Respondents to the survey
- Non-respondents to the survey
Among the 11 MSs that have participated in this survey:

1. Metadata Ignorance

6. Scattered and/or Closed Metadata

1. Open Metadata for Humans

3. Open Reusable Metadata

Survey: Metadata Maturity Level
Overall, it seems that Europe is still at the dawn of metadata management but several Member States are already working to close the gap between their current modest level and the one required for them to be effective and efficient in managing this type of assets.
Agenda

Promoting Open Government Metadata

The need for ADMS

ADMS in practice

Get involved
Problem?

Reuse by Sanction vs. **Reuse by Buy In**

Reduce reuse burden

Make it simple to **find**, **understand** and **reuse** semantic assets
Each organisation works on standards but does not use a standard description to describe their standards.
I am looking for a particular standard
Google is not good enough
There are too many websites
I don’t know what to use

Introduction

Is it the right version?
Is it machine-readable?
What is ADMS?

**Asset Description Metadata Schema (ADMS)**
Standard vocabulary to describe semantic assets.

**Really Simple Syndication (RSS)**
Standard vocabulary to describe a web resource.
Our approach

Distributed approach

ETSI
Digitaliser.dk (Denmark)
OASIS
RIHA (Estonia)
GS1
XRepository (Germany)
W3C
The Open Group
EU Publications Office
Listpoint

Standard vocabulary to describe semantic assets enabling others to find them but keeping the asset in the repository of its publisher

Aggregators

Get ADMS file
Parse ADMS file
1. Analyse your semantic assets

2. You describe your assets following ADMS specification

3. Assets can be searched and retrieved

You describe your assets following ADMS specification.

Typical Use Case

CMS

Wiki

Web page

Asset description
Metadata
Schema
ADMS

Assets can be searched and retrieved

URLs

Assets can be searched and retrieved

URLs

ADMS

Asset description
Metadata
Schema
ADMS

Asset description
Metadata
Schema
ADMS

Asset description
Metadata
Schema
ADMS
How to implement ADMS?

1. MAP ADMS
   - Build a mapping from a native schema to ADMS
   - Use ADMS as a native metadata description schema
   - Annotate assets manually with ADMS and publish them on the web
   - Use the ADMS editor to annotate assets manually on Joinup

2. CREATE ADMS file
   - Use a dedicated exporter that transforms the current representation of the semantic assets into ADMS
   - Create manually the file using a text editor or a form
   - Use the ADMS editor on Joinup: web form to encode the ADMS file directly on Joinup

3. TRANSMIT ADMS file
   - Send the ADMS description (RDF/XML format) to Joinup via REST Web Services
   - Send the ADMS description (spreadsheet or RDF/XML format) to Joinup via e-mail
   - Publish the ADMS Joinup via e-mail description on the Web (RDF/XML format), so Joinup can harvest it
   - Upload manually the ADMS description (spreadsheet or RDF/XML format) on Joinup

Number of assets in the repository
The Commission is creating its own virtual aggregator

Make it easier for everyone to discover the appropriate taxonomy, code list or data model.

Make it easier for everyone to understand semantic assets.
What is the value of ADMS?

Value of a semantic asset

Tipping point

Number of users

More interoperability

+ Adding value to semantic assets and for their users.

Positive externalities (Network effect)
What is the value of ADMS?

- Add value to semantic assets and for their users.

TO BE:

Positive externalities (Network effect)

ADMS/Enabled Federation
- Raising visibility
- Promoting common description

Value of a semantic asset

Number of users

Tipping point
Agenda

Promoting Open Government Metadata

The need for ADMS

ADMS in practice

Get involved
Primary concepts

Repository

Asset
Data dictionaries, data models, code lists, XML schemas and RDF models

Distribution
Web URL

Supporting concepts

Asset Type
Contact Information
Documentation
File Format
Geographical Coverage
Item
Interoperability Level
Language
Licence
Period of time
Publisher
Representation Technique
Status
Theme
Theme Taxonomy

ADMS specification
RDF Schema

- Provided separately and serialised in RDF/XML-with an XSLT-that renders it into an HTML page for human readers.
- The majority of terms used to express ADMS in RDF are taken from existing vocabularies, notably Dublin Core and FOAF.
- The primary classes for ADMS (Semantic Repository, Semantic Asset and Semantic Asset Distribution) are all sub classes of the relevant classes in RADion.
- Terms defined in RADion are not repeated in the RDF schema for ADMS.
- Domain and range restrictions have not been defined for terms borrowed from existing vocabularies and have only been defined sparingly for ADMS’s own terms.

XML Schema

- The backbone of ADMS in XML is a Common Library of information elements provided by the library of the Universal Business Language (UBL).
- The philosophy behind this design is to achieve reusability of information elements defined by the Core Component Technical Specification (CCTS) of UN/CEFACT (the basis of UBL).
Multi disciplinary working group

43 people

20 + Member States and the US.
EU institutions
Standardization bodies
External experts/academia

PwC and W3C facilitation + W3C methodology

Enthusiasm around ADMS
The ADMS specification:

1200 reads
640 downloads

Currently the specification is available in UML, XML, RDF formats for public review
ADMS Lifecycle

- ADMS Specification
  - March 2012
- ADMS Representation
  - April 2012
- ADMS Implementation (Federation)
  - Q4 2012

- V0.7
- V0.8
- V0.9

Public Review (V0.6)

- January 2012
- February 2012
- March 2012
- April 2012

We are here
Who is already involved?

- Digitaliser.dk (Denmark)
- RIHA (Estonia)
- XRepository (Germany)
- International Commission on Civil Status
- Listpoint
- GS1
- Joinup (central repository)
- EU Member States
- W3C
- EU Publications Office
- Third party organisations
- Standardisation bodies
- European institutions
ADMS-enabled federation

With common metadata schema

- Public administrations
- Businesses
- Standardisation bodies
- Academia

ADMS-Asset
Description
Metadata
Schema

FEDERATION

repository

repository

repository

repository
Agenda

Promoting Open Government Metadata

The need for ADMS

ADMS in practice

Get involved
Mission

Provide standards and other information which help governments around the world publish their data as effective and usable Linked Data using Semantic Web technologies.

How to join the WG?
http://www.w3.org/2011/gld/wiki/How_To_Join

Objectives

• Develop W3C Recommendations to guide governments publishing data in which RDF vocabulary terms to use in information about certain common concept areas.

• Gather and publish use cases and requirements for vocabularies to cover each of the following areas:

  **Metadata**
  Suitable for provenance, data, data quality, timeliness of data, status, refresh rate, etc.

  **Statistical "Cube" Data**
  The group will produce a vocabulary, compatible with SDMX, for expressing some kinds of statistical data.

  **People**
  Such as elements of FOAF or vCard in RDF. This is an area for particular attention to privacy considerations.

  **Organizational Structures.**
  Such as the Epimorphics the organization ontology
Join the ADMS-enabled federation and connect your semantic assets on the Linked Data Cloud.

Download

Federation Brochure

ADMS Brochure

Contact

Programme Manager
Vassilios.PERISTERAS@ec.europa.eu

https://joinup.ec.europa.eu/contact
Join ISA Initiatives on Joinup

- Get involved at http://goo.gl/sLHIU
- Join ADMS: Asset Description Metadata Schema at http://goo.gl/3VCZU
- Join CESAR: Community of European Semantic Assets Repositories at http://goo.gl/tY9ty
- Join eGOVERNMENT Core Vocabularies at http://goo.gl/q2RQx
- Join Core Person Vocabulary at http://goo.gl/VODFF
- Join Core Business Vocabulary at http://goo.gl/dqOnm
- Join Core Location Vocabulary at http://goo.gl/03xkP
- Follow SEMICeu on Twitter at http://goo.gl/eK1EY
- Visit Joinup at http://joinup.ec.europa.eu
“It is the unexpected re-use of information which is the value added by the web”

Tim Berners-Lee