



PROPOSED ADMS CONCEPTUAL MODEL AND CONTROLLED VOCABULARIES

Second Working Group DRAFT (15 November 2011)

JOINING UP GOVERNMENTS



TABLE OF CONTENTS

1. Conceptual model	3
1.1 Domain model.....	3
1.2 UML diagram	3
1.3 classes.....	4
1.4 Properties	5
1.5 Relationships	5
2. controlled vocabularies	7
2.1 Interoperability level, Asset type and Resource Type Vocabularies	7
2.2 Domain vocabulary	8
2.3 File format vocabulary	8
2.4 Geographic coverage vocabulary.....	8
2.5 Language vocabulary	8
2.6 Licence vocabulary	8
2.7 Status vocabulary	8
2.8 Subject vocabularies	9
2.9 Other Vocabularies and reference collections.....	9

1. CONCEPTUAL MODEL

1.1 DOMAIN MODEL

In the context of federation of repositories of Semantic Interoperability Assets, a number of entities are relevant. The primary entities to be described by ADMS are the following:

- Repository: a collection of asset descriptions
- Asset: a conceptual resource that can be used to facilitate interoperability
- Release: an implementation of an asset that can be accessed or downloaded

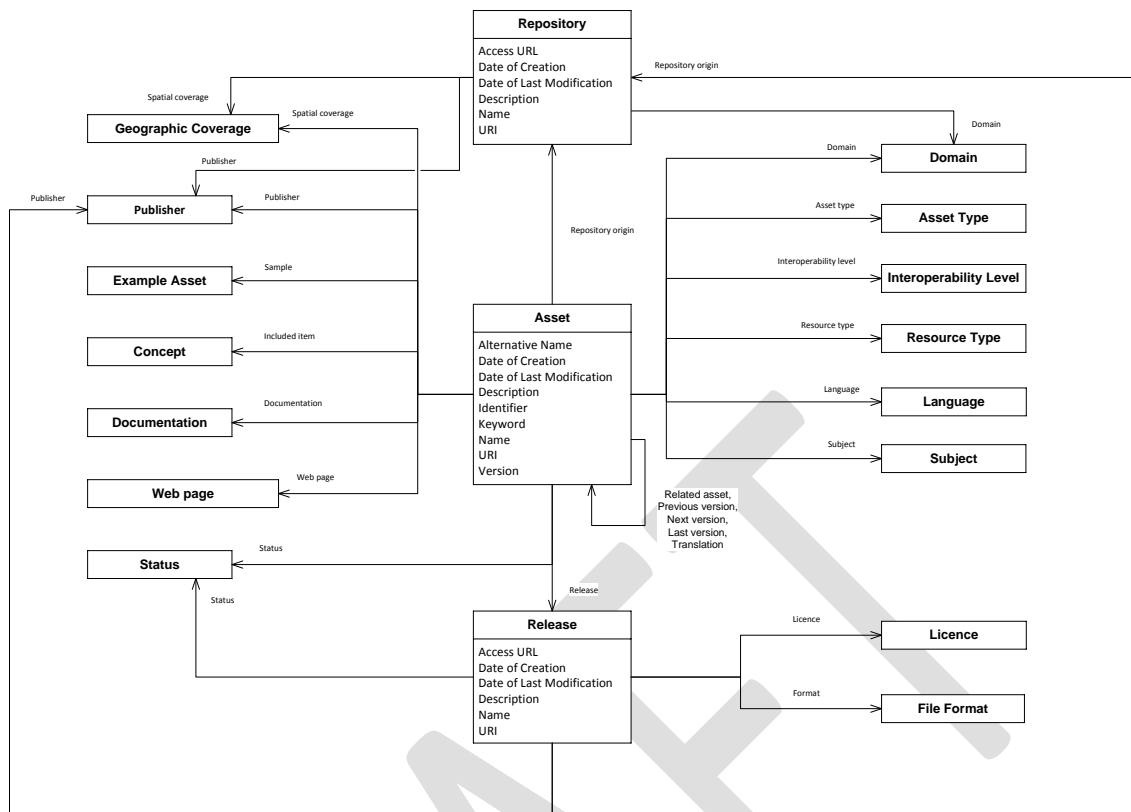
Furthermore, there are a number of secondary or supporting entities:

- Asset type: type of Asset that is being described
- Concept: concepts that are contained in an Asset, e.g. if the Asset is a controlled vocabulary, this would be one of the vocabulary terms
- Documentation: documents that further describe an Asset or give guidelines for its use
- Domain: service area that an Asset or Repository applies to, e.g. “law” or “environment”
- Example asset: sample of an Asset that a user can use to determine whether or not it is relevant for their purposes
- File format: technical format that an Asset is available in, e.g. PDF for a document, XML for a schema
- Geographical coverage, e.g. country or region that an Asset or Repository applies to
- Interoperability level: level according to the European Interoperability Framework (EIF 2.0)¹ that an Asset is related to
- Language: language of an Asset
- License: conditions or restrictions that apply to the use of a Release
- Publisher: organisation responsible for an Asset or Repository
- Resource Type: the ‘genre’ of an Asset .e.g. text or metadata or a dataset
- Status: indication of the maturity of an Asset or Release
- Subject: theme or subject of an Asset, e.g. “elections” or “immigration”
- Web page: Web page associated with an Asset

1.2 UML DIAGRAM

The model presented in the next figure shows the various entities types with their relationships and the descriptive information for the three main objects Repository, Asset and Release. Following the diagram, the classes, properties and relationships will be described in some more detail.

¹ http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf



1.3 CLASSES

Entity	Description
Asset	one of the three main classes of this specification; the class of conceptual resources
Asset Type	the class of types of Assets that can be included in a repository
Concept	The class of items that are contained in Assets
Documentation	the class of documents that describe an Asset
Domain	the class of domains/sectors relevant for eGovernment interoperability (e.g. environment, law enforcement, social services) that an Asset or Repository covers
Example Asset	the class of samples of Assets
File Format	the class of technical formats that a Release can be available in (e.g. the set of tags defined by IANA media types)
Geographic Coverage	the class of geographic locations or jurisdictions to which an Asset or Repository applies
Interoperability Level	the class of interoperability levels according to EIF 2.0
Language	the class of languages of Assets (e.g. the set of tags defined by RFC4646)
Licence	the class of licences that govern (re-)use of Releases (e.g. Creative Commons, EU Public Licence)

<i>Entity</i>	<i>Description</i>
Publisher	the class of organisations (agencies, companies etc.) that are or were involved in the creation and maintenance of an Asset
Release	one of the three main classes of this specification; the class of implementations of Assets in a particular format
Repository	one of the three main classes of this specification; the class of repositories of Assets and Releases
Resource type	the class of genres that an Asset can have
Status	the class of statuses
Subject	the class of subjects of themes that an Asset relates to
Web page	the class of Web documents that contain further information about an Asset (e.g. a Web page)

1.4 PROPERTIES

<i>Property</i>	<i>Description</i>
Access URL	URL of the Repository or Release
Alternative Name	alternative name for the Asset, e.g. nickname, acronym; free text string
Date of Creation	creation date of this version of the Asset, Repository or Release; date expressed using W3CDTF
Date of Last Modification	date of latest update of Asset, Repository or Release; date expressed using W3CDTF
Description	descriptive text for the Asset, Repository or Release; text string
Identifier	any identifier for the Asset; string
Keyword	word or phrase to describe the Asset; text string
Name	name of the Asset, Repository or Release; text string
URI	URI for the Asset, Repository or Release
Version	version number or other designation of the Asset; text string

1.5 RELATIONSHIPS

<i>Relationship</i>	<i>Description</i>
Asset type	type of the Asset
Documentation	further documentation of the Asset
Domain	domain or sector to which the Asset or Repository applies
Format	format that the Release is available in (e.g. PDF, XML, RDF/XML, HTML)

<i>Relationship</i>	<i>Description</i>
Included item	item that is contained in the Asset (e.g. a concept in a controlled vocabulary)
Interoperability level	interoperability level that the Asset is relevant for
Language	language of the Asset
Licence	conditions or restrictions for (re-use) of the Release
Last version	most current version of the Asset
Next version	newer version of the Asset
Previous version	older version of the Asset
Publisher	organisation responsible for the publication of the Asset, Repository or Release
Related asset	unspecified relationship between Assets
Release	implementation of the Asset in a particular format
Repository origin	Repository that contains the primary description of the Asset or Release
Resource type	genre of the Asset
Sample	sample of the Asset
Spatial coverage	geographic region or jurisdiction that the Asset or Repository applies to
Subject	subject or theme that the Asset covers
Status	status of the Asset or Release in the context of a particular workflow process
Translation	translation of the Asset
Web page	Web page associated with the Asset

2. CONTROLLED VOCABULARIES

The section identifies a number of controlled vocabularies to be used for specific entities (Classes) in the ADMS model. In this section, the term “vocabulary” is used as shorthand for various types of controlled vocabularies, including taxonomies (collections of controlled vocabulary terms in a hierarchical structure), thesauri (networked collections of controlled vocabulary terms) and ontologies (controlled vocabularies expressed in an ontology representation language)².

In general, it is recommended to use controlled collections of terms as much as possible. Where such collections do not exist, repository owners should create and maintain such collections to ensure consistent descriptions.

2.1 INTEROPERABILITY LEVEL, ASSET TYPE AND RESOURCE TYPE VOCABULARIES

The table below gives the relationships between the vocabularies for the entities Interoperability Level, Asset Type and Resource Type.

EIF Interoperability Level	Asset Type	Resource Type
Political	Policy document	Text
	Policy Implementation Guideline	Text
Legal	Legislation	Text
	Legal Implementation Guideline	Text
	Licences for re-use	Text
Organisational	Organisational Policy document	Text
	Organisational Implementation Guideline	Text
Semantic	Semantic Policy document	Text
	Metadata Specification	Metadata
	Metadata Schema	Schema
	Controlled Vocabulary	Text
	Mapping Specification	Metadata
	Syntax Specification	Metadata
	Code List	Dataset
Technical	Semantic Implementation Guideline	Text
	Register of organisations	Dataset
	Geographic Reference	Reference
	Technical Policy document	Text
	Technical Implementation Guideline	Text
	Interoperability Framework	Text
	Technical Specification	Metadata
	System/Service	Reference
	API Specification	Metadata

² Based on Woody Pidcock, “What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model?” <http://infogrid.org/wiki/Reference/PidcockArticle>

Additional types can be included later based on observed needs. See section 9 Extension mechanism for further details.

2.2 DOMAIN VOCABULARY

The vocabulary for the entity Domain is the domain level³ of Eurovoc⁴.

2.3 FILE FORMAT VOCABULARY

The vocabulary for the entity File Format is the list of IANA MIME Media Types⁵.

2.4 GEOGRAPHIC COVERAGE VOCABULARY

The vocabulary for the entity Geographic Coverage if related to regions in Europe is NUTS⁶, the Nomenclature of territorial units for statistics maintained by Eurostat. For countries and regions not included in NUTS, DBPedia⁷ references may be used.

2.5 LANGUAGE VOCABULARY

The vocabulary for the entity Language is the code list defined by IETF RFC 4646⁸. As an alternative, a set of URIs is defined by Lexvo.org⁹ that include a URI for each of the languages in RFC 4646.

2.6 LICENCE VOCABULARY

A proposed vocabulary for the entity Licence is the one defined by Creative Commons¹⁰. Other licence vocabularies may be considered depending on needs.

2.7 STATUS VOCABULARY

The proposed values for the terms to be used for the entity Status are:

- Published
- Under development
- Deprecated
- Withdrawn

³ http://eurovoc.europa.eu/drupal/?q=download/subject_oriented&cl=en

⁴ <http://eurovoc.europa.eu/>

⁵ <http://www.iana.org/assignments/media-types/index.html>

⁶ http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction

⁷ <http://dbpedia.org/About>

⁸ <http://www.ietf.org/rfc/rfc4646.txt>

⁹ <http://www.lexvo.org/>

¹⁰ <http://creativecommons.org/>

This list is to be used as a generic list. Statuses related to local workflow processes should be mapped to this list in order to facilitate cross/collection searching.

2.8 SUBJECT VOCABULARIES

Many candidate vocabularies are available for the entity Subject. Examples are the European Commission's ECLAS¹¹ Thesaurus and Eurovoc¹², EIONET's GEMET¹³, FAO's AGROVOC¹⁴, ZBW's STW Thesaurus for Economics¹⁵, the Library of Congress' Subject Headings (LCSH)¹⁶ and Thesaurus for Graphic Materials (TGM)¹⁷ and others. The use of these vocabularies is closely linked to the domain that an Asset is related to, and it should be recommended that a terms should be assigned from a vocabulary that is most widely used in the domain covered.

2.9 OTHER VOCABULARIES AND REFERENCE COLLECTIONS

For the entity Publisher, controlled collections of terms identifying government agencies may be available on a national or regional level which could be used to ensure consistent descriptions.

For the entity Concept, the relationship could be directly to the concepts that are contained in the Asset if they can be referenced separately, or a standardised concept vocabulary, such as the Universal Data Element Framework (UDEF)¹⁸ could be used.

¹¹ http://ec.europa.eu/libraries/doc/catalogues/index_en.htm

¹² <http://eurovoc.europa.eu/>

¹³ <http://www.eionet.europa.eu/gemet>

¹⁴ <http://aims.fao.org/website/AGROVOC-Thesaurus/sub>

¹⁵ <http://zbw.eu/stw/versions/latest/about>

¹⁶ <http://www.loc.gov/aba/cataloging/subject/>

¹⁷ <http://id.loc.gov/vocabulary/graphicMaterials.html>

¹⁸ <https://www.opengroup.org/udef/>