Macedonian Interoperability Building Block: **IOP-S**

(Semantic Interoperability)

Bernhard Krabina, Po-Wen Liu, Peter Reichstädter

Date: 2016

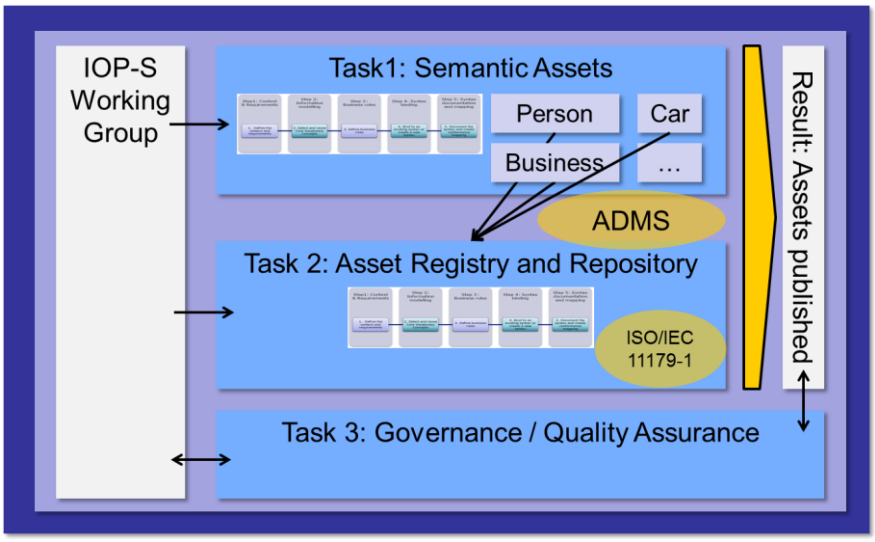
Version: 2.0

This project is funded by the European Union

1 Executive Summary

Semantic Interoperability enables systems to combine received information with other information resources and to process it in a meaningful manner. The process for achieving semantic interoperability is steered by an IOP-S Working Group that is recommended to be initiated by the IOP Committee (outlined in IOP-O).

Figure 1: Overview of the Macedonian Semantic Interoperability Framework



In order to achieve semantic interoperability in Macedonia, a methodology must be set up for specifying, documenting and using **Semantic Assets (Task 1**, see Figure 1**)**, i.e. highly reusable metadata (e.g. xml schemata, generic data models) and reference data (e.g. code lists, taxonomies, dictionaries, vocabularies) which are used for e-Government system development. A semantic asset specifies for example what metadata about a person, a business, a location or a car, or other entities are stored in Macedonian e-Government systems.

In order to raise awareness and facilitate the uptake of the interoperability assets, an **Asset Registry and/or an Asset Repository (Task 2)** needs to be implemented. An Asset Registry is a central location where asset definitions are stored and maintained in a controlled method. A semantic Asset Repository is an infrastructure where semantic assets can be stored, documented and retrieved. It is the physical place where the semantic assets are stored. **Governance and Quality Assurance (Task 3)** have to be established along the metadata lifecycle to ensure sustainability of the chosen approach. Governance addresses a variety of issues like decision making and change management process, policy domains, the enforcement policy, the reuse policy, licencing, updates and version control. The Quality Assurance determines certain criteria like accuracy, availability, completeness, consistency and others in order to enhance the retrieval and the reuse.

Within the European Union Interoperability Framework, methods, specifications and recommendations such as the Asset Description Metadata Schema (ADMS) as well as reusable interoperability assets such as the core vocabularies for persons, organisations, locations and public services are available to be used in a Macedonian context. Furthermore, already existing assets by

Macedonian institutions have to be taken into consideration for an optimal re-use. A mapping to the European standards is recommended.

This document (IOP-S) elaborates on the best practices and gives guidance on how to set up a Macedonian Semantic Interoperability Framework.

2 Table of Contents

3 Introduction

*"Interoperability, within the context of Macedonian public service delivery, is **the ability** of disparate and diverse organisations **to interact** …."*

Looking at the first part of the basic definition of interoperability immediately leads to the question of meaning and understanding. The ability to interact with one another is dependent on having the same view on the meaning of a topic, a sentence, a phrase or in most granularity of a word, sign or symbol. As semantics is the study of meaning, it becomes obvious that semantic interoperability in the context of information systems is the ability to exchange data with unambiguous, shared meaning, i. e. to understand the exchanged data in a similar way.

Scope

The scope of this document (IOP-S) is neither *why* data is shared (IOP-L), nor about *by whom* is it shared (IOP-O) or *how* it is transmitted (IOP-T), but it is about **what** data is shared (IOP-S).

Semantic interoperability needs to simultaneously transmit not only the data itself on an agreed and documented **syntax**, but also the meaning of the data by adding data about the data (**metadata**). Metadata provides information enabling to understand the meaning of **data** (e.g. documents, images, and datasets), **concepts** (e.g. classification schemes) and **real-world entities** (e.g. people, organisations, places, paintings, products). Therefore, this document is closely connected with the other levels of the interoperability and could not be understood without them. The frame for this document is set by the Macedonian Interoperability Framework.

3.1 Definitions

Semantic Interoperability enables systems to combine received information with other information resources and to process it in a meaningful manner.[1]

A **Semantic (Interoperability) Asset** is a collection of highly reusable metadata (e.g. xml schemata, generic data models) and reference data (e.g. code lists, taxonomies, dictionaries, vocabularies) which are used for e-Government system development.[2]

Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information.[3]

[1] EU Commission, Joinup: "What is semantic interoperability", checked on 7.4.2015
https://joinup.ec.europa.eu/asset/page/practice_aids/what-semantic-interoperability

[2] EU Commission, Joinup: "Semantic (Interoperability) Asset", checked on 7.4.2015
https://joinup.ec.europa.eu/category/glossary/semantic-interoperability-asset

[3] NISO Press: „Understanding Metadata", Bethesda, MD 20814 USA, 2004
http://www.niso.org/publications/press/UnderstandingMetadata.pdf

A **Controlled Vocabulary** is a predefined list of values to be used as values for a specific property in the metadata schema.[4]

A **Core Vocabulary** is a simplified, reusable, and extensible data model that captures the fundamental characteristics of an entity in a context-neutral fashion. Well known examples of existing Core Vocabularies include the Dublin Core Metadata Set [DC]. Such Core Vocabularies are the starting point for agreeing on new semantic interoperability assets and defining mappings between existing assets. Semantic interoperability assets that map to or extend such Core Vocabularies are the minimum required to guarantee a level of cross-domain and cross-border interoperability that can be attained by public administrations.[5]

The **Asset Description Metadata Schema (ADMS)** is a vocabulary to describe interoperability assets making it possible for ICT developers to explore and search for interoperability assets. ADMS is a profile of DCAT, used to describe *semantic assets* (or just 'Assets'), defined as highly reusable metadata (e.g. xml schemata, generic data models) and reference data (e.g. code lists, taxonomies, dictionaries, vocabularies) that are used for e-Government system development.[6,7]

4 Basic Concepts for Semantic Interoperability

In order to achieve semantic interoperability, the involved parties have to agree on a controlled and shared **vocabulary** or (more generally) **semantic assets** (i.e. highly reusable metadata e.g. xml schemata, generic data models and reference data e.g. code lists, taxonomies, dictionaries, vocabularies).

4.1 Methodology

To achieve semantic interoperability in Macedonia, a methodology has to put into place in order to identify, specify, use and quality assure semantic assets.

A process needs to be defined through which consensus can be reached among stakeholders and domain experts so that a semantic asset is recognised as meeting its design goals. Having achieved

[4] EU Commission, Joinup: "TM1.4. Introduction to metadata management (en)", checked on 7.4.2015 https://joinup.ec.europa.eu/community/ods/document/tm14-introduction-metadata-management-en

[5] EU Commission, Joinup: "ISA Deliverable: Process and Methodology for Developing Core Vocabularies", checked on 7.4.2015 https://joinup.ec.europa.eu/community/semic/document/isa-deliverable-process-and-methodology-developing-core-vocabularies

[6] EU Commission, Joinup: "Asset Description Metadata Schema (ADMS)", checked on 7.4.2015 https://joinup.ec.europa.eu/asset/adms/description

[7] W3C: "Asset Description Metadata Schema (ADMS)", checked on 7.4.2015 http://www.w3.org/TR/vocab-adms/

consensus, the final part of the process is endorsement of the semantic asset by the different institutions. The process is specified in detail in IOP-O.

A common methodology for developing semantic assets is not in place. However, the methodology of defining semantic assets for the Macedonian Interoperability Framework could be easily derived from the document "D3.1 – Process and methodology for core vocabularies"[8] developed by ISA.

A methodology needs to be defined through which a semantic asset is specified. That is, the best practices for how terms should be selected from existing assets or, where necessary, developed, encoded and presented to its intended audience. ISA proposes to set up objectives for the development:

Objectives (derived from the methodology for core vocabularies and amended)

- to define clearly the kind of problem the semantic asset will solve;
- to make the overall structure of the semantic asset clear, indicating how existing vocabulary terms, and any new ones, work together to comprise the semantic asset;
- to follow technical necessity and gain credibility, by following community convention when naming new terms;
- to follow best practice when publishing the semantic asset.

Clear steps have to been set up for achieving these objectives and to come to semantic assets. For example the first step is to identify stakeholders and the second one is to form workgroups. For more information please refer to the document.

Metadata will be created, changed and deleted - they have different characters in different phases. The next section describes the lifecycle of the metadata.

4.1.1 Metadata lifecycle

Metadata management comprises a set of high-level processes for structuring the different phases of the lifecycle of metadata:

1. **Design and create**
   Metadata creation can be supported by (semi-)automatic processes. Properties like creation date or information from publication workflow could be generated automatically. Some other characteristics require human intervention like licence, quality information, and related information. Appendix 6.4 describes the steps for asset creation.
2. **Document**
   Documentation of the whole process about the designing and creation of the metadata is essential. In this way changes are transparent and the development of metadata could be described and understood.

[8] ISA: "D3.1 – Process and methodology for core vocabularies", https://joinup.ec.europa.eu/community/semic/document/isa-deliverable-process-and-methodology-developing-core-vocabularies

3. **Publish, maintain and update**
   Approaches for maintaining metadata need to be appropriate for the type of data that are published. If data do not change, metadata can be relatively stable. If data change frequently (e.g. real-time data), metadata need to be closely coupled to the data workflow and more frequent changes are required.
4. **Deleting**
   In many cases, metadata must survive even after deletion of the data it describes. There are several causes for the deletion of data, for example: Data are no longer necessary, no longer valid, are wrong, withdrawn by the owner/publisher. In that case the metadata should contain information that the data was deleted, and if it was archived, how and where an archival copy can be requested.

Summing up, the lifecycle of metadata could be seen in different phases. In all these phases a clear metadata governance framework supports the development of metadata. Therefore some aspects of the metadata governance will be explained in order to ensure that metadata will be handled properly.

4.1.2 Metadata governance and management

Metadata governance is about ensuring that the management of structural metadata is conducted properly, i.e. following a set of guiding principles and practices, and in accordance with an organisation's strategic objectives. In this vein, metadata governance comprises well-defined roles and responsibilities, cohesive policies and principles, and decision-making processes that define, govern and regulate the lifecycle of structural metadata. As in IOP-O described the IOP Committee is responsible to set up a permanent working group which fulfil the tasks described in this document.

Metadata Governance addresses a variety of issues to be determined[9]:

- **Governance structure and roles:**
  The IOP-O suggests setting up a permanent working group which is responsible for the semantic interoperability. It also deals with the governance of metadata.
- **Decision making and change management process:** In the Terms of References of this working group the process of taking decisions are determined and described as well as the change management process. ISA described the process and methodology for developing semantic agreements.[10]
- **Policy domain:** The Governance of the metadata could be limited to certain policy domains and exclude some (for example Finance, Health, etc.). At a starting point the exclusion of policy domains is not recommended and should be discussed in a later stage.

[9] ISA: "*Methodology and tools for Metadata Governance and Management for EU Institutions and Member States*", 2014
[10] ISA: "Process and methodology for developing semantic agreements", 2013

- **Enforcement policy:** Determine the consequences if regulations are not followed. It should be clear whether there is a legal applicable context or whether there are other binding mechanisms (like agreements).
- **Reuse**: Metadata facilitate the reuse of the data. It should be clear if there is a legal requirement to reuse the data or if reuse is only voluntary. If the data are not reused the reasons should be explained.
- **Licensing framework:** Licencing is crucial to the use of the data. So the licence should be clear to the parties who want to use the data. There could be several classes of licence:
    - No explicit licence or Class 1: Traditional, proprietary licence; or
    - Class 2: Free to use and redistribute, no modifications allowed or
    - Class 3: Free to use, redistribute, and modify via copy left licence (e.g. CeCILL) or
    - Class 4: Free to use, redistribute, and modify via non-copy left licence (e.g. Academic Free License).

    In this context the European Commission published the ISA Open Metadata Licence v1.1 which defines a licence: A natural or legal person or body of persons corporate or incorporate is granted a worldwide, royalty-free, perpetual, non-exclusive licence to use and re-use any information and / or data offered and any modifications thereof for any commercial and non-commercial purpose under certain conditions like copyright notice, no warranty disclaimer.

- **Update frequency:** In general the update frequency of structural metadata varies case by case and depends on the type of metadata (e.g. reference data is updated more often than data models). The update frequency varies between one to three months to one to three years in some cases. Nevertheless it is advised to plan the updates frequencies in advance otherwise updates are often neglected.
- **Version control:** Related to the updates an effective version control facilitating the retrieval of current documents as well as the identification of changes should be established.
- **Quality controls:** The quality and completeness of the description metadata of datasets directly affect their ability to retrieve and reuse them. Quality criteria are[11]:
    - Accuracy (does the characteristics of the resource correctly reflected?)
    - Availability (can be accessed now and over time into the future?)
    - Completeness (all relevant characteristics of the resource captured?)
    - Conformance (conforming to a specific metadata standard?)
    - Consistency (contain contradictions?)
    - Credibility (based on trustworthy sources)
    - Processability (properly machine-readable?)
    - Relevance (contain the right amount of information for the task at hand?)
    - Timeliness (corresponding to the actual (current) characteristics of the resource?)

  Establishing a quality management process for structural metadata ensures the maintenance of high metadata quality.

[11] PWC: "Introduction to metadata management", presentation, 2014

Having dealt with the management of the metadata the focus changes to the storage and retrieval of the metadata.

4.2 Metadata registry

A metadata registry is a central location where metadata definitions are stored and maintained in a controlled method.

Metadata registries are currently planned and set up in some EU Member States and the European Commission. There are no definite recommendations which standard to apply to set up such a registry.

However, such a registry to be implemented in order to access decentralised resources through a single point of access facilities semantic interoperability. It has to respect the autonomy but at the same time allows the cross-querying and discovery of relevant assets stored in different places. It contains uniform descriptions of data elements used in structural metadata. Each registered item would contain at least an identifier, a name, and a definition (possibly in multiple languages). The metadata registry could be either maintained by a central authority.

Although the experiences with such registries are rare, there is an international standard for modelling a metadata registry. ISO/IEC 11179 specifies the kind and quality of metadata necessary to describe data. It provides for the attributes of data elements and associated metadata to be specified and registered as metadata items in a metadata registry (MDR). It proposes a structure of a metadata registry specified in the form of a conceptual data model. The metadata registry is used to keep information about data elements and associated concepts, generically referred to as “metadata items”. Such metadata are necessary to clearly describe, record, analyse, classify and administer data.

A metamodel supports the modelling of a metadata registry. It describes the precise structure and components of the registry and fosters the understanding of it. The registry metamodel is specified in ISO/IEC 11179 as a conceptual data model. The metamodel will assist registrars in maintaining consistency. The metamodel enables systems tools and information registries to store, manipulate and exchange the metadata for data attribution, classification, definition, naming, identification, and registration. In this manner, consistency of data content supports interoperability among systems, tools and information registries.

The standard also specifies the management and administration of that metadata in a metadata registry. It applies to the formulation of data representations, concepts, meanings, and relationships between them to be shared among people and machines, independent of the organization that produces the data.

4.3 Semantic Asset repository

A semantic asset repository is an infrastructure where semantic assets can be stored, documented and retrieved. It is the physical place where the semantic assets are stored.

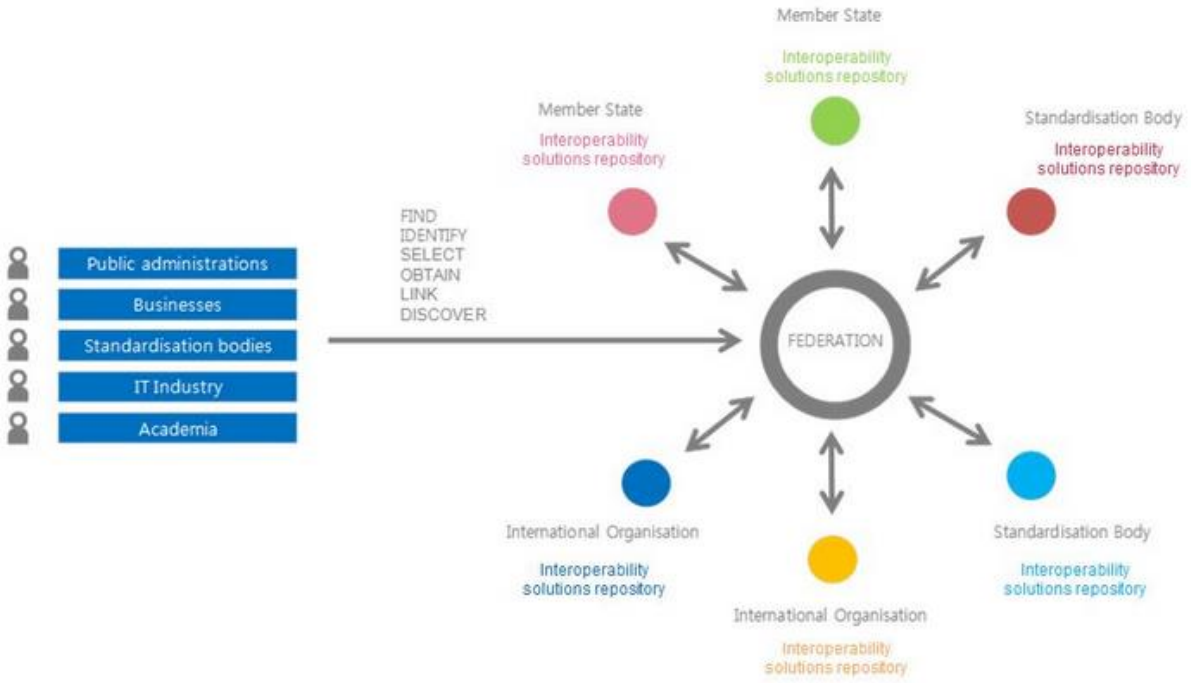
It is recommended that the asset repository is described according the Asset Description Metadata Schema (ADMS, see 4.4.1). In ADMS a concept of a Semantic Asset Repository is proposed:

Table 1: Semantic Asset Repository in ADMS

Property	Description	Data Type	Cardinality
access URL	URL of the Repository	URL	1..*
date of creation	creation date of the Repository	dateTime	0..1
date of last modification	date of latest update of the Repository	dateTime	1..1
Description	descriptive text for the Repository	text	1..*
Name	name of the Repository Note: in cases that a Repository has parallel names, for example if more than one official name exists, or if an organisation or country has more than one official language, this field can be repeated for all name variants	text	1..*
supported schema	Schema according to which the Repository can provide data, e.g. ADMS version 1.0	text	0..*
Includes	An asset for which a description is included in the Repository. This is the reverse relationship of repository origin.	Semantic Asset	0..*
Publisher	organisation making the Repository available	Publisher	0..*
spatial coverage	geographic region or jurisdiction to which the Repository applies	Geographic Coverage	0..*
Theme	theme or sector to which the Repository applies	Theme	0..*

At an international level of interoperability, following this standard facilitates the access to the repositories from other countries. As in Figure 2 shown interested parties could find, identify, select, obtain, link and discover other assets through standardized asset repositories.

Figure 2: Cross-Country Access of a Semantic Asset Repository (Federation)[12]



Best practise examples of repositories in EU Member States could be found in Germany, where XRepository[13] is used as a tool to support semantic interoperability operated by the German Federal Office for Information Technology (BIT). Currently, other EU Members States as Austria (reference.e-government.gv.at[14]), Denmark (Digitalisér.dk[15]), Finland (Yhteentoimivuus.fi[16]) and Estonia (RIHA[17]) have also implemented initiatives in this area. Different EU Member States have different approaches implementing semantic asset repositories. Nevertheless the interoperability among these repositories is ensured.

The following table gives an overview of possible technical solutions that can be used for setting up an asset repository.

Table 2: Technical Solutions for Asset Repositories

Solution	Components	Remark
(Open-Source) **Content-Management Systems**	(different, mostly PHP, MySQL)	e.g. Drupal, Typo3, Wordpress Used in http://reference.e-government.gv.at (Austria) For ADMS support, special extensions or programming effort will be needed.

[12] ISA: „Asset Description Metadata Schema“, https://joinup.ec.europa.eu/elibrary/document/adms-brochure
[13] https://www.xrepository.de
[14] https://reference.e-government.gv.at
[15] https://digitaliser.dk
[16] https://www.yhteentoimivuus.fi
[17] https://www.ria.ee/administration-system-of-the-state-information-system/

Solution	Components	Remark
SEMIC.EU https://joinup.ec.europa.eu/software/repository	Java / Java EE	SEMIC.EU is a participatory platform and a service by the European Commission that supports the sharing of assets of interoperability to be used in public administration and eGovernment. Formerly used by www.semic.eu and currently used by www.yhteentoimivuus.fi
joinup https://joinup.ec.europa.eu/software/joinup/	Drupal (PHP, MySQL)	Used in https://joinup.ec.europa.eu
SMW Semantic MediaWiki http://www.semantic-mediawiki.org	MediaWiki, Semantic MediaWiki (PHP, mySQL)	Can be configured to be used as an asset repository. Used in www.ogdcockpit.eu and http://ogdcockpit.bonn.de (City of Bonn, Germany)
ckan http://www.ckan.org	CKAN (Python, PostgreSQL)	Use for many Open Government Data portals, might be suitable/customizable to be used as asset repository.
dkan the Open Source Data Portal Software https://drupal.org/project/dkan	Drupal (PHP, MySQL)	Rewrite of CKAN in Drupal. Targeted for Open Data Portals, might be suitable/customizable to be used as asset repository. Used in City of Cologne, Germany.
DATA.GOV https://github.com/GSA/data.gov	Wordpress + CKAN	Targeted for Open Data Portals, might be suitable/customizable to be used as asset repository. Used in http://data.gov (US)
DATA.GOV.UK Beta Opening up Government https://github.com/datagovuk	Drupal + CKAN	Targeted for Open Data Portals, might be suitable/customisable to be used as asset repository. Used in http://data.gov.uk (UK)
DataPublic http://datapublic.org/	Drupal 7 + Windows Azure	Targeted for Open Data Portals, might be suitable/customisable to be used as asset repository. Used in: Colombia, Portugal, Cities in Canada, France…
Apache Marmotta http://marmotta.apache.org	Java, SPARQL, Triplestore	Linked Data Platform Used in http://data.enel.com

It is recommended to build up semantic asset repository according to ADMS. Such a repository does not only foster the exchange of data within public administration agencies in Macedonia but also the exchange of data with EU member states. An initial investment will be required by the owners of the data to establish such repository. The development of a technology-neutral architecture is a basic requirement for an interoperable solution.

4.4 Semantic Assets and Vocabularies

4.4.1 Asset Description Metadata Schema (ADMS)

ADMS is a common vocabulary to describe semantic interoperability assets making them easier to search and discover.

https://joinup.ec.europa.eu/asset/adms

The objectives are:

- Provide reusable metadata (e.g. schemas, data models) and reference data (e.g. code lists, taxonomies, dictionaries, vocabularies) used for e-Government
- Facilitate publication of descriptions of such assets
- Descriptions can be shared and aggregated in federated repositories
- The assets themselves remain where they are published and maintained

ADMS is intended as a model that facilitates federation and co-operation. Developers of new repositories or publishers of existing repositories could use ADMS as a format to build their systems containing asset descriptions. It is not the primary intention that repository publishers redesign or convert their current systems and data to conform to ADMS, but rather that ADMS can act as a common layer among repositories that want to exchange data. Build ADMS re-using existing vocabularies as ADMS Asset is a subclass of DCAT Dataset and vocabularies are re-used. ADMS facilitates reuse by providing standardized description of data assets and creation of federated catalogues. A wider base of users may find the assets easier and quicker as standardized description are provided. ADMS also lowers interoperability barriers because a standardized description could be used by more organisations and enhance the interoperability among them.

The current version is 1.0 and published in 2012 that recommends RDF and XML distributions of the model. The RDF representation of ADMS reuses existing vocabularies as far as possible and is aligned with the DCAT vocabulary published by the W3C Government.

The ADMS has three main primary concepts:

1. A Semantic **Asset Repository** is a system or service that provides facilities for storage and maintenance of descriptions of Semantic Assets and Semantic Asset Distributions, and functionality that allows users to search and access these descriptions. A Semantic Asset Repository will typically contain descriptions of several Semantic Assets and related Semantic Asset Distributions. (as described in section 4.3)
2. A Semantic **Asset** in the model is an abstract entity that reflects the intellectual content of the asset and represents those characteristics of the asset that are independent of its physical embodiment. Assets can be versioned. Every time the intellectual content of an asset changes, the result is considered to be a new asset that can be linked to previous and next versions of the Asset.

3. A Semantic **Asset Distribution** in the model represents a particular physical embodiment of a Semantic Asset. A Distribution is typically a downloadable computer file (but in principle it could also be a paper document) that implements the intellectual content of an Asset. A particular Distribution is associated with one and only one Asset, while all Distributions of an Asset share the same intellectual content in different physical formats. Distributions themselves are not versioned.

In the specification the assets of these concepts are identified and described, moreover there are several supporting concepts like Contact Information, Geographic Coverage, File Type and many others.

One of the crucial characteristics of the environment in which ADMS will be deployed is that it is intended to support interoperability in a multilingual environment. The content of Assets, as far as they contain textual information, will be produced in different languages. Repositories, Assets and Distributions are created, maintained and described in different languages. Users of the information will have different linguistic and cultural backgrounds and may expect to be able to search in their own language and find material both in their own and in other languages.

While the conceptual model of ADMS described does not explicitly address the potential requirements for multilingual deployment in a federation of repositories, it does contain a number of capabilities to enable the support of multilingual environments.

Appendix 6.3 shows the UML domain of ADSM 1.0.

Table 3: Properties of an Asset

Property	Description	Data Type	Cardinality
alternative name	alternative name for the asset. Note: this information may be used to provide additional access points, e.g. allowing indexing of any acronyms, nicknames, shorthand notations or other identifying information under which a user might expect to find the Asset	Text	0..*
date of creation	creation date of this version of the asset	dateTime	0..1
date of last modification	date of latest update of asset	dateTime	1..1
description	descriptive text for the asset	Text	1..*
id	URI for the asset	URI	1..1
identifier	any identifier for the asset	identifier	0..*
keyword	word of phrase to describe the asset	Text	0..*
metadata date	date of the most recent update of the metadata for the asset	dateTime	0..1
name	name of the asset.	Text	1..*
	Note: in cases that an asset has parallel names, for example if more than one official name exists, or if an organisation or country has more than one official language, this field can be repeated for all name variants		
version	version number or other designation of the asset	Text	0..1
version notes	description of changes between this version and the previous version of the asset	Text	0..1
asset type	type of the asset	Asset Type	1..*

Property	Description	Data Type	Cardinality
contact point	contact point for further information about an asset	Contact Information	0..*
current version	current or latest version of the asset	Semantic Asset	0..1
home page	a Web page that is fully dedicated to the asset	Documentation	0..*
included asset	an Asset that is contained in the asset being described, e.g. when there are several vocabularies defined in a single document	Semantic Asset	0..*
included item	item that is contained in the asset (e.g. a concept in a controlled vocabulary, an individual code in a code list or any other 'atomic' element)	Item	0..*
interoperability level	interoperability level for which the asset is relevant	Interoperability Level	0..1
language	language of the asset	Language	0..*
main documentation	the main documentation or specification of the asset	Documentation	0..*
metadata language	language of the metadata for the asset	Language	0..*
metadata publisher	organisation making the metadata for the asset available	Publisher	0..*
next version	newer version of the asset	Semantic Asset	0..*
previous version	older version of the asset	Semantic Asset	0..*
publisher	organisation making the asset available	Publisher	1..*
related asset	unspecified relationship from the asset to another asset	Semantic Asset	0..*
related documentation	documentation that contains information related to the asset	Documentation	0..*
asset type	type of the asset	Asset Type	1..*
related web page	a web page that contains information related to the asset	Documentation	0..*
distribution	implementation of the asset in a particular format	Semantic Asset Distribution	0..*
repository origin	repository that contains the primary description of the asset	Semantic Asset Repository	0..1
sample	sample of the asset	Semantic Asset	0..*
spatial coverage	geographic region or jurisdiction to which the asset applies	Geographic Coverage	0..*
status	status of the asset in the context of a particular workflow process	Status	1..1
temporal coverage	time period relevant to the asset, e.g. its validity	Period of Time	0..*
theme	theme or sector to which the asset applies	Theme	0..*
translation	translation of the asset	Semantic Asset	0..*

Source: https://joinup.ec.europa.eu/asset/adms

The third part of ADMS is the semantic assets distribution as a particular physical embodiment of a Semantic Asset.

Table 4: Semantic Assets Distribution

Property	Description	Data Type	Cardinality
access URL	URL of the Distribution Note: more than one URL may be available, for example if mirror sites are maintained.	URL	1..*

Property	Description	Data Type	Cardinality
date of creation	creation date of the distribution	dateTime	0..1
date of last modification	date of latest update of the distribution	dateTime	0..1
Description	descriptive text for the distribution	text	0..*
Id	URI for the distribution	URI	1..1
Name	name of the distribution	text	0..*
Format	format in which the distribution is available (e.g. PDF, XSD, RDF/XML, HTML, ZIP)	File Format	1..1
Licence	conditions or restrictions for (re-)use of the Distribution.	Licence	1..*
	Note: if multiple licenses are given, these licences apply to all files in the Distribution		
distribution of	the asset that this Distribution embodies. This is the reverse relationship of Asset.Distribution	Semantic Asset	0..1
Publisher	organisation making the distribution available	Publisher	0..*
representation technique	language in which the distribution is expressed	Representation Technique	0..*
	Note: this is different from the file format, e.g. a ZIP file (file format) could contain an XML schema (representation technique)		
Status	status of the distribution in the context of a particular workflow process	Status	1..1

Source: https://joinup.ec.europa.eu/asset/adms

4.5 Linked Data and URIs

In Europe, access to government data, and the possibility to freely use it, is seen as an enabler for Open Government and a goldmine of unrealised economic potential. Open Data usually refers to public records (e.g. on transport, infrastructure, education, and environment) that can be freely used and redistributed by anyone - either for free or at marginal cost.

But opening-up data, e.g. in Open Data portals, often happens in an ad-hoc manner, and in many cases thousands of datasets is published without adhering to commonly-agreed data and metadata standards and without reusing common identifiers.

Hence, a fragmented data-scape is created, where finding, reusing, integrating and making sense of data from different sources is a real challenge.

Linked Data can respond to these challenges and can be an enabler of eGovernment transformation, leading to smarter and more efficient government services and applications, and fostering creativity and innovation in the digital economy.

Linked data is a set of design principles for sharing machine-readable data on the Web for use by public administrations, business and citizens.[18]

Tim Berners Lee proposed four design principles of Linked Data in order to facilitate linked data[19]

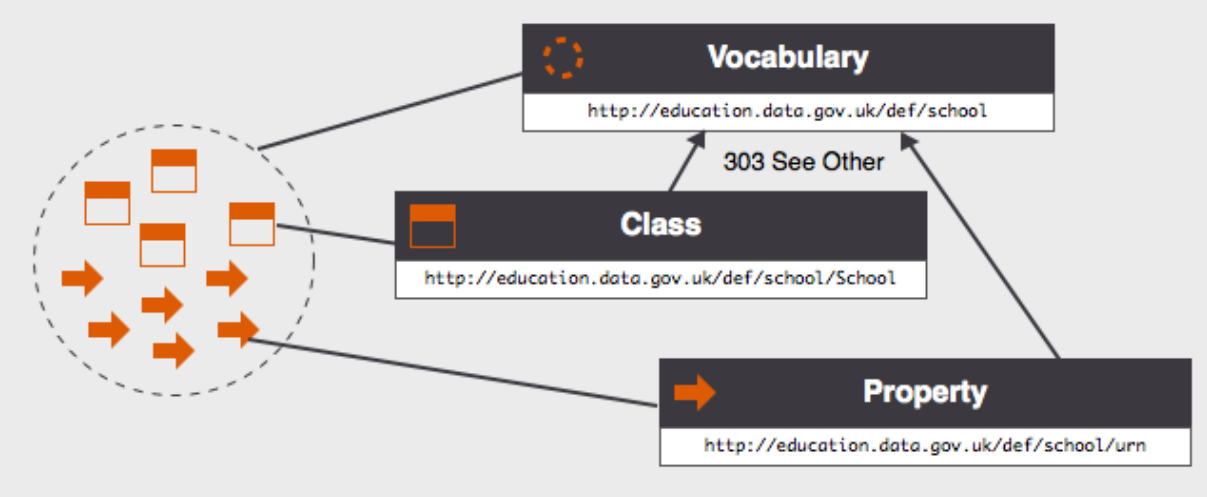
[18] ISA: "How Linked Data is transforming eGovernment" https://joinup.ec.europa.eu/community/semic/document/case-study-how-linked-data-transforming-egovernment

- **Use Uniform Resource Identifiers** (URIs) as names for things.
- **Use HTTP URIs** so that people can look up those names.
- When someone looks up a URI, provide useful information, using the standards (**RDF, SPARQL**).
- Include **links to other URIs** so that they can discover more things.

Linked Data enables public administrations to define links (i.e. relationships) between related datasets of others. So it enables the flexible virtual integration of government data just through linking. The ownership and the design of the information systems are not affected. The collaboration among public administration agencies is facilitated and smarter public services could be created.

In the Macedonian Interoperability Strategy, at least a naming convention for URIs should be developed to identify names of things (first principle). As a reference, it can be referred to the best practice of UK URI conventions.[20] Figure 3 shows an example for URIs in UK for vocabularies, classes and properties. It is important to note that URIs can be used to name documents on the internet as well as real-world things. Putting such URIs in a browser of course would not return anything.

Figure 3: URI Concepts for Vocabularies and Properties



Source: http://data.gov.uk/resources/uris

URIs should as well be used to identify datasets in the Macedonian Open Data Portal (see 5.2.2).

5 Semantic Interoperability Dimensions

5.1 Core Vocabularies

The Core Vocabularies are context-neutral semantic building blocks that can be extended into context-specific data models. The use of the Core Vocabularies as a common building block for developing context-specific data models guarantees a minimum of semantic consistency. When the

[19] See http://www.w3.org/DesignIssues/LinkedData.html
[20] see http://data.gov.uk/resources/uris

Core Vocabularies are extended to create domain models and information exchange data models, additional meaning (semantics) is added to the specifications, due to the contextualisation.[21,22]

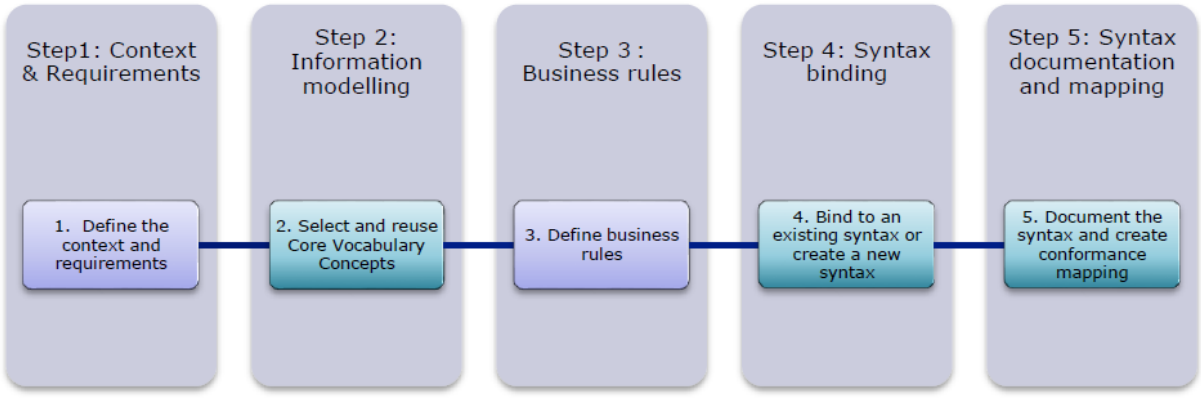
In order to make use of the existing vocabularies, the following basic elements have to be explained:

- A **class** is a description of a set of things that share the same properties, associations, and semantics. For example, "*Person"* represents the set of natural persons.
- An **object** is an instance of a class and is an abstraction of a physical or conceptual thing. In an object, properties and associations have values. For example, the object representing "*Igor Crvenov"* is an instance of the class "Person".
- A **property** is a named, singular characteristic of a class. The data type of a property defines the values that may be assigned to the property in an object. For example, the property "*PersonFullName*" is a property of the class "Person".
- An **association** is a semantic relationship between two classes. The value of an association in an object is an object of the other class. The association describes the role of the former object with respect to the latter object. For example, the association "*PersonAddress*" is an association of the class "Person".

Core vocabularies can be extended by domain specific properties. E. g. for the class "Person" the Ministry of Finance can add a property "TaxNumber" that might not be relevant for other institutions that also use the core vocabulary for "Person". Therefore a common core vocabulary has to be agreed on that can be re-used and extended.

A methodology on creating core vocabularies describes in detail the necessary steps:

Figure 4: Methodology on using the Core Vocabularies[22]



[21] EU Commission, Joinup: "e-Government Core Vocabularies", checked on 7.4.2015 https://joinup.ec.europa.eu/asset/core_vocabularies/description

[22] *PwC EU Services:* "Handbook for using the Core Vocabularies", 2014, https://joinup.ec.europa.eu/site/core_vocabularies/Core_Vocabularies_user_handbook/

The aim of step 1 (Context and requirements) is to define the context and elicit a set of technology-neutral requirements for the data model to be designed. The context could be policy domain, geo-political context, administrative procedures policies and others. The eliciting of information requirements can be done by using common requirement analysis techniques.

Step 2 (Information modelling) is used to create a conceptual data model that covers the information requirements derived from the first step. The output of this step is the conceptual data model aligned with the Core Vocabularies.

Step 3 (Business rules) focuses on action assertions, constraints and derivations concerning some aspects of the conceptual data model that have to be defined. The outcome of the third step is an enhanced data model with the cardinalities and constraints and the lists of sets of values that restrict the possible values for coded elements.

Step 4 (Bind to an existing syntax or create a new syntax) the information requirements are bound to actual elements with a given syntax.

The aim of the last step 5 (Syntax documentation and mapping) is to create documentation of the syntax that allows users to implement it, and at the same time allows the owner to claim conformance of the data model to the Core Vocabularies.

5.1.1 Person

The Core Person Vocabulary is a simplified, reusable and extensible data model that captures the fundamental characteristics of a person, e.g. the name, the gender, the date of birth, etc.

https://joinup.ec.europa.eu/asset/core_person/

A Macedonian core vocabulary for a Person should be developed. As there is the Core Person Vocabulary already available on Joinup.eu, it should be aimed to re-use as much as possible.

Table 5: Core Person Vocabulary

Identifier	Term	Data Type	Definition
PersonIdentifier	Identifier	Identifier	A formally-issued identifier for the person.
PersonFullName	Full Name	String	The complete name of the person as one string.
PersonGivenName	Given Name	String	The denominator(s) that identify the person within a family.
PersonFamilyName	Family Name	String	A name that is usually shared by members of a family.
PersonPatronymicName	Patronymic Name	String	A name referring to the father's given name.
PersonAlternativeName	Alternative Name	String	A name by which the person is known other than her given name and/or full name.
PersonGender	Gender	Code	The gender of the person.
PersonBirthName	Birth Name	String	The full name of the person at the time of her

Identifier	Term	Data Type	Definition
			birth, irrespective of any subsequent changes.
PersonDateOfBirth	Date Of Birth	Date Time	The date on which the person was born.
PersonDateOfDeath	Date Of Death	Date Time	The date on which the person deceased.

Example: Person Vocabulary, Term "Gender", Code List

The Core Person **Vocabulary** lists the Term "Gender" with

- the identifier "PersonGender"
- the type "Property"
- the class "Person"
- the data type "Code"
- the definition "The gender of the person"
- the description "The gender of an individual should be recorded using a controlled vocabulary that is appropriate for the specific context.."

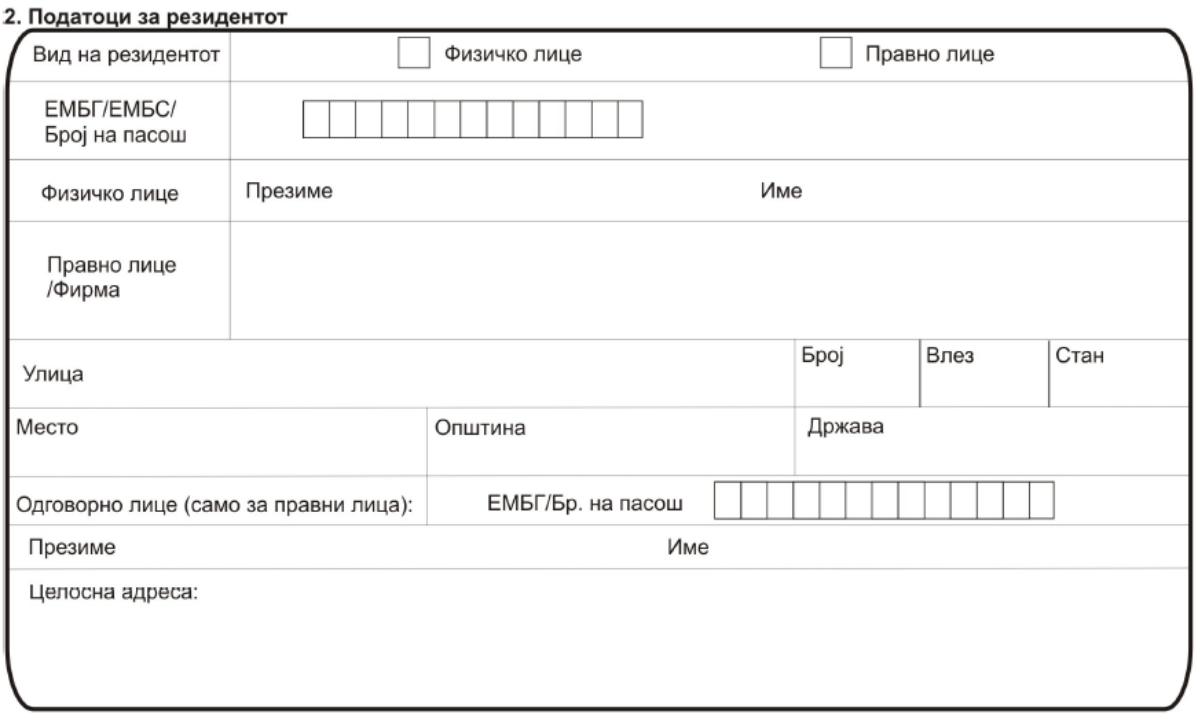
In the EU Semantic Interoperability Catalogue there is a **code list** "SCL – Sex". This classification provides a European structure for the Eurostat standard code list for sex / gender. 5 codes have been defined: Total, Females, Males, Not applicable and Unknown. It is primarily used by all European Union Member States in order to give standard statistics to Eurostat.

https://joinup.ec.europa.eu/catalogue/asset_release/scl-sex

Example: Central Registry Office: Application for Registration of Direct Investments of Foreign Residents

Currently, no detailed information is available on what data about a person is stored in the Central Registry Office. As an indicator, we can reference to an application form, where details about a person has to be filled in. In this example, we look at the Application form for Registration of Direct Investments of Foreign Residents.

Figure 5: Application form from the Central Register



2. Податоци за резидентот

Вид на резидентот	☐ Физичко лице	☐ Правно лице		
ЕМБГ/ЕМБС/ Број на пасош				
Физичко лице	Презиме	Име		
Правно лице /Фирма				
Улица		Број	Влез	Стан
Место	Општина	Држава		
Одговорно лице (само за правни лица):	ЕМБГ/Бр. на пасош			
Презиме	Име			
Целосна адреса:				

Source: http://www.crm.com.mk/DS/default.aspx?MainId=1&CatID=83

If we look at the section „Data on the resident", we can identify what information has to be filed in:

Table 6: Application for Registration of Direct Investments of Foreign Residents

Type of resident	☐ ural person ☐ al person		
Single Identification Number / Passport Number			
Natural person	Family name	Name	
Legal person / Company			
Address	No.	Building	Apartment No.
City	Municipality	State	
Responsible person (only for legal persons)	Single Identification Number / Passport No.		
Family name	Name		
Full address			

Source: http://www.crm.com.mk/DS/default.aspx?MainId=1&CatID=83

The information to be filled in can be matched to the Core Vocabularies:

Table 7: Mapping to Vocabulary Terms to the Application form

Term in application form	Matching vocabulary item	Remark
Type of resident		Identifies whether the data is from class *Person* or class *Legal Entity.* No mapping

Term in application form	Matching vocabulary item	Remark
		is necessary, but based on this selection the mapping will be done e. g. to PersonIdentifier or LegalEntityIdentifier
Single Identification Number / Passport Number	PersonIdentifier	
Natural person		just a section heading
Family name	PersonFamilyName	
Name	PersonGivenName	
Legal person / Company	LegalEntityLegalName	
Address	PersonAddress AddressFullAddress	
No.		
Building		
Apartment No.		
City	AddressPostName	
Municipality	AddressAdminUnitL2	
State	AddressAdminUnitL1	
Responsible person (only for legal persons)		just a section heading
Single Identification Number / Passport No.	PersonIdentifier	
Family name	PersonFamilyName	
Name	PersonGivenName	
Full address	AddressFullAddress	

Example: Birth certificate in Macedonia

Another example is the birth certificate in Macedonia. This could define the starting point of a core vocabulary of a person in Macedonia:

Table 8 Mapping to Vocabulary Terms to the birth certificate

Term in application form	Matching vocabulary item	Remark
Name Surname	PersonBirthName, PersonFullName	
Name	PersonFamilyName	
Surname	PersonGivenName	
Gender	PersonGender	
Day, month, year and hour of birth	PersonDateOfBirth	
Place of birth	Location.PlaceofBirth	
Citizenship	Jurisdiction.Citizenship	
Single Identification Number	PersonIdentifier	
Father Name	PersonFamilyName	
Father Surname	PersonGivenName	
Mother-Name	PersonFamilyName	
Mother Surname	PersonGivenName	

Term in application form	Matching vocabulary item	Remark
Father – Place of residence and address	Address.FullAddress	
Mother – Place of residence and address	Address.FullAddress	
Remarks		

5.1.2 Business

The Core Business Vocabulary (Registered Organisations Vocabulary) is a simplified, reusable and extensible data model that captures the fundamental characteristics of a legal entity, e.g. the legal name, the activity, address, legal identifier, company type, and its activities. The Core Business Vocabulary has been formally published on the W3C standards track as a Public Working Draft.

https://joinup.ec.europa.eu/asset/core_business/

A Macedonian core vocabulary for a Business (legal entity) is recommended to be developed. As there is the Core Business Vocabulary already available on Joinup.eu, it should be aimed to base the Macedonian core vocabulary on the definitions of Joinup.eu.

Table 9: Core Business Vocabulary

Identifier	Term	Data Type	Definition
LegalEntityLegalIdentifier	Legal Identifier	Identifier	The identifier given to the legal entity by the authority with which it is registered.
LegalEntityIdentifier	Identifier	Identifier	A formally-issued identifier for the legal entity, other than the one that confers legal status upon it.
LegalEntityLegalName	Legal Name	Text	The legal name of the business.
LegalEntityAlternativeName	Alternative Name	Text	A recognized name other than the legal name.
LegalEntityCompanyType	Company Type	Code	The type of the business.
LegalEntityCompanyStatus	Company Status	Code	The status of the business.
LegalEntityCompanyActivity	Company Activity	Code	The activity of the business.
LegalEntityRegisteredAddress	Registered Address	Address	The registered address of the business.
LegalEntityAddress	Address	Address	An address related to the business, other than the registered address.
LegalEntityLocation	Location	Location	A location related to the business.

Source: https://joinup.ec.europa.eu/asset/core_business/

Example

The examples in 5.1.1 and 5.1.3 include aspects of the business vocabulary.

5.1.3 Location

The Core Location Vocabulary is a simplified, reusable and extensible data model that captures the fundamental characteristics of a location, represented

as an address, a geographic name, or geometry. The Location Vocabulary is aligned with the INSPIRE data specifications.

https://joinup.ec.europa.eu/asset/core_location/

A Macedonian core vocabulary for a Location should be developed. As there is the Core Location Vocabulary already available on Joinup.eu, it should be aimed to base the Macedonian core vocabulary on the definitions of Joinup.eu.

Table 10: Core Location Vocabulary

Identifier	Term	Data Type	Definition
LocationGeographicName	Geographic Name	Text	A proper noun applied to a spatial object.
LocationGeographicIdentifier	Geographic Identifier	URI	A URI that identifies the location.
LocationAddress	Address	Address	An address representing the location.
LocationGeometry	Geometry	Geometry	A geometry representing the location.

Example

Table 11: Open Data Set of Public Municipal Kindergartens

Municipality	Public Municipal institution - Children's Kindergarten	Address	Telephone	E-Mail
Berovo	August 23	Mladinski Kej No.5	033 471-039	detskagradinka@yahoo.com
Pehchevo	September 7	Dame Gruev bb	033 441 331	detskagradinkapehcevo@yahoo.com
Bitola	Estreja Ovadija Mara	N. N. Borche bb	047 523 232, 047 522 107	estrejamarabt@yahoo.com
…				

Source: www.opendata.gov.mk

Table 12: Mapping to Vocabulary Terms

Term in data set	Matching vocabulary item	Remark
Municipality	-	
Public Municipal institution - Children's Kindergarten	LocationGeographicName	
Address	LocationAddress	
Telephone	-	Not in core vocabulary, but could be matched to foaf:phone: http://xmlns.com/foaf/spec/#term_phone
E-Mail	-	Not in core vocabulary, but could be matched to foaf:mbox: http://xmlns.com/foaf/spec/#term_mbox

It might be more feasible not to use the Location vocabulary for the kindergartens, but the Business vocabulary. Then LegalEntityLegalName, LegalEntityRegisteredAddress and LegalEntityLocation would be used, as Kindergartens cannot be seen only as locations, but as legal entities that (among other properties) also have locations.

5.1.4 Further Core Vocabularies

The Core Vocabularies define more classes of reusable semantic assets. Especially, Addresses and legal assets such as Formal Framework, Jurisdiction and Rule as well as other basic elements such as Channel, Input, Output and Period of Time.

https://joinup.ec.europa.eu/asset/core_vocabularies/asset_release/core-vocabularies-v11

Macedonian core vocabularies for further elements should be developed. As there further vocabulary items available on Joinup.eu (EU wide or from individual member states), it should be aimed to base the Macedonian core vocabulary on the definitions of Joinup.

Table 13: Further Core Vocabularies

	Class	Description
Core Vocabulary covered in separate chapters	Agent	An entity that is able to carry out actions.
	LegalEntity	A business that is legally registered. Extends class „Agent“ See 5.1.2
	Person	A natural person. Extends class „Agent“. See 5.1.1
	Location	An identifiable geographic place. See 5.1.3
	Geometry	A geometry representing a location.
	PublicService	A set of deeds and acts performed by or on behalf of a public agency for the benefit of a citizen, a business or another public agency. See 5.2.1
New	**Address**	An address representing a location.
New	**FormalFramework**	Legislation, policy, or policies lying behind the rules that govern a public service.
	Jurisdiction	A jurisdiction, typically a country, dealing with and making pronouncements on legal matters.
	Rule	A document that sets out the specific rules, guidelines, or procedures that a public service follows.
New	Channel	A medium through which an agent provides, uses or otherwise interacts with a resource.
	Input	A resource to be processed to produce an output.
	Output	An intended result whose required inputs and processes are entirely within the control of the planning organisation.
	PeriodOfTime	An interval of time that is named or defined by its start and end dates.

Source: https://joinup.ec.europa.eu/asset/core_vocabularies/asset_release/core-vocabularies-v11

Example

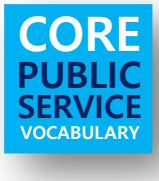
The following terms from the Jurisdiction class could be used to identify the name for the jurisdiction and the URI.

Table 14: Example of Jurisdiction Class

Identifier	Term	Data Type	Definition
JurisdictionName	Name	Text	The name of the jurisdiction.
JurisdictionIdentifier	Identifier	URI	The URI for the jurisdiction.

5.2 Specialised Vocabularies

5.2.1 Public Service



The Core Public Service Vocabulary is a simplified, reusable and extensible data model that captures the fundamental characteristics of a service offered by public administration. Such characteristics include the title, description, inputs, outputs, providers, locations, etc. of the public service.

https://joinup.ec.europa.eu/asset/core_public_service/
http://semanticommunity.info/@api/deki/files/26757/Core_Public_Service_Vocabulary_specification_v1.01.docx

A Macedonian core vocabulary for a Public Service should be developed. As there is the Core Public Service Vocabulary already available on Joinup.eu, it should be aimed to base the Macedonian core vocabulary on the definitions of Joinup.

Table 15: Core Public Service Vocabulary

Identifier	Term	Data Type	Definition
PublicServiceName	Name	Text	The name of the service.
PublicServiceDescription	Description	Text	A free text description of the service.
PublicServiceType	Type	Code	The type of service.
PublicServiceLanguage	Language	Code	The language(s) in which the service is available.
PublicServiceHomepage	Homepage	URI	The Web page through which the service may be available.
PublicServiceChannel	Channel	Channel	A medium through which an agent interacts with the service.
PublicServicePhysicallyAvailableAt	Physically Available At	Location	A physical location at which a user may interact with the service.
PublicServiceRequires	Requires	Public Service	Another public service required by this service.
PublicServiceRelated	Related	Public Service	Another public service related to this service, without being required by this service.
PublicServiceInput	Input	Input	A resource required by the service in order to operate.
PublicServiceProduces	Produces	Output	A resource produced by the service.
PublicServiceFollows	Follows	Rule	A rule under which the public service operates.
PublicServiceSpatial	Spatial	Location	The area in which the service is available.
PublicServiceTemporal	Temporal	Period Of Time	The time frame in which the service is available.

Source: https://joinup.ec.europa.eu/asset/core_public_service/

Example

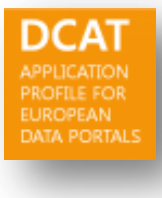
The e-Service "Reservation of name of a legal entity" can be found at http://www.crm.com.mk/namereservation/ for the reservation of name of a legal entity before its registration in the Registers of businesses and other legal entities.

Table 16: Example of Public Service Vocabulary

Identifier	Value	Remark
PublicServiceName	Резервација на име на субјект	Reservation of name of a legal entity Technically, this would be written as: dcterms:title "Резервација на име на субјект "@mk ; dcterms:title "Reservation of name of a legal entity"@en ;
PublicServiceDescription	Резервација на име на правен субјект пред запишување во Регистарот на Трговски друштва и други правни лица.	Reservation of name of a legal entity before its registration in the Registers of businesses and other legal entities.
PublicServiceType		could be linked to http://id.esd-toolkit.eu/function/6
PublicServiceLanguage	mk	dcterms:language <http://id.loc.gov/vocabulary/iso639-1/mk>;
PublicServiceHomepage	http://www.crm.com.mk/namereservation/	foaf:homepage <http://www.crm.com.mk/namereservation/ > ;
PublicServiceChannel	online	
PublicServicePhysicallyAvailableAt	-	
PublicServiceRequires	-	
PublicServiceRelated	-	
PublicServiceInput	Name of Business	cpsv:hasInput <http://cpsv.assetprepository.mk/id/ltu/Input/LegalEntityLegalName>
PublicServiceProduces		The result is a website (HTML output) foaf:homepage <http://www.crm.com.mk/namereservation/>;
PublicServiceFollows	-	
PublicServiceSpatial	Macedonia	dcterms:spatial<http://id.loc.gov/authorities/names/n81038515>;
PublicServiceTemporal	-	

5.2.2 Open Government Data

An Open Government Data portal already exists in Macedonia, with more than 150 datasets already published at www.opendata.gov.mk. However, a metadata description is missing. The EU standard DCAT-AP should be used to describe the published datasets, so other portals can reference the Macedonian datasets.



DCAT Application profile for data portals in Europe (DCAT-AP) is a specification based on the Data Catalogue vocabulary (DCAT) for describing public sector datasets in Europe. Its basic use case is to enable a cross-data portal search for data sets and make public sector data better searchable across borders and sectors. This can be achieved by the exchange of descriptions of data sets among data portals. DCAT is an implementation of the ADMS standard (see **Error! Reference source not found.**), describing assets from Open Data portals.

Example

On the OGD portal www.opendata.gov.mk 154 datasets can be found. The following table shows some examples of datasets:

Table 17: Examples of Datasets from Opendata.gov.mk

Dataset	Category	Catalog	Status	Date	Down-loads	Views
MLSP Directory of children's kindergartens	Open data	Ministry of Labor and Social Policy	Active	07/01/2014	235	0
MLSP Directory of private children's kindergartens	Open data	Ministry of Labor and Social Policy	Active	07/01/2014	86	0
MF Rebalance of budget of the Republic of Macedonia	Open data	Ministry of Finance	Active	02/05/2015	82	277
MLSP Detailed view of persons with disabilities according to municipalities	Open data	Ministry of Labor and Social Policy	Active	07/01/2014	77	0
MIA Border crossings entry/exit	Open data	Ministry of Internal Affairs	Active	07/01/2014	0	347
AEC Transferred numbers	Open data	Agency for Electronic Communications	Active	11/11/2014	59	221
AREC Statistics per municipality	Open data	Agency for Real-Estate Cadastre	Active	01/08/2015	29	204
CA Customs of goods	Open data	Customs Administration	Active	07/01/2014	0	105
…						

Source: www.opendata.gov.mk

In order to make use of DCAT-AP, different levels of information have to be considered. For each of these levels (called “classes”), further specifications for properties exist:

Table 18: Classes and Properties from DCAT-AP

Class	Description	Example from Macedonia
Catalogue	A catalogue or repository that hosts the Datasets being described.	www.opendata.gov.mk Mandatory properties: • *dataset* • *description* • *publisher* • *title*
Agent	An entity that is associated with Catalogues and/or Datasets.	Mandatory properties: • *name*: “Ministry of Information Society and Administration”

Dataset	A conceptual entity that represents the information published.	Mandatory properties: • description • title
Category	A subject of a Dataset.	Mandatory properties: • *preferred label*: "Open data" As all datasets are in category "Open data" – other labels should be used here (see category scheme)
Category scheme	A concept collection (e.g. controlled vocabulary) in which the Category is defined.	Mandatory properties: • *title*: "Macedonian OGD categories" A Macedonian vocabulary could be built up, but it should try to match as close as possible to the publicdata.eu vocabulary[23]: • Employment • Population • Education • Communication • Finance and Budgeting • Geography • Social Questions • Health • Culture and • Agriculture, Fisheries, Forestry • Environment • Transportation • Government Services Politics and Transparency • Economy and Industry

Source: https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/dcat-application-profile-data-portals-europe-final

This example provides a quick overview of how DCAT might be used to represent a government catalogue and its datasets.[24]

First, the catalogue description:

```
:catalogue
    a dcat:Catalogue ;
    dct:title "Open Data Macedonia" ;
    rdfs:label "Open Data Macedonia " ;
    foaf:homepage <http://www.opendata.gov.mk> ;
    dct:publisher :MISA;
```

[23] A detailed comparison for category vocabularies can be found here: http://reference.e-government.gv.at/uploads/media/OGD-Metadaten_2_3_2015_02_19_EN.pdf, p 30

[24] See http://www.w3.org/TR/2013/WD-vocab-dcat-20130312/

```
    dct:language <http://id.loc.gov/vocabulary/iso639-1/mk> ;
        dcat:dataset :MLSP_directory_kindergartens, : MF_state_budget; ….
.
```

The publisher of the catalogue has the relative URI :MISA. Further description of the publisher can be provided as in the following example:

```
:MISA
   a foaf:Organization ;
   rdfs:label "Ministry of Information Society and Administration" ;
   .
```

The catalogue lists each of its datasets via dcat:dataset property. In the example above, an example dataset was mentioned with the relative URI :dataset-001. A possible description of it using DCAT is shown below:

```
:MLSP_directory_kindergartens
   a dcat:Dataset ;
   dct:title "MLSP Directory of children's kindergartens " ;
   dcat:keyword "kindergarten","child care" ,"directory" ;
   dct:issued "2014-01-07"^^xsd:date ;
   dct:modified "2014-01-07"^^xsd:date ;
   dct:publisher :Ministry of Information Society and Administration;
   dct:language <http://id.loc.gov/vocabulary/iso639-1/mk> ;
   dcat:distribution :javni_gradinki.ods;
   .
```

The dataset distribution :javni_gradinki.ods can be downloaded as a 22Kb ODS file. This information is represented via an RDF resource of type dcat:Distribution.

```
:javni_gradinki.ods
   a dcat:Distribution ;
   dcat:downloadURL <http://www.mtsp.gov.mk/content/opendata/javni_gradinki.ods> ;
   dct:title "ODS table directory of children's kindergartens" ;
       dcat:theme :social questions
   dcat:mediaType "ods" ;
   dcat:byteSize "22732,8"^^xsd:decimal ;
   .
```

Apart from the technical necessities of describing a resource, out from the application domain of providing metadata on OGD, there is a need to provide further metadata. In the case auf die Austrian metadata standard[25], the following fields are mandatory of OGD:

- Metadata identifier
- Metadata modified on
- Title
- Description
- Categorization
- Keywords
- Resource URL
- Resource format
- Maintainer

[25] See http://reference.e-government.gv.at/OGD-Metadaten-2-3.3269.0.html

- Publisher
- License
- Begin date and time

This means it is feasible to extend the range of mandatory fields. E. g. it is important to specify a license for a dataset, so users of the dataset know which license applies (see section 4.1.1).

5.3 Domain Specific Vocabularies

Domain specific vocabularies might not only be set up based on the core vocabularies, they might already exist in registers, services or as concepts for certain domains, e. g. e-Health, e-Justice, etc.

Therefore it is advisable to look to international best practices (e. g. on www.joniup.eu) for standardised vocabularies on the EU level or on published vocabularies for specific domains from single member states. Furthermore the Macedonian institutions might already have defined vocabularies in place.

Another approach is to examine existing Macedonian registers/services and look into the type of data and descriptions used there. Examples of this can include

- Population Register
- Company Register
- Real Estate Cadaster, Katactap (the electronic land registry system)
- Multiplatform collaboration system for e-session
- e-Reminder project
- 'See, Report, Fix' project
- System for tracking the status of a document through the use of mobile technologies.
- ESPP (Electronic System for Public Procurement),
- EXIM (Single Window for Export/Import Licenses and Tariff Quotas)
- Online Registration of Employment
- CEMT (Automated System for the Management of International Cargo Transport Licenses),
- e-Tax (Electronic Tax Service)
- Electronic Health Registers
- ORGM (Setting up XML standards in Legislative publishing)
- LDBIS (Legal DataBase Information System),
- Payment of administrative fees by mobile phone
- e-Accounts project
- N-VIS (National Visa Management System)
- integrated system for personalization of documents,
- National Certification Authority
- SWEB ("Secure, interoperable cross-border m-services towards a trustful European cooperation with the non-EU member Western Balkan countries")

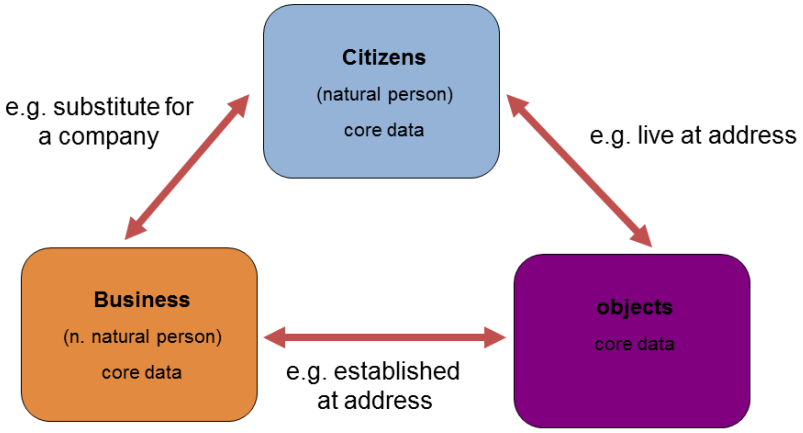
http://www.epractice.eu/files/eGovernmentFYROM.pdf

http://ibarometer.epu.ntua.gr/

5.4 Core Register/Services

Registers are the main basis for many eGovernment applications and a main source to apply core vocabularies interfaces. They offer relevant information on the respective governmental purpose and create valid data within the public sector. This makes it possible for the government to provide fast and dependable action.

Improvement in the quality of the registers is a central theme since only valid data ('quality of data') can promote their use in (electronic) procedures. A core register should be defined for **natural** and **legal persons** with the necessary attributes and optimal processes for entering and updating the data as well as for all kind of **objects** (e.g. codelists, …). Cleaning up, merging and synchronising registers should be drawn up on this basis taken especially into account the approaches provided in the ADMS section of this paper as well as the usage of referenced identifiers/data (URI, …). Updates must be made to be required by law to ensure that the register basis is kept current. Wherever data privacy allows, the transmission of data must be made legal and automated queries must be made possible. Standardised technical interfaces need to be created to be used for all registers according to the approach shown in IOP-T document. With the creation of these kinds of automated mechanisms, it should be a duty of the authorities to support queries to the register. This would remove the need for businesses and citizens to submit information which is already saved in central registers (e.g. Central Register of Residents, Marital status, …). This Central Register would be the key to an optimal procedure flow with the focus on "one-stop".



Especially towards (international / European) administrations the usage of core vocabularies (person, business, location, service, …) and agreed assets of a special domain should be foreseen in new projects or relaunches as a connecting 'interface' out of those core register implementations.

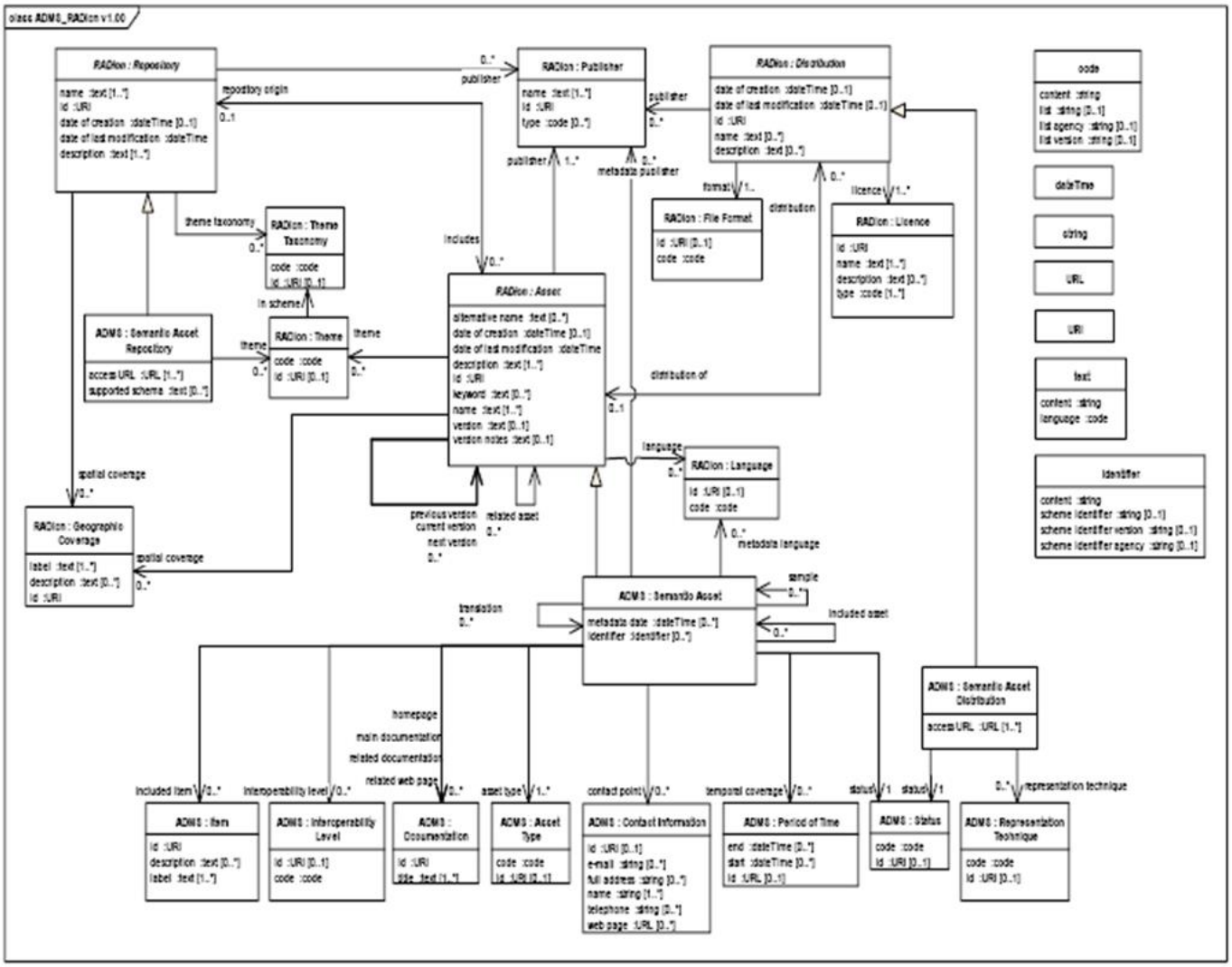
6 Annex

6.1 List of Figures

6.2 List of Tables

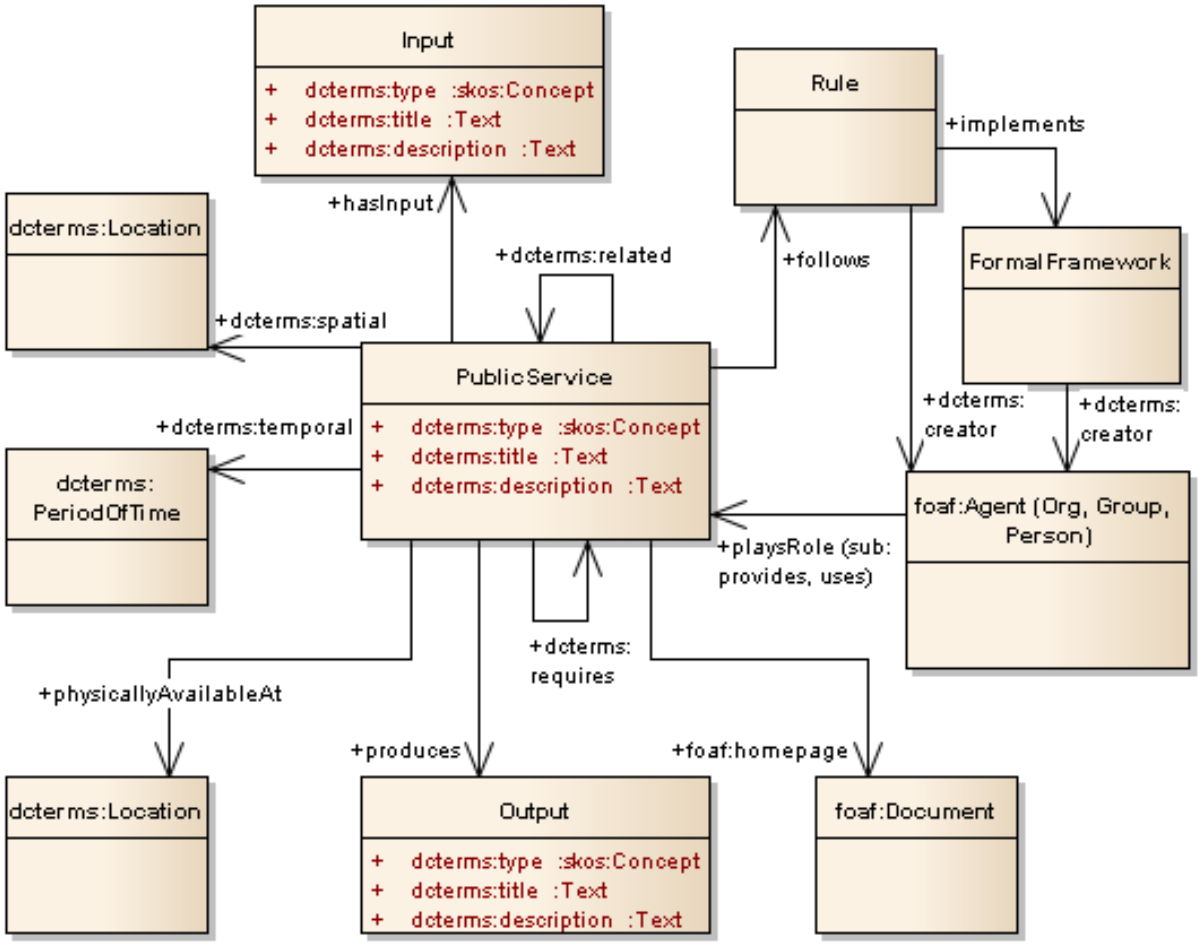
6.3 UML Diagrams

Figure 6: UML Diagram of ADMS



Source: https://joinup.ec.europa.eu/asset/adms/

Figure 7: UML Diagram of Public Service Vocabulary



Source: https://joinup.ec.europa.eu/asset/core_public_service/

6.4 Methodology of asset creation

The following steps illustrate the methodology of asset creation:

Table 19: Methodology of Asset Creation

Step	Description
1	Identify the Core Vocabularies that are likely to meet the most pressing needs of the potential users within Macedonian institutions and for which the relevant expertise is most readily available.
2	The Working Group should research existing vocabularies, their provenance, usage and stability.
3	Research existing published data and services, noting that there should not be any conflicts with the proposed Core Vocabulary.
4	Articulate the problem(s) that the Working Group is trying to solve in the form of a series of use cases.
5	Derive a set of requirements from the use cases.
6	Publish the use cases and requirements in a single document.
7	Unless the vocabulary is very simple with no more than a handful of concepts, create a concept diagram. Consider using UML.
8	Do not impose cardinality rules or domain/range restrictions on vocabulary terms unless strictly necessary.
9	Use words beginning with an upper or lower case letter (A - Z, a - z) or an underscore (_) for all terms in a vocabulary.
10	Use simple nouns for property names.
11	Use verbs for relationship terms.
12	For each relationship, include a definition of its inverse.
13	Use prepositions in vocabulary terms only if necessary.
14	Use a namespace ending with a hash character (#)
15	Keep the namespace as short as possible.
16	Include a portion that identifies the vocabulary for human readers.
17	Do not include any technology-specific component in the namespace (except HTTP)
18	When choosing a namespace, do not restrict the pool of potential users unnecessarily by using a namespace that declares 'ownership' or geographical relevance.
19	If necessary, consider meeting Method Step 18 using PURLs
20	Create and validate the namespace documents in HTML, XML and RDF/XML. Consider also serialising the RDF schema in Turtle.
21	Either the Working Group or the Ministry must make each one available through {namespace}.ext where .ext is the relevant extension.
22	Either the Working Group or the EC must set up content negotiation to handle requests to the namespace itself.
23	When publishing the final version of the Core Vocabulary, link the HTML.

Source: https://joinup.ec.europa.eu/community/semic/document/isa-deliverable-process-and-methodology-developing-core-vocabularies