



SC8DI07171

Methodology and tools for Metadata Governance and Management for EU Institutions and Member States

Date: 28/01/2015



Document Metadata

Property	Value
Date	2014-12-08
Status	Accepted
Version	1.00
Authors	Leda Bargiotti – PwC EU Services Michiel De Keyzer – PwC EU Services Makx Dekkers – AMI Consult Stijn Goedertier – PwC EU Services Nikolaos Loutas – PwC EU Services Brecht Wyns – PwC EU Services
Reviewed by	Pieter Breyne – PwC EU Services
Approved by	Vassilios Peristeras – European Commission, DG Informatics Athanasios Karalopoulos – European Commission, DG Informatics

This study was prepared for the ISA Programme by:

PwC EU Services

Disclaimer:

The views expressed in this report are purely those of the authors and may not, in any circumstances, be interpreted as stating an official position of the European Commission.

The European Commission does not guarantee the accuracy of the information included in this study, nor does it accept any responsibility for any use thereof.

Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by the European Commission.

All care has been taken by the author to ensure that s/he has obtained, where necessary, permission to use any parts of manuscripts including illustrations, maps, and graphs, on which intellectual property rights already exist from the titular holder(s) of such rights or from her/his or their legal representative.

Table of Contents

EXECUTIVE SUMMARY	V
1. INTRODUCTION	1
1.1. CONTEXT: THE NEED FOR COORDINATION	1
1.2. SCOPE OF THIS REPORT	3
1.3. AUDIENCE	4
1.4. APPROACH	4
1.5. GLOSSARY	6
2. SPECIFICATIONS FOR METADATA GOVERNANCE	7
2.1. EXAMPLES OF METADATA GOVERNANCE	7
2.1.1. <i>The Interoperability Solutions for European Public Administrations (ISA) Committee</i>	7
2.1.2. <i>Inter-Institutional Metadata Maintenance Committee (IMMC)</i>	7
2.1.3. <i>INSPIRE Maintenance and Implementation Group (MIG)</i>	8
2.1.4. <i>Germany: IT Planning Council and KoSIT</i>	8
2.1.5. <i>The Netherlands: Knowledge and Exploitation Centre Official Government Publications (KOOP)</i>	9
2.2. EXISTING STANDARDS FOR METADATA GOVERNANCE	11
2.2.1. <i>ISO11179-6 Metadata Registration</i>	11
2.2.2. <i>Data Management Body of Knowledge (DMBOK)</i>	11
2.3. SPECIFICATIONS FOR METADATA GOVERNANCE	11
2.3.1. <i>Determine the scope</i>	12
2.3.2. <i>Set up a governance structure</i>	13
2.3.3. <i>Define decision mechanisms</i>	14
2.3.4. <i>Define procedures to handle requests</i>	15
2.3.5. <i>Ensure that modifications are communicated promptly to relevant stakeholders</i>	16
2.3.6. <i>Set up registry as authoritative source</i>	16
2.3.7. <i>Establish enforcement approach</i>	16
2.3.8. <i>Establish a Licensing framework</i>	17
2.3.9. <i>Set up quality controls</i>	18
3. SPECIFICATIONS FOR METADATA MANAGEMENT	19
3.1. EXISTING STANDARDS FOR METADATA MANAGEMENT	19
3.2. REQUIRED EXPERTISE FOR METADATA MANAGEMENT	21
3.3. METADATA MANAGEMENT PROCESS DESIGN PRINCIPLES	22
3.3.1. <i>Documentation of management processes</i>	22
3.3.2. <i>Tailoring of management processes</i>	22
3.3.3. <i>Managing access to the structural metadata</i>	22
3.3.4. <i>Update cycles</i>	22
3.3.5. <i>Versioning</i>	23
3.3.6. <i>Development environments</i>	23
3.4. PROCESS SPECIFICATION	24
3.4.1. <i>Design structural metadata</i>	25
3.4.2. <i>Manage change of structural metadata</i>	25
3.4.3. <i>Harmonise structural metadata</i>	29
3.4.4. <i>Publish structural metadata</i>	31
3.4.5. <i>Deploy structural metadata</i>	32
3.4.6. <i>Retire structural metadata</i>	34
4. SPECIFICATIONS FOR METADATA TOOLS	36
4.1. EXISTING STANDARDS FOR METADATA REPRESENTATION	37
4.1.1. <i>Unified Modelling Language</i>	37
4.1.2. <i>XML Metadata Interchange</i>	37
4.1.3. <i>XML Schema languages</i>	38

4.1.4.	<i>RDF Schema</i>	38
4.1.5.	<i>Simple Knowledge Organisation System (SKOS)</i>	39
4.1.6.	<i>GeneriCode</i>	40
4.1.7.	<i>HTTP URIs</i>	40
4.1.8.	<i>Asset Description Metadata Schema (ADMS)</i>	41
4.2.	EXISTING METADATA SERVICES AND TOOLS.....	41
4.2.1.	<i>Data model editors</i>	42
4.2.2.	<i>Reference data editors</i>	43
4.2.3.	<i>Publication tools</i>	43
4.2.4.	<i>Deployment tools</i>	46
4.2.5.	<i>Change management</i>	47
4.3.	PROPOSED METADATA TOOLS: GAP ANALYSIS	47
5.	CONCLUSION AND NEXT STEPS	52
	BIBLIOGRAPHY	54
	ANNEX I IDENTIFIED METADATA TOOLS	55
	ANNEX II STAKEHOLDER REQUESTS AND NEEDS	58

List of Figures

Figure 1 – Business case for a common approach to metadata management and governance.....	2
Figure 2 – Levels of metadata management and governance in scope of this report	3
Figure 3 – Requirements gathering and specification approach.....	4
Figure 4 - Organisational structure of IT-Planungsrat	9
Figure 5 – Governance structure OWMS.....	10
Figure 6 – The level of abstraction	12
Figure 7: Metadata Management Lifecycle.....	25
Figure 8: Manage Change Requests	26
Figure 9: Manage Change Release	28
Figure 10: Harmonise Structural Metadata	30
Figure 11: Publish Structural Metadata.....	32
Figure 12: Deploy Structural Metadata	33
Figure 13: Retire Structural Metadata.....	35
Figure 14 - 10 rules for persistent URIs	41
Figure 15 - Tool Categories: generic tool chain	42
Figure 16: Re3Gistry Information Model (Joint Research Centre, 2014)	45
Figure 17: Re3Gistry Schematic Details (Joint Research Centre, 2014).....	45

Figure 18 - Proposed Tool Chain..... 48

List of Tables

Table 1 - Overview of stakeholder groups in the context of this report 5

Table 2 – Glossary..... 6

Table 3 – Comitology procedure 15

Table 4 – Manage changes in structural metadata 28

Table 5 - Mapping of requirements to existing tools 49

Table 6 – Existing tools 55

Table 7 – Metadata governance: stakeholder requests and needs 58

Table 8 – Metadata management: stakeholder requests and needs..... 60

Table 9 – Metadata tools: stakeholder requests and needs 61

EXECUTIVE SUMMARY

This document contains a *generic* specification for metadata governance, metadata management, and metadata tools that may be used by EU institutions and Member States. It is based on requirements identified in the survey 'metadata management requirements and existing solutions in EU Institutions and Member States' [1]¹, as well as on the results of work carried out with the Publications Office of the European Union, the Directorates-General for Competition (DG COMP) and for Maritime Affairs and Fisheries (DG MARE). Also the Publication Office (PO) has contributed its expertise with metadata governance and management to this document. The report is commissioned by the Interoperability Solutions for European Public Administrations (ISA) Programme of the European Commission.

Chapter 1 defines structural metadata as the whole of data models and reference data and provides an introduction to **the need for inter-organisational coordination** to increase semantic interoperability. Proper metadata governance, metadata management and metadata tooling is key for public administrations to enhance coordination and attain more semantic interoperability.

Chapter 2 contains a generic specification for **metadata governance**. It starts with examples of metadata governance such as the Inter-institutional Metadata Maintenance Committee (IMMC) process; the INSPIRE maintenance and implementation group (MIG) and the Dutch Knowledge and Exploitation Centre Official Government Publications (KOOOP). Based on the description of the existing governance models and the requirements **specifications for metadata governance are provided** by enlisting a number of activities that are considered important based on the feedback received by the various stakeholders interviewed.

Chapter 3 contains a specification of **metadata lifecycle management processes**. A number of **design principles** and **process specifications** are given which should be implemented and **tailored to the specific scope** and context of each organisation.

Chapter 4 identifies **metadata standards and metadata tools**, analyses coverage of requirements by existing tools, and formulates recommendations for the further promotion and possible integration of these tools. The report indicates that existing tools are readily available and already used within the EU institutions to support the metadata lifecycle management process. The report indicates that existing tools are readily available and already used within the EU institutions to support the metadata lifecycle management process.

Chapter 5 concludes the report by suggesting a number of next steps. One first step would be to promote the *use* of the generic specification by *tailoring* of the proposed governance, management and tools to the specific needs of EU institutions. Furthermore, implementation experiences should be captured in **lessons learned**, and fed back into the generic specification. Regarding tooling the next steps should be to further promote existing metadata tools as a reference within the EC and

¹ ISA Programme (2014). D4.1 Metadata management requirements and existing solutions in EU Institutions and Member States. <https://joinup.ec.europa.eu/node/78172>

strengthen their integration using the open standards and specifications identified in this report.

1. INTRODUCTION

This report is commissioned by the Interoperability Solutions for European Public Administrations (ISA) Programme of the European Commission. It contains the requirements and specifications for the governance and management of structural metadata as well as for metadata tools that may be used by EU institutions and Member States.

1.1. Context: the need for coordination

The ever increasing volume of information exchanged within and between different organisations at both national and EU level requires setting up solutions that facilitate its automatic processing. Whilst technological developments offer various means to automate the exchange of information, technological developments alone cannot guarantee greater interoperability between information systems. A fundamental aspect is the need for common **structural metadata**²: data models and/or reference data, which can be defined as follows:

A **data model** is a collection of entities, their properties and the relationships among them, which aims at formally representing a domain, a concept or a real-world thing. In practice, data models drive the design and development of information systems, as they can express the different types of information managed by an organisation.

Reference data is a small, discrete set of values that are not updated as part of business transactions but are usually used to impose consistent classification. Reference data normally has a low update frequency. Reference data is relevant across more than one business systems belonging to different organisations and sectors³.

Figure 1 summarises for a common approach to metadata management and governance. To make sure that different entities use the same structural metadata, stakeholders should invest time and effort to *coordinate* among each other. Uncoordinated exchanges among public administrations may lead to among others:

- Limited re-use of structural metadata, such as data models and reference data, that already exist because people are not aware of their existence;
- Use of competing standards;
- Use of different formalisms for encoding structural metadata and incompatible licensing rules;
- Ad-hoc development of structural metadata that do not follow a structured process and methodology.

² See also ‘Understanding metadata’ [Nat_2014]

<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

³ J. Jordan & C. Ellen (2009). Business need, data and business intelligence, Journal of Digital Asset Management Vol. 5, 1, 10–20.

Better coordination requires metadata governance and metadata management, which we define as follows:

Metadata governance comprises well-defined roles and responsibilities, cohesive policies and principles, and decision-making processes that define, govern and regulate the lifecycle of metadata.

Metadata management is the good practice of adopting policies, processes, and systems to plan, perform, evaluate, and improve the use and re-use of data models and reference data.

By setting up proper metadata governance, metadata management and tools, public administrations greatly enhance their potential for coordination and interoperability and ultimately the possibilities for sharing and re-use of metadata thanks to:

- increased quality and traceability of the information exchanged;
- greater re-use of standards;
- reduced risk of duplication;
- increased trust towards the information to be exchanged; and
- savings derived from the re-use of already existing information.

Guidelines on how to develop semantic agreements⁴ already exist, but they do not provide details on how to set up a metadata governance and management for this information.

The purpose of this study is therefore to provide guidelines for the setting up of metadata governance, management and related tools both at the EU and Member States levels, taking stock of best practices and lessons learned from already functioning metadata governance and management initiatives.

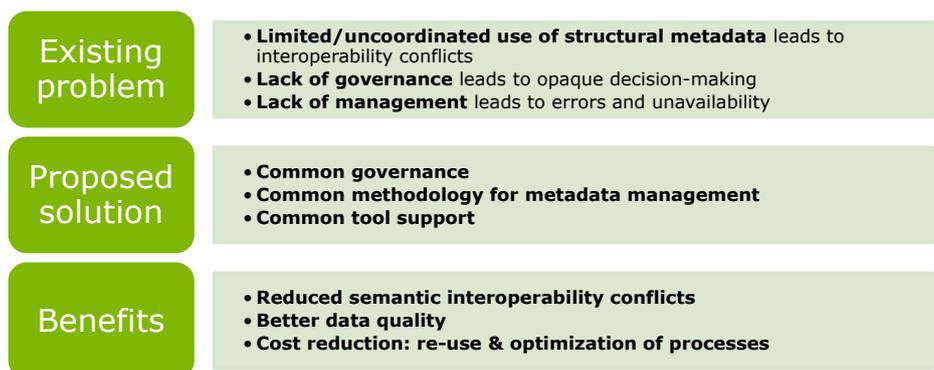


Figure 1 – Business case for a common approach to metadata management and governance

⁴ European Commission. ISA Programme. Process and methodology for developing semantic agreements. 8 February 2013. https://joinup.ec.europa.eu/community/core_vocabularies/document/process-and-methodology-developing-semantic-agreements

1.2. Scope of this report

The scope of this report entails:

- **Local, inter-institutional and trans-European exchanges:** the study is of application for information exchanges that take place at three levels, as depicted in Figure 2: local (i.e. within an EU institution), inter-institutional (between EU institutions) and trans-European (between Member States and the EU institutions).
- **High-level specifications:** the study provides high level solutions that can be applied within a given public administration as well as across various public administrations and domains. Implementation in individual cases needs to be tailored to a specific organisational and technical environment, and therefore a more in-depth implementation guide will be necessary using principles laid down in this document.
- **Structural metadata:** the study focuses on the governance, management and tools for structural metadata⁵ only.

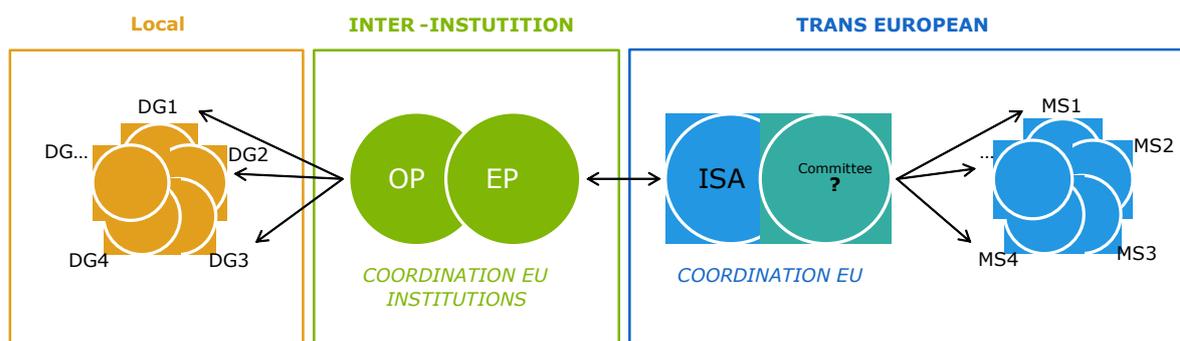


Figure 2 – Levels of metadata management and governance in scope of this report

The following is outside the scope of this report:

- **Metadata design:** This study starts from the assumption that common data models, metadata schemata and reference data have already been agreed upon by the network participants. As a consequence, this study focuses on the structural metadata lifecycle that takes place after a semantic agreement has been set up.
- **Metadata other than structural metadata:** Excluded from the scope are type of metadata other than structural metadata, such as descriptive metadata, i.e. the description of documents, services and other resources that may be created, kept and shared across a network.
- As the management and governance of **persistent URIs** is described in the URI policy [2], the topic is excluded from the scope of this report.

⁵ Structural metadata are: data models or reference data
28/01/2015

1.3. Audience

The main audience of this report is represented by the staff of the EU Institutions and consultative bodies as well as staff from national public administrations involved in metadata governance and management, and tasked to organise and operate the governance structures and maintenance activities.

1.4. Approach

We have decomposed the requirement analysis and specifications for the proposed solution in three parts:

1. Specifications for metadata governance, i.e. decision-making (see Section 2);
2. Specifications for metadata management, i.e. the organisation of the work (see Section 3); and
3. Specifications for metadata tools (see Section 3).



Figure 3 – Requirements gathering and specification approach

For each part, we have undertaken the following steps depicted in Figure 3:

1. First, we identified explicit **stakeholder requests** that emerged in the course of the interviews with stakeholders listed in Table 1. This was then complemented by additional **stakeholder needs** that were not stated explicitly but were validated by the stakeholders. Furthermore, interviews were conducted in the context of the report 'Metadata management requirements and existing solutions in EU Institutions and Member States' [1]⁶. In addition, we conducted three metadata pilots with the European

⁶ ISA Programme (2014). Metadata management requirements and existing solutions in EU Institutions and Member States. <https://joinup.ec.europa.eu/node/78172>

Commission Directorates-General for Competition (DG COMP) and Maritime Affairs and Fisheries (DG MARE), and the Publications Office (PO).

2. Second, we identified **existing solutions** for metadata management, governance, and tools.
3. On the basis of this input, we have elaborated **specifications for metadata governance, metadata management, and metadata tools**. The metadata governance specification consists of a description of scope, organisation structure, decision mechanisms, etc. The metadata management specification consists of a process description. The metadata tooling section consists of a list of standards and tools, and a set of high-level use cases. The pilots also helped validating whether the specifications for metadata governance and management were fit-for-purpose.

The table below lists the stakeholder groups for which the proposed methodology and tools for metadata governance and management may be relevant.

Table 1 - Overview of stakeholder groups in the context of this report

Stakeholder groups	Example of stakeholder organisations
Standardisation organisations	<ul style="list-style-type: none"> • CEN • UN/CEFACT • OASIS • ...
National public administrations	<ul style="list-style-type: none"> • KoSIT (Koordinierungsstelle für IT-Standards), Germany • CISE – Centre for Semantic Interoperability, Spain • Lithuanian Spatial Information Portal (LSIP), Lithuania • Knowledge and Exploitation Centre Official Government Publications (KOOB) , The Netherlands • Local Government Inform (LG Inform / LG Inform Plus), United Kingdom • ...
European Parliament	<ul style="list-style-type: none"> • DG ITEC • ...
Council of the European Union	<ul style="list-style-type: none"> • Archives of the Council of the EU • ...
European Commission	<ul style="list-style-type: none"> • Publications Office (pilot) • European Commission - DG MARE (pilot) • European Commission - DG COMP (pilot) • European Commission - EUROSTAT • European Commission – JRC • ...
Other European institutions	<ul style="list-style-type: none"> • Court of Justice of the European Union • European Court of Auditors • European Central Bank • European Economic and Social Committee • Committee of the Regions • ...

1.5. Glossary

The table below provides common definitions used throughout this report.

Table 2 – Glossary

Term / Acronym	Description
Data model	A data model is a collection of entities, their properties and the relationships among them, which aims at formally representing a domain, a concept or a real-world thing.
Domain	Domain is a specific subject matter area that has government body i.e. ministry or department responsible for that domain e.g. the Ministry of Agriculture, the Ministry of Finance.
Domain model	A domain model is a conceptual view of a system or an information exchange that identifies the entities involved and their relationships.
Interoperability	According to the ISA Decision, interoperability means the ability of disparate and diverse organisations to interact towards mutually beneficial and agreed common goals, involving the sharing of information and knowledge between the organisations, through the business processes they support, by means of the <i>exchange of data</i> between their respective ICT systems.
Metadata	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. (National Information Standards Organization , 2004)
Metadata governance	Metadata governance comprises well-defined roles and responsibilities, cohesive policies and principles, and decision-making processes that define, govern and regulate the lifecycle of metadata.
Metadata management	We define metadata management as the good practice of adopting policies, processes, and systems to plan, perform, evaluate, and improve the use and re-use of data models and reference data.
Metadata alignment	Metadata alignment is the harmonisation of structural metadata either by forging a wide consensus on the use of a common specification for structural metadata or through the creation of mappings between terms of two or more specifications.
Reference data	Reference data is small, discrete sets of values that are not updated as part of business transactions but are usually used to impose consistent classification. Reference data normally has a low update frequency. Reference data is relevant across more than one business systems belonging to different organisations and sectors.
Structural metadata	Data model or reference data.

2. SPECIFICATIONS FOR METADATA GOVERNANCE

This section first analyses examples of metadata governance and existing standards. Then, it formulates generic specifications for metadata governance.

2.1. Examples of metadata governance

This section contains an overview of existing examples of metadata governance. These solutions serve as inspiration for the specifications described in Section 2.3.

2.1.1. *The Interoperability Solutions for European Public Administrations (ISA) Committee*

The European Commission is assisted in the implementation of the Interoperability Solutions for European Public Administrations (ISA) Programme by the **ISA Committee**, which represents the Member States. Furthermore, the **ISA Coordination Group**, nominated by the **ISA Committee**, ensures continuity and consistency at working level.

In the past, the ISA Coordination Group has endorsed structural metadata such as the Core Vocabularies⁷. This governance body may be useful for taking high-level decisions on voluntary, trans-European harmonisation initiatives on structural metadata.

2.1.2. *Inter-Institutional Metadata Maintenance Committee (IMMC)*

The Inter-Institutional Metadata Maintenance Committee (IMMC)⁸ is responsible for the decisions related to key reference data and data models used in the legal decision-making process of EU institutions and the EU Open Data Portal (ODP)⁹.

The IMMC is part of a three-level organisational structure, consisting of an inter-institutional steering committee, a metadata maintenance committee and a metadata registry. The governance methodology applied by the IMMC meets most requirements for inter-institutional governance. However, currently it is limited to inter-institutional governance in the area of the legal decision-making processes of the EU and open data. Given its inter-institutional character, it does not offer the possibility for trans-European governance. Nevertheless, its structure and main principle can be re-used by other entities that want to set up a trans-European and/or local metadata governance mechanism.

⁷ Joinup (30 May 2012), ISA Member State representatives endorse key specifications for e-Government interoperability, <https://joinup.ec.europa.eu/node/48837>

⁸ Publications Office. Proposal for metadata governance on interinstitutional level. http://publications.europa.eu/mdr/resource/core-metadata/IMMC_reu3_adoption_anx3.pdf_A-2011-764293.pdf

⁹ European Union Open Data Portal. <https://open-data.europa.eu/>

2.1.3. INSPIRE Maintenance and Implementation Group (MIG)

The purpose of the INSPIRE Directive is “to lay down general rules aimed at the establishment of the Infrastructure for Spatial Information in the European Community”. Maintenance and evolution of INSPIRE is governed by the INSPIRE Maintenance and Implementation Framework (MIF)¹⁰.

The central role in the governance of metadata management is the INSPIRE Maintenance and Implementation Group (MIG) which is responsible for strategy related to the implementation of INSPIRE. It is chaired by The Joint Research Centre (JRC) of the European Commission and composed of two representatives per country. The INSPIRE Regulatory Committee in which the Member States are represented advises the European Commission on the adoption of the Implementing Rules. Any decisions that require a change in the INSPIRE Regulation are formally taken by the European Commission, the European Parliament and the Council under the Comitology procedure (see Table 3 in section 2.3.3).

The MIG is complemented by a pool of experts drawn from the stakeholder community. The experts in this pool are called upon when MIG sub-groups are formed to address specific implementation or maintenance issues, but will also provide the opportunity to reach out to experts involved or interested in particular aspects of INSPIRE implementation or maintenance.

2.1.4. Germany: IT Planning Council and KoSIT

Since 2009, article 91c of the Basic Law (Grundgesetz), the Constitution of Germany establishes the basis for cooperation between the federal level (Bund) and the state level (Länder) on the implementation and interoperability of IT solutions.

In 2010, the IT Planning Council (IT-Planungsrat)¹¹ was established to coordinate the collaboration between the federal and state levels. The members of the council are the federal state secretary for IT and representatives from the states. In addition, three representatives from local government and the responsible person on the federal level for data protection and freedom of information participate in an advisory role.

Under responsibility of the IT-Planungsrat, KoSIT¹², the Coordination Agency for IT Standards, takes care of the coordination of development and implementation of standards for data exchange. KoSIT manages the XÖV framework (XML in der öffentlichen Verwaltung – XML in public administration) and provides access to several tools, guidelines and XML schemas with code lists, data types and core components.

¹⁰ European Commission. INSPIRE. Maintenance and Implementation.

<http://inspire.ec.europa.eu/index.cfm/pageid/5160/list/mif>

¹¹ IT-Planungsrat. http://www.it-planungsrat.de/DE/Home/home_node.html

¹² IT-Planungsrat. Koordinierungsstelle für IT-Standards. http://www.it-planungsrat.de/DE/Organisation/KoSIT/KoSIT_node.html

All organisations involved in e-Government in Germany can submit requirements for standards to KoSIT. KoSIT submits proposals the IT-Planungsrat which, in its annual meeting, decides on standardisation proposals.

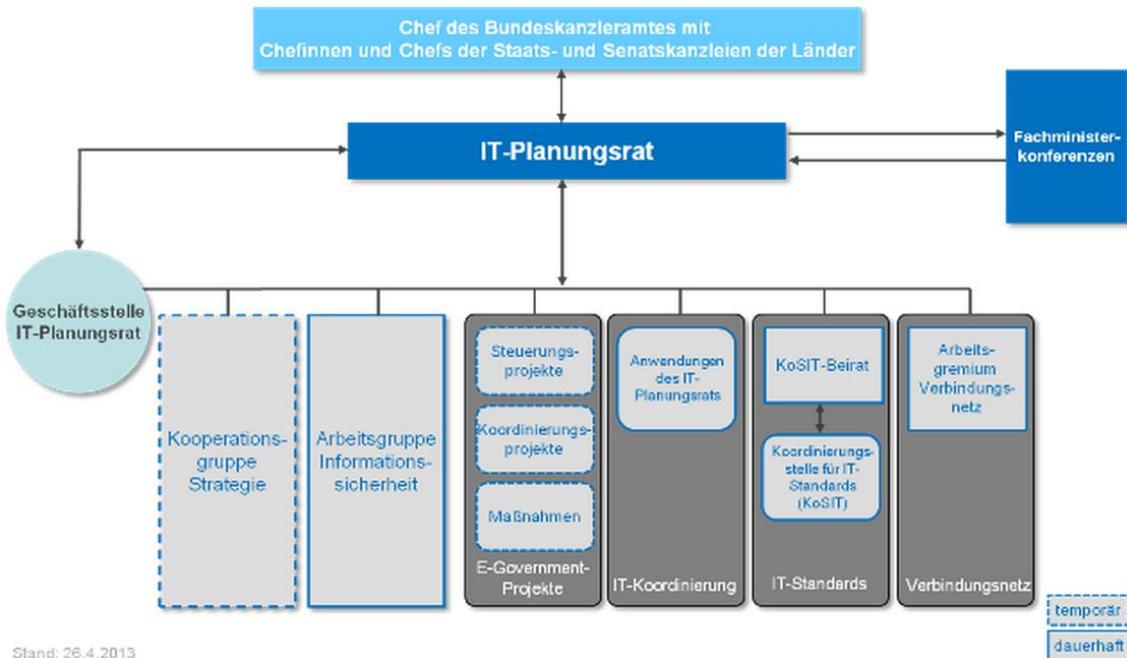


Figure 4 - Organisational structure of IT-Planungsrat

2.1.5. The Netherlands: Knowledge and Exploitation Centre Official Government Publications (KOOP)

The Knowledge and Exploitation Centre Official Government Publications (KOOP) is an autonomous unit under the Ministry of the Interior and Kingdom Relations of The Netherlands. KOOP develops and maintains products and managed services for all levels of government, including central government and provinces, water authorities and municipalities.

KOOP was initially set up to realise the conversion of the three official gazettes (Staatscourant, Staatsblad and Tractatenblad) into electronic publications with the objective to offer the official promulgation of legislation, decrees and treaties exclusively on Internet. These publications are now available at www.overheid.nl.

One of the products developed and maintained by KOOP is the Government Web Metadata Standard OWMS. This national standard, based on the Dublin Core, specifies the metadata properties to be used to provide structured descriptions of unstructured governmental information on the Web, enabling searching, discovering and presentation of such information.

OWMS consists of agreements concerning¹³:

¹³ For more information: <http://standaarden.overheid.nl/owms>

- Properties (descriptors) for describing government information;
- Lists of values to be used for the properties; and
- Syntax of the values to be used for the properties.

The governance structure is as follows¹⁴:

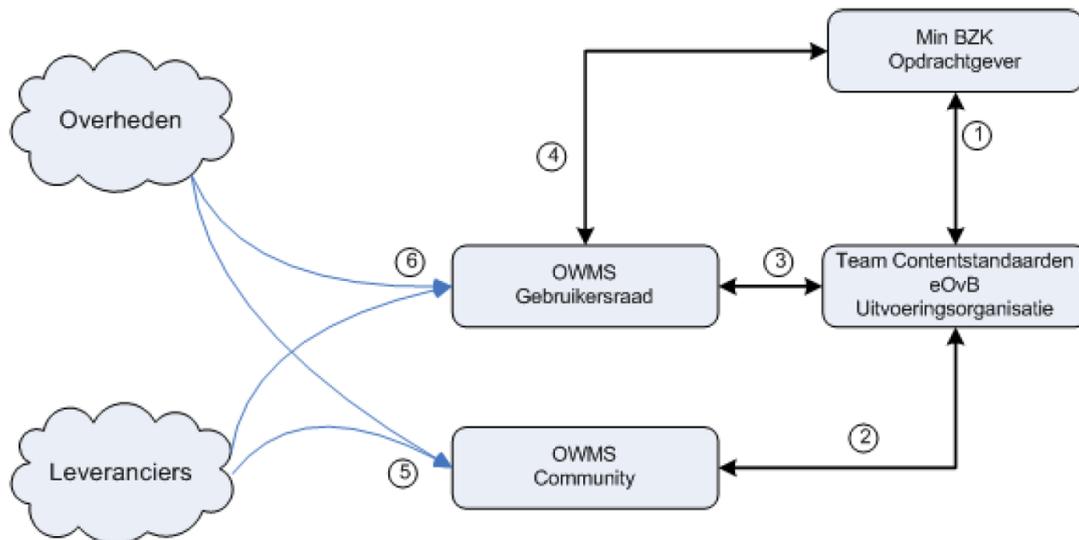


Figure 5 – Governance structure OWMS

1. The Ministry of the Interior and Kingdom Relations (MinBZK) instructs the Team Content Standards (Contentstandaarden) to develop and publish OWMS.
2. Members of the OWMS Community submit change requests to the Team Content Standards. The Team takes the request into consideration and produces a change proposal if the request is feasible and within the scope of its charter.
3. The Team Content Standards submits change proposals to the OWMS User Council (Gebruikersraad) and implements the proposals that are agreed by the User Council.
4. If changes concern normative specifications and would lead to a new version of OWMS, the User Council does not take the decision, but advises the Ministry which then decides on the changes.
5. Anyone with an interest in OWMS can become a member of the OWMS Community.

¹⁴ E-Overheid voor Burgers. Beheerplan OWMS.

<http://standaarden.overheid.nl/owms/beheer/BeheerplanOWMSv1.0.pdf>

6. Membership of the OWMS User Council is open to all Government agencies (Overheden) and vendors (Leveranciers) who apply OWMS in their products and services, subject to consultation with the Ministry.

2.2. Existing standards for metadata governance

This section contains an overview of existing standards, containing specifications for metadata governance. These standards serve as inspiration for the specifications described in Section 2.3.

2.2.1. ISO11179-6 Metadata Registration

A general standard for the registration of metadata items is ISO/IEC 11179. As part of the six-part standard, ISO/IEC 11179-6:2005¹⁵ specifies the procedure by which Administered Items required in various application areas could be registered and assigned an internationally unique identifier. This procedure includes organisations such as the Registration Authority, the Responsible Organisation, and the Submitting Organisation. It also includes roles such as the Registrar, Steward, and Submitter.

2.2.2. Data Management Body of Knowledge (DMBOK)

The Data Management Body of Knowledge (DMBOK)¹⁶ is a general methodology for data management. The DM-BOK guide strives to adoption of a generally accepted view of data management. It provides standard definitions for data management functions, roles, deliverables and other common terminology. The DM-BOK devotes an entire chapter to Reference and Master Data Management.

In terms of Governance Structure, it defines a number of operational roles including the Data Architect, Business Analyst, Data Steward, and Application Architect as responsible roles. It attributes all decision power onto the role of a Data Governance Council.

2.3. Specifications for metadata governance

Based on the description of the existing governance models and the requirements identified above, the next paragraphs provide specifications for metadata governance. It will do so by enlisting a number of activities that are considered important based on the feedback received by the various stakeholders interviewed. Therefore, the following section does not aim to provide an exhaustive list of best practices that are necessary to be applied for the correct functioning of a metadata governance mechanism. Rather, the section aims to extrapolate general best practices from concrete examples coming from the day to day work of a limited number of stakeholders.

¹⁵ ISO/IEC 11179-6:2005. Information technology -- Metadata registries (MDR) -- Part 6: Registration.
http://www.iso.org/iso/catalogue_detail.htm?csnumber=35348

¹⁶ DAMA Data Management Body of Knowledge (DAMA DMBOK). 1st Edition 2009.
<http://www.dama.org/i4a/pages/?pageid=3364>

2.3.1. Determine the scope

When setting up a mechanism for metadata governance it is necessary to determine the scope in which it is applied. The scope comprises among others the following aspects:

The domain or sector: In certain cases it may be limited to one specific domain, such as food safety, defence, healthcare, or finance. In other cases it encompasses a variety of domains such is already the case with regard to the governance of structural metadata in the context of the European Union decision making process. When identifying the policy domain the following elements should be clearly identified: The topics covered; who will be impacted by changes to the structural metadata; the existence of metadata harmonisation efforts for the same instances; the consequences derived from compliance or lack of compliance.

The governance levels: In this report, we have considered the local, inter-institutional and trans-European level.

The level of abstraction: Within one domain or across domains, it is possible to define the extent to which structural metadata is being specified. Consider for instance the following levels of abstraction depicted in Figure 6.

- **Core specifications:** context-neutral structural metadata that defines the fundamental characteristics. The structural metadata can be applied in several contexts. Examples here are the Core Person, Registered Organisation, Location, and Public Service Vocabularies developed by the ISA Programme.
- **Domain specifications:** domain-specific structural metadata that covers a domain to a larger extent. One example here is the HL7 Reference Information Model in the healthcare domain, or the Universal Business Library (UBL) of OASIS.
- **Information exchange specifications:** structural metadata specifications that are specific to one context of information exchange. One example here is the exchange of electronic invoices in Denmark.

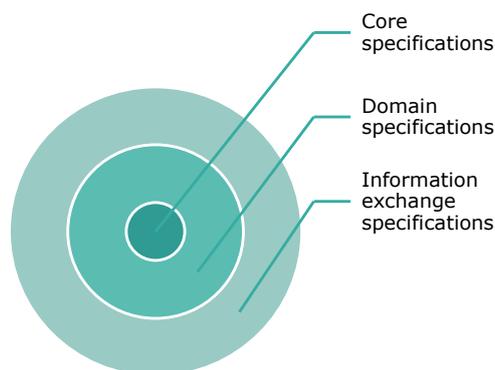


Figure 6 – The level of abstraction

Scope criteria: There must be a clear set of scope criteria that determine whether a specification for structural metadata should be placed under common governance, as this may require considerable coordination costs, but can also entail considerable benefits of interoperability. Some important scope criteria that have been identified as highly relevant in light of the interviewed stakeholders' experience are:

- **The Level of information exchange:** whether the metadata in question will be used exclusively within a given organisation or by two or more institutions/entities for exchanging information;
- **Maturity level:** the stability/maturity of the metadata that an entity wants to share. For example if already at the beginning an organisation knows that a given structural metadata is not mature enough and that will probably still change in the future, it does not make sense to share it with a wider audience, which will start relying on something that actually is not finalised;
- **Potential for use:** A very important criterion is the potential for re-use of a structural metadata. For example, a reference data that may be re-used across various sectors and stakeholders has a greater potential for re-use than a specific data model that only applies to one specific dataset.
- **Commitment for maintenance:** Fundamental is also the degree of commitment to maintain and keep up to date structural metadata. For example, the Publications Office has expressed in different fora its willingness to maintain certain structural metadata, notably the named authority.
- **Commitment for use:** How strong is the commitment of organisations to actually use the specification for structural metadata?

2.3.2. Set up a governance structure

Once the decision to set up a metadata governance mechanism has been taken, it is necessary to put in place the overall structure that the metadata governance should have. The governance should include permanent members, temporary representatives and a secretariat taking care of logistical and coordination matters.

From the concrete cases, it emerged that the optimal solution is to have three-level governance corresponding to:

- **Steering committee:** composed of representatives of the institutions and public administrations that set the strategic directions. It should include representatives from the business and architecture side. This level is driven by people that have the means and vision to take decisions on scope and goals. They meet periodically to review progresses made and intervene sporadically to solve conflicts, and nominate members.
- **Governance committee:** made up of the main stakeholders. It takes decisions on the organisational support required to the operational team. This is for example the role of the IMMC (see section 2.1.2). It oversees the compliance of the operational team and assumes the responsibility to develop, disseminate and enforce the required procedures.
- **Operational team:** It is composed of a single team that carries out the day-to-day work. It deals with various aspects of metadata management and

metadata governance from collection, creation to administration of metadata. It is the level where management is executed. An example is the Metadata Registry (see section 2.1.2).

2.3.3. Define decision mechanisms

Decision mechanisms prescribe how and when to perform tasks related to metadata governance. They are fundamental in achieving the established goals without having to constantly intervene on daily operations.

Decision mechanisms should enable to take decisions such as:

- Whether a metadata specification must be placed under local or inter-institutional governance;
- How to change and improve the metadata management process;
- Whether a change request to a metadata specification must be accepted or rejected (based on an impact analysis; cost-benefit analysis, risk analysis);
- Whether an accepted change request will be released immediately or in a scheduled release;
- Where to store a metadata specification and with which access restrictions (define roles and responsibilities);
- Whether a metadata specification can be published under an open licence;
- Whether a metadata specification can be supplemented with official mappings;
- Which policy is followed to encourage or mandate the reuse of the reference data specification;
- Which method is used for documenting reference data;
- Whether a metadata specification should be deprecated; and
- Which standards and tools to use in the metadata management process.

These decisions can be taken using different modalities:

- Consensus: a decision is taken only when there is a full consensus.
- Majority vote: a decision is taken upon majority vote
- Veto: stakeholders are informed and can raise a strong objection
- Endorsement: asking stakeholders to endorse it after creation / update.

Furthermore, in the context of the European Union, special mentioning should be made with regard to the Comitology procedure. In this context two procedures are particularly relevant for this study:

- The advisory procedure
- The examination procedure

Details of the Comitology procedure are given in Table 3.

Table 3 – Comitology procedure

Illustration: Comitology procedure¹⁷

When the metadata governance involves the EU and Member States and there is a legal instrument requiring uniform conditions for the implementation of structural metadata, then implementing powers can be conferred to the European Commission.

In this case, Member States can control the Commission's exercise of implementing powers. The rules and general principles concerning these control mechanisms are set up in Regulation (EU) No 182/2011¹⁸ of the European Parliament and of the Council of 16 February 2011 laying down the rules and general principles concerning mechanisms for control by Member States of the Commission's exercise of implementing powers.¹⁹

For the purposes of such control, committees composed of the representatives of the Member States and chaired by the Commission are set up. The primary role of these Committees is to provide an opinion on the draft measures that the Commission intends to adopt. These opinions can be more or less binding upon the Commission according to the procedure which has been foreseen by the legislator.

One of the following two procedures is foreseen:

- The advisory procedure: here the Commission shall take the utmost account of the committee's opinion.
- The examination procedure: here implementing acts cannot be adopted by the Commission if they are not in accordance with the opinion of the committee, except in very exceptional circumstances, where they may apply for a limited period of time

In addition, specific procedures are foreseen for measures to apply immediately on imperative grounds of urgency (Article 8). In this case, the Commission adopts an implementing act of immediate application, without its prior submission to a committee.

2.3.4. Define procedures to handle requests

To make sure that the needs of the requestors are taken into account, the metadata governance should establish clear procedures to be followed depending on the case into question. For example, it may be that a requestor submits a request to update a metadata schema. Such a request may have an important impact on several information systems and therefore should be carefully assessed. Here timing may be less relevant than the analysis on the impact that such a request might have. Vice versa a requestor may submit a request for a deprecation and update of a code where the urgency outweighs the impact that such a modification may have. Therefore, when deciding which procedure to apply the structural governance mechanism should take into aspects such as:

- The **justification** behind a given request: is there a real need for taking into account such a request? It may be that the request is made on needs that

¹⁷ See also <https://myintracomm.ec.europa.eu/corp/sg/en/comitology/implementing/pages/tools.aspx> for further information

¹⁸ <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32011R0182>

¹⁹ OJ L 55, 28.2.2011, p. 13–18.

have not really been thought through and therefore the implementation may be postponed or abandoned;

- The **urgency**: does the request need to be implemented as quickly as possible because otherwise several systems will be “blocked”, or stakeholders will be using an outdated version?
- The **impact** of the request in terms of information systems as well as stakeholders involved: it may be that a request for a change of a metadata schema would require an update by several entities and therefore would also impact several systems. In this case, the assessment on the impact should be carried out into details.

An example of good practice in this context comes from the Publications Office. The PO is currently compiling sets of standard requests in order to know already in advance how to treat them based on which category they fall. This approach may save time and help those analysing the various requests in their daily job.

2.3.5. Ensure that modifications are communicated promptly to relevant stakeholders

Once the structural metadata governance mechanism finally takes a decision, it is necessary to ensure that all relevant stakeholders are informed, so that not only they can adapt their systems but can also provide feedback. Therefore, the governance mechanism should establish communication channels through which stakeholders are kept up to date. Depending on the target group and on the way they usually communicate, different solutions may be envisaged including for example: mailing lists, RSS feeds and announcements provided during the plenaries.

2.3.6. Set up registry as authoritative source

When setting up a metadata governance mechanism, it is fundamental to make available an authoritative source on which the metadata is housed. In most cases, the authoritative source is a repository or a file server that is accessible online. It should allow anybody to access code lists, concept schemes, data structure definitions, etc. The existence of an authoritative source increases the confidence of potential re-users because it ensures that everybody has access to the same information as well as the confidence over the quality of the structural metadata.

2.3.7. Establish enforcement approach

The metadata governance mechanism should also establish which enforcement regime should be applied to promote the sharing and re-use of structural metadata and avoid lock-in. Enforcement policy embraces a wide spectrum of activities, going from the drafting of public procurement to the implementation of structural metadata.

With regard to public procurement, it is worthwhile mentioning the best practices and guidelines prepared by the European Commission.²⁰

With regard to the different typologies of enforcement policies, the most common one can be summarised as follows:

- **Legal requirement:** implementation is enforced by law, either by secondary legislation, council conclusions, or by referring to standards, via comitology procedure (as it is the case for state aids and the implementation of the Inspire Directive). Here an important requirement is to make sure that specific structural metadata are not included in the legal instrument, otherwise every time there is a need for an update, then it is necessary to go through the legislative process, which would make it a heavy process not serving users' needs. Details like the values in a code list or the elements of a data model should be specified as part of the implementation documentation and made available from an authoritative source to which the legislation can refer.
- **Comply-or-explain:** implementation is not enforced by law, but public administrations have to comply with the use of a particular specification or standard for metadata otherwise they should explain why they does not fit their needs. In certain cases it may even be requested to contribute to upgrade the model.
- **Voluntary:** implementation is encouraged via information campaigns. What is crucial in this case is that stakeholders share the same goals and are aware of the advantages that an effective and efficient use of the metadata governance may provide. There are several actions that can be undertaken to make sure that this happens.

2.3.8. Establish a Licensing framework

In order to make the metadata available for sharing and re-use purposes, the metadata governance should establish the licensing framework under which the metadata can be exchanged and re-used.

To make sure that metadata are re-used by a critical mass it is recommended to use licences that are as open as possible with protection against misrepresentation.

In addition, in order to increase legal certainty and help potential re-users, it is also recommended to make sure that information related to licensing frameworks is properly conveyed and easily accessible.

Examples of such licences are Creative Commons CCZero (CC0)²¹, Open Data Commons Public Domain Dedication and Licence (PDDL),²² Creative Commons Attribution 4.0 (CC-BY-4.0)²³ and the ISA metadata license²⁴.

²⁰ European Commission, Guide for the procurement of standards-based ICT — Elements of Good Practice, SWD(2013) 224 final, Brussels, 25.6.2013.

²¹ <http://creativecommons.org/publicdomain/zero/1.0/#sthash.dXmnPsbW.dpuf>.

²² <http://www.opendatacommons.org/odc-public-domain-dedication-and-licence/#sthash.cSWz1Guq.dpuf>.

²³ <http://opendefinition.org/licenses/cc-by/#sthash.Hg8dbSEy.dpuf>.

²⁴ <https://joinup.ec.europa.eu/category/licence/isa-open-metadata-licence-v11>

2.3.9. Set up quality controls

In order to ensure that structural metadata is acceptable for publication and use, it is indispensable to apply quality assurance and quality control. The metadata governance should take into account the following aspects:

- **Accuracy:** Structural metadata should enable instance metadata to describe the resources accurately, e.g. a metadata model needs to include all properties and attributes necessary for the applications that use the instance metadata; a controlled vocabulary needs to include all terms necessary.
- **Trustworthiness:** Structural metadata should be made available from an authoritative and reliable source to enhance its potential for re-use and therefore interoperability. If structural metadata is derived from an external source, such as a respected international standards body, this provenance information needs to be provided so that anybody wanting to re-use it can check the origin of the metadata itself.
- **Integrity:** Structural metadata should be protected against unauthorised alteration.
- **Timeliness:** Structural metadata should be kept up-to-date and promptly available when users want to access it or use it. The frequency with which changes are applied should find the right balance between stability and flexibility. A main challenge is to make sure that the governance procedure put in place allows the processing of requests fast enough for users to actually be able to use the metadata when needed.
- **Completeness:** Structural metadata should be created and maintained in conformance with an agreed standard, respecting common rules for identifiers, names and descriptions. This is an example of something that can relatively easily be checked by tools.
- **Validity:** Structural metadata may have restricted validity, for example in specific time periods or geographical areas. This information needs to be readily available to users.
- **Accessibility:** Structural metadata should be easily accessible, understandable and usable, for consumption both by humans and by machines.

In addition, control processes should be in place in order to validate and guarantee the quality of the metadata. Consistency and completeness of structural metadata may be imposed by the tools for change management or checked before publication through automated checks (e.g. whether the metadata conforms to common standards, or whether newer versions have later dates of modification) and human intervention, e.g. peer review.

3. SPECIFICATIONS FOR METADATA MANAGEMENT

This chapter first identifies existing good practices from existing standards for metadata management. Then, it specifies *generic* lifecycle management processes for structural metadata.

3.1. Existing standards for metadata management

For the management of metadata, and in particular the registration of metadata in registries, several standards exist. A general standard for the registration of metadata items is ISO/IEC 11179. There are also domain-specific standards; an example is ISO 19135 for geographic information.

ISO/IEC 11179²⁵ specifies the kind and quality of metadata necessary to describe data, and it specifies the management and administration of that metadata in a metadata registry (MDR). 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. It does not apply to the physical representation of data as bits and bytes at the machine level. As part of the six-part standard, **ISO/IEC 11179-6:2005**²⁶ specifies the procedure by which Administered Items required in various application areas could be registered and assigned an internationally unique identifier. For each Administered Item to be registered, ISO/IEC 11179-6:2005 defines the type of information that is specified, the conditions that are met, and the procedure that is followed.

ISO 19135:2005²⁷ specifies procedures to be followed in establishing, maintaining and publishing registers of unique, unambiguous and permanent identifiers, and meanings that are assigned to items of geographic information. In order to accomplish this purpose, ISO 19135:2005 specifies elements of information that are necessary to provide identification and meaning to the registered items and to manage the registration of these items.

The Data Management Association's guide to the **Data Management Body of Knowledge (DMBOK)** recommends that changes to controlled vocabularies and their reference data sets be conducted by following a **change request process**:

1. Create and receive a change request;
2. Identify the related stakeholders and understand their interests;
3. Identify and evaluate the impacts of the proposed change;
4. Decide to accept or reject the change, or recommend a decision to management or governance;
5. Review and approve or deny the recommendation;

²⁵ ISO/IEC 11179-1:2004. Information technology -- Metadata registries (MDR).
http://www.iso.org/iso/catalogue_detail.htm?csnumber=35343

²⁶ ISO/IEC 11179-6:2005. Information technology -- Metadata registries (MDR) -- Part 6: Registration.
http://www.iso.org/iso/catalogue_detail.htm?csnumber=35348

²⁷ ISO 19135:2005. Geographic information -- Procedures for item registration.
http://www.iso.org/iso/catalogue_detail.htm?csnumber=32553

6. Communicate the decision to stakeholders prior to making the change;
7. Update the data;
8. Inform stakeholder that the change has been made.

ITIL²⁸ is the abbreviation for the guideline **IT Infrastructure Library**. The main focus of the development was on mutual best practices for all British government datacentres to ensure comparable services. Today ITIL is the **worldwide de-facto-standard** for service management and contains broad and publicly available professional documentation on how to plan, deliver and support IT service features. In the meantime ITIL is already 20 years old and is now at its fourth release of the publications. The core publications are:

- Service Strategy
- Service Design
- Service Transition
- Service Operation
- Continual Service Improvement

These core publications describe 26 processes starting from the strategic orientation of the IT to the continual improvement of Services.

ITIL is a systematic approach to the delivery of quality IT services. It provides a basic vocabulary and a number of processes that are relevant in managing the lifecycle of IT services such as change management, release management, and service validation and testing.

²⁸ <http://www.itil-officialsite.com/>

3.2. Required expertise for metadata management

In the process of metadata management, a number of essential competences can be distinguished. The following areas of expertise should be included in the team that is responsible for the management of the metadata. It is not necessary that every metadata management team consists of at least five members; individual team members may provide one or more of these roles.

- **Domain expertise:** knowledge about the semantics of the data for which the metadata is used and the applications in which the data is used.

This expertise ensures that the team has a good understanding of the functionality that the metadata is supporting. This allows the team to identify potential problems that could be generated by changes in models, schemas and reference data.

- **Information management expertise:** knowledge about theory and practice of information management, e.g. information and library science.

This expertise ensures that approaches to definitions of metadata elements and expression of relationships between metadata elements – e.g. hierarchies in controlled vocabularies – are sound and based on best practices in the domain of information science.

- **Technical expertise:** knowledge about the technical approaches to be used for the technical implementation in the environment in which the metadata is used.

This expertise ensures that the implementation conforms to the technical environment, e.g. using the protocols, schema language and mark-up languages used across the technical and networking infrastructure.

- **Documentation and publication expertise:** knowledge about the documentation rules and publication processes used in the environment in which the metadata is used.

This expertise ensures that the metadata and changes are documented in the format and language that are appropriate for the users of the metadata, and that the metadata is published in the formats (human- and machine-readable) that allow easy integration in applications and services.

- **Standardisation expertise:** knowledge about standardisation rules and procedures if the metadata and/or management approaches are intended to be submitted to standards bodies for national, regional or international standardisation.

This expertise ensures that submission to the appropriate standards body conforms to the format and procedures used in the standardisation process.

3.3. Metadata management process design principles

Having described existing standards and required expertise, we also advise to use a number of design principles. Design principles explain how a certain process or system works and are meant to give guidance in decision making.

3.3.1. Documentation of management processes

To ensure an orderly development of the metadata resources to be managed, it is necessary that the methodology, including practices, processes, principles, roles and responsibilities, is clearly documented and regularly reviewed. An efficient change request process with minimal delivery delays should be part of such a methodology.

The management methodology should also determine the process by which data quality is maintained in the operational environment.

In cases where resources are managed across organisations, it is important that there is agreement on a common management approach to ensure that the different parts remain interoperable without great efforts in transposition or translations.

3.3.2. Tailoring of management processes

The management processes described in section 1.2 provide a generalised view on the steps to be taken in managing structural metadata. In practice, application of the approach in individual cases will require tailoring of the processes to the organisational and technical environment of such cases.

3.3.3. Managing access to the structural metadata

In cases where parts of the structural metadata are confidential, an access policy needs to be defined that governs who can get access to it.

For example, it might be unwanted that external actors get access to data models that are used in military applications, or that enemies can derive information about military capabilities from controlled vocabularies for classifications of weapons systems or for military locations.

In such cases, the authoritative source where the structural metadata is housed, need to be able to assign access credentials and permissions to users.

3.3.4. Update cycles

There are differences in the requirements for the periodicity of changes for data models on one hand and reference data on the other hand. These differences are linked to the different needs for stability versus flexibility.

Data models are strongly linked to the interoperability of applications and therefore changes in a data model have a direct effect on the applications that are based on it. In many cases, software systems will need to be rebuilt importing the new model and upgrading the functionality before they can interoperate with others. In practice, changes in data models will be relatively infrequent (less than annual) and changes

will be accompanied by a strongly managed implementation plan aligned with a software upgrade cycle.

Reference data is usually more loosely linked to the basic functionality of applications. Changing or adding a code in a code list will not have a disruptive effect on the existing functionality. These types of changes may also occur with a higher frequency (one or more times per year) than model changes, and are usually easier to propagate through a network.

3.3.5. Versioning

As part of the lifecycle, the change management process will lead to the creation of a set of versions of the structural metadata.

While the latest version of a data model or reference data collection is clearly the most important resource to be re-used as this supports the functionality at that particular point in time, it is also necessary that previous versions are still available for inspection. This makes it possible to determine what functionality was available in the past. In relation to that, it is also important that the documentation of previous versions as well as change logs are kept available.

Identification of versions of structural metadata can be done by time-stamping the versions, by assigning version numbers or by combining those two approaches.

3.3.6. Development environments

Changes to structural metadata will, as a principle, not be made directly in the production environment. In software development, four environments are usually foreseen:

- **Development:** all changes are developed on this environment.
- **Testing:** after development different types of users will need to test the change on an environment dedicated to them.
- **Acceptance:** this is a separate environment for user acceptance
- **Production:** the live environment

In the management of structural metadata, such a strict separation of environments might not always be necessary. For example, if the change involves adding a new code to a code list, a full acceptance test may not be necessary; if on the other hand, fundamental changes are made to a core data model, it may be necessary to link such a change to the software development cycle, including formal testing and acceptance.

3.4. Process specification

This section contains a specification for the lifecycle management processes of data models and reference data, based on the requirements identified in Table 8 of Annex II. Structural metadata as covered in this document comes in two types:

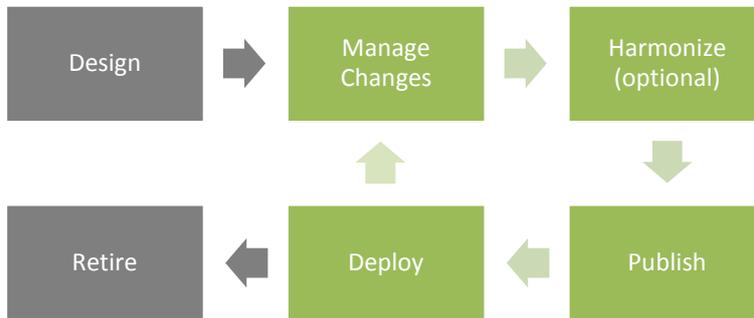
- **Data models:** schemas; and
- **Reference data:** for example controlled vocabularies, name authority lists, code and value lists.

Both these types can either be managed in an XML-based environment using XML Schema Definitions, or in a Linked Data environment using RDF-based formats.

Almost all structural metadata will evolve over time, either because of changes in the environment (e.g. emergence of new subject areas) or because of changes in functionality that must be supported (e.g. new services). The lifecycle of the structural metadata in this section is structured in six main phases:

- **Design structural metadata:** to support a new service or applications, structural metadata needs to be designed, implemented and subsequently used in applications to support interoperability;
- **Manage change of structural metadata:** while requirements and functions of applications evolve, structural metadata needs to change to support changing applications;
- **Harmonise structural metadata:** optionally, and in particular for reference data, cross-references may be defined between a common reference data collection used in the network and other collections of reference data (for example, linking a local list of languages to the ISO standard ISO639), and between local reference collections and the common reference data;
- **Publish structural metadata:** after changes have been applied to structural metadata, the resources and associated documentation need to be released to the stakeholders;
- **Deploy structural metadata:** when a new version of the structural metadata has been released, the changes need to propagate to the operational systems used by the stakeholders;
- **Retire structural metadata:** when applications are no longer supported or migrate to new data models or reference data collections, the structural metadata is no longer relevant and may be decommissioned.

Figure 7: Metadata Management Lifecycle



The following sections describe the high-level administrative processes that are included in the management of these six lifecycle stages of structural metadata. Although there are different levels of metadata governance, the processes described below are generic and are applicable to all.

3.4.1. Design structural metadata

Creation and design of structural metadata entails the processes of agreeing on the syntax and the semantics, and encoding the structural metadata in different formats. This phase is out of scope of this work. Documents like 'process and methodology for developing semantic agreements'²⁹ provide a description of the steps that need to be taken for developing common data models and reference data that can be a basis for information exchanges between systems.

3.4.2. Manage change of structural metadata

3.4.2.1. Manage change requests

Request	Request a change to structural metadata
Goal	To create a change request for a desired modification to the structural metadata (data model or reference data collection).
Preconditions	<ul style="list-style-type: none"> Structural metadata has been designed and published. The structural metadata has been implemented in a production system. An authoritative source is available where stakeholders can access the structural metadata.
Success End Condition	The creation of a change request, which triggers the "Build" phase
Failed End Condition	Decision not to create a change request

²⁹ ISA Programme. Process and methodology for developing semantic agreements, June 2013.

<https://joinup.ec.europa.eu/node/67006>

Primary Actor	<ul style="list-style-type: none"> • Governance Committee – receives feedback and decides on creation of change ticket • Operational Team – performs analysis
Secondary Actors	Stakeholders – submit feedback
Frequency	<ul style="list-style-type: none"> • Ad hoc: when receiving feedback from users and/or when (new) legal obligations arise; or • Periodic: when carrying out periodic reviews of structural metadata to ensure conformance to reused standards.
Trigger	<ul style="list-style-type: none"> • User feedback; • Periodic review; • Legal obligation; • Release of a new version of a reused standard.

Flow:

Step	Description	Actor
1	Receive request	Governance Committee
2	Initial evaluation	Operational Team
3	Accept request	Governance Committee
4	Propose solution with impact analysis and roll-out plan	Operational Team
5	Review proposal	Governance Committee, Stakeholders
6	Accept proposed solution	Governance Committee

Figure 8: Manage Change Requests



3.4.2.2. *Release management of structural metadata*

Build	Create, modify or delete an item in existing structural metadata
Goal	To transpose into structural metadata under the remit of the Governance Committee the accepted change requests, leading to a new release of the structural metadata and the documentation that accompanies it.
Preconditions	Accepted change request
Success End Condition	The structural metadata is successfully updated.
Failed End Condition	The change request is not incorporated.
Primary Actors	<ul style="list-style-type: none"> Operational Team – handles the change request and develops the solution Governance Committee – accepts the solution
Secondary Actors	Stakeholders – are involved in the process to make sure that the solution meets their requirement
Frequency	<ul style="list-style-type: none"> Ad-hoc; or Periodic.
Trigger	When a change request has been accepted and the stakeholders have been informed of an upcoming change, optionally the “Harmonise” phase is executed followed by the “Release” phase.
Comment(s)	<p>Activities in this phase may consider the incorporation of individual changes in structural metadata, or group changes together into pooled releases, depending on the urgency and impact of the changes.</p> <p>Activities and frequency are different for changes to data models and changes to reference data collections.</p>

Flow:

Step	Description	Actor
1	Plan change	Operational Team
2	Apply changes	Operational Team
3	Test solution	Operational Team, Stakeholders
4	Prepare documentation	Operational Team
5	Accept change	Governance Committee

Figure 9: Manage Change Release



The sub-workflow for applying changes to the different types of resources is outlined in the table below. The steps are further described in the text following the table.

Table 4 – Manage changes in structural metadata

Step	XSD ³⁰ model (data model)	XSD code list (reference data)	RDF schema ³¹ (data model)	SKOS ³² vocabulary (reference data)
2.1	Determine whether the element is already defined in an existing XML schema. If so, import if possible.	Determine whether a code is already available in an existing code list. If so, use the same code if possible.	Determine whether the element (class, property) is already defined in an existing RDF namespace. If so, re-use; if not continue.	Determine whether the concept is already available in an existing SKOS concept scheme. If so, re-use; if not continue.
2.2	Identify XSD where element needs to be added, changed or deleted.	Identify XSD where code needs to be added, changed or deleted.	Identify existing namespace for new element.	Identify existing SKOS concept scheme for new concept.
2.3	Create new or modified element	Create new code in context of code list	Mint URI and create definition	Mint URI and create definition
2.4	Add new element, make change to existing element, or delete element.	Add new code, change meaning of existing term, or delete term.	Add element to namespace.	Add new concept in concept scheme.

Step 2.1:

³⁰ W3C. W3C XML Schema Definition Language (XSD) 1.1 Part 1: Structures.

<http://www.w3.org/TR/xmlschema11-1/>

³¹ W3C. RDF Schema 1.1. <http://www.w3.org/TR/rdf-schema/>

³² W3C. SKOS Simple Knowledge Organization System Reference. <http://www.w3.org/TR/skos-reference/>

In the development of data models and reference data, standard schemas and vocabularies should be re-used as much as possible; when local schemas and vocabularies are used, map those to standard elements as much as possible.

Step 2.2:

In all cases, it needs to be determined in which file the element needs to be added, changed or deleted. If the metadata is part of an XML Schema Definition, it is the XSD to be amended; if the metadata is managed as an RDF schema, it is the RDF namespace; if it is a controlled vocabulary expressed using SKOS, it is the SKOS concept scheme.

Step 2.3:

For an element in an XSD model, the element needs to be defined with its element name and structural definition. For a code to be included in an XSD-based code list, the name, attributes and location in a hierarchy need to be defined.

For elements (class, property) in an RDF schema and for a concept in a SKOS concept scheme, a URI needs to be minted in the context of the RDF namespace or SKOS concept scheme, together with an unambiguous definition of the element or concept.

Step 2.4:

In this step the element that was prepared in the previous step is incorporated in or deleted from the existing schema.

In XSD, a new element or code is added; a change in an existing element or code overwrites the old version; a deletion is simply removed from the schema definition.

In RDF namespaces and SKOS concept schemes, a change in semantics can only be applied if this does not affect existing applications. In general, semantic meaning may be broadened (as existing data remains valid) but never narrowed (which could make existing data invalid).

Deletion of elements or vocabulary terms should be avoided unless it can be verified with complete certainty that such items are not used anywhere; otherwise, items should be annotated, e.g. with an status property (e.g. adms:status) with value "Deprecated".

3.4.3. Harmonise structural metadata

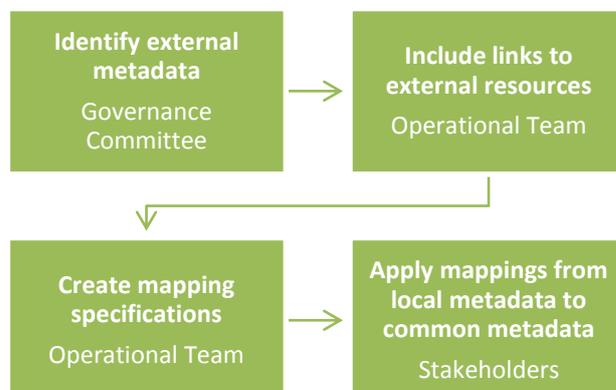
Harmonise	Create links to internationally standardised or widely used structural metadata and mapping specifications for local structural metadata
------------------	---

Goal	To establish equivalence links between structural metadata under remit of the Governance Committee and other structural metadata, either harmonising common structural metadata with internationally standardised or otherwise widely used structural metadata, or providing information to enable mapping from local structural metadata to common structural metadata.
Preconditions	Change implemented in structural metadata,
Success End Condition	The structural metadata is successfully harmonised.
Failed End Condition	Equivalence links and mapping specifications are not available.
Primary Actors	<ul style="list-style-type: none"> • Governance Committee – decide which external metadata collections to use for linking • Operational Team – creates the links to external metadata collections and prepares mapping specifications
Secondary Actors	<ul style="list-style-type: none"> • Stakeholders – receive mapping specifications • Owners of external structural metadata – may be contacted to create appropriate links
Frequency	Ad-hoc
Comment(s)	It needs to be decided on the level of the Governance Committee to which external metadata collections links are established

Flow:

Step	Description	Actor
1	Identify external structural metadata to be linked to	Governance Committee
2	Include links to external resources, if necessary contact owners of external resources	Operational Team
3	Create mapping specifications	Operational Team
4	Apply mappings from local metadata to common metadata	Stakeholders

Figure 10: Harmonise Structural Metadata



If appropriate, RDF classes or properties should be linked to other items in the namespace (e.g. to express sub-class or sub-property relationships) or to items in other namespaces (e.g. to indicate equivalent classes or properties); SKOS concepts can be linked to other concepts in the concept scheme (e.g. to link the concept to broader or narrower terms) or to concept in other concept schemes (e.g. similar concepts).

3.4.4. Publish structural metadata

Release	Document and publish the changed structural metadata
Goal	To document a new version of the structural metadata and to publish it on the authoritative source.
Preconditions	N/A
Success End Condition	The release of the structural metadata is published and documented on the authoritative source, including, the public documentation
Primary Actors	<ul style="list-style-type: none"> • Operational Team – prepares and issues the release of the structural metadata and documentation • Governance Committee – oversees the release and the announcement
Secondary Actors	Stakeholders – are informed of the release
Frequency	<ul style="list-style-type: none"> • Ad-hoc; or • Periodic.
Trigger	Completion of build of an updated (and optionally harmonised) version of the structural metadata.
Comment(s)	<p>Metadata should be documented in human- and machine-readable formats. The machine-readable documentation metadata should follow a standard vocabulary, such as ADMS. In addition to the machine-readable data, it is helpful to provide guidance documentation, for example outlining which standards and methods have to be used in specific cases.</p> <p>Models and model elements, as well as the items in controlled vocabularies should be assigned URIs and those should be maintained persistently. Descriptions of the metadata should follow unambiguous guidelines, in order to facilitate search and retrieval.</p> <p>Wherever possible, metadata should be made available under an open licence on an open platform such as Joinup. However, if some parts of documentation are sensitive, those should be protected by appropriate access control.</p> <p>All the resources managed should be published in such a way that they can be re-used easily by other systems, for example as plugins, via web-services, via API, or using a dedicated client. It is important to make sharing of and accessing of the shared model and reference data easy because sharing is the basis for interoperability.</p>

Flow:

Step	Description	Actor
1	Issue release documentation	Operational Team
2	Move release to production environment	Operational Team
3	Notify stakeholders of release and roll-out plan	Governance Committee

Figure 11: Publish Structural Metadata



3.4.5. Deploy structural metadata

Deploy	Roll out changed structural metadata to distributed systems used by stakeholders
Goal	To effectively implement the changes in structural metadata in the operational systems used by stakeholders while protecting the live environment of their systems through planning, testing, building and implementing a grouped set of changes.
Preconditions	<ul style="list-style-type: none"> • New version of structural metadata published on authoritative source • Roll out plan established
Success End Condition	Changes successfully implemented in all systems that use the structural metadata
Primary Actors	<ul style="list-style-type: none"> • Operational Team – provide project management and support for the roll-out. • Stakeholders – execute the roll-out in their systems
Secondary Actors	Governance Committee – oversees the roll-out
Frequency	<ul style="list-style-type: none"> • Ad-hoc; or • Periodic.
Trigger	Release of an update version of the structural metadata
Comment(s)	For releases with low impact (e.g. regular releases of reference data collections) roll-out might be done using a fixed script, while for release with higher impact (e.g. restructuring in a data model) a detailed implementation plan needs to be developed and agreed with stakeholders.

Flow:

Step	Description	Actor
1	Monitor roll-out	Operational Team

2	Apply changes in local systems	Stakeholders
3	Report progress to Governance Committee and Stakeholders	Operational Team

Figure 12: Deploy Structural Metadata



For **changes in data models**, two situations can occur:

- Changes are not backward compatible. This situation arises when there are fundamental rearrangements in the data model or changes in existing elements.
- Changes are backward compatible. This situation arises when minor changes to the data model are made, such as addition of new elements that do not affect existing model elements.

In case changes are not backward compatible and cannot work with the software that used to previous version of the model or schema, the deployment of these changes need to be accompanied by a software upgrade process. Especially in cases were multiple software vendors are involved, such upgrades need to be carefully planned and executed with ample time for testing and verification. To avoid disruption of the operational system, testing and verification should be conducted in a separate test and acceptance environment.

For changes that are backward compatible, the process does not rely on all systems in the operational environment installing the changes at the same time. Existing systems can continue to operate unchanged, but before they upgrade they will not be able to access functionality that is provided by the new model elements. This means that in the environment of interconnected systems the availability of the new functionality will become available gradually over a certain period of time. To maintain interoperability, two conditions need to be met:

- Systems that still operate with the old version of the model need to be able to ignore the additional model elements in the new version of the model; and
- Systems that have already upgraded to the new version of the model need to be able to process data using both versions of the model.

Even in the case of backward compatibility, it is recommended to organise the upgrade across the network as a well-planned and well-communicated project so that all communication partners are aware of the status of the propagation of the new functionality across the network at all times during the transition period.

The way that **changes in reference data** affect interoperability, and therefore the way those changes propagate, depends on the technical implementation.

If the reference data is implemented as an enumerated list of string values in an XML schema, changes in reference data are in fact changes in the metadata model and schema and therefore the approaches described in the previous section apply.

Otherwise, if reference data is implemented in a Linked Data approach, for example as a SKOS Concept Scheme, every item in the collection is identified by a URI. If implemented this way, changes in reference data are generally non-disruptive. Using the example of a collection of references to organisations that participate in a network, the following changes may occur:

- Addition of an organisation to the network: the addition of a new item in the reference data can be done without disruption as long as systems ignore items that they do not recognise. The new item will be identified by a new URI that enables all systems in the network to access the new item and its characteristics whenever they need it.
- Deletion of an organisation from the network: in general it is not a good idea to delete items that are no longer needed. As long as a certain item is used as a reference in instance data, physically deleting the item from the reference collection would make that instance data invalid. As discussed in the previous section, a better approach is to give the item that is no longer needed a status of 'deprecated' or 'withdrawn' so that further use is discouraged.
- Amendment of the information of a particular organisation, such as names, addresses etc.: if the reference data is implemented in SKOS, such changes do not affect the interoperability as these characteristics are properties of the organisation that continues to be persistently identified by its URI. However, if some of the characteristics that have changed are being used, for example for indexing, systems that refer to the item may need to re-ingest the data for the item to be able to update the indexes.

3.4.6. Retire structural metadata

Retire	Delete or deprecate structural metadata
Goal	To mark structural metadata as no longer relevant for applications at the level of the data model or a collection of reference data.
Preconditions	Structural metadata is no longer relevant
Success End Condition	Structural metadata is marked as deprecated.
Failed End Condition	N/A
Primary Actors	<ul style="list-style-type: none"> • Governance Committee – decides on retirement of structural metadata • Operational Team – takes actions to delete or deprecate
Frequency	Ad-hoc

Trigger	Observation that data model or reference data collection is no longer in use.
Comment(s)	Before retiring structural metadata, a complete impact analysis should be done to verify that indeed the metadata is no longer used in production environments. In general, it is recommended not to physically delete structural metadata but to mark it as deprecated.

Flow:

Step	Description	Actor
1	Assess the impact of deprecation	Operational team
2	Review for approval	Governance Committee
3	Approach all consumers of the data	Operational team
4	Clearly mark reference data as deprecated	Operational team
5	Ensure backwards compatibility	Operational team

Figure 13: Retire Structural Metadata



4. SPECIFICATIONS FOR METADATA TOOLS

This chapter assesses the requirement coverage of a set of existing tools for metadata management, identified during the pilots with DG COMP, DG MARE, and the Publications Office of the EU. Of course there are many other tools, some of which are listed in Annex I. Public administrations should evaluate and select appropriate tools depending on their own contextual and evaluation criteria. When we created this list, we have taken into account preferences such as the following:

- **Tools that are already used by public administrations:** tools that are already used by public administrations have a proven value and can be more beneficial because they are standards within an existing ecosystem. In this section, tools already used by DG COMP, DG MARE, and the Publications Office are analysed;
- **Tools that implement standards:** tools that are based on standards are more likely to reduce ICT vendor lock-in [3]³³; and
- **Open-source tools:** although there is no policy that mandates the use of open-source software tools; it is often recommended because it can allow contributions by the public sector to be used by others. The European Interoperability Framework (EIF) for example, recommends openness in developing software systems *allowing European public administrations generate results that can be interconnected, reused and shared, which also improves efficiency* [4]³⁴.

The tools mentioned in this chapter should support stakeholder requests and needs, but also existing standards for metadata management. Before looking at standards, we should note that in the context of this report the following categories of tools should be considered:

- **Tools for data modelling:** to support the design and change of data models
- **Tools for editing reference data:** to support the design and change of reference data;
- **Tools for managing data models and reference data changes:** managing changes and releases of reference data including the use of an authoritative source;
- **Tools for meta data deployment:** tools for implementing data models and reference data in information systems; and
- **Tools for meta data publication.**

³³ European Commission, Staff Working Document. (2013). *Guide for the procurement of standards-based ICT — Elements of Good Practice. Against lock-in: building open ICT systems by making better use of standards in public procurement*
<https://ec.europa.eu/digital-agenda/node/67038>

³⁴ European Commission, ISA Programme (2010). European Interoperability Framework.
http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf

4.1. Existing standards for metadata representation

This section lists a number of metadata standards that should be supported by metadata tools:

- Standard exchange formats for reference data;
- Standard exchange formats for data models;
- Standards for documenting metadata specifications.

4.1.1. Unified Modelling Language

The Unified Modelling Language (UML) is a standard by the Object Management Group (OMG) that can be used for data modelling. UML allows capturing the fundamental characteristics of the classes, properties and relations. Its primary purpose is to enable *humans* to understand the meaning of the data model. It is not used as a physical data model for information exchange per se.

UML has the following characteristics:

- **Graphical representation:** UML has become a de-facto standard for the graphical representation of a data model in the form of a class diagram.
- **XML Exchange format:** UML model can be serialised and exchanged with other tools using the XML Metadata Interchange (XMI) – even though XMI conformance and interoperability is a known weak point of UML tools³⁵.
- **Local data elements:** In the UML language attributes and associations are local data elements that are encapsulated within the classes in which they are defined. This encapsulation prevents attributes and associations from being reused independently from the classes in which they are defined. Unlike properties in RDF Schema, UML attributes and relationships are not global data elements.
- **UML profiles:** The use of UML profiles allows extending the UML language with a number of method-specific stereotypes, tagged values, and constraints. UML profiles are useful to adhere to a specific design patterns, and use model-driven development aids for the generation of XML and RDF Schemas.

4.1.2. XML Metadata Interchange

The XML Metadata Interchange, or XMI, is a standard for the exchange of metadata information, using XML. The standard is managed by the Object Management Group (OMG). In principle, XMI splits models in two parts:

- Abstract models, the representation of the semantic information; and
- Concrete models, the representation of the visual diagrams.

Most commonly, XMI is used for the exchange of UML models. As described in chapter 4.2.1, XMI is implemented in many UML modelling tools as a standard for exporting or importing UML models, and thus exchanging data between those tools. However,

³⁵ See also the work of the OMG Model Interchange Working Group (MIWG)
<http://www.omgwiki.org/model-interchange/doku.php>

currently the implementation of XMI in those tools is often incompatible [5], which makes it difficult in practice to exchange metadata between different tools.

4.1.3. XML Schema languages

An XML schema is a description of a type of XML document, typically expressed in terms of constraints on the structure and content of documents of that type, above and beyond the basic syntax constraints imposed by XML itself. There are several different languages available for specifying an XML schema such as XSD, WXS and Schematron. Each language has its strengths and weaknesses.

Schema-validity assessment has three aspects:

1. Determining local schema-validity, that is whether an element or attribute information item satisfies the constraints embodied in the relevant components of an XSD schema;
2. Determining an overall validation outcome for the item by combining local schema-validity with the results of schema-validity assessments of its descendants, if any; and
3. Determining the appropriate augmentations to the info set (and, if desired, exposing them to downstream applications in some way, to record this outcome).

As mentioned in chapter 3.4.2 structural metadata can be managed in an XML-based environment using XML Schema Definitions, or in a Linked Data environment using RDF-based formats

4.1.4. RDF Schema

In the Resource Description Framework (RDF) data is organised in graphs around subject-predicate-object statements (called triples). RDF has come to be used as a general method for conceptual description or modelling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

RDF has among others the following characteristics:

- Flexible data integration
- Global data elements
- Uniform Resource Identifiers
- Multiple formats

4.1.5. Simple Knowledge Organisation System (SKOS)

SKOS³⁶, the Simple Knowledge Organisation System, is a common data model for sharing controlled vocabularies such as code lists, thesauri, and taxonomies via the Web in a machine-readable format. SKOS is a W3C Recommendation and commonly used in open-source and proprietary tools. In the Core Vocabularies³⁷ specifications of the ISA Programme, SKOS is the recommended vocabulary for the representation of code lists. The Publications Office already uses SKOS as the official format of EuroVoc, the EU's multilingual thesaurus, and the Named Authority Lists.

SKOS provides a standard way to represent knowledge organization systems using the Resource Description Framework³⁸ (RDF). Encoding this information in RDF allows it to be passed between computer applications in an interoperable way.

Using RDF also allows knowledge organization systems to be used in distributed, decentralised metadata applications. Decentralised metadata is becoming a typical scenario, where service providers want to add value to metadata harvested from multiple sources.

SKOS represents the terms in a controlled vocabulary as instances of the class `skos:Concepts`. SKOS also defines properties for multi-lingual labels (`skos:prefLabel`), associated codes (`skos:notation`), and definitions (`skos:definition`). The publication of controlled vocabularies represented in SKOS on the Web brings the following advantages:

1. **De-referencing:** the principles of [Linked Data](#) requires each term in the controlled vocabulary to be identified by a corresponding term URI based on the HTTP protocol. The term "Taxonomy" in the "Asset Type" scheme has for example the following term URI: <http://purl.org/adms/assettype/Taxonomy>. This means that when someone else encounters such an URI, he can look up its meaning by entering the URI in the address bar in his browser. This is called *de-referencing*. This is a simple yet powerful feature of the Web.
2. **Machine-readability:** In the example of "Taxonomy", the user can use the term URI to retrieve both a machine-readable and human-readable file containing definitions, labels, and related concepts for this term expressed in [SKOS](#). Well-known thesauri such as [EuroVoc](#) have been defined using an ontology that extends SKOS.
3. **Multilingualism:** SKOS allows to associate labels and definitions in multiple languages to any concept. This means that we can associate the labels "taxonomie"@FR, "Taxonomie"@DE, or "taxonomia"@PT to the concept identified with URI <http://purl.org/net/mediatypes/application/OWL+XML> to include the French, German, and Portuguese labels.

³⁶ <http://www.w3.org/2004/02/skos/vocabs>

³⁷ https://joinup.ec.europa.eu/system/files/project/Core_Vocabularies-Business_Location_Person-Specification-v1.00_0.pdf

³⁸ <http://www.w3.org/RDF/>

4. **Metadata alignment:** SKOS provides mapping properties like `skos:closeMatch`, `skos:exactMatch`, `skos:broadMatch`, `skos:narrowMatch` and `skos:relatedMatch`. These properties are used to state mapping alignment links between SKOS concepts in different concept schemes, where the links are inherent in the meaning of the linked concepts.
- a. The properties `skos:broadMatch` and `skos:narrowMatch` are used to state a hierarchical mapping link between two concepts.
 - b. The property `skos:relatedMatch` is used to state an **associative mapping link between two concepts**.
 - c. The property `skos:closeMatch` is used to link two concepts that are sufficiently similar that they can be used interchangeably in some information retrieval applications. In order to avoid possibilities of "compound errors" when combining mappings across more than two concept schemes, **`skos:closeMatch` is not declared to be a transitive property**.
 - d. The property `skos:exactMatch` is used to link two concepts, indicating a high degree of confidence that the concepts can be used interchangeably across a wide range of information retrieval applications. **`skos:exactMatch` is a transitive property**, and is a sub-property of `skos:closeMatch`.

SKOS is an extensible vocabulary. One popular extension is SKOS-XL, which extends SKOS with labels (SKOS eXtension for Labels).

4.1.6. GeneriCode

The OASIS Code List Representation format, GeneriCode³⁹, is a single model and XML format (with a W3C XML Schema) that can encode a broad range of code list information. The XML format is designed to support interchange or distribution of machine-readable code list information between systems.

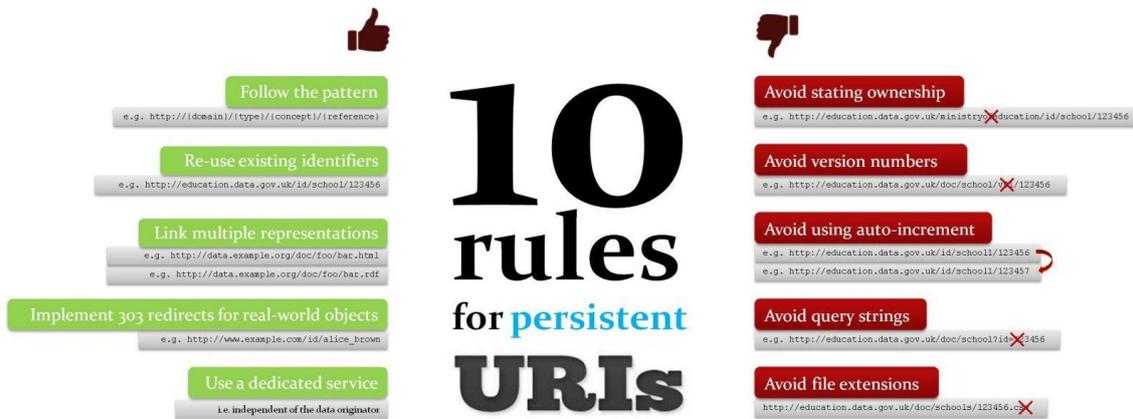
4.1.7. HTTP URIs

In order to facilitate its sharing and reuse across systems and organisation, structural metadata needs to have persistent unique identifiers. As we are experiencing the era of the Web of Data, it is recommended that such identifiers come in the form of HTTP URIs. The ISA Programme as well as W3C have created good practices and guidelines for the design and management of well-formed, persistent URIs [6], e.g. see ISA's 10 Rules for Persistent URIs⁴⁰ as represented in Figure 14. Moreover, the ISA Programme has set up a persistent URI Task Force, which works on a persistent URI policy for EU institutions [7].

³⁹ <http://docs.oasis-open.org/codelist/ns/genericcode/1.0/>

⁴⁰ <https://joinup.ec.europa.eu/community/semic/document/10-rules-persistent-uris/>

Figure 14 - 10 rules for persistent URIs



4.1.8. Asset Description Metadata Schema (ADMS)

The Asset Description Metadata Schema (ADMS) is a common vocabulary for descriptive metadata, used to describe interoperability solutions [8] making it possible for ICT developers to explore and search for interoperability assets. ADMS⁴¹ was developed by the ISA programme in 2011 and 2012. ADMS was subsequently published by the World Wide Web Consortium (W3C) as a W3C Note⁴² in 2013.

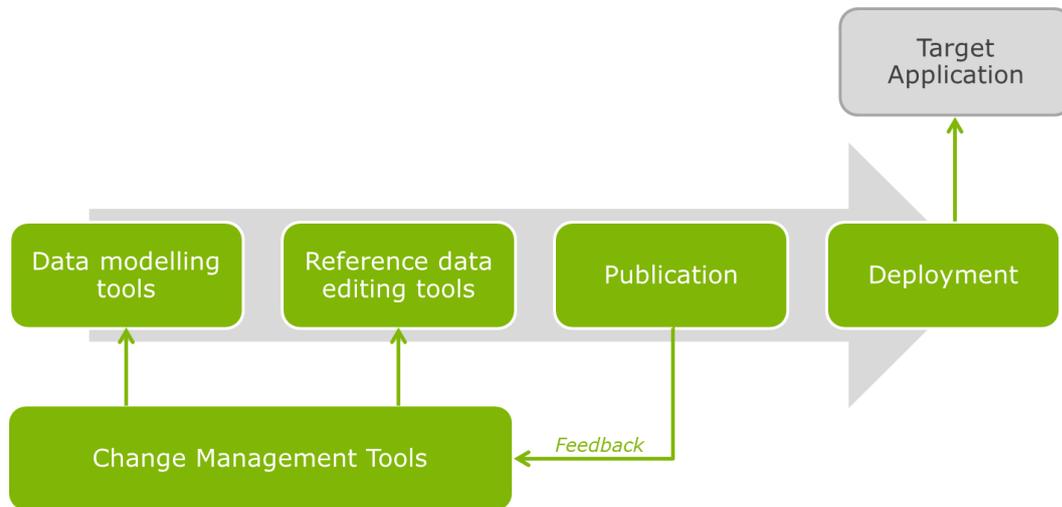
4.2. Existing metadata services and tools

This section lists a number of metadata tools that could support the management of data and metadata. The sections below contain an overview of tools that are most commonly used. The tools are categorised following the generic tool chain, as represented in Figure 15.

⁴¹ European Commission. Joinup. The Asset Description Metadata Schema (ADMS). <https://joinup.ec.europa.eu/asset/adms/home>

⁴² W3C. Asset Description Metadata Schema (ADMS). W3C Working Group Note 01 August 2013. <http://www.w3.org/TR/vocab-adms/>

Figure 15 - Tool Categories: generic tool chain



4.2.1. Data model editors

4.2.1.1. Tool: Sparx Enterprise Architect

Sparx Enterprise Architect is a commercial-licensed tool for data modelling and visualisation. The built-in visualization function of Enterprise Architect can be used for representing data models in a human readable format. Enterprise Architects allows modifying the script that is used for exporting data models to HTML. By slightly adapting the script, we can link the human-readable representation with the repository of machine-readable distributions of the metadata and the code lists used, and even with the issue tracker on JIRA.

4.2.1.2. Alternatives

Sparx Enterprise Architect is already successfully used in different Institutions of the European Commission. However, for UML modelling, multiple open source and commercial alternatives are available, for instance:

- **ArgoUML:** ArgoUML is an open source application for designing UML models, running on Java and providing support for all UML 1.4 diagrams, licensed under the Eclipse Public License (EPL). Through the XMI standard ArgoUML supports the exchange with other UML modelling tools.
- **Umbrello:** Umbrello is an open source diagram tool, licensed under GNU Public License (GPL). Amongst other features, the tool supports XMI import and export.
- **Open ModelSphere:** OpenModelSphere is a data, process and UML modelling tool. The software is available as open source under the GPL.
- **UMLet:** UMLet is an open source tool for quickly creating UML models. The software has been licensed under GPL. Since no underlying dictionary or directory for reusable design objects is used, UMLet is more of a drawing tool rather than a modelling tool.

- **PowerDesigner:** SAP Sybase PowerDesigner is a commercial modelling for process, data and metadata management, making it. PowerDesigner has features for collaborative modelling.
- **MagicDraw:** MagicDraw is a commercial UML, BPMN, SysML and UPDM modelling tool, allowing for collaborative design of models.

4.2.2. Reference data editors

4.2.2.1.Tool: VocBench

VocBench⁴³ is a web-based editing and workflow tool for managing thesauri, authority lists and glossaries based on SKOS and RDF. The tool was developed by the Food and Agricultural Organisation (FAO) of the United Nations. VocBench supports collaborative editing, multilingual terminologies and administration functions that allow assigning roles for maintenance, validation and quality assurance. The Publications Office of the European Commission uses VocBench to manage its EuroVoc thesaurus.

4.2.2.2.Tool: PoolParty: Thesaurus Management

PoolParty Thesaurus Server⁴⁴ is a software tool for creating and maintaining taxonomies, thesauri, ontologies and knowledge graphs. The tool manages metadata based on standards like RDF and SKOS. Designing code lists can be done via the graphical interface or by importing existing lists in formats like XML, Excel, etc. Moreover, the tool carries out automatic quality checks based on SKOS.

For system integration purposes, PoolParty provides an API which is based on the SPARQL standard, an RDF database query language.

4.2.2.3. Tool: Reference Data Component (RDC) - Editor

The generic Reference Data Component (RDC), which will evolve into the Reference Data Deployment Adaptor (REDDA), is developed by DG COMP with the intention to be also used by other EU Institutions.

RDC has the main purpose of automatically deploying reference data into information systems, offers a basic reference data editor feature. For the REDDA editor to be compliant to the proposed management and governance methodology, further development of the tool is however needed.

4.2.3. Publication tools

4.2.3.1.Service: Joinup

Joinup is an online platform which was developed by the ISA programme of the European Commission for documenting and disseminating semantic assets such as ontologies, data models, code lists, XML schemas, reference data, etc. Publishing

⁴³ <http://aims.fao.org/tools/vocbench-2>; <http://vocbench.uniroma2.it/>

⁴⁴ <http://www.poolparty.biz/portfolio-item/poolparty-thesaurus-server/>

reference data on Joinup would allow users to easily find the data, download it and provide feedback.

4.2.3.1.Tool: Metadata Registry of the Publications Office (MDR)

The Metadata Registry (MDR) of the Publications Office⁴⁵ of the EU is the authoritative source for definition data – metadata elements, named authority lists, schemas, etc. – and authority data used for exchanging data between institutions involved in the legal decision making process. Many of the definition data sets contained in the MDR are governed by the Inter-Institutional Metadata Maintenance Committee (IMMC).

The Publications Office uses a tool chain and some scripts to edit the Named Authority Lists. For each NAL, the Publications Office publishes a set of distribution which can be downloaded from the MDR website. These sets are composed of a SKOS, XML, XSD and HTML version.

A publication package is also available as a zip file. It contains the distribution of changed NALs (XML, SKOS, ATTO-XML), a comparison file allowing to identify differences between the previous and the current version, and the release notes listing the changes to the NALs included in the publication.

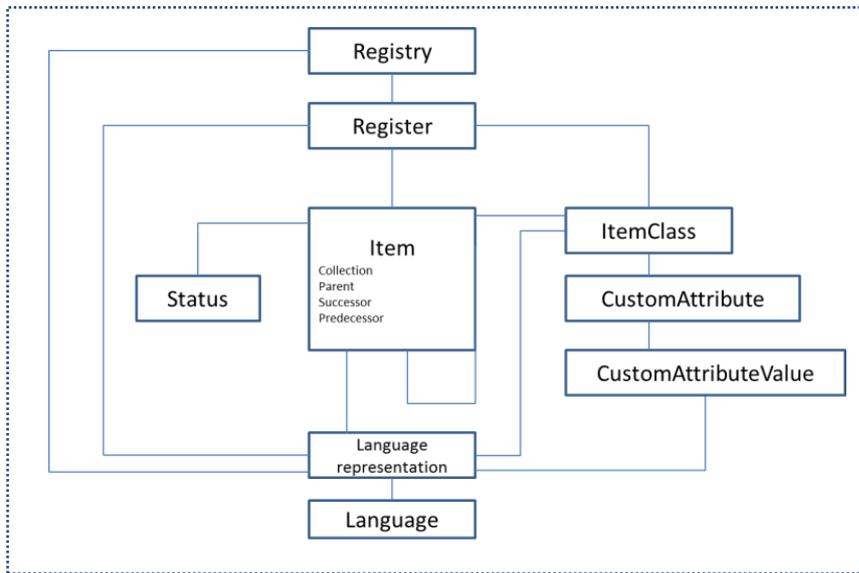
4.2.3.1.Tool: The Re3Gistry

The Re3Gistry is an online platform for managing metadata registers, which was initially developed as a solution for publishing INSPIRE⁴⁶ code lists. The first release of the Re3Gistry supports multilingual representations, browsing and downloading functionalities via its online interface. The registry's content can be made available for download in several formats, including HTML, XML, Atom, JSON and RDF/SKOS [9]. Moreover, the tool allows the owner of structural metadata to assign statuses (Valid, Invalid, Submitted, Superseded, Retired) following the ISO 19135 standard [10]. Figure 16 represents the information model of the Re3Gistry.

⁴⁵ <http://publications.europa.eu/mdr/>

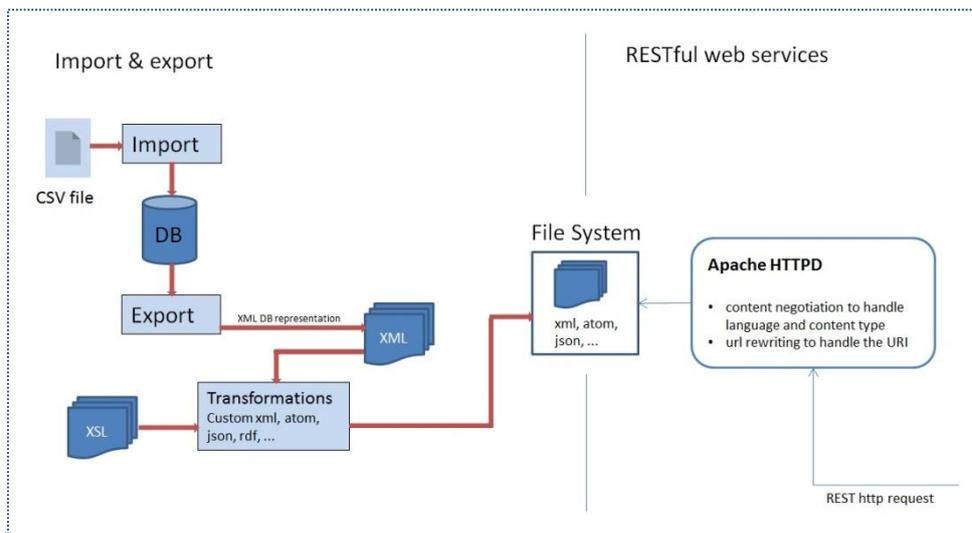
⁴⁶ Infrastructure for Spatial Information in the European Community: <http://inspire.ec.europa.eu/>

Figure 16: Re3Gistry Information Model (Joint Research Centre, 2014)



The software consists of two main parts: a data handling system and a web interface. Figure 17 represents the schematic details of the system:

Figure 17: Re3Gistry Schematic Details (Joint Research Centre, 2014)



Future releases could support metadata mapping, SPARQL querying and integration with collaborative development tools for maintaining registers and registered items. The Re3Gistry is made available as open source software under the EUPL licence via Joinup: <https://joinup.ec.europa.eu/software/re3gistry/>.

4.2.3.1.Tool: Apache Subversion

On the server side, Apache Subversion (SVN) is used as the versioning tool in Joinup and JIRA. In order to support collaborative development, Subversion releases above 1.2 can be directly integrated in Enterprise Architect, allowing multiple developers to work on specific parts of the model, A client side SVN tool is however required for implementing Enterprise Architect with SVN. Many executable files for Subversion, such as Visual SVN's Apache Subversion command line tools⁴⁷, are listed on the Subversion website⁴⁸. After creating an initial copy of the online SVN repository, completing the installation is limited to linking the executable Subversion file to the Enterprise Architect document.

4.2.3.2.Tool: GIT

GIT is a version control system that offers a more extensive set of features than Apache SVN. Although both solutions aim at supporting the versioning process, there are some key differences in their approach and features. The first distinct feature of GIT is that the document repositories are distributed, meaning that every developer has a full clone of the entire repository. As a consequence, GIT offers more flexibility and a better integration with several development work flows⁴⁹ compared to non-distributed solutions like SVN. A second unique feature of the GIT is its branching model. The tool allows users to create many local branches of the versioned assets, which offers endless possibilities to modify and test ideas in a separate branch without interfering with another branch. Moreover, branches can easily be merged, deleted or restored to previous versions.

4.2.4. Deployment tools

4.2.4.1.Tool: Reference Data Component (RDC)

In the context of the Generic Interoperable Notification Services (GENIS) project, funded by the ISA programme, a GENIS Reference Data Component (GENIS RDC) was built. The GENIS RDC provides two main features:

- A basic editor for creating, updating and managing reference data; and
- A reference data deployment adaptor that propagates the reference data into operational systems.

The tool allows to

- Import reference data from a file;
- Create, read, update, delete reference data using the Web-based graphical user interface;
- Export reference data to an XML file; and
- Deploy reference data directly into operational systems (DG COMP).

⁴⁷ <http://www.visualsvn.com/downloads/>

⁴⁸ <http://subversion.apache.org/packages.html>

⁴⁹ Please refer to <http://git-scm.com/about/distributed> for a more information on work flow support.

In 2014, the ISA Programme funded the further development of the GENIS RDC. The reference data editing features were further extended, but the development efforts were mainly focused on the deployment feature, such as supporting international metadata standard formats as SKOS or GeneriCode. Therefore, the GENIS RDC was renamed to REDDA: the Reference Metadata Deployment Adaptor.

4.2.4.2. Alternative: Manual or semi-automated deployment

When the update frequency of reference data is low and when the number of systems using the reference data is limited, there might be little added value in automatically deploying reference data. In such cases, one might opt for a manual or semi-manual – e.g. scripting based – deployment of reference data into the relevant IT systems. The metadata governance and management specifications as described in chapters 2 and 3 do not require the deployment to be automated.

4.2.5. Change management

4.2.5.1. Tool: JIRA and confluence

JIRA is an online ticket tracking system that supports organising and following up on issues, assigning work packages and monitor team activity. JIRA can be used for following up on change requests and to support the development and maintenance of reference data.

Confluence can be used as authoritative source. Confluence is already widely used within the European Institutions as an online, wiki-based collaboration platform. Confluence can be easily integrated with other open source Atlassian products, such as JIRA and supports fine grained access control.

4.3. Proposed metadata tools: gap analysis

This section compares the requirements for metadata tools, as elicited in Figure 18 of Annex II, with existing tools, already being used within the European Commission.

Figure 18 - Proposed Tool Chain

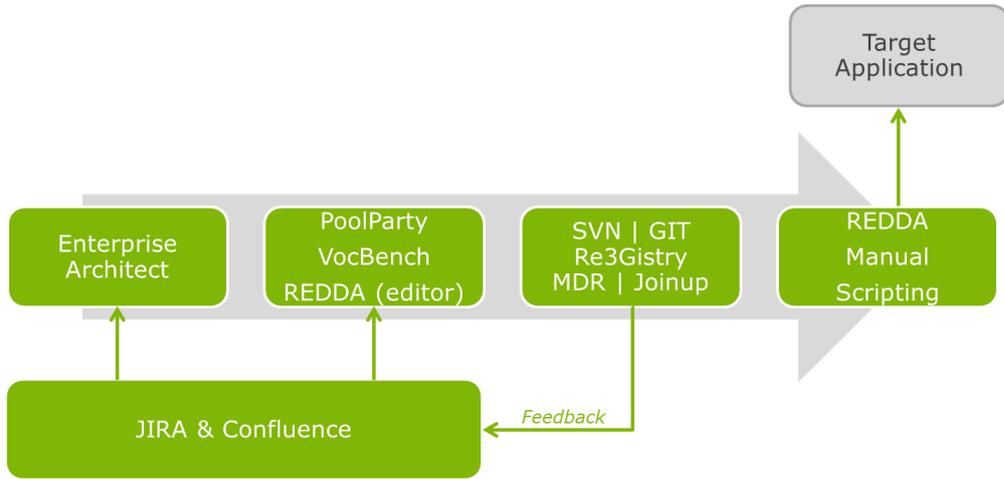


Table 5 - Mapping of requirements to existing tools

Requirement	Enterprise Architect	Apache SVN	JIRA	Joinup	Confluence	GENIS RDC	VocBench	Poolparty	MDR (PO)
T1 data modelling									
Visualisation	X						X		
Navigability	X								
Class & property search	X								
Multilingualism			X	X	X				
IT and business guidelines									
Facilitate feedback			X	X	X				
T2 Access									
Machine-readable	X						X		
Integration with other tools		X	X						
Exchange formats	X					X	X		
Reuse of existing data models	X								
HTTP URIs				X			X		
T3									
Public repository		X		X					
Private authoritative source		X			X				
Versioning integration	X	X	X	X		X	X		
Ticketing system			X						
User Notification				X					
Access management		X	X	X	X	X	X		
T4 Edit ref. data									

**Methodology and tools for Metadata Governance and Management
for EU Institutions and Member States**

Requirement	Enterprise Architect	Apache SVN	JIRA	Joinup	Confluence	GENIS RDC	VocBench	Poolparty	MDR (PO)
Import reference data from external source				x		x	x		
CRUD ConceptScheme						x	x	x	x
multilingualism						x	x		
Order of concepts						x	x		
Versioning						x	x	x	x
Export						x	x	x	x
T5 Changes To ref data.									
Log changes			x						
Keeping track of impact analysis			x						
Log decisions			x						
Create release notice			x						
Linking change requests to release notes			x						
Linking change requests to versions			x						
T6 Deploy ref. data									
Deploy as a service						x			
Deliver services while disconnected						x			
Provision all versions						x			
T7 Publication of ref. data									
File-based read-access over HTTP				x					
Write-access over WebDAV or Subversion.				x					

Conclusion on metadata tools

The analysed tools cover the identified requirements for metadata management. However, no single tool can cater for *all* requirements. It is important that tools are interoperable; the proposed standards for metadata representation in Section 4.1 may contribute to achieving this.

Therefore, we propose the following next steps:

- Identify lessons learned from the use of these tools by public administrations;
- Explore the integration possibilities of all these tools based on open standards;
- Look into the further development of existing tools such as GENIS RDC;
- Define a pilot for the integrated use of these tools preferably on inter-institutional level;
- In terms of commercial products it should be investigated how re-usable they are and what the consequences of the licensing models are. In some cases open source licenses are offered.

5. CONCLUSION AND NEXT STEPS

This report provides generic guidelines for putting in place metadata governance, management and tools both at the EU and Member States levels. It takes stock of best practices and lessons learned from already functioning metadata governance and management initiatives.

Metadata governance

The generic specifications for metadata governance in this document are based on existing standards. They should be tailored based on the specific situation of a public administration. This comprises:

- Determining the right scope;
- Setting up a governance structure;
- Defining a decision mechanism;
- Tailoring a management process to handle requests;
- Defining communication channels;
- Setting up a registry as an authoritative source;
- Establishing an enforcement approach;
- Establishing a licence framework; and
- Setting up quality controls.

Metadata management

The generic principles and lifecycle management processes for structural metadata are based on best practices from the Publications Office of the EU, DM-BOK, and ITIL. This comprises agreeing on processes to:

- Design structural metadata;
- Manage change of structural metadata
- Harmonise structural metadata;
- Publish structural metadata;
- Deploy structural metadata; and
- Retire structural metadata.

Metadata tools

An evaluation of existing tools has shown that the lifecycle management processes can be supported and that requirements can be matched by existing tools. Metadata tools in the following categories should be chosen:

- Data model editor;
- Reference data editor;
- Tools for change requests;
- Tools for the deployment of structural metadata; and
- Publication tools.

Next steps

Public administrations are invited to implement and tailor the proposed generic governance and management models and use tools to support them. The guidelines provided here could apply at local (i.e. within an EU institution), inter-institutional (between EU institutions) and trans-European (between Member States and the EU institutions) levels. For example:

- **At local level** EU institutions such as DG Informatics of the European Commission could use this document as a guideline to better manage its data and ensure consistency among its systems.
- **At inter-institutional level** institutions such as the Publications Office of the European Union have demonstrated an interest in becoming a hub for coordinating the development and maintenance of specifications for structural metadata among the EU institutions in other domains; hereby expanding the work of the Inter-Institutional Metadata Maintenance Committee (IMMC) beyond the legal decision-making process of EU institutions. The Publications Office could play a leading role in the process "harmonise structural metadata".
- **At the Trans-European level** this document can be of guidance to all sorts of situations where EU institutions have to coordinate with public administrations in the Member States on structural metadata. The latter is often a consequence of an EU initiative (e.g. an EU project), or an EU legal act.

It is important to promote this report and gain experience with the use of these models and tools tailoring them to specific needs in the context of local or inter-institutional environments. The lessons learned should inform future versions of this document.

In terms of governance and management, each public administration should decide whether they need to implement each formal group within the governance model and each step proposed in the management model or if for instance a less formal or broad approach is needed. These experiences should be captured in lessons learned and could potentially lead to updating the present document.

In terms of tools, where these are readily available and used within public administrations, it is recommended to promote their usage and strengthen their integration using the standards identified in Section 4.1.

BIBLIOGRAPHY

- [1] European Commission, ISA Programme, “Metadata management requirements and existing solutions in EU Institutions and Member States,” [Online]. Available: <https://joinup.ec.europa.eu/node/78172>.
- [2] European Commission - ISA Programme, “Common Approach for the Management of Persistent URIs by EU Institutions,” European Commission, Brussels, 2014.
- [3] European Commission, Staff Working Document, “Guide for the procurement of standards-based ICT — Elements of Good Practice,” 2013.
- [4] European Commission, ISA Programme, “European Interoperability Framework (EIF) for European public services,” European Commission, Brussels, 2010.
- [5] H. Eichelberger, Y. Eldogan and K. Schmid, “A Comprehensive Survey of UML Compliance in Current Modelling Tools,” Gesellschaft für Informatik, Bonn, 2009.
- [6] European Commission, ISA Programme, “D7.1.3 - Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC,” 2012. [Online]. Available: <https://joinup.ec.europa.eu/community/semic/document/10-rules-persistent-uris/>.
- [7] European Commission, ISA Programme, “Towards a common policy for the management of persistent HTTP URIs by EU Institutions,” 2014. [Online]. Available: <https://joinup.ec.europa.eu/node/94830>.
- [8] G. Shukair, N. Loutas, V. Peristeras and S. Sklarß, “Towards semantically interoperable metadata repositories: The Asset Description Metadata Schema,” *Computers in Industry*, no. 64, pp. 10 - 18, 2013.
- [9] R. S. Smith, M. Lutz, A. Perego and R. Sgaolin, “Building a missing item in INSPIRE: The Re3Gistry,” European Commission - Joint Research Center, Ispra, 2013.
- [10] European Commission - Joint Research Centre, “Re3gistry software v0.3,” European Commission, Ispra, 2014.
- [11] National Information Standards Organization, “Understanding Metadata,” 2004.

ANNEX I IDENTIFIED METADATA TOOLS

The table below lists a number of commercially licensed and open-source metadata tools.

Table 6 – Existing tools

Tool	Licence	Comments
Enterprise Architect ⁵⁰	Commercial	<ul style="list-style-type: none"> • Data Modelling • Comparing data • Data visualisation • Import existing modelling techniques such as Archimate. • Has a licensing model with yearly renewal, costs are relatively low.
CAM editor	Open source	The CAM editor is a toolkit for building and deploying XML exchanges and Open Data APIs with JSON/XML/SQL. The OASIS CAM is a public open standard. The CAM editor can import, analyze and refactor existing exchange XML Schema for better compatibility and use in middleware, including generating model compliant XML Schema consistent with enterprise integration patterns.
ArgoUML ⁵¹	Open source	ArgoUML is an open source UML designing application running on Java and providing support for all UML 1.4 Diagrams, licensed under the Eclipse Public License (EPL). Through the XMI standard ArgoUML supports the exchange with other UML modelling tools.
Umbrello ⁵²	Open source	Umbrello is an open source diagram tool, licensed under GNU Public License (GPL). Amongst other features, the tool supports XMI import and export.
Open ModelSphere ⁵³	Open source	OpenModelSphere is a data, process and UML modelling tool. The software is available as open source under the GPL.
UMLet ⁵⁴	Open source	UMLet is e open source tool for quickly creating UML models. The software has been licensed under GPL. Since no underlying

⁵⁰ www.sparxsystems.com/products/ea/

⁵¹ <http://argouml.tigris.org/>

⁵² <http://umbrello.kde.org/>

⁵³ <http://www.modelsphere.org/>

⁵⁴ <http://www.umlet.com/>

Tool	Licence	Comments
		dictionary or directory for reusable design objects is used, UMLet is more of a drawing tool rather than a modelling tool.
PowerDesigner ⁵⁵	Commercial	SAP Sybase PowerDesigner is a commercial modelling for process, data and metadata management, making it. PowerDesigner has features for collaborative modelling.
MagicDraw ⁵⁶	Commercial	MagicDraw is a commercial UML, BPMN, SysML and UPDM modelling tool, allowing for collaborative design of models.
GENIS Reference Data Component (RDC)	Open source	<ul style="list-style-type: none"> • Focus on reference data • Access management Versioning
VocBench	Open source	<ul style="list-style-type: none"> • Focus on reference data • Access management • Versioning • Supports SKOS • Only RDF
Poolparty	Commercial	<ul style="list-style-type: none"> • Linked data management • Creating and maintaining taxonomies, thesauri, ontologies and knowledge graphs. • Uses standards like SKOS and RDF • Has a licensing model
Atlassian JIRA ⁵⁷	Commercial	<ul style="list-style-type: none"> • Teamwork platform • Manage the software development lifecycle • SVN Implementation • Has a regular licensing model and an open source project license
Atlassian Confluence	Commercial	<ul style="list-style-type: none"> • User Feedback • Any format supported • Access management + User profiles • Has a regular licensing model and an open source project license
Apache Subversion (SVN) ⁵⁸	Open source	<ul style="list-style-type: none"> • Version management • Access control

⁵⁵ <http://www.sybase.be/products/modelingdevelopment/powerdesigner>

⁵⁶ <http://www.nomagic.com/products/magicdraw.html>

⁵⁷ <https://www.atlassian.com/software/jira>

⁵⁸ <http://subversion.apache.org/>

Tool	Licence	Comments
		<ul style="list-style-type: none">• Command line or visual solutions (CollabNet⁵⁹) available.
Joinup	Open source	<ul style="list-style-type: none">• User Feedback• Different formats supported• Versioning (clear listing + statuses)• Access management + User profiles

⁵⁹CollabNet Subversion Edge provides a web interface to create and import repositories and manage them : <http://www.collab.net/downloads/subversion>

ANNEX II STAKEHOLDER REQUESTS AND NEEDS

In the context of the report 'Metadata management requirements and existing solutions in EU Institutions and Member States' [1]⁶⁰, and during the pilots with DG COMP and DG MARE a number of stakeholder requests and needs have been identified. These are summarised in the below tables:

- Metadata governance: see Table 7;
- Metadata management: see Table 8; and
- Metadata tools: see Table 9.

The table below lists the stakeholder requests and needs for metadata governance.

Table 7 – Metadata governance: stakeholder requests and needs

ID	Stakeholders requests and needs
ORGANISATION	
G1	Involve direct stakeholders in the governance process: direct stakeholders should be involved in the metadata governance process to ensure that the interests of the stakeholders are taken into account.
G2	Involve operational staff in functional meetings: representatives from the operational level should participate in functional-level meetings.
SCOPE	
G3	Local and inter-institutional governance: the mechanism for governance should encompass both Local and inter-institutional data exchange.
G4	Identify a core set of metadata to be standardised first.
G5	Mappings should be considered as a solution of last resort. It is recommended to try as much as possible to come up with a common agreement and only if it is not possible to reach such an agreement, then the governance should consider mapping as a solution of last resort.
DECISION MECHANISM	
G6	The mandate should be clear The governance mechanism should clearly state the mandate of the governance body with regard to taking decisions on: <ul style="list-style-type: none"> • Changes to reference data; • Intellectual property rights linked to reference data; and • Enforcement, i.e. implementation of reference data specifications in systems.
G7	Decision process should take into account time constraints Decision processes should be linked to time constraints which are dependent on the nature of the decision to be taken.

⁶⁰ ISA Programme (2014). Metadata management requirements and existing solutions in EU Institutions and Member States. <https://joinup.ec.europa.eu/node/78172>

ID	Stakeholders requests and needs
G8	<p>Describe how agreements are reached The decision making processes should describe how agreements are reached – e.g. via a qualified majority or via consensus building.</p>
ENFORCEMENT POLICY	
G9	<p>Legislation should be formulated at high level and should not specify details like the values in a code list or the elements of a data model.</p>
G10	<p>Comply or explain approach: it is recommended to deploy a “comply-or-explain” enforcement policy for the implementation of standards and specifications for structural metadata, irrespective of whether the implementation is realised through procurement or in-house development..</p>
G11	<p>Increase awareness on benefits of sharing and re-use: the solution should foresee to increase the awareness among stakeholders on sharing and reuse benefits by means of clear arguments aligned with specific business cases.</p>
G12	<p>Take into account legal instruments: the information which is exchanged between Member States and the European Commission is often specified in EU legislation. When including metadata governance in the decision-making process efficiency and speed should be taken into account.</p>
PROCESS FOR CONTINUOUS IMPROVEMENT	
G13	<p>Decisions should be documented Specific decision making processes which are depending on the context in which a decision is required should be developed, documented and shared with all relevant stakeholders.</p>
G14	<p>Licensing framework: the governance should also take care of decisions related to the licensing framework.</p>
G15	<p>Adaptations to the needs of the users should be delivered timely. At the same time it is necessary to guarantee stability</p>
G16	<p>Quality Assurance The reference data management and governance methodology should take into account quality controls when</p> <ul style="list-style-type: none"> • designing reference data; • updating or importing reference data; and • propagating reference data to production systems.
G17	<p>Risk mitigation Risks related to the deployment of changes to reference data into operational systems, should be mitigated by governance processes.</p>

Table 8 – Metadata management: stakeholder requests and needs

ID	Stakeholders requests and needs
M1	<p>Access rights management</p> <p>As certain parts of the data models can contain sensitive information, for example in the case of the data models for the Marine and Maritime environment, the metadata management and governance methodology should foresee processes to:</p> <ul style="list-style-type: none"> • manage access rights for the data model; and • manage the access request procedure.
M2	<p>Provide unambiguous guidelines for metadata use</p>
M3	<p>Accuracy: Standards and methods should be documented in a clear and straightforward manner.</p>
M4	<p>Documentation</p> <p>To all documentation a version number should be assigned; it should be made easily accessible for all stakeholders. Metadata should be documented in both human- and machine-readable formats.</p>
M5	<p>Continuous improvement</p> <p>The management processes should allow for continuous improvement for the purpose of remaining responsive to the needs of the users.</p>
M6	<p>Alignment with external bodies</p> <p>The lifecycle and management processes of the data model should be aligned with procedures from external standardisation bodies, especially when data models from those bodies are reused.</p>
M7	<p>Data Model Lifecycle Support</p> <p>The proposed management processes should support the lifecycle management of metadata models, which includes:</p> <ul style="list-style-type: none"> • Designing a data model; • Assuring the quality; • Manage access requests to the data model; • Request a change to the data model; • Update a data model; and • Document and publish a data model.
M8	<p>Reference Data Lifecycle Support</p> <p>The metadata management and governance methodology should support the lifecycle of reference data management, which includes:</p> <ul style="list-style-type: none"> • Designing reference data; • Managing reference data changes; • Sharing and using reference data; and • Harmonizing reference data.

Table 9 – Metadata tools: stakeholder requests and needs

ID	Requests and needs
Data modelling and visualisation	
T1	<ul style="list-style-type: none"> • Support visualisation of the UML diagram; • Support navigability in the visualised diagram; • Support search at the metadata element level, i.e. the level of granularity of search should be at the level of classes and properties; • Support multilingualism; • Provide complete documentation and guidelines for IT and business users; and • Facilitate feedback from users and work in collaborative way.
Accessibility and interoperability	
T2	<ul style="list-style-type: none"> • Make structural metadata available in a machine-readable form and via a Web tool; • Support integration with other tools (used during the design, maintenance and documentation phases); • Support fit-for-purpose data exchange formats; • Facilitate reuse of existing data models, both those developed internally by DG MARE and the CISE Cooperation Project and those owned by external standardisation organisations. • Have an integrated solution which provides public HTTP URIs for the different data entities.
Storing and versioning	
T3	<ul style="list-style-type: none"> • Offer a publicly available repository of reusable structural metadata – for those parts of structural metadata that are not subject to privacy/confidentiality constraints; • Offer a private authoritative source for the structural metadata, with restricted access; • Support the connection to an SVN for keeping track of internal drafts and published versions of the structural metadata; and • Allow integration with a ticketing system, e.g. JIRA. • Notify users on new releases of the data model.
Reference data editor	
T4	<p>Feature list</p> <p>DG COMP needs a tool that is capable of editing reference data and support the design of reference data in the context of one or more information systems. The tool should support tasks in the following processes:</p> <ul style="list-style-type: none"> • Design reference data; • Manage reference data changes.

	<p>The tool should have the following features:</p> <ul style="list-style-type: none"> • Import reference data from an external source and detect changes; • Create, read, update, or delete a concept scheme; • Create, read, update, or delete concepts in a concept scheme; • Add multilingual labels to a concept scheme; • Foresee a possibility to define the order of concepts in a concept scheme; • Version concept schemes; • Version concepts; • Version the labels of concepts; • Export one or more versions of a concept scheme.
Tools for managing reference data changes (ticketing/workflow)	
T5	<p>Feature list DG COMP needs a tool that is capable of</p> <ul style="list-style-type: none"> • Keeping a log of change requests; • Keeping track of impact analyses; • Keeping a log of decisions on change requests; • Creating release notice; • Linking change requests to release notes and vice-versa; and • Linking change requests to version and vice-versa.
Tools for reference data deployment	
T6	<p>Feature list The tool should allow:</p> <ul style="list-style-type: none"> • Deploy versioned reference data-as-a-service to an information system; • Deliver services while disconnected (local cache); and • Provision all versions (full versioning of temporal changes and language versions).
Tools for publication of reference data	
T7	<p>Feature list The tool should provide:</p> <ul style="list-style-type: none"> • File-based read-access over HTTP; • Write-access over WebDAV or Subversion.