SC17DI06692

Metadata management requirements and existing solutions in EU Institutions and Member States

Document Metadata

Property	Value
Release date	2014-05-08
Status	Final
Version	0.46
Authors	Makx Dekkers – AMI Consult Stijn Goedertier – PwC EU Services Audrius Leipus – PwC EU Services Nikolaos Loutas – PwC EU Services
Reviewed by	Margarida Abecasis – European Commission Pieter Breyne – PwC EU Services Athanasios Karalopoulos – European Commission Vassilios Peristeras – European Commission
Approved by	Margarida Abecasis – European Commission

This report was prepared for the ISA Programme by:

PwC EU Services

Disclaimer:

Table of Contents

List of Tables

List of Figures

EXECUTIVE SUMMARY

This report is commissioned by the Interoperability Solutions for European Public Administrations (ISA) Programme of the European Commission in the context of its Action 1.1. One of the prime objectives of Action 1.1[1] of ISA Programme is to document what is available in each Member State with regards to metadata policies and the management of structural metadata, i.e. data models and reference data, and raise awareness on the importance of metadata governance management.

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

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 management and governance ensures the coordinated development, use and maintenance of metadata, whilst ensuring also the sustainability of the metadata.

In this vein, one of the aims of this report is to identify and document best practices concerning metadata management requirements and existing solutions in EU Institutions and Member States.

Another goal of the report is to make conclusions and provide recommendations for the improvement of metadata management and governance practices and methodologies, also to identify opportunities for the reuse of metadata tools, repositories. The study is focused on data models and reference data within the inter-organisational information exchanges at both national and pan-European levels.

The key method of the study was in-depth analysis of selected cases conducted via structured interviews. The analysis framework was organised according to the three dimensions, namely metadata governance, metadata management and tool support. The cases for the detailed research were selected upon the predefined criteria such as regional diversity, information accessibility among others. The following cases were selected at pan-European level:

- The Statistical Office of the European Union - Eurostat, which process and publish comparable statistical information at European level.
- Joint Research Centre - INSPIRE, which enables the sharing of environmental spatial information among public sector organisations.
- Inter-institutional Metadata Management Committee (IMMC), Publications Office Metadata Registry (MDR), which registers and maintains metadata elements, named authority lists, schemas, etc.

To cover national dimension of the analysis those cases were examined:

- KoSIT (Koordinierungsstelle für IT-Standards), which coordinates the development and operation of IT standards for data exchange in the public administration in Germany.

[1] Action 1.1 of ISA Programme :
http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-1action_en.htm

- CISE – Centre for Semantic Interoperability, which carries out projects related to information and communication technologies (ICTs) in accordance with the strategic priorities in Spain.
- Lithuanian Spatial Information Portal (LSIP), which makes conditions for the provision of centralized spatial data.
- Knowledge and Exploitation Centre Official Government Publications (KOOP), which manages the metadata standard for information published by the Dutch government.
- Local Government Inform (LG Inform/LG Inform Plus), which provides benchmarking data service of Local Government Administration for councils in the UK.

The in-depth analysis of selected cases was supported by the results of the online survey, which was distributed among organisations involved in metadata governance and/or management. The survey was conducted in order to bring a wider perspective and better understand the research objectives and the related context.

Metadata governance

The findings of the report show that there are many similarities among organisations in terms of metadata governance goals and structure. Metadata governance is commonly aiming at interoperability between existing applications or systems and also at management of common reference data. We recommend composing the governance structure of at least 3 layers with the top being responsible for strategy, the middle taking decisions on structural metadata and the last one playing the operational role. It is also highly recommended to leave governance structure open and include stakeholders in the decision making process.

Enforcement policies vary depending upon the context and goals of metadata governance, also upon the policy domains. If the enforcement policy is voluntary, it is very important that the stakeholders would be made aware of the advantages of sharing and reuse of the metadata. On the other hand, if sharing and reuse is enforced by law, it is highly recommended not to describe structural metadata in the law, as in such case it would be very difficult to make amendments. The good practice would be to only indicate the links to the authoritative source in the legislation.

Open licences are recommended with protection against misrepresentation. Besides, the in-depth analysis of selected case studies suggests that EU Institutions are more aware of licencing frameworks than national organisations of Member States. Even though specific licences, such as EUPL, are used, sometimes they are not explicitly visible and accordingly stakeholders might be not aware of them. Thus improved conveyance of information related to licencing frameworks is needed.

Metadata management

The analysis shows that metadata management process is more standardized in EU Institutions than in organisations of Member States. Some of the national organisations do not apply any standards neither for metadata management nor for

documentation and representation. Our recommendation is to use at least metadata management standards that give advantages of cost effectiveness and reusability. The selection of specific standard should depend upon the policy domain.

The metadata update frequency and change management process vary case by case and depends on the type of metadata. We suggest using a well-balanced approach towards change management. Although, the organisation should maintain a consistent implementation towards defined goal, at the same time it should be agile enough to address the needs evolving from changes in organisational and technological environment.

Most of the organisations under the review have established processes for mappings between related metadata sets and for assessing alternatives, yet they are commonly used only for reference data and not for other sources such as vocabularies. There is usually authoritative source where metadata is published and the publication processes are commonly well documented and uniform. Our analysis suggests that XSD could be used for the publication of metadata schemas and the transmission of protocol schemas, while SKOS files could be used to manage value and code lists.

The findings of the study show that there are usually few full time equivalents FTEs responsible for metadata management and it is emphasized that benefits notably offset the costs. The most common benefits include better interoperability, better reuse of metadata, more efficient use of resources, access to information, as well as other aspects.

Identification of reusable tools

In all of the reviewed cases human- and machine-readable formats are available for retrieval, yet human- and machine-readable distributions of the structural metadata are published separately. Publishing and retrieval is usually supported by various tools that are domain specific. The possibility to reuse the tool depends not only on its functionality, but also on the type of licence and owner/ vendor. Commercial and open-source products are easily reusable, while in-house built require efforts for development and adaptation.

Summary of recommendations and good practices

Concluding, the study identified a number of observations of good practices put forward by the selected case studies and a number of recommendations.

The identified good practices are:

1. In a good governance structure, the roles concerning legislation, strategy, functionality and operations are clearly distinguished and assigned to designated bodies.
2. The involvement of direct stakeholders in the metadata governance process ensures that the interests of the stakeholders are taken into account which maximises buy-in and take-up.
3. Voluntary sharing and re-use works best if stakeholders are aware of the advantages of collaboration and of the benefits for interoperability.

4. Application of a standard for metadata management creates a well-structured management environment based on existing good practice. We defined two families of standards:
    a. Standards for metadata management, such as ISO/IEC 11179, ISO 19135 and ISO 19115.
    b. Standards for documentation and representation, such as SMDX, ADMS and SKOS.
5. Good change management processes are based in stability where possible without sacrificing flexibility where needed and take into account an alignment between the life cycles of structural metadata development and software development.
6. Changes in structural metadata are well planned and tracked, preserving backward compatibility as much as possible; in cases where disruptive changes are unavoidable, these changes should be planned and communicated well in advance.
7. Structural metadata is managed in formats that are appropriate for the type of use. Metadata describing the structural metadata is expressed or exported using the Asset Description Metadata Schema.
8. Standard reference data is used wherever appropriate; if locally defined reference data is used, this is mapped to standard reference data to enable wider interoperability.
9. Structural metadata is distributed in machine-readable formats that can be processed by the tools available by the reusers.
10. Content negotiation is used to manage and provide different types of formats from the same URI.
11. Metadata governance and management ensures the sustainability of structural metadata.

The recommendations are:

1. Legislation should be formulated on a sufficiently high level and should not specify details like the values in a code list or the elements of a data model; these details should be specified as part of the implementation and made available from an authoritative source to which the legislation can refer.
2. The structural metadata management processes should be documented.
3. Owners of structural metadata should be made aware of the importance of clear licensing arrangements that specify unambiguously under which conditions the metadata can be reused. Open reusable metadata is recommended.
4. Stakeholders should be aware of the expected benefits of metadata sharing and reuse.
5. Management processes and publication frequencies should be different for changes to data models on one hand, and reference data on the other hand.
6. Structural metadata should have persistent unique identifiers.

7. Tools used for supporting metadata governance and management should be based on open standards and should be interoperable.

1. INTRODUCTION

This report presents the findings of the study on existing metadata management and governance practices in EU Institutions and Member States.

1.1. Context & scope

The study on metadata management requirements, governance and existing solutions in EU Institutions and Member States is commissioned by the Interoperability Solutions for European Public Administrations (ISA) Programme of the European Commission, in the context of its Action 1.1 on improving semantic interoperability in European e-Government systems.

The scope of the study covers metadata management and governance practices, methodologies and tools used by EU Institutions and Member States, focusing on two types of inter organisational (within or across borders) structural metadata that is used in the context of **inter-organisational information exchanges** within and/or across borders:

- Data models;
- Reference data.

Metadata management and governance ensures the coordinated development, use and maintenance of metadata, whilst ensuring also the sustainability of the metadata.

The study builds upon the previous assessment of metadata management which was conducted by the ISA Programme in 2011-12 (ISA Programme of the European Commission, 2012).

What is a 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. 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.

What is 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[2].

Examples of reference data include:

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

- Type codes, i.e. codes and values used to categorize an object by its type, e.g. file formats, business activity types.
- Status codes, i.e. codes and values used to describe the lifecycle of an object, e.g. 'completed' or 'under development'.
- Descriptive taxonomies and vocabularies, i.e. (relatively) stable lists of codes and values of real-world things, e.g. a list of countries or a list of currency codes.

1.2. Objectives

The main **objectives** of this study are the following:

a. To compile existing requirements, methodologies, practices, procedures and reusable tools for metadata governance and management in the EU Institutions and the Member States.

b. To elaborate on the costs, benefits and feasibility constraints, of the different alternatives.

c. To document lessons-learnt and good practices.

1.3. Scope of the report

In order to address its objectives, this study provides answers to the following research questions:

1. **Existing solutions:** How do the Member States and EU Institutions currently govern and manage their metadata (including tools, roles, processes and methodologies, costs and benefits – if any)?
2. **High-level requirements:** Which are the common principles and requirements for metadata management and governance in EU Institutions and Member States?
3. **Reusability of existing solutions:** Is it feasible to reuse existing solutions for metadata management and governance and what are the main constraints if so?

1.4. Structure

The remainder of this report is structured as follows:

Chapter 2 outlines the methodology that was followed for carrying out this study.

Chapter 3 indicates a long list of relevant cases as a result of desk research and shortened list of cases for further analysis. This chapter also provides results and analysis of the online survey.

Chapter 4 summarises the results of the online survey.

Chapter 5 outlines the key findings of case studies and online survey.

Chapter 6 represents the good practices, whereas chapter 7 outlines recommendations for metadata management and governance.

Chapter 8 lists conclusions arising from the completion of the previous parts of this deliverable.

Annex II includes an in-depth analysis of eight case studies that consists of information collected from a desk research and interviews. The long list of cases identified is available in Annex I.

1.5. Glossary

This section provides common definitions used throughout the study.

Table 1 - Glossary

Term / Acronym	Description
ADMS	A common metadata vocabulary to describe standards, so-called interoperability assets, on the Web. ADMS is a W3C Working Group Note[3].
Content negotiation	A mechanism defined in the HTTP specification that makes it possible to serve different versions of a document (or more generally, a resource representation) at the same URI; so that user agents can specify which version best fit their capabilities.
CSV	Comma-separated (or character-separated) values file stores tabular data (numbers and text) in plain-text form.
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
DCMI	Dublin Core Metadata Initiative.
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.
ICT	Information and Communication Technology
Interoperability	According 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 *exchange of data* between their respective ICT systems.
IRR	The internal rate of return (IRR) or economic rate of return (ERR) is a rate of return used in capital budgeting to measure and compare the profitability of investments
GUI	Graphical User Interface
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.

[3] W3C Working Group Note: http://www.w3.org/TR/vocab-adms/

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.
NISO	National Information Standards Organization
OWL	The Ontology Web Language (OWL) is a set of markup languages which are designed for use by applications that need to process the content of information instead of just presenting information to humans.
RDF	The Resource Description Framework is a general-purpose language for representing information in the Web
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.
SKOS	Simple Knowledge Organization System – RDF Vocabulary for the representation of key reference data such as code lists, and taxonomies.
Structural metadata	Data model or reference data
Trans European Systems	Trans-European ICT solutions contribute to the realisation of a Digital Single Market in Europe and the free movement of people, information and goods across the Member States. They are set up to support an EU policy, often – but not necessarily – as a direct consequence of new EU legislation. Examples of Trans-European ICT Solutions are the VAT Information Exchange System (VIES), the European Criminal Record Information System (ECRIS), the Emissions Trading System (ETS), the Visa Information System, the Internal Market Information System (IMI) and the Electronic Exchange of Social Security Information (ESSI).
EFTA	The European Free Trade Association
EEA	The European Economic Area
W3C	World Wide Web Consortium
XML	Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.

2. STUDY METHODOLOGY

The study consists of six major phases including: scope definition, desk research for selecting relevant cases, identification of eight cases for further analysis and interviews, preparation of analysis framework, carrying out an online survey, analysis of the findings and elicitation of conclusions. Figure 1 presents the sequence of the main tasks to be completed and outcomes of each of them.

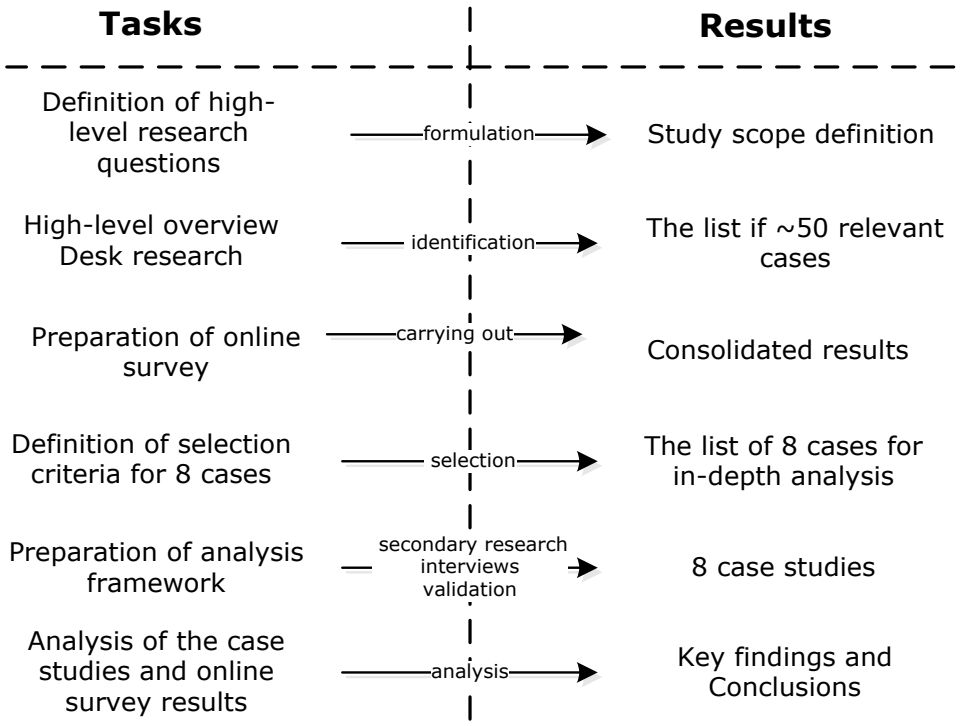

Figure 1. Study methodology concept

1. Definition of high-level research questions (definition of the study scope). Firstly, a number of research questions are formulated to support the objectives of this study. The research questions are outlined in section 1.3 of this report.

2. High-level overview. We investigate the metadata governance and management initiatives existing in EU Institutions and Member States by the means of desk research and an online survey.

a. **EU Institutions.** A number of EU Institutions that had put in place metadata governance, management methodologies and tools were identified in the business case (ISA Programme of the European Commission, 2013) This study takes a closer look into these initiatives; and
b. **Member States.** Building on the results of the survey of 2011-2012 (ISA Programme of the European Commission, 2012), we reviewed the current state of affairs with regards to metadata management and governance in the Member States that helped in identifying requirements, tools and best practices in relation to metadata governance and management.

In parallel, an online survey was also prepared and sent out, which aimed at building a better understanding on the main aspects of metadata governance and management across EU Institutions and Member States.

3. Definition of case selection criteria. By following the objectives of this study we defined a number of case study selection criteria that were later used to shortlist eight case studies for further analysis. Section 2.1.1 describes the selection criteria.

4. Detailed analysis of the selected case studies. First, the analysis framework was prepared for further investigation of each case study. The analysis framework was later used as the basis for structuring each of the 8 interviews (1 interview per case study). Prior to the interviews, desk research was used for collecting available information; then the interviews were run to gather additional information as well as to confirm our initial understanding of the case studies. After debriefing the interviews, a summary report was sent to the interviewees for their review, asking them to validate the findings. Section 2.1 explains the steps taken for the detailed analysis for each of the eight cases studies.

5. Analysis and summary of the findings. The last phase of our study summarised the findings of the eight cases study and the online survey which is targeted at wide audience across Europe. Conclusions are defined in this last stage. Section 2.2 describes the approach of this phase.

2.1. Selection and analysis of case studies

The study's approach for the selecting case studies consisted of the following steps:

- **Creation of a long list of candidate cases:** First we identified a long list of case studies (see section 3.1).
- **Selection and detailed analysis of eight cases studies:** Then based on the selection criteria (see section 2.1.1) we selected and further analysed 8 of them.
    a. **Analysis framework**: We outlined a guide to help carrying out case studies in a consistent and harmonised manner listed in section 2.1.2. Next, we selected 8 case studies as the most valuable in terms of metadata governance, management, tools applied and information gathered.

    b. **Secondary research:** By using the analysis framework we gather publicly available case related information, which was later validated in the interview.

    c. **Interviews and elicitation of requirements:** We conducted one interview per case study and collected specific information requirements for metadata management, governance and tools that was not publicly available; and

    d. **Quality control.** We invited the organisations participating in each case study to validate the gathered information and findings.

2.1.1. Case selection criteria

Cases studies for further analysis were shortlisted based on the following criteria:

- **Inter- / intra-organisational dimensions**: We selected cases with a metadata management and governance with inter / intra – organisational dimension;
- **Administrative dimension:** We selected cases with a cross-border/-sector information exchanges in the context of pan-European and/or national systems.
- **Accessibility of information:** We selected cases where information regarding metadata management and governance was easily accessible, and there was the willingness of the organisation involved to contribute to the study.
- **Regional diversity:** We selected cases coming from various EU Institutions and Member States aiming to provide a balanced geographical coverage of the EU.

2.1.2. Case study analysis framework

We organise our analysis of eight selected case studies according to the following three dimensions that map directly to the objectives of this work:

- Metadata governance;
- Metadata management; and
- Tool support.

2.1.2.1. Metadata governance

Tailoring TOGAF's[4] definition of governance in the context of this study, we argue that structural metadata governance is about ensuring that the management of metadata is conducted properly, i.e. following a set of guiding principles and practices, and in accordance with an organisation's strategic objectives. In this vein, structural metadata governance comprises well-defined roles and responsibilities, cohesive policies and principles, and decision-making processes that define, govern and regulate the lifecycle of the structural metadata.

- **Goals**: What are the long-term goals of cross-border/sector governance, management and reuse of common structural metadata?
- **Governance structure:** What are the roles responsible for structural metadata governance and what is their mandate?
    - **Roles;**
    - **Openness** of the process;
    - **Decision making process:** What is the process of taking decisions on the lifecycle of structural metadata, e.g. on updating a definition, releasing a new version or deprecating an existing one?
- **Context**: What is the context in which structural metadata is developed, applied, shared and reused:
    - **Organisational dimension**:
        - **Inter-organisational;** or
        - **Intra-organisational.**
    - **Administrative dimension:**
        - **Pan-European;**

[4] The Open Group Architecture Framework (TOGAF): http://www.opengroup.org/togaf/

        - **National;**
        - **Regional/Local.**
- **Policy domain:** Is structural metadata governance restricted to specific policy domains? Which policy domains?
    - **Not restricted;** or
    - **Restricted** (e.g. Finance, Healthcare, etc.).
- **Enforcement policy:** Where relevant, the applicable legal context must be completed.
    - **Sharing**: which governance model is used to encourage/ensure sharing of structural metadata?
        - **Legal requirement**: sharing is enforced by law; it is an official requirement;
        - **Voluntary**: sharing is not enforced, but encouraged on a voluntary basis.
    - **Reuse**: which governance model is used to encourage/ensure reuse of structural metadata?
        - **Legal requirement**: reuse is enforced by law; it is an official requirement;
        - **Comply-or-explain**: reuse is not enforced by law, but public administration have to comply with the use of a particular specification or standard for metadata, or if they do not comply, explain publicly why they do not;
        - **Oversight board**: reuse is encouraged via project review committees;
        - **Voluntary**: reuse is encouraged via information campaigns.
- **Authoritative source:** Is there an authoritative source, e.g. a repository or a file server, on which the structural metadata is housed?
- **Licensing framework**: Under which licensing framework are structural metadata shared and reused?
    - **No explicit licence;**
    - **Class 1**: Traditional, proprietary licence;
    - **Class 2**: Free to use and redistribute, no modifications allowed;
    - **Class 3**: Free to use, redistribute, and modify via copy left licence (e.g. CeCILL);
    - **Class 4**: Free to use, redistribute, and modify via non-copy left licence (e.g. Academic Free License).

The ISA Open Metadata Licence v1.1

ISA Open Metadata Licence v1.1[5] is a licence published by the European Commission which states that any use of the work other than authorised under this Licence or copyright law is prohibited.
In accordance to Commission Decision 2011/833/EU, natural or legal person or body of persons corporate or incorporate are herewith granted a worldwide, royalty-free, perpetual, non-exclusive licence to use and re-use any information and / or data offered under this Licence and any modifications thereof for any

[5] ISA Open Metadata Licence: https://joinup.ec.europa.eu/category/licence/isa-open-metadata-licence-v11

commercial and non-commercial purpose are allowed by the law and provided that the following conditions are met:

- Distributions or communication to the public must retain the above Copyright Notice;
- Distributions must retain the following "No Warranty" disclaimer;
- Where possible and practical, distributions or communication to the public will provide a link to the Joinup platform;
- Acts directed to mislead others or misrepresent the Work, its content or source are prohibited;
- Name of the European Commission and that of its contributor(s) won't be used to endorse or promote products and services derived from the use of the work without specific prior written permission.

Between 30 August 2013 and 22 November 2013, the European Commission run a consultation on guidelines on recommended standard licences, datasets and charging for the reuse of public sector information[6]. In this context input was collected on the methods by which public sector bodies should establish rules governing the reuse of data, on the choice of the possible conditions to be laid down and on other important aspects of licensing, such as interoperability and standards. The vast majority of the respondents indicated Creative Commons[7] as the preferred licence, with its variants CC0[8] and CC-BY[9] being the most popular choices.

- **Quality controls:** What is the quality management process for structural metadata?
- **Metadata Schema:**
  - **Vocabulary:** Is there a common vocabulary in place, such as the Asset Description Metadata Schema - ADMS, for documenting metadata?
  - **Identifiers scheme:** Are there common guidelines and design patterns for creating identifiers for the metadata?
  - **Schema documentation:** Is there schema documentation in place, which explains commonly agreed definitions about the meaning of the data?
  - **Multilingualism:** Are different languages supported for the metadata properties and/or values?

What is ADMS?

The Asset Description Metadata Schema (ADMS)[10] is a common metadata vocabulary to describe standards, so-called interoperability assets, on the Web. ADMS was created by the EU's Interoperability Solutions for European Public Administrations (ISA) Programme of the European Commission to help publishers of

[6] http://ec.europa.eu/digital-agenda/en/news/consultation-guidelines-recommended-standard-licences-datasets-and-charging-re-use-public

[7] http://creativecommons.org/

[8] http://creativecommons.org/about/cc0

[9] http://creativecommons.org/licenses/by/4.0/

[10] https://joinup.ec.europa.eu/asset/adms/description

standards document what their standards are about (their name, their status, theme, version, etc.). So ADMS helps people make sense of the complex multi-publisher environment around standards and in particular the ones which are semantic assets such as ontologies, data models, data dictionaries, code lists, XML and RDF schemas. ADMS is currently a W3C Working Group Note[11].

What is ISO 11179[12] Metadata Registry standard?

ISO/IEC 11179 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 organisation that produces the data. It does not apply to the physical representation of data as bits and bytes at the machine level.

What is ISO 25964 Thesauri and interoperability?

The ISO 25964 standard[13] is applicable to thesauri and other types of vocabulary that are commonly used for information retrieval. It describes, compares and contrasts the elements and features of these vocabularies that are implicated when interoperability is needed. It gives recommendations for the establishment and maintenance of mappings between multiple thesauri, or between thesauri and other types of vocabularies.

2.1.2.2. Metadata management

In this work, we define metadata management as a set of high-level processes for structuring the different phases of the lifecycle of structural metadata. The following phases comprise the lifecycle of metadata:

- **Design and development:** which 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 as it has already been treated in previous work of ISA Action 1.1, resulting in the specification of a process and methodology for developing commonly agreed structural metadata*** (ISA Programme of the European Commission, 2013b).
- **Update:** which entails the processes of updating the structural metadata, and deleting/deprecating outdated versions.
    - **Update frequency:** What is the average update frequency of structural metadata?
    - **Change management process:** Is there a defined procedure in place?
    - **Version control:** Is there a version control system in place?

[11] http://www.w3.org/TR/vocab-adms/

[12] Metadata registries – Framework (ISO/ IEC 11179) standard: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=35343

[13] http://www.niso.org/schemas/iso25964/

- **Harmonisation:** which covers the processes that have to be put in place for:
    - **Creating mappings between related structural metadata sets:** for example, what are the process and the criteria to follow when aligning two code lists of language codes or public service types?
    - **Assessing alternative structural metadata sets:** what is the process to be followed for selecting between alternative structural metadata sets that serve the same purpose? For example, how can one select between alternative country code lists or data models for describing a person?
- **Documentation:** which covers the processes that facilitate the sharing and reuse of structural metadata:
    - **Publication** of structural metadata documented according to a common vocabulary on an authoritative source which is accessible and supports search capabilities for both humans and machines.
    - **Retrieval** of structural metadata either by humans (e.g. downloading a file) or by machines (e.g. by consuming the metadata via an API or a service).
    - **Supported formats of publication tool(s):** the formats supported bya particular solution.
        - **Human-readable only (H):** formats readable by humans but not readable by machines e.g. unstructured text presented in pdf, jpeg, png, etc;
        - **Machine-readable only (M):** formats readable by machines but not readable by humans e.g. binary representations as e-Signatures, Unicode text, barcodes;
        - **Human and Machine readable (HM):** formats which can be read by humans and by machines e.g. highly structured languages such as XML and OWL, but also CSV etc. Human and machine readability is not a crisp concept; human readability much depends on the convenience by which a format can be read and machine-readability on the structure and semantics that can be conveyed.

As part of metadata management, we also investigated:

- **Tool support:** Which functionality from the above (in terms of metadata management) is supported by a particular tool.
- **Standards**: the standards used for structural metadata management? For example, ISO 11179 Metadata Registry standard, ISO 25964 Thesauri and interoperability, the Simple Knowledge Organisation System - SKOS, the Asset Description Metadata Schema – ADMS, etc.
- **Costs**
    - the number of FTEs occupied with structural metadata governance and management in particular organisation.
- **Benefits**
    - The key benefits of structural metadata governance and management in a particular case?
    - Do the benefits offset the costs?

2.1.2.3. Tool support

Structural metadata governance and management can be supported by software tools. Such tools implement the metadata governance principles and the metadata management processes.

- **Name of the tool(s) and documentation:** The tools used for structural metadata governance and management in particular case.
- **Reusability**:
  - **Open source:** Is a tool developed using open source technologies?
  - **Open licence:** Is a tool released and distributed under an open licence?
  - **Owner/vendor:** Who has developed and/or funded the development of the tool?
  - Has the tool already been reused? How long did it take to adopt/adapt it?
  - What are other constraints for reusing a tool?
- **Key functionality:** which are the key functionalities of the tool that allow it to effectively support the particular metadata governance and management process?
- **Costs:**
  - What were the costs of metadata-related tools (e.g. out of the shelf) and implementation?
  - Are there any other costs (licences, maintenance)?

2.2. Analysis and summary of the findings

On the basis of our high-level overview, the detailed analysis of the selected eight cases studies and results of the online survey, we derived:

a. **Recommendations and good practices** for structural metadata governance and management methodologies;

b. **A list of reusable tools** that can support metadata governance and management.

2.3. Online survey

The online survey questionnaire comprised 4 simple questions that aimed to cover the main points of structural metadata governance and management surveyed in this study. The questionnaire can be found in Annex III – Survey Questionnaire.

The master version of the questionnaire was uploaded onto the online survey tool – called Interactive Policy Making (IPM)[14]. We then sent invitations by email to the contact persons of cases listed in section 3.1. The links to the survey were also communicated on Joinup.eu[15] and on social media, such as SEMIC's LinkedIn group[16] and twitter account. The outcomes of the online survey are discussed in section 4.

[14] Interactive Policy Making (IPM): http://ec.europa.eu/yourvoice/ipm/index_en.htm

[15] Joinup: https://joinup.ec.europa.eu/

[16] SEMIC LinkedIn Group: http://www.linkedin.com/groups/SEMIC-2736596?trk=my_groups-b-grp-v

3. IDENTIFICATION AND SELECTION OF CASE STUDIES

In this chapter, we provide an overview of the information that was collected in the scope of this survey with regards to the practices, procedures, methodologies and tools that EU Institutions and Member States use for structural metadata governance and management.

3.1. Identification of cases

We identified cases with presence of structural metadata governance and management methodologies or tools by reviewing our former studies and conducting a desk research of publicly available information. We started by selecting several cases in each of the Member State to achieve regional diversity and afterwards shortlisted them by the criterion of inter / intra – organisational metadata governance / management, also by the criteria of information accessibility. The full list of the 49 identified cases is given in Annex I. Identified representatives of those cases were invited to take part in the online survey, findings of which can be found in section 4. The list of cases were further analysed and shortlisted based on the selection criteria.

3.2. Evaluation and selection of cases for in-depth analysis

After identifying and analysing the full list of cases we selected the most relevant and most value adding cases in the context of this study. The selection was based on the information found and criteria listed in section 2.1.1. While selecting cases studies we balanced them between cases of EU Institutions and Member States. At the end we selected three case studies coming from EU Institutions and five from Member States.

Table 2 – List of selected cases for detailed analysis

Line No.	Country	Name	Responsible institution	Description
1.	EU	Eurostat	Eurostat	The main role of Eurostat is to process and publish comparable statistical information at European level.
2.	EU	EU spatial data infrastructure (INSPIRE)	European Commission	INSPIRE aims to create EU spatial data infrastructure, which will enable the sharing of environmental spatial information among public sector organisations. It also helps to facilitate better public access to spatial information across Europe.
3.	EU	The Metadata Registry (MDR) of EU	Inter-organisational Metadata Management Committee	MDR registers and maintains definition data (metadata elements, named authority lists, schemas, etc.) used by the different EU Institutions.
4.	DE	Finanzen Bremen - KoSIT - XÖV Framework	KoSIT	The coordinating body for IT standards KoSIT has the task to coordinate the development and operation of IT standards for data exchange in the public administration.
5.	ES	Centro de Interoperabilidad	CISE	CISE is the instrument defined in the Spanish National Interoperability

Line No.	Country	Name	Responsible institution	Description
		Semántica (CISE)		Framework to publish data models and encodings associated with the exchange of data between the different administrations.
6.	LT	Lithuanian Spatial Information Portal (LSIP)	The state enterprise National Centre of Remote Sensing and Geoinformatics "GIS-Centras"	The scope of the LSIP is to make good conditions for centralized data (spatial data) provision among the data users. The shared data is provided and could be used by those who create and use spatial data and metadata.
7.	NL	KOOP - Knowledge and Exploitation Centre Official Government Publications	Knowledge and Exploitation Centre	KOOP is a government organisation that develops and manages all levels of government, both central government and the provinces. It manages the metadata standard for information published by the Dutch government on the Web, including schemas, code lists and value syntaxes.
8.	UK	Local Government Inform (LG Inform)	Local Government Association	LG Inform is the benchmarking data service of Local Government Administration for councils, fire and rescue authorities based on e-Government Metadata Standard.

4. THE OUTCOMES OF THE ONLINE SURVEY

The purpose of the survey was to identify the current practices and existing solutions in the area of inter-organisational metadata governance and management in European public administrations. The survey results will support our findings deriving from the case studies and will help the ISA Programme to better understand the current situation and needs of metadata governance and management and more effectively target its initiatives. Five topics were discussed in the survey:

- The level at which organisations apply metadata governance and/or management;
- The solutions that are currently used for metadata management, including methodologies, policies, standards and tools;
- The extent to which organisations share and reuse structural metadata;
- The benefits of applying metadata management and governance;
- The incurred roadblocks to inter-organisational metadata governance and management.

4.1. Respondents

The survey was launched on January 7, 2014 and closed on January 30, 2014. During that period, 20 respondents completed the survey. The respondents represent public sector organisations (16 out of 20), covering a differentiated set of domains including finance, archiving, statistics, ICT and public administration. Geographically, the institutions that participated in the survey are spread across 15 EU Member States and the US.

Table 3 – List of survey respondents

Institution	Country	Sector
Federal Chancellery of Austria	Austria	public sector
Flemish ICT Organisation - V-ICT-OR	Belgium	public sector
Ministry of Public Administration	Croatia	public sector
Croatian Bureau of Statistics	Croatia	public sector
Ministry of Finance	Finland	public sector
Institut National de la Statistique et des Etudes Economiques	France	public sector
Ministry of Finance and Economic Affairs	Iceland	public sector
Agency for Digital Italy (AgID)	Italy	public sector
Publications Office of the EU	Luxembourg	public sector
National Statistics Office	Malta	public sector
Malta Information Technology Agency (MITA)	Malta	public sector
AMA – Agency for Public Services Reform	Portugal	public sector
Ministry of Finance of the Slovak Republic	Slovak Republic	public sector
Statistics Sweden	Sweden	public sector
Ministry of the Interior	The Netherlands	public sector
Stichting ICTU (Dutch government)	The Netherlands	public sector
Actea Consulting AB	Sweden	private sector

Institution	Country	Sector
Europeana	The Netherlands	private sector
Liberata UK Ltd	United Kingdom	private sector
Wolters Kluwer	United States	private sector

4.2. Results

Level of metadata management and governance

The aim of this question was to identify the context in which metadata governance and management are executed. This can be within and across organisations, in the same country or across borders.

Six respondents indicated that they have not applied any level of metadata management or governance in their organisation. In our population, six institutions apply metadata management and governance on an intra-organisational level. Intra-organisational metadata governance and management systems are often specifically created for and applied by the institution itself. This however does not always entail that the metadata is isolated from other institutions. Three respondents who apply intra-organisational metadata management and governance indicated that they exchange metadata with other institutions, or participate in the development of international metadata management and governance processes. From eleven institutions applying cross-institutional metadata governance and management, eight do so on a national level by developing, applying or sharing guidelines and laws with other public institutions. Five institutions indicated that they have metadata management and governance processes at EU-level. The main reason for doing so is to share metadata with a dedicated EU body, such as Eurostat or the European Geoportal (INSPIRE). Only two organisations apply metadata management and governance at international level.

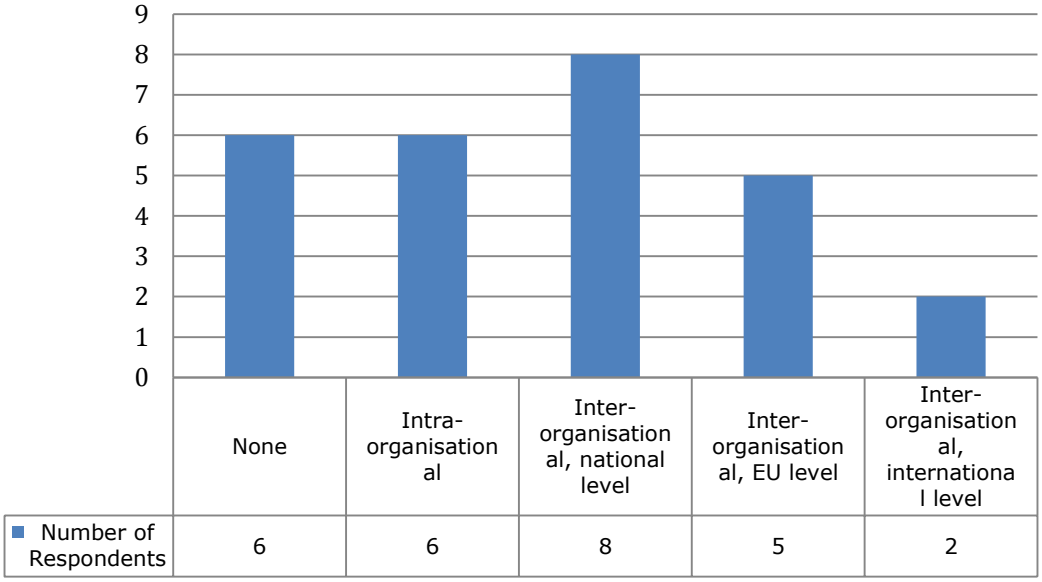

Figure 2. Distribution of respondents by their metadata management and governance

Solutions for metadata management

Many of the respondents use a combination of tools, guidelines and standards for metadata governance and management. Two respondents indicated that their metadata governance system, for example, was based on ISO/IEC 11179, three respondents referred to Dublin Core and four institutions use SKOS. ADMS is very well known among the participants of our survey. From the eight respondents indicating to be familiar with the metadata schemas, five claimed to be (partly) using ADMS and two are planning to do so in the future. Two of the respondents referred to the INSPIRE Directive as the standard for exchanging spatial information.

Whereas the main standards are used by different institutions, as described above, tools, policies and methodologies are rather specific per each institution. Methodologies like UN/CEFACT or CENBII are referred only once. Tools being used include JIRA for governance workflow, SVN for version control and PERL or XSLT for conversion.

Sharing and reuse of structural metadata

75% of the respondents claim they are sharing and re-using structural metadata. These are mainly international vocabularies, standards and classifications are reused.

Benefits of applying metadata management and governance

The respondents of the survey identified several benefits of applying metadata management and governance. "Improved data quality" is seen as the main benefit by ten respondents, "more efficient administrative processes" by six. Only one respondent indicated the reduced system and development costs as main driver for applying metadata management and governance. Most of the respondents however noted that all possible benefits are interrelated and relevant for their institution.

Roadblocks to inter-organisational metadata management and governance

Figure 3 below summarizes the most important roadblocks to inter-organisational metadata governance and management:

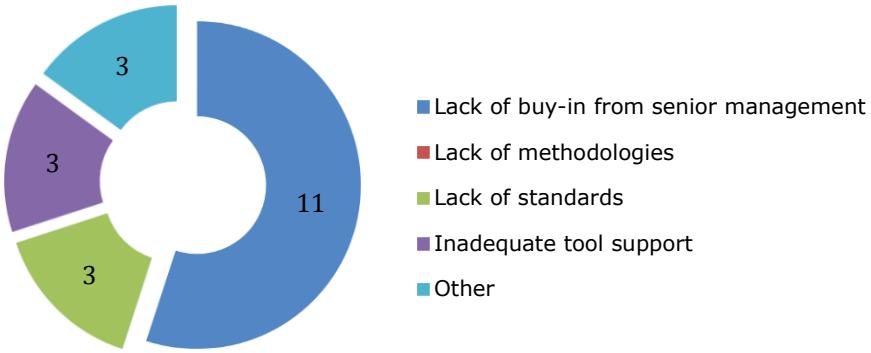

Figure 3. Most common and important roadblocks

The survey indicates that the **lack of buy-in from senior management is the main roadblock to inter-organisational metadata management and governance**. The complexity of metadata management makes it difficult to convince senior management of its importance. Three respondents selected "lack of standards" as the main roadblock. More specifically, the respondents explain that the limited use of uniform standards between institutions forms a barrier to efficient exchange of metadata. "Inadequate tool support" was identified as the main roadblock by three institutions. These institutions mentioned lacking tools able of generating and maintaining documentation, modelling tools and platforms for collaboration. Out of other barriers identified by the respondents include the lack of internal policies, the lack of willingness to apply a unified process, the problem of keeping up to date data of all public bodies and the limited knowledge as well as availability of experts for dealing with metadata management and governance issues.

The table below lists the metadata management solutions collected from the respondents of the survey. Notably, 8 organisations reported using or planning to use ADMS for documenting their metadata, while 4 of them use SKOS for representing code-lists. The use of Dublin Core descriptors is also common. 4 respondents are using ISO 11179 for metadata management. OWL and XML seem to be the most popular languages for encoding structural metadata.

Table 4 – Respondents' solutions for metadata management

Solution	Comment
Standards - Data	
Core Vocabularies	Simplified and extensible data model
DDI	Metadata standard for statistical and social science data
FOAF	Vocabulary for people and information linking
Standards – Metadata	
ADMS	Metadata documentation
DCAT	Vocabulary for interoperability facilitation
Dublin Core	The metadata governance system
ESMS	Statistical data documentation

Solution	Comment
ISO 19115	XML implementation schema standard
ISO/IEC 11179	Metadata registry standard
METS	Metadata encoding and transmission standard
RDA	Library cataloguing standard
SDMX	Data and metadata exchange standard
SKOS	Standard for representing controlled vocabularies
Languages	
OWL	Web information processing language
XMI	Metadata information exchange standard
XML	Documentation in human and machine readable format
Methodologies and Policies	
CEN/BII	Business Interoperability Interfaces for Public procurement in Europe
IMMC	Metadata elements for data exchange related to the legal decision making process between the institutions
UN/CEFACT	Electronic Data Interchange standards for electronic trade documentation in XML
Tools	
JIRA	Governance Workflow
PERL, XSLT	Conversion
SVN	Version Control

5. KEY FINDINGS

In this chapter we analyse the collected information to provide the key findings on the requirements, alternatives, feasibility constraints, and high-level features of metadata management, governance mechanisms and relevant tools for EU Institutions and Member States. The analysis is structured following the analysis framework of section 2.1.2 and therefore comprises of three major sections: metadata governance, metadata management and tools applied. The findings of the case studies and online survey are discussed and supported by good practice examples that form a base for high-level requirements in section 6.4.

5.1. Metadata governance

Metadata governance is about ensuring that the management of structural metadata is conducted properly, i.e. following a set of guiding principles and practices, and in accordance with an organisation's strategic objectives. In this vein, metadata governance comprises well-defined roles and responsibilities, cohesive policies and principles, and decision-making processes that define, govern and regulate the lifecycle of structural metadata.

Goals

After the in depth-analysis of all eight case studies we understand that there are many similarities among them in terms of metadata governance goals. The most common goals of structural metadata governance and management are:

- To achieve and ensure interoperability between existing applications and systems (e.g. KoSIT, Eurostat, INSPIRE, IMMC/MDR, etc.);
- To provide an authoritative source from where structural metadata could be made available, stored and managed, e.g. see the cases of Eurostat and LGI).

The main challenges faced are the following:

- The goals and the importance of metadata management are **often not clear to the management of the organisation and there is a lack of support from the leadership**. The complexity of metadata and the lack of tangible result make it difficult to convince senior management of its importance and the necessity of support. This seemed to be a common issue in most of the case studies analysed especially the ones with national administrative dimension where metadata management is voluntary rather than legal requirement.
- It is important to keep track of new extensions meaning i.e. well thought versioning process and manage links of the common models i.e. that updates or new extensions would be properly aligned with the existing model. For example INSPIRE and KOOP apply this in practice for their common models.
- The alignment of structural metadata managed locally and structural metadata managed in the central authority. There is an intention to manage structural metadata while at the same time managing links to local extensions and local models e.g. the Netherlands KOOP case.

Governance structure

Roles. We have observed a generic four level governance structure.

- **Legal level** establishing the legal environment, e.g. involving parliament passing laws as with INSPIRE and CISE (not always there);
- **Strategic level** on which an Authority sets the high-level goals (often Ministries or bodies like IMSC);
- **Functional level** where decisions are taken on (changes to) models and vocabularies, such as OWMS User Group, IMMC);
- **Operational level** where the agreed changes are implemented such as MRT, JRC, KOOP, CISE etc.

In overall, metadata governance is usually more complex within EU Institutions and spans across all four levels, while in the Member States the differentiation between the legal and the strategic levels is not always clear.

Metadata governance in the EU institutions

The governance structure within the EU Institutions is usually inter-institutional. It covers all four layers defined above and engages implementation groups or a steering committee. For example, the INSPIRE MIG is responsible for strategy related to the implementation of INSPIRE and also has a role in the governance of structural metadata. In the case of IMMC/MDR there is the IMSC that takes necessary strategic decisions, while Eurostat is part of ESS which mandate includes also the governance of metadata. The lowest layer of metadata governance consists of committees and working groups responsible for collection of proposals, evaluation and approval of the changes and implementation of the approved changes (e.g. in case of Eurostat - working groups, in case of IMMC/MDR – IMMC and MRT, in case of the INSPIRE – metadata working group).

Metadata governance in the Member States

The metadata governance structures of the analysed case studies in the Member cases comprise three levels. At the top of the structure (usually a combination of the legal and the strategic levels) there are one or two institutions responsible for steering and final decision making as organisational and legal aspects. Usually the institution responsible is a Ministry e.g. in the case of KOOP there is the well Ministry of the Interior. There can also be an institution reporting to Ministry with some designated governance responsibilities in the structure e.g. in case of LSIP there is NLS.

Then, at the functional level there are operational units, e.g. in the case of LSIP there is 'GIS-centras' responsible for development of structural metadata and the lifecycle of metadata, in the KOOP case provisions of operations service is done by KOOP. There might be additional bodies dealing with change requests and proposals in the structure such as the OWMS Community in case of KOOP.

It is important to note that while the roles comprising the governance structure are responsible for the overall success of the end result, we can argue that governance is most important in the design and development phase of structural metadata and also when change requests have to be processed. In the remaining phases of the lifecycle, part of the day-to-day operations – well-structured and documented management processes are the key.

Openness. Summarising our findings on the openness of structural metadata governance and management in EU Institutions and Member States, we can conclude that participation in the process is rarely completely open to the public and is usually restricted to the key stakeholders. We have observed two ways of stakeholder involvement, either the person/organisation that wants to participate needs to ask for an invitation/permission by the authority responsible for metadata governance, or the governance body invites the people/organisations that they believe that should/have a stake in participating. For example in the case of KoSIT any stakeholder from the government structure can be involved as long as they ask for it and on condition that they are related to IT or a related subject matter area, whereas in the case of IMMC/MDR the key stakeholders are invited.

Decision making process. The decision making process can be distributed through three different inter-organisational levels with certain responsibilities:

- Steering body or so called 'authority' that decides on vision and strategy. In case of Member States cases it is usually the Ministry (e.g. The Ministry of the Interior in case of KOOP, the Ministry of Agriculture in case of LSIP) and in cases of EU institutions it is usually a central organisation (e.g. IMSC in the case of INSPIRE);
- Organisational tasks are usually assigned to the institution reporting to the Steering body (e.g. NLS in case of LSIP);
- Collection of change requests and implementation of decisions is usually assigned to communities and working groups (i.e. OWMS community in case of KOOP). They also usually run operations and are responsible for the implementation of changes (e.g. MRT in case of IMMC/MDR).

The proposal for change collection mechanism (i.e. collecting change requests) varies per case, but usually aims at one of the following:

- Collection of proposals coming from stakeholders (agencies represented in the "governing body" (e.g. IMMC. MIG, LSIP, KoSIT, etc.);
- Collection of proposals coming from the community (e.g. KOOP, LG Inform, etc.).

Context.

The analysed case studies are positioned in an **inter-organisational** context, since all cases analysed have a scope where more than one (semi-)independent organisations share and re-use structural metadata that is centrally maintained. The results of the survey show that in some cases intra-organisational dimension exist, although the majority of the respondents referred to inter-organisational dimension.

The **administrative dimension** depends on the case study, i.e. EU institution versus Member State cases:

- There are three case studies under **Pan-European administrative dimension**: IMMC/MDR, INSPIRE and Eurostat;
- The remaining five case studies are under **National administrative dimension:** CISE, KOOP, LSIP, LG Inform and KoSIT.

The categorisation of the cases into the two levels of the administrative dimension is not absolute. The national solutions can be promoted for adoption in other countries, and on the other hand European initiatives can be applied to the national, regional or local level. KoSIT and LG Inform cases (or some parts of them) are being considered for adoption in other Member States. KoSIT XOV framework is being considered to be adopted outside Germany, although, some mandatory requirements (i.e. use of German language) create barriers for cross-border adoption.

Policy domain

Most of the case studies are **not restricted** to any policy domain and have a cross-domain approach. However, **three cases are restricted** to specific domains i.e. INSPIRE and LISP are restricted to the environment domain, whereas Eurostat is restricted to the statistics domain. In this context, ISA Action 1.17 "Re-usable INSPIRE Reference Platform[17]" is trying to identify and implement missing components from the implementation of the INSPIRE Directive that can be re-used to other domains as well. The Re3gistry, a cross-domain re-usable data model for metadata management, is an example of such a component.

Enforcement policy

The enforcement policy depends on the domain and the country. The way it is implemented varies depending on the purpose and the context of use of the metadata (e.g. provision of statistics, geographic information, provision of information about councils' performance). In the case of INSPIRE (pan-European dimension) and LSIP (national dimension) the INSPIRE Directive defines the rules that should be followed. In the case of LG Inform (national dimension) there is no enforcement policy and sharing as well as reuse is done on a voluntary basis. The results of the online survey indicate that 75% of respondents share and reuse structural metadata, especially at the international level.

- **Sharing:**
  - **Legal requirement**: CISE, INSPIRE, LSIP, KoSIT, Eurostat.
  - **Voluntary**: KOOP, IMMC/MDR, LG Inform.
- **Reuse**:
  - **Legal requirement**: Eurostat, INSPIRE.
  - **Comply-or-explain**: CISE, KOOP, LSIP, KoSIT
  - **Voluntary**: IMMC/MDR, LG Inform.

[17] ISA Action 1.17 : http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-17action_en.htm

There is trade-off between stability and flexibility when it comes to choosing the enforcement policy for structural metadata governance. Having enforcement on the basis of legal requirement caters for stability since the use of the structural metadata is obligatory and changes can only be made by amending the legal environment. On the other hand, not having strict legal requirements gives more flexibility and provides conditions for faster adaptation to changing requirements to the upcoming needs and also for simplified change process that requires fewer resources.

Authoritative source

There is an authoritative source (e.g. a repository or a file server, on which the metadata is housed) for structural metadata in all eight case studies. In most cases the authoritative source is a tool (registry) that is accessible online where structural metadata is publicly available, e.g. in the case of KoSIT code lists are publicly accessible and some reachable by registered users only. Eurostat allows to view or edit artefacts (such as: concept schemes, code lists, data structure definitions, etc.) for registered users only. In some cases the organisations do indexing and listing of external metadata that exist and is used. In the case of KOOP, registry of registries is being considered for linking vocabularies and value lists hosted elsewhere in the near future.

Licensing framework

In all case studies licensing is not clearly identified on authoritative sources of information i.e. models or reference data. In the case of IMMC/DMR and INSPIRE they refer to the general European Commission copyright statement. Also KOOP states that they do have a licence in place in a document describing their framework, although the licence does not appear in the documentation of every single resource. EUPL is applied in some cases for the case of open source tools (e.g. Eurostat). Others are currently considering possible licensing options (e.g. KoSIT). In case of Eurostat where SDMX registry is used, there is also SDMX licence applied. However, in general insufficient attention to licensing can be noticed in the majority of analysed cases. To summarise, we can state that EU Institutions seem to be more aware of licensing than the Member States.

Quality controls

The quality control process is usually manual and ad-hoc (e.g. KOOP, CISE). In-house developed scripts or tools are being used (e.g. INSPIRE validator) and in most cases the process itself is not clearly defined. Thus, quality control process is not formalised and is ad-hoc rather than based on some well-known standards, such as ISO 9000 series [18].

[18] ISO 9000 series: http://www.iso.org/iso/iso_9000

Metadata Schema

Vocabulary. A set of descriptors are used for the structural metadata in each case study reviewed. These vocabularies are usually very simple limited to three or four descriptors. In almost all cases ADMS is either being considered internally for the metadata descriptions (e.g. KOOP), or ADMS exports are available, e.g. in the case of the Publications Office and Eurostat.

Identifiers scheme. There are identifiers schemes in place at Eurostat, CISE, LSIP, other analysed institutions are currently considering establishment of common guidelines for it (e.g. IMMC/MDR, INSPIRE).

Schema documentation. All of the cases have documentation of their metadata in place. INSPIRE has guidelines and xml dictionary, KoSIT generates schema documentation directly from UML model.

Multilingualism. Documentation of all pan-European initiatives is grounded on an authoritative language i.e. English, and labels (on a few cases together with other parts of structural metadata e.g. LSIP) are translated into different languages European initiatives do support other languages (in terms of tools functionality). The tools used by LSIP are capable of German, French and Russian languages as an addition to Lithuanian and English. In the case of KoSIT the law requires to have German as mandatory language although it is possible to describe some structural metadata in English. The solutions of national initiatives such as CISE or KoSIT, are harder to reuse due to their limited support of different languages.

5.2. Metadata management

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

- Design and Create *(out of scope of our work)*;
- Documentation;
- Maintain and Update; and
- Share and Reuse.

Update

It entails the processes of updating the metadata, and deleting/deprecating outdated versions.

Update frequency. In general the update frequency of structural metadata varies case by case and depends on the type of metadata (e.g. reference data is updated more often than data models). The update frequency varies between one to three months to one to three years in some cases. For example in the case of LG Inform metadata is being reviewed once a year.

Change management process. A generic change management process is in place and consist of change prioritisation as well as communications management in some cases (e.g. KOOP, IMMC/MDR) while in other cases it is being discussed and

hasn't been implemented yet (e.g. CISE, INSPIRE). Changes are being discussed at different governance levels depending on the possible impact of change.

Version control. Version control is established in all of the case studies, although it is done in different ways. For example KOOP uses three-level version numbering approach X.Y.Z where so far only X and Y versions have been published. In the case of IMMC/MDR date stamps are used for version as well as sequence numbering.

Harmonisation

Most organisations do create mappings for their reference data, especially to authoritative code lists. For example the Publications Office is creating mappings of their Named Authority Lists to related code lists published by the Library of Congress. Mappings between data models are not pursued. Organisations that do not do yet create mappings for their metadata element sets (analysed in the Table 5 below) said that they are considering this in their future plans (e.g. INSPIRE is considering integration of code lists).

Table 5 – Overview of creation of mappings and assessment of alternative metadata sets.

Case Study	Create mappings	Assessment of alternatives
Eurostat	Yes, they do this.	Yes, they do this.
Inspire	Not in place, is being considered.	Not in place, they do not need that.
IMMC/MDR	Yes, they do this.	Yes, they do this.
KoSIT	The process is not defined.	Yes, they do this.
CISE	Not in place, is being considered for the near future.	Not in place, is being considered for the near future.
LSIP	Yes, they do this.	Yes, they do this.
KOOP	Not in place, is being considered.	Not in place, is being considered.
LG Inform	Yes, they do this.	Not in place.

Documentation

It covers the processes that facilitate the publication and retrieval of metadata as well as tool support for these processes. Structural metadata is documented in all eight case studies, but the documentation processes themselves are not formal.

Publication of metadata. There is usually an authoritative source where metadata is published (see also section 5.1). Only in few cases like Eurostat or INSPIRE a common vocabulary is used for publication. Searching inside the structural metadata is not always possible. Hence, structural metadata is usually published as black boxes. Few exceptions exist however. In the case of CISE search is unavailable, but it is possible to navigate through the definitions of data models. INSPIRE and LSIP have search capabilities inside their structural metadata. INSPIRE in particular has an elaborate data dictionary available online. Others are considering this functionality to be implemented in the future.

Retrieval of metadata. In all of eight case studies the structural metadata can be retrieved in human- and machine-readable formats. The structural metadata can be downloaded manually (e.g. KOOP) and/or automatically (e.g. KoSIT, LSIP). In the case of KOOP a complete dump of the schema is also available, while INSPIRE allows to retrieve data using content negotiation.

Supported formats of publication tool(s). In the majority of the case studies human- and machine-readable distributions of the structural metadata are published separately. Only in some of them, both human- and machine-readable formats are available in one URL (e.g. Inspire HTML and three machine readable formats (xml, JSON, Atom)).

Supported formats are usually of three types i.e. human-readable only, machine-readable only or human- and machine-readable. The most popular human-readable formats are pdf and html for text, while the most popular human- and machine-readable formats are XML, CSV, OWL, etc. In the case of LSIP published metadata is available in HTML format as well as in XML, CSV formats. In the case of KOOP metadata schema files are published in XSD format while value lists are available as XML and SKOS. IMMC/MDR use XSD for core metadata and transmission protocol schemas and SKOS, XML for authority tables.

Tool support

Publishing and retrieval of human-, machine- and human- and machine- readable formats are usually supported by various tools that can be domain specific. For example an adopted version of ESRI Geoportal is used as well as Ad-hoc solutions in the case of LSIP, whereas Eurostat tools are being developed in house. Tools that are used for design and development of structural metadata (e.g. editing of XML, creation of UML) are different in each organisation and are in-house developed, developed by third parties or bought of the shelf and adopted to the needs of organisation. Tools used in each of the case studies are analysed in the tools section below.

Standards

We can see from the case studies that standards can be applied to two areas:

- Standards for metadata management; and
- Standards for documentation and representation.

In many of the cases, as shown also in Table 6 below, ISO standards for metadata management, such as ISO/IEC 11179, are put in practice , while for documentation and representation SKOS and/or ADMS are being used at INSPIRE, IIMC/MDR, CISE, LG Inform and KOOP. Additionally, the respondents of the online survey indicated that a variety of documentation and representation standards such as ADMS, SKOS are already used, or will be used in the near future. Table 6 below summarises the standards used in each of the case study.

Table 6 – Standards used in the analysed case studies.

Case Study	Metadata management standard(s)	Description
Eurostat	ISO 17369[19]	
INSPIRE	ISO 19135[20] is used for registration process.	Code list are implemented in SKOS. ADMS is used for metadata descriptions.
IMMC/MDR	ISO/IEC 11179[21] is used in some restricted way.	Source format of authority tables is XML with subset exported to SKOS for CELLAR. Publication is done using ADMS.
KoSIT	None.	AMDS is used for publishing metadata provided by the xRepository.
CISE	None.	SKOS might be considered with XSD as the publication standard.
LSIP	ISO 19115[22], ISO 19119[23], ISO 19139[24]	N/A
KOOP	Management is based on BOMOS approach.	Value list are set up using SKOS, RDF/OWL.
LG Inform	None.	Vocabularies conform to the SKOS standard and are available via Joinup.

Costs

In almost all cases there are between 1 and 2 FTEs responsible for metadata management, although some have more than 2 FTEs. For instance, IMMC/MDR has 5 FTEs but this covers also the effort required for operating the governance structure. LG Inform has approximately 4 FTEs dedicated to metadata management.

Benefits

In all case studies it is emphasized that benefits offset the costs however, there is no quantitative evidence to support this. The key benefits common for all of the cases analysed include:

- **Better interoperability:** processes based on well-known standards and methodologies increase the level of accessibility, understanding, sharing, and reusability of instance metadata;
- **Better reuse of metadata:** application of same standards for documentation and representation of metadata increases its reusability.
- **More efficient use of resources:** when there is a clearly defined process based on well-known methodology, less resources have to be involved in the management process;
- **Access to information (metadata, data):** application of tools with the

[19] ISO 17369 : http://www.iso.org/iso/catalogue_detail.htm?csnumber=52500

[20] ISO19135: http://www.iso.org/iso/catalogue_detail.htm?csnumber=32553
[21] ISO/IEC 11179: http://metadata-standards.org/11179/
[22] ISO 19115: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020
[23] ISO 19119: http://www.iso.org/iso/catalogue_detail.htm?csnumber=39890
[24] ISO 19139: http://www.iso.org/iso/catalogue_detail.htm?csnumber=32557

user oriented front-end increases the accessibility to metadata;
- **Standardised and centralised metadata management tools:** unified tools allow smoother and less expensive access to it as well as adoption to the future needs;
- **Improvement in metadata/data quality:** well defined processes and application of quality management practices result in the increase of the metadata/data quality. This was also indicated as the main benefit of metadata management by the online survey respondents.

It is important to stress out that all possible benefits are interrelated and relevant for the majority of institutions regardless the country or dimension. This was also noted by the survey respondents.

5.3. Tools

Structural metadata governance and management is supported by software tools. Such tools implement the metadata governance principles and the metadata management processes.

The tools used can be grouped in the following categories depending on their intended use:

- Tools for creating, editing, processing and visualising structural metadata;
- Tools for storing, management and accessing structural metadata;
- Tools for searching, reporting and publication structural metadata;
- Tools for versioning structural metadata; and
- Tools for supporting the governance of structural metadata.

In terms of development models, tools could be categorized as commercial products or products based on the specific needs and requirements of the organization. Commercial products among others include TopBraid Composer, Maestro edition from TopQuadrant used by KOOP, XMLSpy developed by Altova. One of the examples of ad hoc products is esd-toolkit developed by Porism Ltd. We discuss the products in more detail in the remainder of this section.

Reusability

Due to the fact that the majority of the tools are developed based on the individual needs of each organisation, reusability is not common and does require adaptation. Thus, it is rarely considered in practice, apart from domain-specific cases, e.g. tools used by Eurostat can easily be reused by other National Statistical institutions. We list the identified tools and relevant information related to them in Table 7.

Table 7 – Tools for structural metadata management

Tool	Key functionality	Type of licence	Owner / vendor
Creating, editing, processing, visualisation			
XSLT/PERL scripts (IMMC/MDR)	Designed for file conversion, generation of different files.	N/A	N/A
Java Libraries (Apache XMLBeans, Apache Xerces, Xalan, Saxon) (LSIP)	These tools are used for compiling XML schemes to Java objects, XSLT transformation, XSTL and XQuery processing.	All Java Libraries and some parts of ESRI Geoportal toolkit is based on open source although it is a commercial product.	Developed by Apache Software Foundation and Saxonica.
Xgenerator (KoSIT)	It is the central tool for production. It also validates individual artefacts according by rules defined in the framework.	There is no specific licence.	Developed in-house and by external suppliers.
Genericoder (KoSIT)	This tool is used for the creation of code lists and similar tasks.		
Production environment components (KoSIT)	They are used for the technical production of the standards, documentation and artefacts related to the standards (e.g. to produce the DocBook chapters which are directly generated from the model to XML).	These tools are mainly developed under open source technologies.	N/A
XSLT and PERL scripts (CISE)	This tool is used for harvesting and processing.		
XMLSpy (INSPIRE)	This tool provides XML development environment for modelling, editing, transforming, and debugging XML technologies. It also provides XML editor and graphical schema designer, code gen, file converters, debuggers, profilers, database integration, chart creation, support for XSLT, XPath, XQuery, WSDL, SOAP, XBRL, JSON and Open XML (OOXML), plus Visual Studio and Eclipse plug-ins among others.	Commercial product.	Developed by Altova.
Shapechange (INSPIRE)	Java tool that takes application schemas constructed according to ISO 19109 from a UML model and derives implementation representations.	Available under the GNU General Public Licence.	Tool was originally written by interactive instruments GmbH, and expanded by The MITRE Corporation.

Enterprise Architect (INSPIRE)	This tool supports high performance modelling, visualization and design platform based on the UML 2.4.1 standard.	Commercial product.	Developed by SparxSystem.
ESRI ArcIMS (LSIP)	This tool allows LSIP to efficiently serve maps over the Internet webpage.	All Java Libraries and some parts of ESRI Geoportal toolkit is based on open source although it is a commercial product.	Tools and the platform are created by ESRI.
TopBraid Composer, Maestro Edition (KOOP)	Develops RDF/S and OWL ontologies, SPARQL queries and Semantic Web rules based on standards; Contains visualization and diagramming tools including visual construction of queries and auto-generation of SPARQL; Usability, extensibility and robustness of its underlying technologies – Eclipse and Jena; Seamless round-tripping between XML and RDF/OWL (import – export); SPARQL-based HTML and XML document generation using built-in JSP engine; Ability to convert Emails into OWL, supporting semantic analysis and classification of emails; Supports rapid iterative construction and evolution of semantic web applications.	Commercial product.	The tool is developed by TopQuadrant.
OpenRDF Sesame (KOOP)	De-facto standard framework for processing RDF data. This includes parsers, storage solutions (RDF databases a.k.a. triplestores), reasoning and querying, using the SPARQL query language. It offers a flexible and easy way to use Java API that can be connected to all leading RDF storage solutions.	Open source.	The tool is developed by Sesame.
Storing, management, accessing			

GENIS Reference Data Component (GENIS RDC)	In the context of the Generic Interoperable Notification Services (GENIS) project, funded under Action 1.11 of the ISA programme, a GENIS Reference Data Component (GENIS RDC) was built. The Reference Data Component can be used for the management of changes to reference data and the deployment of reference data as a service available to other information systems. It has the following features: • Import reference data from a file and detect changes with previous versions; • Create, read, update, delete reference data using the Web-based graphical user interface. The RDC supports versioning of concept schemes and concepts and multilingual labels; • Export reference data to a file; and • Deploy reference data as a service to other information systems.	Open source (to be released)	The tool is developed by the European Commission.
SDMX Registry (Eurostat)	This tool provides information about the following: what and where data sets and metadata sets are available; how often they are updated, what their contents are, how they can be accessed; what is data sets structure. It allows applications to sign up (or subscribe) for notifications. When a data set or metadata set of interest becomes available, the application will be automatically alerted.	Open source, under EUPL.	Developed in-house.
The Re3gistry (INSPIRE)	This tool is used to manage the data and metadata contained in the INSPIRE Registry. It is made up of an import function (it loads the data provided (in .csv format) into the database), an export function (it prepares the data saved in the database in order to be ready for the web service) and a web service (it has a RESTful interface to access the data). It is possible to access different formats and languages using content negotiation or directly calling the desired file.	Open source (EUPL).	Developed in-house with support from ISA.
ESRI Geoportal extension (LSIP)	This suite of modules provides the capability to build and custom-configure Geoportals to meet each entity's particular style, needs, and use objectives.	All Java Libraries and some parts of ESRI Geoportal toolkit is	Tools and the platform are created by ESRI.

		based on open source although it is a commercial product.	
LG Inform (based on esd-toolkit) (LG inform)	• Develops reporting tools that compare councils with one another; • Identifies councils in each region of the country and how each council area can be broken down for local communities; • Summarises and compares metrics across (pre-defined and user-defined) 'comparison groups' of areas; • Creates reports for sharing as HTML, PDF or MS Word documents or for embedding in web sites; • Allows municipalities and areas within them to be compared via data on 1,800+ metrics.	This tool is distributed under the terms of the General Public Licence.	These tools are developed in-house and built on a variety of languages, including C# / .NET, Java and PHP. Some features are developed by Drupal.
Searching, reporting, publication			
RAMON - Eurostat's metadata server (Eurostat)	• There is possibility to present information in all languages in which it exists. • It is a powerful search engine. When a search is performed, the search engine of RAMON goes through all the objects loaded in the database and returns all the results found, whatever the metadata category in which it was found. • Layout is adapted to the type of metadata presented (information displayed on the screen depends on the kind of metadata - narrow terms, broad terms, etc.).	Open source, under EUPL.	Developed in-house.
XRepository (KoSIT)	It is infrastructure component for publication, documentation and similar functions.	The standard for the structure of the repository is open source but the product itself (the publishing platform) is in proprietary development.	Developed in-house and by external suppliers.
Semantic Asset Manager (CISE)	This tool is used for publication and retrieval.	N/A	Developed in-house.

Drupal (KOOP)	This tool is used for content management.	Open source, distributed under the terms of the General Public Licence.	This tool was maintained and developed by a community of 630,000+ users and developers.
Versioning			
Apache Subversion (INSPIRE)	Version control.	Open source for re-use available under the Apache Licence (version 2.0).	Developed by Apache Software Foundation.
Apache Subversion (SVN) (IMMC/MDR)	This tool is designed for versioning of assets maintained in MDR.	Open source for re-use available under the Apache License (version 2.0).	Developed by Apache Software Foundation.
Governance			
Atlassian JIRA (IMMC/MDR)	Governance workflow.	This tool is a commercial product.	Developed by Atlasssian.

6. IDENTIFIED GOOD PRACTICES

After having analysed eight selected case studies, we elicit requirements for metadata governance, management and tools. These requirements should be perceived as observations for the organisational structure for metadata governance and management, also as guidelines for selection of the appropriate tools. These are understood as a good practice that is identified in the analysed cases and the outcomes of the online survey and can be supported by examples. Good practices are valuable for the development and improvement of existing metadata management and governance practices. They can be also used to support the definition or improvement of the metadata governance and management in EU Institutions and Member states.

1. **In a good governance structure, the roles concerning legislation, strategy, functionality and operations are clearly distinguished and assigned to designated bodies.**

   It is a good practice to separate metadata governance roles. The proposed structure could comprise of the authorities responsible for strategy setting and legislation (shall the metadata be enforced by law), the governance body(-ies) responsible for taking decisions on structural metadata and a number of operational agencies responsible for implementing these decisions. Such a governance structure is already in place in various cases, such as INSPIRE, KOOP, LG Inform and LSIP.

2. **The involvement of direct stakeholders in the metadata governance process ensures that the interests of the stakeholders are taken into account which maximises buy-in and take-up.**

   The goals, the expected benefits and the importance of metadata management have to be made clear to all direct stakeholders, including the management of the organisation, in order to ensure commitment and support. This was observed in the 8 selected case studies, but also came as input from the respondents of the online survey. Hence, it is a good practice to raise awareness and involve representative of the direct stakeholders of all four levels – legal, strategic, functional and operational - (e.g. those responsible for the management of metadata, also the ones that are known as users) in the governance process. This will allow consensus on various aspects to be built early, already in the design and development phase, and will ensure that an optimal, commonly-agreed direction is followed in the later stages.

3. **Voluntary sharing and re-use works best if stakeholders are aware of the advantages of collaboration and of the benefits for interoperability.**

   There is trade-off between stability and flexibility when it comes to choosing the enforcement policy for structural metadata governance. Having enforcement on the basis of legal requirement caters for stability since the use of the structural metadata is obligatory and changes can only be made by amending the legal

environment. In such case amendments would be judicious and rare, as the changes in law are time and resource consuming.

On the other hand, not having strict legal requirements gives more flexibility and provides conditions for faster adaptation to changing requirements to the upcoming needs and also for simplified change process that requires fewer resources. Voluntary enforcement requires that the stakeholders would be made aware of the advantages of sharing or reuse of metadata. In this case clear arguments aligned with the case specifics should be formed to encourage sharing and/or reuse.

4. **Application of a standard for metadata management creates a well-structured management environment based on existing good practice.**

   The use of standards is recommended. This was observed in the 8 selected case studies, but also came as input from the respondents of the online survey. The selection of the standard to be used depends on the domain and the particularities of the structural metadata.

   We defined two families of standards:
   - Standards for metadata management. Any of the following standards can be considered for reuse in this case: ISO/IEC 11179, ISO 19135, ISO 19115 or some other.
   - Standards for documentation and representation, such as SMDX, ADMS and SKOS.

5. **Good change management processes are based in stability where possible without sacrificing flexibility where needed and take into account an alignment between the life cycles of structural metadata development and software development.**

   A change management process needs a balanced approach concerning flexibility i.e. on the one hand, maintaining a consistent implementation towards the goal, and on the other hand, being flexible i.e. allowing specifications to evolve with changing organisational and technological environment. It is necessary to synchronise the life cycle of structural metadata development and that of software development. Thus, coherence of the legislation, structural metadata development life cycles, metadata strategy and organisational as well as technological capabilities is necessary.

   We also observed that version control is also always in place. A formal version control approach, with clear naming rules for structural metadata elements and documents, aligned with the change management process allows for changes to be tracked internally and changes in versions are visible and understandable by the external users of structural metadata.

6. **Changes in structural metadata are well planned and tracked, preserving backward compatibility as much as possible; in cases where**

disruptive changes are unavoidable, these changes should be planned and communicated well in advance.

Changes in structural metadata should be judicial and well-planned. A change request should be issued only when the metadata cannot support an application scenario, because of a new requirement. For example, a new Member State is entering the EU, hence the Countries Named Authority List has to be updated.
It is necessary to make a distinction among processes and publication frequencies for changes in data models, which are usually more stable and changes there may also impact the operation of production systems, e.g. services and databases, and changes in code lists, which are more volatile and easier to update.

Disruptive changes, if necessary, need to be well-planned and announced not to negatively affect day-to–day operations. Backwards compatibility of changes should be an important consideration.

7. **Structural metadata is managed in formats that are appropriate for the type of use. Metadata describing the structural metadata is expressed or exported using the Asset Description Metadata Schema.**

   Structural metadata is managed in formats that are appropriate for the type of use, for example XML Schema Documents for data models and SKOS for reference data.

   ADMS is also used for exchanging the description of structural metadata between different systems and organisations. This way, structural metadata can be made accessible via different portals and infrastructures, hence improving its visibility and potentially its reuse. For example, INSPIRE, Eurostat, IMMC and CISE are already sharing the ADMS descriptions of the structural metadata with the European Federated Interoperability Repository on Joinup.

   It is also possible to use ADMS as the native metadata description vocabulary. In the analysed cases some of those who have not applied ADMS yet are considering doing so in the near future, e.g. KOOP.

8. **Standard reference data is used wherever appropriate; if locally defined reference data is used, this is mapped to standard reference data to enable wider interoperability.**

   The creation of mappings is not a common practice among the analysed case studies. However, whenever mappings between structural metadata element sets are created, such as in the case of the IMMC, these mappings are usually limited to exact match mappings between synonymous classes. Partial mappings are also be explored, but their precision and accuracy are usually questionable.

9. **Structural metadata is distributed in machine-readable formats that can be processed by the tools available by users.**

Case studies demonstrate that XSD could be used for publication of metadata schemas as well as transmission of protocol schemas. Therefore, it is suggested to manage data models as XSD with appropriate tools that fit into local infrastructure.

10. Content negotiation is used to manage and provide different types of formats from the same URI.

It is important that both humans and machines can make sense of the meaning of structural metadata. Hence, documentation of structural metadata should be available in both human- and machine-readable formats.

In the case of structural metadata served on the Web, content negotiation is used for delivering both human-readable and machine-readable descriptions of the metadata from the same URI, given that multiple formats are available. This simplifies the access to the resources and the related information, and broadens also the possibilities for reuse, as different distribution formats are supported, e.g. XML and RDF.

11. Metadata governance and management ensures the sustainability of structural metadata.

The implementation of metadata management and governance ensures the sustainability of the metadata. The pioneers in the field, like the organisations interviewed in the context of this survey, understand this. Defining clear roles, responsibilities and processes enables the coordinated development, evolution, use and reuse of the metadata. It also makes the value proposition of well-managed structural metadata clear to all stakeholders involved, particularly then to the senior management of an organisation who decides on the allocation of budget for such tasks.

7. RECOMMENDATIONS

This chapter introduces recommendations regarding structural metadata governance and management, and reusable tools.

7.1. Recommendations for structural metadata governance and management

In this section, we form recommendations in order to better organise day-to-day operations related to metadata governance and management. These recommendations derive from our experience based on the interviews and online survey results, but entail also a certain degree of subjectivity.

1. **Legislation should be formulated on a sufficiently high level and should not specify details like the values in a code list or the elements of a data model; these details should be specified as part of the implementation and made available from an authoritative source to which the legislation can refer.**

   In case of enforcement by law, it is recommended to mention a link in the law referring to the authoritative source of structural metadata, rather than citing the structural metadata itself in the law. This way amendments in the structural metadata would be more flexible as the law itself would not have to be changed.

2. **The structural metadata management processes should be documented.**

   In inter-organisational contexts, where different stakeholders – often from different countries – participate in the governance and management of structural metadata, the need for clear and complete documentation of practices, process, principles, roles and responsibilities is highly recommended. Having stable and transparent processes is expected to improve the efficiency of the process itself, and to set a common ground for operating metadata management, taking decisions, and resolving conflicts.

   In order to foster reuse, the documentation of both the management process and the structural metadata itself should be made available in different languages.

3. **Owners of structural metadata should be made aware of the importance of clear licensing arrangements that specify unambiguously under which conditions the metadata can be reused. Open reusable metadata is recommended.**

   Owners of structural metadata need to be made aware of licensing options and requirements and the importance of licensing their structural metadata. A clear license is considered as an enabler for sharing and reusing structural metadata. If no licence is in place, users will hesitate to use the metadata, since it is effectively not clear to them if they can do so and under which conditions. **Open licences for structural are recommended** with protection against

misrepresentation. Our findings show that, especially in the case studies of the EU Institutions, EUPL and the EU legal notice are usually applied, but it is not always explicitly publicly visible. The ISA Open Metadata Licence is another option.

4. **Stakeholders should be aware of the expected benefits of metadata sharing and reuse.**

   In order to gain management support and encourage the sharing and reuse of metadata, the direct stakeholders should be made aware of the expected benefits. In section 5.1, we observed that there is an agreement both in findings of the case studies and the online survey on the expected benefits of structural metadata governance and management.

   Sharing and reuse of metadata contributes to the realisation of the expected benefits, particularly the ones related to interoperability. Especially in cases when enforcement is not mandatory, stakeholders should be made aware of the advantages of sharing or reusing of metadata, and should be persuaded to do so. In this case clear arguments supported practical evidence will be required.

5. **Management processes and publication frequencies should be different for changes to data models on one hand, and reference data on the other hand.**

   Changes in structural metadata should be justified and well-planned. A change request should be issued only when the metadata cannot support an application scenario, because of a new requirement. For example, a new Member State is entering the EU, hence the Countries Named Authority List has to be updated.

   It is necessary to make a distinction among processes and publication frequencies for changes in data models, which are usually more stable and changes there may also impact the operation of production systems, e.g. services and databases, and changes in code lists, which are more volatile and easier to update.

   Hence, changes in the structural metadata should also take into account the impact that these may have in running software that is using this metadata. In these cases, the structural metadata lifecycle – and its change management process - should be aligned with the software development lifecycle.

   Disruptive changes, if necessary, need to be well-planned and announced not to negatively affect day-to–day operations. Backwards compatibility of changes should be an important consideration.

6. **Structural metadata should have persistent unique identifiers.**

   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, e.g. see ISA's 10 Rules for Persistent URIs[25].

7.2. Recommendations for the reuse of tools

We observed that **no single tool exists that can cater for all different requirements of metadata governance and management**. Hence, a collection of different tools needs to be deployed. **It is important that these tools are based on open standards**, especially in terms of interfaces and formalisms used for managing and representing structural metadata, and are **interoperable**.

In terms of tools used for metadata management and governance there are three groups to mention:

- General purpose commercial tools that can be easily reused (given that the organisation can afford them);
- General purpose open source tools developed by third parties that can be easily reused;
- Open source built in-house tools that would require adoption and configuration to cater for new needs and requirements.

Software available on the market, either under an open source or a commercial licence, is easily reusable, while custom-made software addressing the particular needs of an organisation will require more efforts for development and adaptation. Therefore, cost-benefit analysis should be considered on the basis of the specific case to define the feasibility for reuse.

We present the list of tools that could be reused in other metadata governance and management practices in Table 8 below by estimating the effort required for reuse (in the reusability column).

The toolset that will be deployed by an organisation in order to support metadata management and governance should, as a minimum, cater for the following functionalities:

- Provision of an authoritative source for the structural metadata, which allows for storing the metadata and its documentation, and offers search, visualisation and browsing functionalities;
- Issue tracking and ticket management to manage metadata changes;
- Version control;
- Maintenance and linking to support metadata harmonisation.

It is also important to mention that since XML is the prevailing technology, a good editor is essential; however there are a number various open source and commercial products for this that can be considered for implementation. This should be supported also by XML schema maintenance and validation tools.

[25] https://joinup.ec.europa.eu/community/semic/document/10-rules-persistent-uris

Table 8 – List of reusable tools

Tool	Key functionality	Type of licence	Owner/ Vendor	Reusability
Commercial				
Enterprise Architect (INSPIRE)	This tool supports high performance modelling, visualization and design platform based on the UML 2.4.1 standard.	Commercial product.	Developed by SparxSystems.	Possible to reuse due to the fact that it is commercial product if organisation can afford them.
Atlassian JIRA (IMMC / MDR)	Governance workflow.	Commercial product.	Developed by Atlasssian.	Possible to reuse.
ESRI Geoportal extension (LSIP)	This suite of modules provide the capability to build and custom-configure Geoportals to meet each entity's particular style, needs, and use objectives	All Java Libraries and some parts of ESRI Geoportal toolkit are based on open source although it is a commercial product.	Tools and the platform are created by ESRI.	Possible to reuse.
ESRI ArcIMS (LSIP)	This tool allows LSIP to efficiently serve maps over the Internet webpage			Possible to reuse.
TopBraid Composer, Maestro Edition (KOOP)	• Develops RDF/S and OWL ontologies, SPARQL queries and Semantic Web rules based on standards; • Contains visualization and diagramming tools including visual construction of queries and auto-generation of SPARQL; • Usability, extensibility and robustness of its underlying technologies – Eclipse and Jena; • Seamless round-tripping between XML and RDF/OWL (import – export); • SPARQL-based HTML and XML document generation using built-in JSP engine ;	Commercial product. Also free edition with limited functionality exist.	Developed by TopQuadrant.	Possible to reuse.

Tool	Key functionality	Type of licence	Owner/ Vendor	Reusability
	• Ability to convert Emails into OWL, supporting semantic analysis and classification of emails ; • Supports rapid iterative construction and evolution of semantic web applications.			
Open source developed by the third parties				
Shapechange (INSPIRE)	Java tool that takes application schemas constructed according to ISO 19109 from a UML model and derives implementation representations.	Available under the GNU General Public Licence.	Tool was originally written by interactive instruments GmbH, and expanded by The MITRE Corporation.	Possible to reuse with some adoption.
Apache Subversion (INSPIRE)	Version control.	Open source for re-use available under the Apache Licence (version 2.0).	Developed by Apache Software Foundation.	Possible to reuse.
Apache Subversion (IMMC / MDR)	This tool is designed for versioning of assets maintained in MDR.	Open source for re-use available under the Apache License (version 2.0)	Developed by Apache Software Foundation.	Possible to reuse.
Drupal (KOOP)	This tool is used for content management.	Open source, distributed under the terms of the General Public Licence	This tool was maintained and developed by a community of 630,000+ users and developers.	Possible to reuse.
OpenRDF Sesame (KOOP)	De-facto standard framework for processing RDF data. This includes parsers, storage solutions (RDF databases a.k.a. triplestores), reasoning and querying, using the SPARQL query language. It offers a	Open source.	The tool is developed by Sesame.	Possible to reuse.

Tool	Key functionality	Type of licence	Owner/ Vendor	Reusability
	flexible and easy to use Java API that can be connected to all leading RDF storage solutions			
Open source built in-house				
XRepository (KoSIT)	It is infrastructure component for publication, documentation and similar functions.	The standard for the structure of the repository is an open standard but the product itself (the publishing platform) is in proprietary development	Developed in-house and by external suppliers.	Possible to reuse.
Production environment components (KoSIT)	They are used for the technical production of the standards, documentation and artefacts related to the standards (e.g. to produce the DocBook chapters which are directly generated from the model to XML).	These tools are mainly developed under open source technologies.	Developed in-house and by external suppliers.	Possible to reuse.

8. CONCLUSIONS

8.1. The metadata governance and management landscape

In the process of conducting the desk research and interviews, we have been able to gain an understanding of the current landscape of activities around metadata governance and management.

We have seen that there are many activities underway, both on the European and the national level, which proves the importance of this task for public administrations. We found no lack of potential cases for investigation, as demonstrated by the long list of 49 identified cases listed in section 3.1, including cases from six European institutions and all Member States except Hungary. In the selection of the cases for more in-depth consideration reported in Annex II, we focused in particular on cases with an inter-organisational focus, i.e. those cases that involved cooperation across different organisations, and sought to find a balance between pan-European initiatives (undertaken by European institutions) and national activities from different parts of Europe.

A common concern for all activities studies was that the importance, the expected benefits and the relevance of metadata management and governance is often not clear to the management of the organisation. Therefore there is a lack of support from the leadership.

In general, the activities around metadata governance and management appear to be in an early phase: in most cases investigated, governance structures have been established relatively recently, and some of the management processes and procedures are still in the process of being defined and implemented. In many of the cases considered, the organisations responsible for the implementation of these processes and procedures expressed a vivid interest in sharing their experiences and learning from others.

8.2. The relevance of this deliverable and its recommendations

This deliverable fits into this landscape as a document that can help organisations around Europe to build towards stable and sustainable structural metadata. On one hand, the publication of this document is not too early in the sense that it can already build on emerging practice; on the other hand, it is not too late as many organisations still have issues and questions where this deliverable can help.

In line with the objectives listed in section 1.2, this deliverable:

a. compiles existing requirements, methodologies, practices, procedures and reusable tools for metadata governance and management in the EU Institutions and the Member States by bringing together information about the current activities in metadata governance and management in the eight cases and through the overview of responses in the online survey in chapter 4;

b. elaborates on the costs, benefits and feasibility constraints, of the different alternatives as part of the key findings in chapter 5; and

c. documents lessons learnt and observations on good practice in chapter 6 and formulating recommendations in section 7.

8.3. Further work

It is the intention of this deliverable that the recommendations given in section 6 can serve as guiding principles by organisations that are in the process of defining structures, procedures and workflows for governance and management.

This deliverable feeds into the future work. In particular, the lessons learnt and recommendations will be used in the preparation of deliverable 4.2 *Methodology and tools for Metadata Governance and Management for EU Institutions and Member States* that proposes the necessary mechanisms that should be put in place for the coordination of metadata management at the EU level. In that same report, high-level specifications for reusable metadata government and management tools are developed, based also on the recommendations of this report.

The application of the recommendations of this report in three pilots, focusing on different aspects of metadata governance and management that took place between January and April 2014, in collaboration with DG COMP, DG MARE and the JRC, and the Publications Office, allowed for testing the recommendations and enabled their further refinement towards a set of procedures and approaches that will help organisations around Europe and beyond to ensure adequate metadata management in support of efficient, effective and sustainable metadata applications.

In a more general context, given the **importance that is attributed by public administrations around Europe and beyond on the governance and management of structural metadata** (as an enabler to interoperable information exchange and cross-border/-sector public services), we expect that this topic will **rank higher on the policy agendas** of different countries. In this vein, works like this one, will play a significant role in ensuring that public administrations can learn and benefit from international good practices and recommendations, avoid re-inventing the wheel, and receive guidance in adapting reusable metadata governance and management practices and tools in their own contexts.

ACKNOWLEDGEMENTS

Specific acknowledgement is due to:

Person	Organisation
Marco Pellegrino	The Statistical Office of the European Union - Eurostat
August Götzfried	The Statistical Office of the European Union - Eurostat
Michael Lutz	Joint Research Centre of the European Commission
Andrea Perego	Joint Research Centre of the European Commission
Roberto Sgnaolin	Joint Research Centre of the European Commission
Jesus M. Hermida	Joint Research Centre of the European Commission
Polyxeni Mylona	The Publications Office of the European Union
Willem van Gemert	The Publications Office of the European Union
Lutz Rabe	Koordinierungsstelle für IT-Standards (KoSIT)
Francisco José Martín Lizar	CISE – Centre for Semantic Interoperability, Spain
Raminta Vitkauskienė	National Land Service under the Ministry of Agriculture of the Republic of Lithuania
Giedrė Beconytė	The National Center of Remote Sensing and Geoinformatics „GIS-Centras"
Danas Motiejauskas	The National Center of Remote Sensing and Geoinformatics „GIS-Centras"
Hans Overbeek	Knowledge and Exploitation Centre Official Government Publications (KOOP)
Tim Adams	Local Government Association (LG Inform)

REFERENCES

Bechhofer, S., & Miles, A. (2009). *SKOS Simple Knowledge Organization System Reference.* W3C.

Coates, A., & Watts, M. (2007). *Code List Representation (Genericode) Version 1.0.* OASIS.

Council of the European Union. (2009). Council Decision 2009/316/JHA of 6 April 2009 on the establishment of the European Criminal Records Information System (ECRIS) in application of Article 11 of Framework Decision 2009/315/JHA. *Official Journal L 093*, 33-48.

De Leenheer, P., de Moor, A., & Christiaens, S. (2010). *Business Semantics Management at the Flemish Public Administration.*

Dekkers, M., & Goedertier, S. (2013). *Metadata for Public Sector Administration.* NISO/DCMI.

European Commission. (2010). *Evaluation of the 2004 action plan for electronic public procurement.* Brussels: European Commission.

European Commission. (2011, December 12). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions. *Open data - An engine for innovation, growth and transparent governance*. Brussels, Belgium.

European Commission. (2013). *INSPIRE Directive*. Retrieved December 10, 2013, from European Commission: http://inspire.jrc.ec.europa.eu/

European Union. (2011). Regulation no. 182/2011 laying down the rules and general principles concerning mechanisms for control by Member States of the Commission's exercise of implementing powers.

Government On-Line Metadata Working Group. (2006). *Records Management Application Profile.* Canada: Government of Canada.

Graux, H. (2009). *Study on electronic documents and electronic delivery for the purpose of the impementation of Art. 8 of the Services Directive.* Brussels: Timelex.

International Organization for Standardization. (2009). *ISO/IEC 11179-1:2004 - Metadata registries.*

ISA Programme of the European Commission. (2011, May 5). e-Government Core Vocabularies: The SEMIC.EU approach. Brussels, Belgium: European Commission.

ISA. (2010). *European Interoperability Framework (EIF) for European public services.* Brussels: European Commission.

ISA Programme of the European Commission. (2011). *Towards Open Government Metadata.* Brussels.

ISA Programme of the European Commission. (2012). *Metadata Management Survey Results.*

ISA Programme of the European Commission. (2013). *Towards harmonised governance and management of data models and reference data - Business case.* Brussels.

ISA Programme of the European Commission. (2013). *Process and methodology for developing semantic agreements.* https://joinup.ec.europa.eu/community/core_vocabularies/document/process-and-methodology-developing-semantic-agreements.

Kurt Salmon. (2013). *Assessment of TESs supporting EU policies.*

LeBlanc, P., & Smith, B. L. (2002). *A Workshop on Managing Horizontal Issues.* Retrieved 12 24, 2013, from Managing Horizontal Issues: http://www.thinkwell.ca/groupwork/managingHorizontalIssues/documents/MHIWkshopOutlinev2.pdf

National Information Standards Organization . (2004). *Understanding Metadata.*

NIEM. (2013). *NIEM Tools Catalog*. Retrieved 11 25, 2013, from NIEM | National Information Exchange Model: https://www.niem.gov/tools-catalog/Pages/tools.aspx

Object Management Group, Inc. (2012, October). *Unified Modeling Language™ (UML®)*. Retrieved November 20, 2013, from Object Management Group: http://www.omg.org/spec/UML/

Portal Administración Electrónica. (2013). *Technical standards for interoperability* . Retrieved December 09, 2013, from Portal Administración Electrónica: http://administracionelectronica.gob.es/pae_Home/pae_Organizacion/pae_DGMAPIAE.html?idioma=en

Publications Office of the European Union. (2011). *Proposal for metadata governance on interinstitutional level.*

Spanish Ministry of Finance and Public Administration. (n.d.). Decision of 19 February 2013 of the secretary of state for public administration approving the technical interoperability standard for the reuse of information resources.

UN/CEFACT. (2009). *XML Naming and Design Rules Technical Specification - Version 3.0.*

ANNEX I: LONG LIST OF CASES

Table 9 – Descriptions of identified cases

Line No.	Country	Name (website)	Description
1.	AT	Statistics Austria, Integrated Metadata System (IMS)	IMS, a system developed by Statistics Austria, aims to achieve planned storage and use of meta data, to establish norms for data management in general and to provide support to users via general-use tools.
2.	AT	Digital Austria	Digital Austria develops e. services for the Austrian citizens.
3.	BE	Federal Public Service for Information and Communication Technology (FEDICT)	FEDICT uses innovative information and communication technology to help the various federal public services to improve their service portfolios and tailor them to meet the needs of the general public, businesses and civil servants.
4.	BG	State Agency for Information Technologies and Communications (SAITC)	SAITIC pursues the state policy in the field of IT and communications aiming at the development of Information Society.
5.	CY	Department of Information Technology Services (DITS)	DITS is responsible for developing e- services taking the public's needs, mentality and culture into consideration.
6.	CZ	Czech Statistical Office, Statistical metainformation system (SMS)	SMS is used for meta-information inside and outside the Statistical Information System as a tool for internal and external integration.
7.	CZ	Ministry of Interior - Information system of data elements (ISPD)	ISPD provides official information on data elements of information systems used by public administrations. It also serves in publishing data elements and code lists.
8.	DE	Federal Statistical Office (Destatis), Standardisation of Production (SteP)	SteP is a joint initiative to standardise statistical production. Under this initiative a sub working group for metadata has been established. The idea behind this sub working group was to develop a metadata portal. This web portal allows users to access the metadata stored in already existing IT-systems.
9.	DE	Data models database (XRepoditory)	XRepository provides information and access to freely available subject-specific and interdisciplinary data models and XML standards of public administration.
10.	DE	Finanzen Bremen - KoSIT - XÖV Framework	The coordinating body for IT standards KoSIT has the task to coordinate the development and operation of IT standards for data exchange in the public administration.
11.	DK	Digitaliser.dk – one stop shop to the digitization of Denmark	Digitaliser.dk brings together key resources, recommendations and guidance on IT, communication and IT development. The goal is to strengthen cooperation on digitalization between public and private parties. It is also a social networking platform.
12.	EE	RIHA - The Administration system of Estonia	RIHA serves as a catalogue for the State's information system. It is also a procedural and administrative environment via which the comprehensive and balanced development of the State's information system is ensured.
13.	ES	Centro de Interoperabilidad Semántica (CISE)	CISE is the instrument defined in the Spanish National Interoperability Framework to publish data

Line No.	Country	Name (website)	Description
			models and encodings associated with the exchange of data between the different administrations.
14.	ES	The Technology Transfer Centre of the Spanish Government (CTT)	CTT has the goal of encouraging the reuse of solutions on all levels of government. As a result CTT has developed a catalogue that may include several semantic interoperability standards (example SICRES 3.0).
15.	ES	Spanish National Government data portal (SNGDP)	SNGDP is the national portal that organises and manages the Public Information Catalogue of the General State Administration. Portal gives access to practical, informative and educational resources that are useful in developing products and services with a high social and economic value based on the reuse of public sector information.
16.	EU	European Environment Agency (EEA)	EAA is a major information source for those involved in developing, adopting, implementing and evaluating environmental policy, and also for the general public.
17.	EU	Eurostat	The main role of Eurostat is to process and publish comparable statistical information at European level.
18.	EU	Europeana	Europeana portal is a database of cultural heritage. The metadata for all the objects in the Europeana portal is open. In 2012 a large subset of this data was transformed into linked data and represented in the Europeana Data Model (EDM).
19.	EU	The Metadata Registry (MDR) of EU	MDR registers and maintains definition data (metadata elements, named authority lists, schemas, etc.) used by the different EU Institutions.
20.	EU	EU spatial data infrastructure (INSPIRE)	INSPIRE aims to create EU spatial data infrastructure, which will enable the sharing of environmental spatial information among public sector organisations. It also helps to facilitate better public access to spatial information across Europe.
21.	EU	European Multi-Stakeholders Platform on ICT Standardisation	European Multi-Stakeholders Platform on ICT Standardisation is Advisory Expert Group on all matters related to European ICT Standardisation and its effective implementation.
22.	FI	Library Network Services of the National Library, National Metadata Repository project	The aim of the project was to create a joint database for all library sectors in Finland. The expansion to cover all library sectors should have brought more library systems into the structure.
23.	FI	The National Archives Service of Finland, SÄHKE	SÄHKE is one of the national specifications for electronic records management system. It has over 120 metadata elements, many of which can be used at several levels in archival hierarchy.
24.	FI	Yhteentoimivuss.fi – national portal	Yhteentoimivuss.fi is a publishing platform and national portal, which gathers together information for supporting interoperability and enterprise architecture work by Finnish Public Administration.
25.	FR	The Interministerial Department of Information and Communication Systems (DISIC)	The purpose of SISIC is to define a coherent strategic framework for the evolution of information systems management. SISIC also develops a framework that helps to implement and monitor performance.

Line No.	Country	Name (website)	Description
26.	GR	Datasets of all public bodies of Greek government (data.gov.gr)	IT is the central catalogue of public data which provides access to datasets of all public bodies of Greek government.
27.	HR	Commission for Public Administration Informatisation (CPAI)	CPAI performs administrative and professional tasks related to the development of the state administration IT system and establishment of information technology and security infrastructure in the state administration bodies.
28.	HR	Croatian Bureau of Statistics (CBS), the central metadata repository (CROMETA)	CROMETA is the essential part and the core of the Integrated Statistical Information system. The CROMETA model contains Reference ModelTM concepts extended and customized for CBS needs as well as specifics of a previous CBS metadata model.
29.	IT	Public Connectivity and Cooperation System (SPCData)	SPCData is the data space of the Italian Public Connectivity and Cooperation. It consists of several datasets in Linked Data format such as IPA (Index of Public Administration). SPCData aims to become the hub of the data of the Linked Italian public administration. Currently, SPCData is linked to other Linked Data for some agencies such as National Research Council, the City of Florence and the Piedmont Region.
30.	LT	e-Government gateway	The e-Government gateway is created as a single window for e-services. It is using metadata structure in order to make e. service available.
31.	LT	Lithuanian Spatial Information Portal (LSIP)	The goal of the LSIP is to make good conditions for centralized data (spatial data) provision among the data users. The shared data is provided and could be used by those who create and use spatial data and metadata.
32.	LU	Statistics portal of the Gran-Duchy of Luxembourg (STATEC)	This portal aims to implement new innovation and communication technologies for the benefit of citizens and administrative reform. It's a database of various data and statistics from other governmental institutions altogether.
33.	LU	Centre of Information Technologies of the State (CTIE)	The CTIE aims to support government departments in their process of reorganisation and optimization tasks.
34.	LV	Ministry of the Environmental Protection and Regional Development - State Information System (SIS)	SIS is a structured set of an information technology and a database (data/information, which is considered as a one logical unit) aggregate. There is a secured initiation, creation, compilation, collection, processing, use and a liquidation of information required for the public function implementation.
35.	MT	On-Line Statistical database (StatDB), National Statistics office of Malta (NSO)	StafDB provides statistics on a wide range of social and economic matters, covering the population in general, Government and the business sectors. In 2012 NSO implemented changes so that StatDB would be compatible with metadata standards of European Statistical System.
36.	MT	Malta Information Technology Agency (MITA)	MITA manages the implementation of IT programmes to enhance public service delivery. It also provides the infrastructure needed for Government institutions to develop and provide e. services.
37.	NL	KOOP - Knowledge and Exploitation Centre Official	KOOP is a government organisation that develops and manages all levels of government, both central

Line No.	Country	Name (website)	Description
		Government Publications	government and the provinces. It manages the metadata standard for information published by the Dutch government on the Web, including schemas, code lists and value syntaxes.
38.	NL	Stelselcatalogus	Stelselcatalogus gives users, customers, suppliers and others a comprehensive picture of the available data and concepts within the System of Basic Registrations and their meaning.
39.	PL	Ministry of Administration and Digitization	The ministry is concerned with various aspects of administration, Internet and telecommunication in Poland. It helps to develop the broadband infrastructure, supports the development of web content and services and promotes digital literacy among its citizens.
40.	PT	The National Statistical System of Portugal (NSS)	The purpose of NSS is to record, refine, coordinate and disseminate official data. Indicators and associated metadata disseminated on the website are registered in the Variables System.
41.	PT	Agency for Administrative Modernisation (AAM)	AAM aim is to contribute to the profound transformation of the relationship between the Portuguese Public Administration and those who justify it, the Citizens and Businesses, structuring its activity around one principle: provide the correct information, in the right format, to the right person at the right time.
42.	RO	National Centre of Digital Romania (CNRD)	CNRD mission is focused on optimizing the performance of government in order to increase user accessibility to e. services. CNRD also ensures informational content management and information services in the eRomania portal, manages and operates systems through which services are provided by e-Government.
43.	SE	Official Statistics of Sweden (SCB)	SCB contains all statistical surveys, historical statistics and regional statistics of Sweden. It is possible to find tools for further processing and analysis of statistics. Here metadata has the role of driving and delivering information between different process steps. The metadata development is included in the overall strategy for Statistics Sweden's data architecture.
44.	SI	National Interoperability Framework (NIO)	NIO is a national portal that organizes and manages the interoperability assets, Open data and applications. It is a single point of access to repository, data sets and applications tools of the General State Administration.
45.	SK	Central Metainformation System of Public Administration (Meta IS), Ministry of Finance	Meta IS is designed to provide users with available government metadata. It also provides maintenance of all planned and implemented public administration objects at a single level.
46.	UK	Local Government Inform (LG Inform)	LG Inform is the benchmarking data service of Local Government Administration for councils, fire and rescue authorities based on e-Government Metadata Standard.
47.	UK	The National Archive of UK / PRONOM	The National Archives is a government department, which is a centre of expertise in every aspect of creating, storing, using and managing official information. PRONOM is the National Archives' technical registry and holds the data, which is in a linked open data format.

Line No.	Country	Name (website)	Description
48.	UK	Listpoint - open data platform of public and private sector	Listpoint is the essential data management service and toolkit that unleashes the valuable insights locked up in big data and open data.
49.	UK	The Government Digital Service (GDS)	GDS purpose is to ensure the Government offers digital products that meet people's needs. GDS has an Open Standards Board to help them decide which open standards to use in government technology.

ANNEX II: ANALYSIS OF SELECTED CASE STUDIES

This chapter discusses the detailed analysis of the eight selected case studies, summarised in Table 2. The analysis comprises both desk research and interviews with representatives from each case study.

EU Institutions

Case study 1: The Statistical Office of the European Union - Eurostat

Eurostat[26] is the statistical office of the EU. Its task is to provide the EU with statistics at European level that enable comparisons between countries and regions. Eurostat offers a whole range of important and interesting data that governments, businesses, the education sector, journalists and the public can use for their work and daily life.

Eurostat does not work alone. Since the early days of the EU it was realised that decisions on and planning and implementation of Commission policies must be based on reliable and comparable statistics. So the European Statistical System[27] (ESS) was built up gradually with the objective of providing comparable statistics at EU level.

The ESS is a partnership between the Community statistical authority, which is the Commission (Eurostat), and the national statistical institutes (NSIs) and other national authorities responsible in each Member State for the development, production and dissemination of European statistics. This Partnership also includes the EEA and EFTA countries. Member States collect data and compile statistics for national and EU purposes. The ESS functions as a network in which Eurostat's role is to lead the way in the harmonization of statistics in close cooperation with the national statistical authorities. ESS work concentrates mainly on EU policy areas but, with the extension of EU policies, harmonization has been extended to nearly all statistical fields.

Due to this partnership, transparent and integrated descriptions of information flows within and outside Eurostat are vital. The use of technology for data collection, interactive communication with users, and dissemination of statistics, calls for a coherent and well-functioning metadata management and governance.

The whole statistical production process from survey design over data collection and processing to dissemination takes place independently of other domains, and each has its own data suppliers and user groups. Yet, the goal is to produce statistics as integrated parts of comprehensive production systems (the so-called data warehouse approach) for clusters of statistics.

The figures below describe the way statistical data is gathered in specific domains, through the use of structural metadata governance and management. This process is now shifting from a distributed to an integrated one as presented in the next figure.

[26] The statistical office of the European Union (Eurostat): http://ec.europa.eu/yourvoice/ipm/index_en.htm
[27] The European Statistical System (ESS): http://epp.eurostat.ec.europa.eu/portal/page/portal/ess_eurostat/introduction

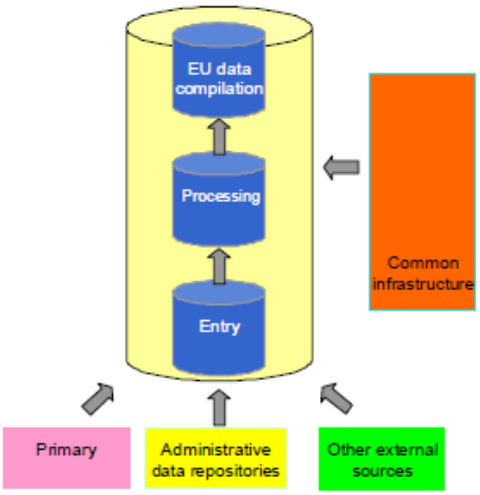

Figure 4. The integrated model with non-restricted domains[28]

The structural metadata governance structure has several levels i.e. councillors of Parliament, Eurostat working groups. However the primary governance role is played by Metadata Working Group..

Eurostat put the SDMX registry[29] in place with the purpose of storing structural metadata and data for querying which can be used by any other application in the network and can be seen as the index of a distributed database of metadata.

The off-line mode of SDMX registry is intended to be used for the creation and maintenance of the following SDMX objects: Data Structure Definitions, Code Lists, Concept Schemes, Data Flows, Hierarchical Code lists, Category Schemes and Organisation Schemes. In the on-line mode, users can perform the same operations as in off-line mode plus the possibility to interact with any standard-compliant SDMX Registry.

What is SDMX?

SDMX is an specification to foster standards for the exchange of statistical information It started in 2001 and aims at fostering standards for Statistical Data and Metadata eXchange (SDMX).
SDMX provides:

- A logical model to describe statistical data, together with guidelines on how to structure the content;
- A standard for automated communication from machine to machine; and
- A technology supporting standardised IT tools that can be used by all parties involved in the data exchange and processing.

Euro-SDMX Metadata Structure (ESMS)[30] is used for describing the statistics at Eurostat which also has methodologies, quality and production processes

[29] SDMX registry: https://webgate.ec.europa.eu/sdmxregistry/index.do

documented. ESMS also provides statistical domain manager with more detailed guidelines based on which information has to be delivered.

ESMS are based on the Euro SDMX Metadata Structure. It aims at documenting methodologies, quality and the statistical production processes in general. ESMS uses 21 high-level concepts, with a limited breakdown of sub-items, strictly derived from the list of cross domain concepts in the SDMX Content Oriented Guidelines (2009). The quality criteria used are taken from the European framework for data quality.

Eurostat has implemented a number of ISO metadata management and exchange standards for the statistics domain. One of them is the Statistical data and metadata exchange standard ISO 17369:2013[31], applicable to any organisation that has a need to manage the reporting, exchange and dissemination of its statistical data and related metadata. The information model at the core of ISO 17369:2013 has been developed to support statistics as collected and used by governmental and supra-national statistical organisations, and this model is also applicable to other organisational contexts involving statistical data and related metadata.

Another standard, implemented by Eurostat is the Metadata Registry (MDR) standard ISO/IEC 11179[32]. The ISO/IEC 11179 model is a result of two principles of semantic theory, combined with basic principles of data modelling. The first principle from semantic theory is the thesaurus type relation between wider and more narrow (or specific) concepts, e.g. the wide concept 'income' has a relation to the more narrow concept 'net income'. The second principle from semantic theory is the relation between a concept and its representation e.g., 'buy' and 'purchase' are the same concept although different terms are used.

The main tools used by Eurostat for metadata management (SDMX Registry and RAMON[33]) are open source, some under EUPL V1.1[34], and are developed in-house.

In the remainder of this section, we present the summary of the interview with Eurostat conducted in the context of this case study. The summary is structured according to the analysis framework of section 2.1.2.

Interview date	17 January 2014
Interviewee	Marco Pellegrino, August Götzfried
Interviewer	Audrius Leipus

[30] More information : http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/metadata/metadata_structure

[31] ISO 17369:2013 standard: http://www.iso.org/iso/catalogue_detail.htm?csnumber=52500

[32] ISO/IEC 11179 standard: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=50340

[33] More information: http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC

[34] EUPL V1.1: http://ec.europa.eu/idabc/eupl.html

Title: The statistical office of the European Union – Eurostat	
Governance	
Goals	The goal of structural metadata governance is to document data in order to provide the EU with a high-quality statistical information service. Shorter-term goals include: • Better integration of the ESS in terms of IT tools, data quality, metadata, methodology etc. (both in terms of horizontal and vertical integration); • Shift from statistical 'stove pipes' to more integrated statistical production processes; • Broader use of administrative data sources in the statistical data production processes.
Governance structure	**Roles:** The governance structure consists of several levels. The highest level is the ESS level, councillors of Parliament, while the lowest level includes Eurostat internal working groups. Eurostat is also maintaining a Metadata Working Group, a group with ESS member States to discuss issues concerning management and standardisation of statistical metadata and a CIRCA (Communication and Information Resource Centre Administrator) interest group on Statistical metadata where working papers, technical documents and other web references are regularly posted, and are available to the general public. **Openness:** The governance structure is open to new requirements and needs of both Eurostat and national statistical offices. **Decision making process:** The decision making process is the same as in ESS. It consists of several levels where decisions can be made at any of the levels depending on the importance of the decision to be made.
Context	**Organisational dimension:** • **Inter-organisational.** Eurostat belongs to ESS, which is the partnership between the Community statistical authority, and the national statistical institutes (NSIs) and other national authorities responsible in each Member State for the development, production and dissemination of European statistics. **Administrative dimension:** • Pan-European; and • National.
Policy domain	Policy domain is restricted to statistics domain only. On the other hand, statistics domain could be broken down to specific themes, like agriculture and fisheries, international trade, economy and finance etc.
Enforcement policy	**Sharing and reuse:** There is no specific legal act enforcing sharing and reuse of ESS structural metadata, but there are several, each statistic area specific, legal acts defining data structure definitions, dimensions, etc. In addition harmonised structural metadata is part of data structure definition (DSD) maintenance agreements making the use of these structural metadata compulsory.

Authoritative source	There are several metadata repositories dealing with different types of metadata and provide ESS and Eurostat metadata producers with a single entry point for handling different types of statistical metadata. One of them is the Statistical Data and Metadata eXchange (SDMX) Registry, the other one is RAMON, Eurostat's Metadata Server. . The SDMX Registry is also an authoritative source of Eurostat and an application which stores metadata for querying, and which can be used by any other application in the network with sufficient access privileges[35]. It can be seen as the index of a distributed database or metadata repository which is made up of all the data provider's data sets and reference metadata sets within a statistical community.
Licensing framework	All pages on the Eurostat Website refer to a copyright notice[36], encouraging free reuse with attribution, based on the **general legal notice** of the European Commission[37].
Quality controls	The quality control process is very complex, especially for the descriptive metadata, as it comprises several responsibility layers and depends on national statistical institutions. Yet the key body in the process is a separate directorate unit (comprising15-20 people), which collaborates with other statistical institutions.
Metadata schema	**Vocabulary:** There is a central glossary of compilation practises, technical vocabularies etc. **Identifiers scheme:** Identifiers are used for various objects (identifiable artefacts) in SDMX[38]. For example there is agency scheme used for describing agencies. URN's are used for identifiable artefacts. **Schema documentation:** Schema documentation is available on the SDMX Registry website**. Multilingualism:** Most of the metadata schemas are in English. Documentation of technical aspects (e.g. label names, publications) is available in 3 main languages: French, German and English.
Management	
Update	**Update frequency:** The update frequency depends on the type of metadata and statistical domain where they have around 80 of them. There is no fixed update frequency. **Change management process:** Eurostat has a defined change management process for the SDMX statistical working group[39]. **Version control:** Version control is used for structural metadata.

[35] SDMX Registry: https://webgate.ec.europa.eu/sdmxregistry/start.do

[36] http://epp.eurostat.ec.europa.eu/portal/page/portal/about_eurostat/policies/copyright_licence_policy

[37] http://ec.europa.eu/geninfo/legal_notices_en.htm

[38] SDMX Registry specification: http://sdmx.org/wp-content/uploads/2011/08/SDMX_2-1-1_SECTION_5_RegistrySpecification_201108.pdf

[39] Governance and management of SDMX: http://sdmx.org/?page_id=11

Harmonization	**Creating mappings between related metadata sets:** Mapping process is based on the SDMX standard. The Mapping Assistant is used to facilitate the mapping between the structural metadata provided by an SDMX-ML Data Structure Definition (DSD) and those that reside in a database of a dissemination environment. The mapping process with the Mapping Assistant tool can be described in four steps: Step 1 – loading of the SDMX structures – Category Scheme, Data Flow, Data Structure Definition – from SDMX-ML structure files; Step 2 – loading of the local non-SDMX database schema and the creation of the Dataset; Step 3 – mapping of local concepts to SDMX Concepts of the Data Structure Definition; Step 4 – transcoding of local codes to SDMX Codes of the Codelists referenced in the Data Structure Definition. **Assessing alternative metadata sets:** Assessment of alternative metadata sets is done by the expert of the domain.
Documentation	**Publication:** Metadata is always published in Eurostat website according to a common vocabulary namely SDMX. **Retrieval:** Retrieval can be done by both humans and machines. E.g. the Eurostat SDMX Registry is accessible to human users via the GUI and to machines via the Web services of the registry. **Supported formats of publication tool(s):** Eurostat publication tools support human and machine readable formats (e.g. XLS, CSV, SPSS, PDF, TSV, PC AXIS formats).
Standards	ISO 17369:2013 (Statistical data and metadata exchange (SDMX)) provides an integrated approach to facilitating SDMX, enabling interoperable implementations within and between systems concerned with the exchange, reporting and dissemination of statistical data and related metadata. ISO/IEC 11179 is used for metadata representation in the metadata registry. ESMS is used for describing the statistics released by Eurostat.
Tool support	Most of available tools deal with the whole process of the metadata management, i.e. from data creation to its validation and publication.
Costs	There are 5-6 people fully dedicated for the metadata management, who play coordinating role.
Benefits	Benefits offset the costs, but there are no measures used to estimate the benefits of metadata management. Some of the benefits include: • Better quality of collected data; • Better interoperability and reuse of metadata; • Centralised metadata management tools.
Tool support	
Name of the tool(s) and documentation	SDMX Registry[40] RAMON - Eurostat's metadata server[41]

[40] SDMX Registry: http://sdmx.org/wp-content/uploads/2009/10/2_Eurostat_registry_41682274.pdf

[41] Eurostat's metadata server - RAMON: http://ec.europa.eu/eurostat/ramon/foreword/index.cfm?targetUrl=DSP_FOREWORD

Reusability	**Open source:** Tools are open source, some released under EUPL. Tools can be reused by other National Statistical institutions. The Eurostat SDMX Registry has been published as Open Source Software under the EUPL. It can be downloaded via the tools page of the SDMX website[42].
Key functionality	The main functionality of RAMON is the following: • Multilingual character of the information disseminated. Whenever possible or available, the information is presented in all languages in which it exists. This is especially true for some classifications which sometimes exist in 20 languages. • Its powerful search engine. When a search is performed, the search engine of RAMON goes through all the objects loaded in the database and returns all results found, whatever the metadata category in which it was found. • Layout adapted to the type of metadata presented; the way the information is displayed on the screen depends on the kind of metadata; for instance thesauri will respect the main international standards used for presentation of these data (Narrow terms, broad terms, etc.). An Eurostat SDMX Registry performs a number of functions: • It provides information about what data sets and metadata sets are available, and where they are located. • It provides information about how the data sets and metadata sets are provided: how often they are updated, what their contents are, how they can be accessed, and similar questions. • It provides information about the structure of data sets and metadata sets, answering questions like: What code lists do they use? What concepts are involved? • It allows applications to sign up (or subscribe) for notifications, so that when a data set or metadata set of interest becomes available, the application will be automatically alerted.
Costs	Tools are developed in-house. External experts take minimum part in tool creation. Cost of the tools depends on the cost of internal resources.

Case study 2: Joint Research Centre – INSPIRE

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle. Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners[43].

The JRC acts as the overall technical co-ordinator of INSPIRE.

[42] SDMX tools page: http://www.sdmx.org/index.php?page_id=13

[43] Joint Research Centre: http://ec.europa.eu/dgs/jrc/

The INSPIRE Directive[44] aims to create a European Union (EU) spatial data infrastructure, enabling the sharing of environmental spatial information among public sector organisations and better facilitate public access to spatial information across Europe.

A European Spatial Data Infrastructure will assist in policy-making across boundaries. Therefore the spatial information considered under the directive is extensive and includes a great variety of topical and technical themes.

INSPIRE is based on a number of common principles[45]:

- Data should be collected only once and kept where it can be maintained most effectively.
- It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications.
- It should be possible for information collected at one level/scale to be shared with all levels/scales; detailed for thorough investigations, general for strategic purposes.
- Geographic information needed for good governance at all levels should be readily and transparently available.
- Easy to find what geographic information is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used.

The INSPIRE Directive defines 'infrastructure for spatial information' as metadata, spatial data sets and spatial data services; network services and technologies; agreements on sharing, access and use. In this document, the focus is on the structural metadata, i.e. the structure of the metadata and the permitted values as defined in the INSPIRE Implementing Rules[46].

The INSPIRE Directive came into force on 15 May 2007 and is implemented in various stages, with full implementation required by 2019, based on the technical guidelines that are published by JRC as INSPIRE Implementing Rules.

The metadata elements that are used for INSPIRE are laid down in the INSPIRE Regulation (Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata)[47].

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

[44] INSPIRE Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:en:PDF

[45] Infrastructure for Spatial Information in the European Community (INSPIRE): http://inspire.ec.europa.eu/index.cfm/pageid/48

[46] INSPIRE Implementing Rules. http://inspire.ec.europa.eu/index.cfm/pageid/47

[47] COMMISSION REGULATION (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:326:0012:0030:EN:PDF

Community". All infrastructures, and INSPIRE is no exception, require maintenance and evolution if they want to remain relevant for serving the purposes for which they have been put in place. Lessons learned by implementing the infrastructure need also to be taken into account to further optimise its performance in order to meeting its policy objectives. Setting up a framework for such maintenance, further implementation and evolution for INSPIRE is therefore a logical initiative.

INSPIRE Directive
The JRC ensures the viability and evolution of the technical infrastructure for INSPIRE and guarantees the liaison with the European and international research community. JRC also initiates and monitors the work with international standardisation bodies for the purposes of INSPIRE and will be responsible for the technical coordination with other relevant international initiatives.

Other organisations involved in the coordination of INSPIRE are Directorate-General Environment as overall legislative and policy co-ordinator and the European Environmental Agency (EEA) that takes on tasks related to monitoring and reporting, and data and service sharing[48].

INSPIRE Maintenance and Implementation Framework (MIF)
The Commission, in agreement with the Member States, is therefore setting up the INSPIRE Maintenance and Implementation Framework (MIF), which is based on the same principles as those applied for its development. This MIF will have to address the following main challenges:

- Be fully aligned with and interfaced to the ongoing development of the remaining IRs.
- Be supportive to the further implementation of the IRs in the Member States.
- Be responsive to lessons learned from the implementation (which may require modifications to the Legal Acts and Technical Guidelines and associated registers and tools).
- Be comprehensive to ensure the cross-cutting coherence of the components of the infrastructure - some of the issues resulting from implementation of the IRs may affect more than one INSPIRE component, e.g. data specifications and network services, and it is crucial that these are resolved and applied in a consistent manner.
- Be flexible for taking into account requirements emerging from environmental policies and policies or activities which may have an impact on the environment.
- Be adequately resourced and organized for dealing with event-driven requests and needs for maintenance and evolution.

INSPIRE Maintenance and Implementation Group (MIG)
In April 2013, the INSPIRE Committee agreed to set up a Commission expert group called INSPIRE Maintenance and Implementation Group (MIG), chaired by the European Commission (JRC) with representatives of the INSPIRE national contact

[48] Structure of INSPIRE: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/481

points and representatives from relevant Commission Directorates-General. The tasks of the INSPIRE MIG are:

- to bring about an exchange of experience and good practice related to the implementation of the INSPIRE Directive and the Implementing Rules;
- to identify and give advice about the priority issues to be addressed in the maintenance of the INSPIRE Directive, Implementing Rules and/or Technical Guidance documents;
- to identify issues related to INSPIRE implementation (including, but not limited to, technologies, standards, methods, coherence across INSPIRE chapters and communication measures to be adopted) and advise the Commission on how to address them.

The basis of the work of the MIG and its sub-groups is a common work programme based on issues and change requests submitted by INSPIRE stakeholders.

Pool of Experts

The MIG will be complemented by a pool of experts drawn from the stakeholder community. The experts in this pool will be 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[49].

In the remainder of this section, we present the summary of the interview with JRC conducted in the context of this case study. The summary is structured according to the analysis framework of section 2.1.2.

The INSPIRE Registry[50]

The INSPIRE infrastructure involves a number of items, which require clear descriptions and the possibility to be referenced through unique identifiers. Examples for such items include INSPIRE themes, code lists, application schemas or discovery services. Registers provide a means to assign identifiers to items and their labels, definitions and descriptions (in different languages). The INSPIRE registry provides a central access point to a number of centrally managed INSPIRE registers. The content of these registers are based on the INSPIRE Directive, Implementing Rules and Technical Guidelines[51].

The INSPIRE Registry is a live instance of the Re3istry[52], a tool that was developed under Action 1.17 of the ISA Programme: A Reusable INSPIRE Reference Platform[53].

[49] More information about INSPIRE maintenance and implementation: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5160

[50] INSPIRE Registry: http://inspire.ec.europa.eu/registry/

[51] INSPIRE Technical specifications listed at http://inspire.ec.europa.eu/index.cfm/pageid/2/list/1

[52] The Re3gistry tool: https://joinup.ec.europa.eu/software/re3gistry/description

[53] A Reusable INSPIRE Reference Platform. https://joinup.ec.europa.eu/community/are3na/description

Interview date	14 January 2014
Interviewee	Michael Lutz, Andrea Perego, Roberto Sgnaolin, Jesús Hermida
Interviewer	Makx Dekkers, Stijn Goedertier

Title: Joint Research Centre – INSPIRE	
Governance	
Goals	The **goal** of the metadata management activity is to support the distributed implementation with a central registry.
Governance structure	**Roles:** 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 JRC 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[54] procedure. There is a proposal for a work item[55] to update the metadata Technical Guidelines for metadata, and a MIG sub-group may be formed to define whether and how the Technical Guidelines need to be updated. **Openness:** In addition to the formal structure above, a group of self-registered experts is involved on a per-case basis to give advice. **Decision making process:** Processes for **decision making** are under development.
Context	**Organisational dimension:** • **Inter-organisational.** **Administrative dimension:** • **Pan-European**, including the EFTA and accession countries.
Policy domain	The **primary policy domain** covered is **environment**, but it also covers any other domain that has a relationship with the environment, and therefore includes geospatial data, health, safety etc.
Enforcement policy	**Sharing and reuse:** The INSPIRE Implementing Rules are **mandatory**. Centrally managed schemas can be extended for local use.
Authoritative source	The INSPIRE Registry is the **authoritative source**.
Licensing framework	There is **no explicit licence** for the re-use of the schemas and code lists. All pages on the INSPIRE Website refer to the **general legal notice** of the European Commission.
Quality controls	A three-cycle process was used for **quality control** with both internal and external review in the development phase. Formal quality control procedures for both schemas and code lists are under discussion. Validation of metadata is done using the INSPIRE Geoportal Metadata Validator[56].

[54] Europa. Summaries of EU legislation. Glossary. Comitology. http://europa.eu/legislation_summaries/glossary/comitology_en.htm

[55] MIG. Proposal for Update of Metadata TG. https://ies-svn.jrc.ec.europa.eu/issues/2130

[56] INSPIRE Geoportal Metadata Validator: http://inspire-geoportal.ec.europa.eu/validator2/

Metadata schema	**Vocabulary:** The **vocabulary** for description of schemas is currently very simple (label, theme, status). **ADMS is used** for the description of Annex I specifications; those are made available through Joinup. **Identifier scheme:** An **identifier scheme** is being proposed for identification of objects and datasets[57]. HTTP URIs are assigned to any item in INSPIRE, including metadata records. **Schema documentation** is available in the implementing rules for metadata and in the INSPIRE UML repository[58]. **Multilingualism**: support of **multilingualism**, the legal act includes labels in all official languages. Documentation is provided centrally in English; member states in some cases translate and adapt them to the national situation, sometimes using automatic translation. A new activity is considered to at least translate the overview and summaries of other documents.
Management	
Update	**Update frequency:** updates are not foreseen more often than once every few years. As much of the INSPIRE schema and code lists are in the legal text, changes require changes in the law. **Change management** processes are under discussion. **Versioning** is implemented in the INSPIRE Registry.
Harmonization	There is currently **no harmonisation** activity. There is flexibility for implementation. As long as information is classified according to the INSPIRE themes, any other classification vocabulary can be used additionally. Integration of local code lists is an issue that is being considered.
Documentation	**Publication** is done on the INSPIRE Registry that has a search **function**. **Retrieval** from the INSPIRE Registry uses content negotiation. **Supported formats** of publication are both: human-readable (HTML) and machine-readable (XML, JSON and Atom).
Functionality of the tool(s)	Tools used for the management of the schemas and code lists include: • Re3gistry software for the **Registry service** • XMLSpy for **XML editing** • Shapechange for the **creation of implementation representations from UML models** • Enterprise Architect for **UML analysis and design** • Subversion for **document management**
Standards	The **standard** applied for the registration process is ISO 19135. Code lists are implemented in SKOS. ADMS is used for export of descriptions contributed to Joinup.
Costs	**Cost** for the Registry is estimated at 1.5 FTE, with a further 1 FTE for the content.
Benefits	The **benefit** of the Registry is that without it, content is locked in legal text and is not easily accessible. The Registry is used by vendors to create mapping. The benefits are obvious but there is **no quantitative data**.
Tool support	
Name of the tool(s) and	**The Re3gistry[59]**

[57] More about implementation of identifiers: http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5120

[58] INSPIRE UML repository: https://inspire-twg.jrc.ec.europa.eu/index-noredir.html

[59] The Re3gistry tool: https://joinup.ec.europa.eu/software/re3gistry/description

documentation	
Reusability	**Developed in-house** with support from ISA. It is available as **open source** for re-use from the link in the cell above under EUPL
Key functionality	The Re3gistry is a tool used to manage the data contained in the INSPIRE Registry. It is made up of an import function, an export function and a web service. The import procedure loads the data provided (in .csv format) into the database. The export procedure prepares the data saved in the database in order to be ready for the web service. The web service has a RESTful interface to access the data. You can access different formats and languages using content negotiation or directly calling the desired file. The INSPIRE Registry is a live instance of the Re3gistry.
Costs	None
Name of the tool(s) and documentation	**XMLSpy**[60]
Reusability	**Commercial product** from Altova
Key functionality	XML development environment for modelling, editing, transforming, and debugging XML technologies including XML editor and graphical schema designer, code gen, file converters, debuggers, profilers, database integration, chart creation, support for XSLT, XPath, XQuery, WSDL, SOAP, XBRL, JSON and Open XML (OOXML), plus Visual Studio and Eclipse plug-ins and more.
Costs	Starting at €399 for Professional edition and €799 for Enterprise Edition
Name of the tool(s) and documentation	**Shapechange**[61]
Reusability	Source code available under the GNU General Public Licence
Key functionality	Java tool that takes application schemas constructed according to ISO 19109 from a UML model and derives implementation representations.
Costs	None
Name of the tool(s) and documentation	**Enterprise Architect**[62]
Reusability	**Commercial product** from SparxSystems
Key functionality	High performance modelling, visualization and design platform based on the UML 2.4.1 standard
Costs	Corporate edition between US$239 for single-user to US$185 per user for more than 100 users.
Name of the tool(s) and documentation	**Apache Subversion**[63]
Reusability	**Open source**
Key functionality	Version control

[60] XMLSpy tool: http://www.altova.com/xmlspy.html

[61] Shapechange tool: http://shapechange.net/

[62] Sparx Systems Enterprise Architect. http://www.sparxsystems.com/products/ea/

[63] Apache Subversion: http://subversion.apache.org/

Costs	None

Case study 3: Inter-institutional Metadata Management Committee (IMMC), Publications Office Metadata Registry (MDR)

The Publications Office of the European Union (Publications Office) is an inter-institutional office whose task is to publish the publications and disseminate the information of the institutions of the European Union.

The Publications Office publishes the daily Official Journal of the European Union in 24 languages and produces (or co-produces) publicity for EU initiatives and activities. It publishes or co-publishes the publications in the context of the communication activities of the institutions.

Moreover, the Publications Office offers a number of online services giving free access to information on EU law (EUR-Lex)[64], EU publications (EU Bookshop)[65], public procurement Tenders Electronic Daily (TED)[66], EU research and development portal Community Research and Development Information Service (CORDIS)[67] and the EU open data portal (EU ODP)[68].

In the past, all domains in the Publications Office were silos with their own metadata. Thus a need was identified to harmonise metadata across domains (e.g. publications, tenders, scientific, legal information) and to create a horizontal content and metadata layer to support search and access in a coherent way. When confronted with the results of this harmonisation effort, the other EU Institutions with similar problems asked the Publications Office to share their proposed solution to this problem which led to the establishment of the Metadata Registry (MDR)[69], maintained by the Publications Office, in which reference data (metadata elements, named authority lists, schemas, etc.) used by the different European Institutions involved in the legal decision making process is registered and maintained in a controlled manner.

The scope of the data managed in the MDR covers reference data relevant for in the Inter-institutional Metadata Maintenance Committee (IMMC) and for the Publications Office of the EU in its production and dissemination process.

The following datasets are maintained in the Metadata Registry:

- Named Authority Lists (Common Authority Tables/Value lists)
- IMMC Core Metadata element set
- EuroVoc thesaurus and alignments (SKOS/XML distributions)

[64] EUR-Lex. http://eur-lex.europa.eu/en/index.htm

[65] EU Bookshop. https://bookshop.europa.eu/en/home/

[66] TED – Tenders Electronic Daily. http://ted.europa.eu/TED/main/HomePage.do

[67] CORDIS - Community Research and Development Information Service. http://cordis.europa.eu/home_en.html

[68] European Union Open Data Portal. https://open-data.europa.eu/en/data/

[69] Metadata Registry (MDR). http://publications.europa.eu/mdr/

- Style sheets for presentation

Involved organisational entities[70]

IMMC

The Inter-institutional Metadata Maintenance Committee (IMMC) consists of representatives of the following EU bodies: European Parliament, Council of the EU, European Commission, Court of Justice of the EU, European Court of Auditors, European Economic and Social Committee, Committee of the Regions. As decided in the first meeting of the committee, the Publications Office of the EU assures the presidency and secretariat of the IMMC. The role of the IMMC is to evaluate and approve metadata elements and authority data that are relevant for the exchange of data between the Institutions involved in the legislative procedures and the Publications Office with the aim to publish this information and make it available for the European citizen. By its nature, the IMMC forms an inter-institutional platform for collaboration and knowledge exchange in the metadata domain. It will also monitor progress of technical implementation and, when necessary, set up working groups to that effect.

IMSC

The Interinstitutional Metadata Steering Committee (IMSC) is the Management Board of the Publications Office (Comité de direction. Membres suppléants). Its role is to provide guidance and to take the necessary strategic decisions to assure the metadata governance on the interinstitutional level and the implementation and planning of actions resulting from registration proposals of the IMMC. The IMSC can be called to decide in case of differences of opinion that cannot be solved on the IMMC level.

MRT

The Metadata Registry Team (MRT) consists of members of the Enterprise Architecture unit of the Publications Office of the EU. The MRT keeps track of proposals for registration, manages the approval workflow and implements approved changes in the Metadata Registry. Upon request, the MRT can provide functional and technical support to other Institutions for core metadata related issues.

Infrastructure

The infrastructure used for managing the reference is provided by the Metadata Registry (MDR). The tool where the metadata definitions and related authority data are registered and maintained is called the Metadata Registry. The Registry is hosted and managed by the Publications Office.

The Metadata Registry consists of a back-office for the maintenance of the metadata elements and related authority data and a front-office application (website) for consulting all MDR content. Access to the back-office is restricted to the Metadata Registry Team (MRT). The website provides read-access to all MDR content to all users and offers the possibility to provide feedback.

[70] http://publications.europa.eu/mdr/resource/core-metadata/IMMC_reu3_adoption_anx3.pdf_A-2011-764293.pdf

Proposals for registration are registered by the MRT in the Metadata Registry and submitted to the Interinstitutional Metadata Maintenance Committee (IMMC) for approval. Upon approval the necessary changes are made in the Metadata Registry. A new version of the updated asset is generated and published on the MDR website. In the future the dissemination of the MDR assets will take place through the CELLAR, the common content and metadata repository of the Publications Office.

Decision procedure

Core metadata

After the validation of the proposed metadata, each Institution will evaluate if a new metadata element or a new authority value is part of the so-called core metadata and therefore relevant on interinstitutional level. If it is the case, it will submit the proposed item to the IMMC for discussion and approval. The participating Institutions will commit themselves to implementing the approved metadata elements and authority values in the agreed time frame.

The MRT keeps track of proposals for registration, manages the approval workflow and implements approved changes in the Metadata Register.

Institution specific metadata

If the proposed item is not relevant on interinstitutional level, but Institution specific, the Institution is free to manage the proposed item as it wishes without informing the IMMC.

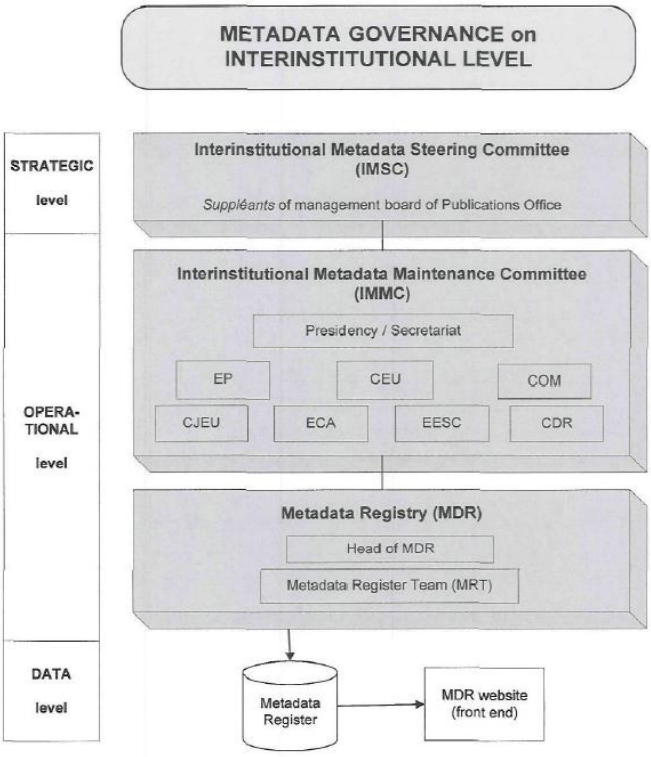

Figure 5. Interinstitutional metadata governance

In the remainder of this section, we present the summary of the interview with the Publications Office conducted in the context of this case study. The summary is

structured according to the analysis framework of section 2.1.2.

Interview date	15 January 2014
Interviewee	Polyxeni Mylona, Willem van Gemert
Interviewer	Makx Dekkers

Title: Inter-institutional Metadata Management Committee (IMMC), Publications Office Metadata Registry (MDR)	
Governance	
Goals	The **objectives** of the metadata governance are to: • Harmonise metadata across sectors and across institutions. • Improve interoperability. • Improve metadata quality as a result of maintaining schemas and values lists in one place. • Make use of resources more efficient. • Align with common and international standards. • Enable re-use within and between EU Institutions and beyond.
Governance structure	**Roles**: three **roles** are defined: • The Inter-institutional Metadata Steering Committee (IMSC), which provides guidance and takes necessary strategic decisions. • The Inter-institutional Metadata Maintenance Committee (IMMC), which evaluates and approves metadata elements and authority data, and monitors progress of implementation. • The Metadata Register Team (MRT), which keeps track of proposals, manages the workflow and implements approved changes. **Openness**: In terms of **openness**, this governance structure is open to all EU Institutions. The European Central Bank has recently joined, while the European Investment Bank has expressed the interest. Agencies that are part of the European Commission (e.g. JRC/INSPIRE, Eurostat) are not directly members of the IMMC but are represented by the Commission. **Decision making process:** The **decision process** has the following steps: • A proposal for changes and additions is submitted by the member institutions to IMMC through their representative in the IMMC. • The IMMC forwards the proposal to the MRT. • After analysis, the MRT submits a recommendation to IMMC. • After positive decision by IMMC, a new property, new value or even a new authority table is created. • The addition is then added to CELLAR and published on the MDR Website.
Context	**Organisational scope:** • **Inter-organisational.** **Administrative scope:** • **Pan-European**. Most of it is focused on EU Institutions but also national entities are involved, for example the national parliaments through the European Parliament. Also

	requests could come from activities with national participation, such as the Open Data Portal and the DCAT Application Profile.
Policy domain	The policy domain is **not restricted.** For example, address schemas are out of scope. Initially, the focus was on the legal domain because the legacy system was phased out and there was an urgent need for a new agreement for exchange of information.
Enforcement policy	**Sharing** is not enforced by law but is entirely **voluntary**. **Re-use** of the core metadata schema and authority tables is likewise **voluntary**. The member institutions of IMMC have a clear commitment to apply to contribute to and use the solutions that are based on the agreements in the IMMC, and in particular the IMMC Core Metadata, the Transmission protocol and the Value lists. Further action is necessary in order to raise awareness across European institutions.
Authoritative source	The Metadata Registry (MDR) is the **authoritative source** for the IMMC Core metadata schema, the associated NALs (Name Authority Lists) and also for the institution-specific metadata schemas.
Licensing framework	There is currently **no explicit licence** mentioned in the schemas or the documentation but it is made explicit in the ADMS feed to Joinup and in the description on the EU Open Data Portal. All information published on MDR is governed by the general licence of the European Commission[71] according to which reuse is authorised for commercial and non-commercial purposes subject to attribution.
Quality controls	**Quality control** is based on a four-eye principle where changes made by one person are reviewed by someone else. Automatic checks are also employed, for example to check that dates are consecutive. Furthermore, all changes are checked in a diff file[72] that is published alongside a new version.
Metadata schema	**Vocabulary:** There is no internally used metadata **vocabulary** to describe the schemas and the value lists. The descriptions are provided in ADMS to Joinup and in the description on the EU Open Data Portal. **Identifier scheme:** URI policies exist for each table but those policies are not yet public. **Schema documentation:** is available on the MDR site in HTML. **Multilingualism**: almost all labels of the schemas and value lists are in all official languages of the European Union, with the exception of the NAL for file types. Documentation and comments in the schemas are in English only.
Management	
Update	**Updates** are frequent in the initial phase of an authority table (e.g. when translations are being added) and less frequent later on. In practice, there have been 11 publications in 2013. For Core metadata a cycle of every two to three months is foreseen. **Change management process:** The **change management**

[71] For more details about this licence: http://europa.eu/geninfo/legal_notices_en.htm

[72] In computing, **diff** is a file comparison utility that outputs the differences between two files. It is typically used to show the changes between one version of a file and a former version of the same file.

	process defines three types of updates with their associated processes (ordinary, urgent and in-depth). Users are notified of changes by e-mail and in the future through notifications sent out from CELLAR. **Versioning:** The schemas are maintained in SVN. Versioning of schemas is indicated by a date plus sequence number (e.g. 20140101-0, 20140101-1 etc.). The sequence number is mostly used for immediate bug fixes.
Harmonization	**Mappings** are created in new authority tables to the old internal codes, and links to external terms are also considered (e.g. links to Library of Congress ISO 639 languages). During development of schemas and authority lists **existing standards are assessed** with the objective to align as much as possible with those.
Documentation	**Publication of metadata** about the schemas and authority tables is not currently done on MDR and search functionality is not available. This is being considered for the year 2015. **Retrieval** is through manual download from the pages of the MDR. **Supported formats of publication** of the schemas and authority tables are: XSD for the core metadata and transmission protocol schemas; SKOS, XML, XSD and HTML for authority tables.
Functionality of the tool(s)	The following tools are used : • JIRA for the **Governance workflow.** • SVN for **versioning of schemas.** • XSLT/PERL scripts for the **technical workflow and conversions.** The current toolset allows changes to be made quickly and is not locked in to a particular product. However, in the future selection of a database product may be considered, and the Publications Office is interested to see if there are any existing registry products that could be used.
Standards	Processes are inspired by the ISO/IEC **standard** 11179 (Information technology -- Metadata registries (MDR)) but implemented in a restricted way as the whole approach was considered too complicated. The source format of authority tables is XML with a subset exported to SKOS for CELLAR. In the future, an approach based on a richer RDF format may be considered. Metadata is exported in ADMS for Joinup and in the description on the EU Open Data Portal, and is generated for every publication.
Costs	Total human resources for metadata management are equivalent to 5 FTEs which include the role of IMMC presidency and secretariat.
Benefits	Benefits include improvement in quality, better interoperability, more efficient use of resources through single-point maintenance. However, a challenge is that the collaboration across institutions leads to more interdependency. Overall, the benefits of the approach to metadata management offset the cost. However there is no quantitative information to support this.
Tool support	
Name of the tool(s) and documentation	**Atlassian JIRA**[73]

[73] https://www.atlassian.com/software/jira

Reusability	**Commercial product**
Key functionality	Governance workflow
Costs	From US$10 for 10 users to US$42.000 for 10.000+ users
Name of the tool(s) and documentation	**Apache Subversion (SVN)[74]**
Reusability	**Open source**
Key functionality	Versioning of assets maintained in MDR
Costs	None
Name of the tool(s) and documentation	**XSLT/PERL scripts**
Reusability	No
Key functionality	File conversion, generation of diff files, validation
Costs	

Member States

Case study 4: KoSIT[75] (Koordinierungsstelle für IT-Standards), Germany

In the past Germany faced a crucial need of uniform data interchange between the 5,400 German municipal citizen registers and many electronic data exchange processes between public authorities. So OSCI-XMeld[76] was created, which aim was to efficiently develop a specification for the standardized, XML-based electronic data exchange between all German municipal citizen registers. This project was so successful, that the German federal and state governments officially started to develop XML-based e-government data interchange specifications (so-called XÖV standards) an official recommendation in 2006.

To address the interoperability issues that came up naturally with a growing number of XÖV specifications, the standardization of e-government standards came into focus. In September 2010, the federal and state government established the permanent IT Planning Council[77] (IT-Planungsrat), whose IT standards coordination office KoSIT (Koordinierungsstelle für IT-Standards) is now (among other responsibilities) in charge of assuring and operating a uniform model-driven development method and corresponding tools for currently 17 XÖV specifications in several public administration domains, with further projects on the horizon.

XÖV framework[78] functionality is based on specific architecture which is constructed from these elements:

- Several UML models[79] and further semantic artefacts, some belonging to the owners of the individual XÖV standards, some being shared using semantic artefacts maintained by the IT coordination office;

[74] http://subversion.apache.org/

[75] For more information about KoSIT: http://www.xoev.de/sixcms/detail.php?gsid=bremen83.c.8159.de

[76] In German, `Meld' abbreviates Meldewesen, the domain of citizen registration

[77] For more details about IT-Planungsrat: http://www.it-planungsrat.de/DE/ITPlanungsrat/itPlanungsrat_node.html

[78] More info about XÖV Framework: http://www.xoev.de/sixcms/detail.php?gsid=bremen83.c.4995.de

[79] Unified Modelling Language (UML): http://www.omg.org/spec/UML/2.5/Beta2/PDF/

- UML profiles defining domain-specific metadata and governing that the models conform to given interoperability criteria for XÖV;
- A common, configurable open source tool, the XGenerator[80], that automatically validates the models against the profiles and transforms them into various artefacts that make up a data interchange specification (e.g., documentation fragments, XML Schema files, web service description files);
- A central, web-based repository, the XRepository[81], which holds the various structural metadata.

XÖV framework provides some standards which are interesting in a pan-European context. However, there are rules and recommendations in that framework which are seen as barriers for those standards to act internationally, such as the requirement that the main language should be German. Also some standards in the e-procurement area could be used in an international context, but such standards would have to be available also in other languages, at least in English.

KoSIT has the general role for maintaining the XÖV framework. This organisation is also responsible for gathering the requirements of different stakeholders, monitoring the specification landscape and for performing XÖV certification[82]. Before a specification can be released as an XÖV specification, it has to be verified whether the specification meets all organisational, semantic, and technical requirements.

Nevertheless KoSIT has no authority to impose XÖV standards. It is the IT Planning Council who decides which standards to make mandatory in case they are not domain specific.

In the remainder of this section, we present the summary of the interview with KoSIT conducted in the context of this case study. The summary is structured according to the analysis framework of section 2.1.2.

Interview date	23th January 2014
Interviewee	Lutz Rabe
Interviewer	Audrius Leipus, Nikolaos Loutas

Title: Koordinierungsstelle für IT-Standards (KoSIT)	
Governance	

[80] For more details: http://www.xoev.de/sixcms/detail.php?gsid=bremen83.c.11551.de

[81] XRepository: https://www.xrepository.de/

[82] More information about XÖV certification: http://www.xoev.de/sixcms/detail.php?gsid=bremen83.c.5032.de

Goals	• KoSIT aims to increase the degree of reuse of all building blocks that they provide, to increase efficiency in their development and to increase the quality, because they use concepts that are already in use and proven by the practise. • The main idea is to increase the interoperability between different standards such as the overall interoperability and particular areas of e-Government.
Governance structure	**Roles:** • **KoSIT** operates the XÖV Framework for the IT-Planungsrat with collaboration with different stakeholders from the government (maintains XÖV framework, gathers the requirements of different stakeholders, and monitors the specification landscape, responsible for certifications of XÖV). KoSIT also adopts their products and building blocks in respect to requirements coming from the projects. • **A third party** is responsible for XÖV certification. They certify new standards, which describe how to reuse shared semantic components and provide the technical representation of the models as XML schemata, by testing their conformity against KoSIT's namingand design rules (conformance testing). In the end standard usability is verified by KoSIT, based on results of certification process. Even if a standard has an XÖV certificate it does not mean that the usage of this standard is mandatory. • **IT-Planungsrat (IT Planning Council)** is in charge of providing and operating a uniform model-driven development method and corresponding tools for currently 17 XÖV specifications in several public administration domains, e.g. making a particular character set mandatory for all the administrations in Germany. If the area is domain specific (some specific information exchange, i.e. financial) the decision is made by the conference of ministers. **Openness:** There is no formal restriction to particular stakeholders from the government. Any stakeholders can make a request to be part of this structure and participate in defining requirements and similar activities, as long as they are involved in IT system development, infrastructure or other related areas. **Decision making process:** Stakeholders define requirements, which are then assessed by KoSIT. This process is now being made public to do a pre-assessment of the defined requirements and publish possible solutions which then could be assessed by other stakeholders of the community and based on that, changes would be made within the next release of the particular part of the framework. Part of this process is published on the website. The whole framework will be described and defined in the XÖV handbook[83] which will be published with the complete overhauled framework at the beginning of May, 2014. The involvement of the stakeholders is not always present on particular complex problems. This makes it hard to decide if the solution is good or not, before it is launched to the projects.

[83] The XÖV handbook: http://www.xoev.de/sixcms/media.php/13/X%D6V-HandbuchV1_1.pdf

	Feedback from projects comes only after it is published officially.
Context	**Organisational dimension:** • Inter-organisational. **Administrative dimension:** • National and regional / local.
Policy domain	Metadata governance is not restricted to a policy domain.
Enforcement policy	**Sharing and Reuse:** It is defined in the criteria that XÖV conform with to share constructs and concepts and also to reuse the structural metadata that they provide. It is neither a legal requirement, nor it is voluntary. For particular standards developed in XÖV framework sharing and reuse are determined by legal requirements. For example, standards on civil registration are mandatory by law. Therefore sharing and reuse is partially enforced by law. **Comply-or-explain** model is used if there are no legal requirements.
Authoritative source	XRepository
Licensing framework	**No explicit licence**: There are none explicitly defined licences, except some requirements that the standards and metadata should belong to the German administrations and should be freely available for everyone, cost free, published on the XRepository as well as other similar requirements.
Quality controls	Quality controls are not well defined. KoSIT is operating the framework while standards are operated by other stakeholders. There are 5 XÖV standards, used in the process of quality testing.
Metadata schema	**Vocabulary:** Depends on the artefact. KoSIT is using a few different metadata descriptors. The descriptions of all structural metadata available on XRepository can also be exported in ADMS in order to enable the interchange with other national repositories. Metadata on the code lists is defined by OASIS[84] (based on the Genericode standard). Also they use this metadata standard for publishing information on code lists but they extended the metadata schema with some descriptors particular to XÖV. **Identifiers scheme:** KoSIT does not have guidelines for creating identifiers for the metadata. They use specific schemas which can identify particular artefacts in standards. **Schema documentation:** This documentation is generated directly from the UML model. The documentation is in the form of a PDF document and is located on the XRepository[85]. **Multilingualism:** The primary language for all the metadata properties and values is German. Support of other languages is only in its initial stage, but in some cases it is possible to use English to describe structural metadata.
Management	
Update	**Update frequency:** Based on the demand. A release of the complete framework is done every one and a half or two years. Also there is specific metadata within the standard that is not

[84] Organization for the Advancement of Structured Information Standards (OASIS): http://docs.oasis-open.org/codelist/cs01-ContextValueAssociation-1.0/doc/context-value-association-us.pdf

[85] XRepository: https://www.xrepository.de/

	under the control of KoSIT. This metadata is controlled by other organisations that decide on their own, how many releases they need. **Change management process:** Decisions on what to change and how to change are made based on the feedback from the projects and from the key system providers. **Version control:** There is version control for separate parts of the framework. The core components have a particular versioning system. Other parts like code lists, rules, handbooks, individual products have individual versioning systems.
Harmonization	**Creating mappings between related metadata sets:** No explicit process defined for that. It mainly depends on the standards. **Assessing alternative metadata sets:** As long as the sets are compliant to their regulation they are not involved in assessing the alternatives. It is done by the users (system vendors or stakeholders). Cases where there are two or more standards for the same application area or scenario are possible. KoSIT do not deal with the quality of the content of the metadata, but instead they concentrate on the quality and the processes of operating and related topics. Depending on the specific requirements of an information exchange scenario they also have recommendations on possible code lists and other similar topics.
Documentation	**Publication:** Structural metadata is documented according to a common vocabulary and located on the XRepository, which supports search functionality of it. **Retrieval:** The retrieval of metadata is possible by both humans and machines. However, search capabilities for machines are limited. **Supported formats of publication tool(s):** • Human-readable only (H): pdf. • Human and Machine readable (HM): UML, XML Schema, ADMS, Genericode.
Standards	XÖV standards are used for metadata management with the aim to ensure consistent XML-based e-government data interchange. Genericode is used to define standard model for defining code list.
Functionality of the tool(s)	XRepository is used for publication. XGenerator is used to create different artefacts of the standard (Schema, DocBook documentation used for creating pdf, artefacts based on WSDL standard and Genericode code list, etc.)
Costs	At least one and a half or two FTEs are directly employed. Also they have the budget from the IT-Planungsrat that can be used for hiring external experts. These external experts add up about three FTEs.
Benefits	The key benefit is the interoperability. The benefits have to be measured on two levels: • The first level: the particular area of the individual standard and the benefits in data exchange in this application area. • The second level: benefits across application areas are expected; for example, the benefit of increasing the interoperability between the German administrations, increasing the quality of the data objects and tools they are providing. They believe that the benefits offset the costs.

Tool support	
Name of the tool(s) and documentation	**Xgenerator**[86]
Reusability	Xgenerator is configurable open source tool, which is developed by KoSIT and external suppliers. This tool is distributed under European Union Public Licence (EUPL) and GNU General Public License (GPL) 2.0, so it can be reused. The only restriction is that current version of the XGenerator only works with MagicDraw 16.5.
Key functionality	This tool automatically validates the models against the roles and transforms them into various artefacts that make up a data interchange specification (e.g., documentation fragments, XML Schema files and web service description files).
Name of the tool(s) and documentation	**Genericoder**[87]
Reusability	Genericoder is developed by KoSIT and external suppliers under European Union Public Licence. This tool is open source and can be reused without any constrains.
Key functionality	It is the central tool for production. It converts CSV code lists into OASIS Genericode Version 1.0 compliant XML files and validates individual artefacts according by rules defined in the framework.
Name of the tool(s) and documentation	**XRepository**[88]
Reusability	XRepository is developed by KoSIT in collaboration with external suppliers. There is no explicit license on the use and the XRepository platform software is open source (it will be available for download on Joinup in the near future). There are no known constraints for reuse.
Key functionality	It is infrastructure component for publishing, documenting and providing both artefacts provided by the XÖV-Framework and the artefacts of the XÖV-standards. It simply holds the various semantic assets of all standards.
Name of the tool(s) and documentation	**Production environment components**
Reusability	Production environment components are developed by KoSIT and mainly developed under open source technologies, so they are open and can be reused without any constrains.
Key functionality	These components are used for the technical production of the standards, documentation and artefacts related to the standards (e.g. to produce the DocBook chapters which are directly generated from the model to XML).
Total cost of tools	Development and operation of the tools cost around 150 000 euros per annum including administration of all the users and artefacts.

Case study 5: CISE – Centre for Semantic Interoperability, Spain

The Centre for Semantic Interoperability (Centro de Interoperabilidad Semántica – CISE) is the instrument defined in the Spanish National Interoperability Framework

[86] XGenerator: https://joinup.ec.europa.eu/software/xgenerator/description

[87] Genericoder: https://joinup.ec.europa.eu/software/genericoder/description

[88] XRepository: https://joinup.ec.europa.eu/sites/default/files/Case%20Study%20-%20XRepository%20semantic%20asset%20repository.pdf

to publish data models and encodings associated with the exchange of data between the different administrations.

The legal basis of the work of CISE is in three documents:

- The Law 11/2007, of 22 June, on electronic access to Public Services for members of the public[89];
- Royal Decree 4/2010, of 8 January, which regulates the National Interoperability Framework within the e-government scope[90]; and
- Resolution of the Secretary of State for Public Administration of 28 June 2012, giving approval to the Technical Interoperability Standard for Data Models.[91]

The role of CISE is specified in the Resolution above as the organisation responsible for the publication of data models used in the public administration, including:

a) Structural metadata (data models) in XSD (XML Schema Definition) format, classified by service or business unit of the corresponding agency.
b) Explanation guides in PDF (Portable Document Format) format, in compliance with the Technical Interoperability Standard for Standard Catalogues, for the various exchange systems or services, including:
    i. Description of exchanged data types and definitions under the data model in question, and functional description of the operations that can be performed.
    ii. Brief description of the security measures applicable to the data model's exchanges.
    iii. Requirements to be met by recipients of the information the model applies to.
    iv. Examples of service implementation under the data model in question.
    v. User's manuals for exchange services and test kits (optional).

Working towards a longer-term goal is to identify common data models for use across the public administration, three phases have been defined[92]:

- In the current first phase of its operation, CISE works to collect and disseminate different data models of different administrations and public bodies in the Semantic Asset Manager (GAS). Some of these models may

[89] For more details about this law: http://www.boe.es/buscar/doc.php?id=BOE-A-2007-12352, English translation available at: http://administracionelectronica.gob.es/pae_Home/dms/pae_Home/documentos/Documentacion/pae_BIBLIOTECA_NORMATIVA_ESTATAL_Leyes/LAW_11-2007_22Jun2007_e-Gov_Spain_NIPO_000-10-075-0.pdf

[90] For more details about this law: http://www.boe.es/buscar/act.php?id=BOE-A-2010-1331, English translation available at: http://administracionelectronica.gob.es/pae_Home/dms/pae_Home/documentos/Documentacion/pae_BIBLIOTECA_NORMATIVA_ESTATAL_Leyes/Royal_Decree_4_2010_Interoperability_framework_NIPO_000-10-058-X.pdf

[91] For more details about this law: http://www.boe.es/diario_boe/txt.php?id=BOE-A-2012-10050; English translation available at:http://administracionelectronica.gob.es/pae_Home/dms/pae_Home/documentos/Documentacion/pae_BIBLIOTECA_NORMATIVA_ESTATAL_Otras_disposiciones_relevantes/2012-10050_TIS-data-models_NIPO-630-12-208-1.pdf

[92] http://administracionelectronica.gob.es/pae_Home/pae_Estrategias/pae_Interoperabilidad_Inicio/pae_Centro_Interoperabilidad_semantica/pae_Centro_de_Interoperabilidad_Semantica_(CISE)_-_Cual_es_la_filosofia_del_CISE_.html

contribute to a set of common models to be defined in the second phase; however, until a critical mass of data models is reached, it cannot be determined which of these models can be considered common.

- In the second phase, after a critical mass of data models has been achieved, the models published through CISE will be analysed and, with the participation of the different bodies involved, a definition of common data models will be proposed to serve as references for the exchange of information between administrations.
- In a subsequent third phase, the models that have been accepted as common and published as such by CISE will become mandatory as stipulated in article 10.2 of the Royal Decree 4/2010 which reads:

  Public Administration bodies or Public Law Entities linked or depending on them, holders of competences with regard to information exchange with citizens and with other Public Administrations, as well as in terms of common infrastructures, services and tools, will establish and publish the corresponding interchange data models that will be of mandatory application for information interchanges in Public Administrations.

In order for CISE to be able to publish a data model, the service provider needs to provide additional information[93]:

1. Whether the service is capable of exchanging information using a specific data model.
2. What the conditions are under which the information can be used. This includes information on the potential users and the technical requirements for use.
3. Which data structures can be obtained from the service provided and how the exchange of data is organised. To do this, it is sufficient to know the XSDs with appropriate explanations.

In the remainder of this section, we present the summary of the interview with the CISE conducted in the context of this case study. The summary is structured according to the analysis framework of section 2.1.2.

Interview date	14 January 2014 (face-to-face)
Interviewee	Francisco José Martín Lázaro
Interviewer	Makx Dekkers

Title: CISE – Centre for Semantic Interoperability, Spain	
Governance	
Goals	The ultimate **goal** of CISE is to implement the vision of the legislator to define and publish common data models that will be mandatory across the public administration to improve interoperability across all sectors and services.

[93] For more details about CISE:
http://administracionelectronica.gob.es/pae_Home/pae_Estrategias/pae_Interoperabilidad_Inicio/pae_Centro_Interoperabilidad_semantica/pae_Centro_de_Interoperabilidad_Semantica_CISE_1_QUE_ES_EL_CENTRO_DE_INTEROPERABILIDAD_SEMANTICA_CISE.html

Governance structure	**Roles:** The **role** of CISE as the organisation that collects and publishes the data models is defined in the Resolution of the Secretary of State for Public Administration of 28 June 2012. **Openness**, the various public bodies will be involved in the definition of common models in a future second phase. **Decision making processes:** Currently, the **decisions** on making changes to data models are taken by the creators and the owners of the models that are harvested by CISE.
Context	**Organisational dimension:** • **inter-organisational**. **Administrative dimension:** • **National**.
Policy domain	The policy domain is **not restricted**. However, it is expected that an important role in establishing the common models will be played by the domains that exchange the most data, in particular in the areas of taxes and social security.
Enforcement policy	**Sharing** of data models is **legally required** on the basis of the Resolution of the Secretary of State for Public Administration of 28 June 2012 which requires public administrations to provide their data models to CISE. **Re-use** of the common models that will be identified in phase 2 is **legally mandated** in the Royal Decree 4/2010. It is expected that implementation of this legal obligation will be based on a **'comply-or-explain'** approach, to take into account sectors that may be required by law to use other models (e.g. INSPIRE)
Authoritative source	CISE publishes the data models in the Semantic Asset Manager (GAS[94]) which is the **authoritative source** for the common data models. For the non-common models that are submitted to CISE, the authoritative source is the institution that created and maintains those models.
Licensing framework	Models published by CISE have **no explicit licence**, but a specific licence will be specified for the common models.
Quality controls	As the responsibility for the models published by CISE at the moment lies with the originating institution, the **quality control** is also a distributed responsibility. CISE makes sure, through frequent harvesting, that the models published at CISE are synchronised with the source.
Metadata schema	**Vocabulary:** The **vocabulary** if the description of data models is defined in the Resolution of the Secretary of State for Public Administration of 28 June 2012: name of the data model, body that makes the data model available, statistical relevance, location (URI and Web service) and contact e-mail. Extension of this list is being considered. The Technology Transfer Centre (CTT) exports descriptions in ADMS to Joinup. **Identifiers scheme:** identifiers for public sector bodies are assigned in the Common Directory of Organical Units and Offices (DIR3[95]) where all bodies need to be registered. **Schema documentation** for all data models is available in PDF as required by the Resolution. **Multilingualism:** There is **no multilingualism**: All information is

[94] CISE's Semantic Asset Manager GAS: http://cise.redsara.es/SGAS/Visor.jsp

[95] More information about DIR3: http://administracionelectronica.gob.es/ctt/verPestanaGeneral.htm?idIniciativa=dir3#.UtrdixA1iUk

	available only in Spanish.
Management	
Update	**Update frequency:** depends on the particular data model as these are managed by the originating bodies. CISE harvests the data models frequently. **Change management process and versioning:** are not yet in place and will be developed for CISE's second phase.
Harmonization	**Harmonisation** is not yet done. It will be the objective of phase 2 when common data models will be proposed.
Documentation	**Publication** of the data models is done on the Semantic Asset Manager. This tool has no search functionality but allows navigation through the definitions of the data models. **Retrieval** is supported by GAS as it gives access to the XSD documents as well as access to the documentation in PDF for manual download. The material is available both in **human-readable (PDF) and machine-readable (XSD)** formats.
Functionality of the tool(s)	Semantic Asset Manager for **publication and retrieval**. XSLT and PERL scripts for **harvesting and processing**.
Standards	**No specific standard** is used for metadata management. For value lists, SKOS may be considered but with XSD as the publication format.
Costs	Human resources for the management of the data models at CISE are between 1 and 2 FTEs**.**
Benefits	Currently, the benefit is in the central publication of the data models so that they can be accessed at a central location. The benefit of the common models to be established in phase 2 is improved interoperability.
Tool support	
Name of the tool(s) and documentation	**Semantic Asset Manager, Gestor de Activos Semánticos (GAS)**
Reusability	**Open source, developed in-house.** The tool is written in J2EE with Oracle as database. In principle**,** it could be used by others but it is not currently foreseen to make it available. This could change in the future if considered useful.
Key functionality	Publication and retrieval
Costs	Tool was developed for an earlier project (INDALO[96]) and was used for various applications which make it impossible to attribute cost to CISE.
Name of the tool(s) and documentation	**XSLT and PERL scripts**
Reusability	**Developed in-house**
Key functionality	Harvesting and processing
Costs	Not available

[96] Proyecto INDALO: http://administracionelectronica.gob.es/pae_Home/pae_Estrategias/pae_Interoperabilidad_Inicio/pae_Centro_Interoperabilidad_semantica/pae_Centro_de_Interoperabilidad_Semantica_(CISE)_-_ANTECEDENTES_DEL_CISE__PROYECTO_INDALO.html #.UuaOoRA1iCg

Case study 6: Lithuanian Spatial Information Portal (LSIP), Lithuania

The LSIP[97] initiative has started as part of Phare[98] project in 2003 when the Lithuanian government has decided that the national solution for collecting and sharing spatial data is a priority and has to become one of the national goals. After a long legislation process it became Lithuania's national geo portal.

Later on (15 May 2007), Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE Directive) came into force. In an effort to ensure the implementation of the provisions of the Directive, the National Land Service under the Ministry of Agriculture of Lithuania (hereinafter 'the NLS') was designated as a body in charge of the creation of infrastructure facilities for ensuring the functionality of metadata corresponding to the themes referred to in INSPIRE Directive, data sets, network services, sharing services and access via the LSIP. This is why the creation and implementation of LSIP was initiated.

In terms of metadata governance structure at LSIP, there are three levels starting from the highest - The Ministry of Agriculture of the Republic of Lithuania, then middle the NLS and the lower one – Geoinformatics GIS-Centras (hereinafter 'GIS-Centras').

The National Centre of Remote Sensing and 'GIS-Centras' is a state enterprise and the LSIP manager provided for in the Law of the Republic of Lithuania on Geodesy and Cartography.

'GIS-Centras' manages the information structure and general contents, collects and processes metadata, stores it and initiates its regular updating, creates electronic data and metadata services and provides them to users. The LSIP manager ensures correct registration, management and transfer of all information published on the LSIP and its protection from illegal actions. The LSIP manager has the right to remove false information or information not meeting requirements from the LSIP without warning.

LSIP performs the following functions:

- collect metadata of spatial data of state cadastres, registers, state and municipal authorities and other organisations according to a methodology approved by the portal manager;
- provide metadata and spatial data sets to users on the Internet;
- provide services related to spatial data sets to users on the Internet;
- publicise and promote the use of spatial data for solving public administration sector problems; to provide information on the possibilities for the public and private business entities to effectively use spatial data controlled by the public sector;
- create technological conditions for linking spatial data sets managed by different authorities and using them together.

[97] Lithuanian Spatial Information Portal (LSIP): www.geoportal.lt
[98] Link to Phare Programme: http://europa.eu/legislation_summaries/enlargement/2004_and_2007_enlargement/e50004_en.htm

Apart from the INSPIRE Directive, LSIP has applied a number of ISO standards ISO19115[99], ISO19119[100], ISO19139[101]. Geographic data providers have to prepare metadata using National Metadata Profile corresponding to ISO 19115 (Geographic information – Metadata) and to ISO 19119 (Geographic information – services) standards. The metadata answers to main questions, such as: what (data name and description), when (date of data creation and update), who (data author and supplier), where (coverage of data), how (data purchase possibilities). LGII metadata editor is based on ISO 19139 (Geographic information – Metadata – XML schema implementation) standard.

What is ISO 19139 (Geographic information Metadata XML schema implementation)?

It provides the XML implementation schema for ISO 19115 specifying the metadata record format and may be used to describe, validate, and exchange geospatial metadata prepared in XML.
The standard is part of the ISO geographic information suite of standards (19100 series), and provides a spatial metadata XML (spatial metadata eXtensible Mark-up Language (smXML)) encoding, an XML schema implementation derived from ISO 19115, Geographic information – Metadata. The metadata includes information about the identification, constraint, extent, quality, spatial and temporal reference, distribution, lineage, and maintenance of the digital geographic dataset.

Metadata could be submitted to the Geoportal in one of three ways:

- manually entering information using online metadata editor application;
- uploading XML file directly using online interface of LGII metadata system; or
- by harvesting (automated procedure that connects directly to providers' database and reads metadata file that is copied over old metadata file identified by the same unique identifier).

The types and the frequency of the updates of the metadata depend on each data resource and are performed by the data provider. Significant updates have to be registered in the Geoportal via a harvesting procedure or manually.

The metadata editor has integrated tools for metadata core element validation. After initial automated validation procedure metadata administrator checks metadata for semantic correctness against the national directive 1P-(1.3.)-295, which is based on the ISPIRE directive, and approves or rejects it correspondingly.

Search based on the structural metadata is available on the portal. Numerous search parameters, such as provider, data theme, territory or time period can be used. Once geographic data product is discovered, the user is given options to browse short summary or full metadata for that product. Metadata can be included in a list for comparison.

[99] Geographic information – Metadata standard (ISO19115): http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020

[100] Geographic information – Services standard (ISO19119): http ://www.iso.org/iso/catalogue_detail.htm ?csnumber=39890

[101] Geographic information – XML schema implementation standard (ISO19139): http://www.iso.org/iso/catalogue_detail.htm?csnumber=32557

Structural metadata are provided free of charge and publicly. For specific cases, there is a possibility for data provider to limit metadata access to particular users' groups.

The major difficulties related to the sharing of structural metadata are due to methodological and organisational problems, not due to technological limitations. According to the Geographic information – Metadata standard (ISO 19115), metadata extent is defined by a single rectangle that misleads the user in case of fragmented datasets. Although it is time and resource consuming to create legal acts and organisational procedures for sharing structural metadata, such procedures are necessary to ensure correctness and timely update of metadata.

LSIP is using some of commercial tools e.g. ESRI Geoportal and in the near future plan to implement Portal for ArcGIS and ArcGIS for INSPIRE.

What is ArcGIS for INSPIRE?

The INSPIRE Directive sets out a framework and timetable for sharing spatial data within the European Community to help address pan-European issues in a multinational and multiagency spatial data infrastructure (SDI). ArcGIS provides a powerful and comprehensive SDI solution that now includes capabilities to ensure INSPIRE compliance supporting data, services, and metadata, which are delivered in the new ArcGIS for INSPIRE.
ArcGIS for INSPIRE helps meet INSPIRE compliance in a timely manner by extending the ArcGIS platform that already exists in your organisation.

In the remainder of this section, we present the summary of the interview with LSIP conducted in the context of this case study. The summary is structured according to the analysis framework of section 2.1.2.

Interview date	21st January 2014 (face-to-face)
Interviewee	Raminta Vitkauskienė, Giedrė Beconytė, Danas Motiejauskas
Interviewer	Audrius Leipus

Title: Lithuanian Spatial Information Portal – LSIP, Lithuania	
Governance	
Goals	• To increase the efficiency of state's spatial data use by creation of new interactive online services; • To allow easy access or search for spatial data sets or related e. services; • Provision of the structure of spatial data set or service in advance so that evaluation of properties terms of use is possible; • Expansion of the network of data providers and online services, increase in efficiency of state's spatial data use by creation of new interactive online services.

Governance structure	**Roles:** There are three governance levels in general. The Ministry of Agriculture of the Republic of Lithuania (MoA) is responsible for assuring the compliance of LSIP with the INSPIRE Directive. NLS is in charge of the legal and organisational procedures for metadata management. 'GIS-Centras' is responsible for spatial data infrastructure, new technological solutions development of metadata documents and the lifecycle of metadata. **Openness:** The governance structure has a limited access and those related and affected can be invited on a request basis. **Decision making process:** Depends on the decision to be made where it can be taken to MoA or left at the lower level NLS or 'GIS-centras'.
Context	**Organisational dimension:** • **Inter-organisational;** **Administrative dimension:** • **Pan-European** (data provision based on INSPIRE Directive); • **National.**
Policy domain	Restricted to the national policy domain covered by the INSPIRE Directive, i.e. environment and related domains.
Enforcement policy	**Share:** Legal requirement – provision and requirements of infrastructure of spatial data are approved by the National Metadata Profile[102] (NMDP) (metadata is provided free of charge and openly) which applies to all LSIP metadata providers. **Reuse:** Comply-or explain. MoA was charged with drafting a law amending the Law of the Republic of Lithuania on Geodesy and Cartography and implementing legislation for transposing the provisions of the INSPIRE Directive (Directive 2007/2/EC of the European Parliament and of the Council establishing an Infrastructure for INSPIRE) into the legal system of the Republic of Lithuania.
Authoritative source	LSIP (www.geoportal.lt).
Licensing framework	There is no defined licence.
Quality controls	'GIS-Centras' and data suppliers are responsible for the quality of data and network services in accordance with the NMDP. Quality management is a part of 'GIS-Centras' spatial data service creation process.
Metadata schema	**Vocabulary:** Exists and is defined in NMPD which is based on ISO19115:2003, ISO19119:2005, ISO19139:2007. **Identifiers scheme:** Exists and is based on ISO19115. **Schema documentation:** Exists, based on XSD[103] **Multilingualism:** Available languages Lithuanian and English according to INSPIRE Directive requirements. However, the tool allows using other languages such as German, French and Russian.
Management	

[102] Constructed National Metadata Profile (NMDP) corresponds to ISO 19115: http://www.geoportal.lt/wps/portal/!ut/p/c0/04_SB8K8xLLM9MSSzPy8xBz9CP0os_gAQwNnc09LYwMLA3dzA08D8yB_E4NAA3dLQ_3gnBL9gmxHRQCF_ZHY/

[103] W3C XML Schema Definition Language (XSD): http://www.w3.org/TR/xmlschema11-1/

Update	**Update frequency:** Metadata is updated only when spatial data is updated, in most cases once a month (e.g. Lithuania territory M 1:10 000 geo-spatial data set (GDR10LT)). **Change management process:** Changes on metadata can be made by provider and must be approved in the system by metadata administrator. **Version control:** Does not exist, documents are simply updated. However version control might be on the metadata provider's side.
Harmonization	**Creating mappings between related metadata sets:** Exists and is based on the INSPIRE Directive. **Assessing alternative metadata sets:** There are two metadata profiles (NMPD and INSPRE[104]) which are applied according to the intended use. Other metadata sets are converted to one of these two profiles.
Documentation	**Publication:** Metadata is documented according to a common vocabulary and search functionality in metadata is available. Publication of metadata is done in manual and semi-manual ways. **Retrieval:** Retrieval can be done by both humans and machines. **Supported formats of publication tool(s):** Human readable (e.g. in PDF, JPEG, PNG formats), Machine readable (e.g. xml, csv, etc.)
Standards	NMDP and INSPIRE profiles that are based on ISO 19115:2003, ISO 19119:2005 and ISO 19139:2007.
Functionality of the tool(s)	All functionalities as per above are supported by the tools (LSIP metadata editor based on ESRI Geoportal).
Costs	The cost is about 1 FTE.
Benefits	Exact evaluation of benefits was not made, possible IRR (internal rate of return) is 30% and thus benefits offset the costs. The main benefits are: • For the public: access to relevant spatial data accumulated by public authorities and improved quality of public services. • For the business: possibilities to include the latest data in a common information system, for example, local restaurants, hotels and filling stations. In addition, supply various value-added products to the market. • For the public sector: possibilities to use data stored by other authorities and agencies as well as increase work efficiency in performing public administration functions and providing services to citizens.
Tool support	

[104] Draft Implementing Rules for Metadata (Version 3): http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/INSPIRE_Metadata_ImplementingRule_v3_20071026.pdf

Name of the tool(s) and documentation	The tools are developed based on geoportal.lt needs using ESRI Geoportal extension[105] and ESRI ArcIMS[106] programming tools, API[107]. Same of java Libraries that are used: • Apache XMLBeans (compile XML schemes to Java objects); • Apache Xerces (XML parser); • Xalan (XSLT transformation); • Saxon (XSTL and XQuery processor). In the near future (2014-2015) they plan to implement Portal for ArcGIS[108] and ArcGIS for INSPIRE[109].
Reusability	**Open source:** Java libraries, some parts of ESRI Geoportal Toolkit. **Open licence:** No open licence. **Owner / vendor:** Creation of LSIP (including tools) was funded by EU Structural Funds for Lithuanian geographical information infrastructure (Project No. BPD2004-ERPF-3.3.0-02-04/0014[110]) Tools are accessible online on LSIP and used by different subjects (i.e. metadata editor). There has been no need so far to be reused somewhere else.
Key functionality	The main functionality supported by the tools : • Create, manage and publish metadata documents; • Manage the search in metadata; • Automatically upload metadata from other remote servers; • Verify metadata compliance with NMPD metadata profiles;
Costs	The tools for metadata management were developed in parallel with other LSIP tools and systems, so there is no detail evaluation however it did cost approximately 0.6M EUR.

Case study 7: Knowledge and Exploitation Centre Official Government Publications (KOOP)[111], The Netherlands

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.

[105] ESRI Geoportal Extension (Toolkit): http://webhelp.esri.com/geoportal_extension/9.3.1/

[106] ESRI ArcIMS: http://webhelp.esri.com/arcims/9.3/General/arcims_help.htm

[107] Application Programming Interface (API): http://resources.esri.com/help/9.3/Geoportal/JavaDoc/index.html?help-doc.html

[108] ArcGis: http://resources.esri.com/help/9.3/Geoportal/JavaDoc/index.html?help-doc.html

[109] ArcGis for INSPIRE: http://resources.arcgis.com/en/communities/arcgis-for-inspire/

[110] LSIP project: http://www.gis-centras.lt/gisweb/index.php?pageid=338

[111] Link to KOOP: http://koop.overheid.nl/over-koop

In addition, KOOP has developed various products designed for local authorities that enable them to publish their legislation, announcements and permits electronically. The use of national standards allows this to be done in a uniform and cost-effective way. The data is easily accessible and reusable for other applications (open data).

KOOP intends to continue its role as the driving force in the area of government publication services, improving the electronic publication of legislation. KOOP does this by developing innovative products and services that contribute to increased transparency and service-orientation of the government.

OWMS (*Overheid.nl Web Metadata Standard*)[112]
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 concerning[113]:

- 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.

Other metadata standard implementations
For the description of datasets on Overheid.nl KOOP is involved in a project that uses an implementation of the DCAT vocabulary.

Geo-information is described using the Dutch national standard NEN-19115 which is the national version of ISO 19115. The schemas and documentation for the use of this standard are maintained by another organisation, Geonovum.
A standard for record management is under development.

In the remainder of this section, we present the summary of the interview with KOOP conducted in the context of this case study. The summary is structured according to the analysis framework of section 2.1.2.

Interview date	13 January 2014
Interviewee	Hans Overbeek
Interviewer	Makx Dekkers

Title: Knowledge and Exploitation Centre Official Government Publications (KOOP), the Netherlands	
Governance	
Goals	The **goal** of the metadata governance and management activities is to improve coherence and findability of government information

[112] More about Overheid.nl Web Metadata Standard: http://koop.overheid.nl/producten/owms-overheidnl-web-metadata-standaard

[113] For more information: http://standaarden.overheid.nl/owms

	to save costs, improve quality and thereby trust of citizens, companies and other government bodies in this information.
Governance structure	The governance and management is described in the OWMS maintenance plan[114]. **Roles:** The main **roles** are: • The Ministry of the Interior as the overall authority with final decision power • KOOP as the operational service • The OWMS User Council comprised of agencies and vendors that use and/or implement OWMS which takes decisions • The OWMS Community, which is the source for change requests **Openness:** The **openness** of the process is ensured through the OWMS Community that is open to all interested organisations and individuals. **Decision making process:** The **decision process** starts with proposals originating in the OWMS community, followed by discussion in the OWMS User Council. Most decisions are taken in the User Council, except for main decisions that imply major changes in the standard which are taken by the Ministry. In practice, the process is mostly initiated by KOOP. After a period of inactivity, the Community was discontinued in 2013.
Context	**Organisational dimension:** • **Inter-organisational.** **Administrative dimension:** • **National**.
Policy domain	**Not restricted** by domain, but covering all information objects from Dutch government agencies as specified by the national archive law. In practice, coverage is not complete for all sectors.
Enforcement policy	**Sharing** of models and reference data developed in various agencies is **voluntary** and not enforced, although this is encouraged by the Forum Standaardisatie[115], a government agency under the Ministry of the Interior and the Ministry of Economic Affairs that provides advice to several ministries on interoperability and maintains a comply-or-explain[116] list of standards relevant for the public sector. Whether or not value lists need to be published and maintained centrally alongside the OWMS standard schemas is still under discussion. **Reuse:** application of OWMS is stimulated by a **comply-or-explain** approach. Value lists are being re-used by various agencies but KOOP does not have a list of their users.
Authoritative source	OWMS is published on standaarden.overheid.nl/owms[117] which is the **authoritative source**. Vocabularies are only partly published centrally and available for re-use. In the future, a registry of

[114] Link to OWMS maintenance plan: http://standaarden.overheid.nl/owms/beheer/BeheerplanOWMSv1.0.pdf

[115] Link to Forum Standaardisatie: http://www.forumstandaardisatie.nl/organisatie/forum-standaardisatie/

[116] Link to this list: https://lijsten.forumstandaardisatie.nl/lijsten/open-standaarden?lijst=Pas%20toe%20of%20leg%20uit&status%5B%5D=Opgenomen&pagetitle=pastoeof

[117] Link: http://standaarden.overheid.nl/owms

	registries may be set up to link to vocabularies and value lists hosted elsewhere.
Licensing framework	OWMS is available under the Dutch **CC BY-ND 3.0 licence**[118] (a **class 2** licence). This is the preferred approach for reasons of quality control. Value lists are published with **no explicit licence** but should also be under a class 2 licence.
Quality controls	**Quality control** is mainly done with ad-hoc manual checks.
Metadata schema	**Vocabulary:** There is currently no metadata **vocabulary** in place for describing schemas and value lists. Use of ADMS or DCAT-AP will be considered when in the future a federated repository is set up. **Identifier scheme**: a URI-strategy is in preparation[119] that is based on the 10 Rules for Persistent URIs[120] at Joinup and on the UK strategy[121] adapted to the Dutch situation. **Schema documentation** including the specification of OWMS and related documentation, is available freely on standaarden.overheid.nl/owms. The specification of the implementation for the DCAT implementation is not yet available on-line. **Multilingualism** is not supported: the metadata values and all documentation are only available in Dutch.
Management	
Update	**Update frequency** is irregular. In practice, a new version was published every two or three years. The **change management process** is differentiated: minor changes (e.g. addition of a term to a value list) are done by KOOP directly, while for major changes, the User Council is involved. In practice, there have been almost no change requests after a new version was published. **Versioning**: a three-level version numbering approach X.Y.Z is defined. So far, only X and Y versions have been published. Backward compatibility is preserved as much as possible.
Harmonization	**Harmonising** and **mapping between metadata sets and value lists** are a growing concern. Solutions are under consideration. For the **assessment** for selection **between alternative schemas and value lists**, the current thinking is that a registry of registries could be set up that would show the various value lists with some form of status information (e.g. indicating which is the most authoritative).
Documentation	**Publication:** Currently, for **publication** there is no common vocabulary for metadata and there are no search capabilities. **Retrieval** of schemas, value lists and associated documentation is through manual access and download. Schema files can be downloaded manually as bulk on request to the maintainer at KOOP. The **supported formats for publication** are HTML and PDF for

[118] CC BY-ND 3.0 licence: http://creativecommons.org/licenses/by-nd/3.0/nl/

[119] URI-strategy description: http://www.pilod.nl/wiki/Bestand:D1-2013-09-19_Towards_a_NL_URI_Strategy.pdf

[120] Rules for Persistent URIs: https://joinup.ec.europa.eu/community/semic/document/10-rules-persistent-uris

[121] For more about design of URI sets for the UK public sector: https://www.gov.uk/government/publications/designing-uri-sets-for-the-uk-public-sector

	human-readable material. The metadata schema files are published as XSD, while the value lists are available as XML and SKOS (RDF/XML and N3).
Functionality of the tool(s)	The following tools are used in the management of OWMS: • Top Braid Composer for **maintenance of the semantics of the value lists** • OpenRDF Sesame for the **triple store** • Drupal for **publication**
Standards	The management approach is set up according to BOMOS[122]. The lists of values are set up using SKOS, RDF/OWL
Costs	Human resources assigned to OWMS and DCAT are between 1 and 2 FTEs.
Benefits	Interoperability of the systems for official government publications by hundreds of central and local government bodies on overheid.nl, supporting harvesting systems, search engine, geo-viewer and announcement service. Coherent presentation of information from different suppliers would simply not be possible at all without standardised metadata.
Tool support	
Name of the tool(s) and documentation	**TopBraid Composer, Maestro Edition[123]**
Reusability	**Commercial product** from TopQuadrant
Key functionality	• Standards-based, syntax directed development of RDF/S and OWL ontologies, SPARQL queries and Semantic Web rules • Import/export-from / to a variety of data formats including RDBs, XML and Excel • Visualization and diagramming tools including visual construction of queries and auto-generation of SPARQL • Triggered execution of SPARQL-based business rules and constraint checking using SPIN (SPARQL Inferencing Notation) • Seamless integration with inference engines including OWLIM, Jena Rules, Oracle Rules and SPARQL Rules • Choice of leading RDF stores • Support for re-factoring and model evolution • Usability, extensibility and robustness of its underlying technologies – Eclipse and Jena • Unique capabilities to develop/test/deploy/manage applications • Visual creation of RDF processing chains using SPARQLMotion™ scripts • Seamless round-tripping between XML and RDF/OWL (import – export) • SPARQL-based HTML and XML document generation using built-in JSP engine • Business intelligence reports can be generated and inserted into web pages for semantically enriched applications • Ability to convert Emails into OWL, supporting semantic analysis and classification of emails

[122] Link to the article: https://joinup.ec.europa.eu/sites/default/files/ckeditor_files/files/BOMOS2i(1).pdf

[123] TopBraid Composer tool: http://www.topquadrant.com/products/TB_Composer.html

	• Supports rapid iterative construction and evolution of semantic web applications:
	  ○ Built-in personal TopBraid Live web server
	  ○ Direct integration with TopBraid Ensemble

Costs	TBC costs US$3.450 for a one-time per-user licence fee. KOOP has three licences. Implementation of the complete tool set was outsourced at a cost of around €60.000.
Name of the tool(s) and documentation	**OpenRDF Sesame**[124]
Reusability	**Open Source**
Key functionality	De-facto standard framework for processing RDF data. This includes parsers, storage solutions (RDF databases a.k.a. triplestores), reasoning and querying, using the SPARQL query language. It offers a flexible and easy to use Java API that can be connected to all leading RDF storage solutions.
Costs	None
Name of the tool(s) and documentation	**Drupal**[125]
Reusability	**Open source**
Key functionality	Content Management
Costs	None

Case study 8: Local Government Inform[126] (LG Inform / LG Inform Plus), United Kingdom

The Local Government Association[127] (LGA) is a politically-led, cross-party organisation that works on behalf of councils to ensure local government has a strong, credible voice within national government. It is a membership organisation. In total, 421 authorities are members of the LGA for 2013/14. These include English Councils, the 22 Welsh councils via the Welsh LGA, fire authorities, passenger transport authorities, national parks via corporate membership through the English National Park Authorities Association (ENPAA) and one town council.

Local Government Inform is the LGA's benchmarking, data analysis and performance management service for councils and fire and rescue authorities. Since the LGA's successful campaign to reduce the burden of central government inspection and assessment of authorities, it has been working with local government to develop an approach to improvement which is based on the sector's learning about what works best.

LG Inform brings together a range of key performance data for authorities, alongside contextual and financial information, in an online tool[128] (LG Inform). Users can view data from over 1000 individual items, make comparisons between their authority and other councils or groups of councils, or construct their own

[124] OpenRDF Sesame tool: http://www.openrdf.org/

[125] Drupal tool: https://drupal.org/https://drupal.org/

[126] Local Government Inform (LG Inform): http://www.local.gov.uk/about-lginform

[127] Local Government Association: http://www.local.gov.uk/

[128] Local Government Inform (LG Inform); http://www.local.gov.uk/about-lginform

reports bringing several data items together. Importantly, the data is updated quickly after being published at its source. It is a flexible site that can be personalised to each authority's needs. The home page dashboard can be tailored to monitor important indicators and, in future, the ability to collect local data will be available.

With the focus on transparency and the desire to increase local accountability, local government has an obligation to be more open. LG Inform can help facilitate this. Yet, LG Inform is not an end in itself. It gives local government the ability to start thinking about its own data needs, rather than simply responding to requests for data from central government.

LGA publish all relevant publically available data to LG Inform within three days of first publication. In some cases LGA is the first to publish this data. Users in the sector tell LGA what is relevant and therefore what LGA should collect.

Alongside work on the online tool, LGA is working with a number of regional groups to develop the 'benchmarking club' function of LG Inform. This will allow local authorities to agree a set of data items they all want to collect, either more frequently than collected by central government or because they are not currently collected centrally at all. LG Inform will give those local authorities the ability to submit the data and then compare with others who have also submitted data.

To support councils with their commitments to transparency and increased local accountability, from November 2013 publically available data held within the new LG Inform is open to the public. With this, councils can create reports from with LG Inform that include the relevant context to help residents understand the information and the reasons for notable performance.

LG Inform is based on esd-toolkit[129]. The metadata governance for LG Inform is handled as part of the esd-toolkit programme which is also facilitated by the LGA. The esd-toolkit is an online resource service provided by a third party (Porism Ltd) and those who need access to it are paying an annual fee which varies depending on the functionality they use or the amount of data used.

Esd-toolkit's toolset allows examining the performance of the councils and provides tools for improvement. It is a framework of tried and tested resources, guidance and practical examples to support innovation in public service delivery. Esd-toolkit's community maps show specific metrics broken down within different areas. As councils define and share their own metrics, these are input via esd-toolkit pages.

Metadata governance and management

There was change in governance and management roles since 10 years ago Web managers and information managers in the councils had most responsibility for information management. Now, when applications and tools are available to use metadata standards, services heads have as much interest and authority for the governance of metadata. Esd-toolkit has as a well-established governance structure that is applied to LG Inform too.

The starting point of the governance is the vision for esd-toolkit. This is established by the programme board at the end of the previous year. The Knowledge & Infrastructure (K&I) group is asked to identify 'developments' of functionality and

[129] Eds-toolkit: http://esd.org.uk/esdToolkit/default.aspx

controlled lists which might meet the vision. The Engagement & Involvement (E&I) is asked to identify 'activities' in order to ensure engagement and involvement of the local authority officers. The controlled list group, working groups and TLC's are asked by K&I and E&I to feed in their thoughts for this vision and programme. The programme board will prioritise and plan the initiatives from K&I and E&I into the three projects to meet the vision. The following diagram sets out the structure.

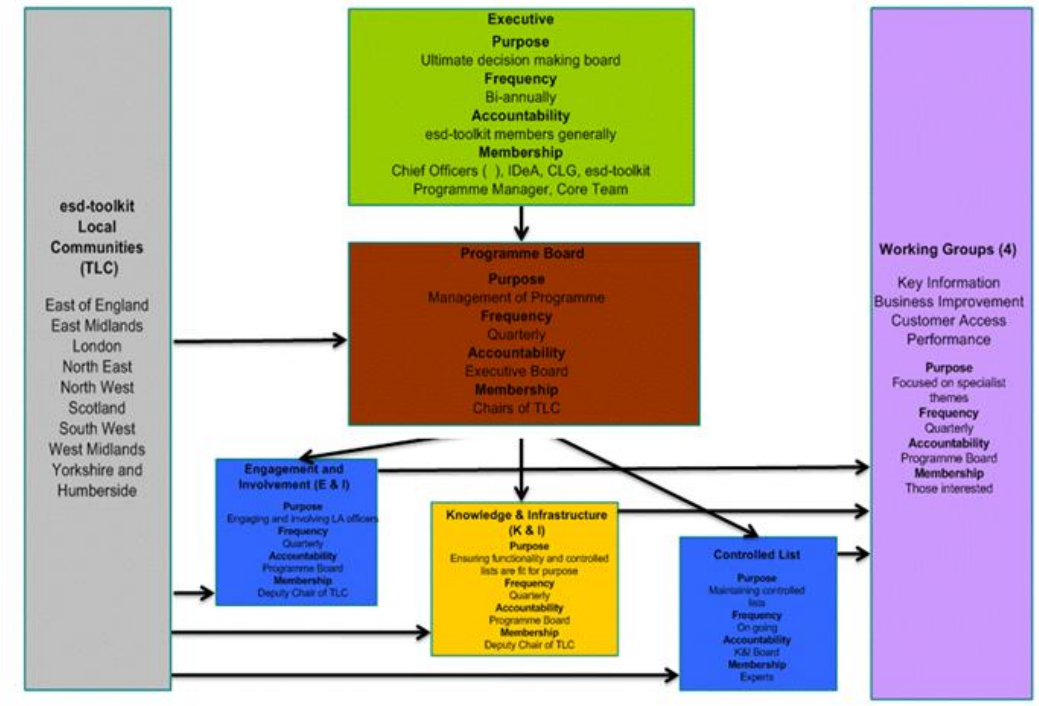

Figure 6. The esd-toolkit governance structure and the flow of influence

LG Inform takes data and the metadata definitions for that data from many other data suppliers – e.g. Office for National Statistics, NOMIS, Department of Works and Pensions, Department for Communities and Local Government. They absorb the metadata for these data into the esd standards and "repeat" their definitions on the esd standards pages. They are also represented on the UK Open Data Panel.
LG Inform is using UK Open Government Licence (OGL)[130] whereas metadata is free to use and distribute.

What is UK Open Government Licence (OGL)?

The Open Government Licence is a copyright licence for Crown Copyright works published by the UK government. It may also be applied to publications by other public sector bodies in the UK. It was developed, and is maintained, by The National Archives.

[130] UK open government licence (OGL): http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/

The OGL was developed as part of the UK Government Licensing Framework, which also includes a non-commercial Government licence that restricts the commercial use of licensed content, as well as a charged licence for situations where charging for the re-use of content is deemed appropriate.
The licence is interoperable with the Creative Commons Attribution licence, and OGL-licensed work could be used in a CC licensed work, however it should be clear that the material used is being used under the OGL and it should still be linked to the OGL.

In the remainder of this section, we present the summary of the interview with LG Inform conducted in the context of this case study. The summary is structured according to the analysis framework of section 2.1.2.

Interview date	16th January 2014
Interviewee	Tim Adams
Interviewer	Audrius Leipus, Nikolaos Loutas

Title: Local Government Inform - LG Inform	
Governance	
Goals	• Consistent collection and publication of key performance data for authorities (councils, etc.). • Creation of a single list of every service that UK government deliver (there are ~1500 services that UK Gov. delivers, e.g. from emptying the waste to running car parks in city centre).
Governance structure	**Roles:** The metadata governance for LG Inform facilitated by the LGA. Metadata governance is undertaken in accordance with the process for managing changes and update to the esd-standards. Municipalities usually appoint Information Managers and Data specialists for metadata governance. Working groups are composed with one group being responsible for semantics and definitions, another group being responsible for accuracy etc. Different experts are involved for a few hours a month to consult on the suggested changes. **Openness:** LG Inform and esd-toolkit are primarily local government owned and led initiatives and so are managed by local government representatives. They (local government) do invite experts and advisers from wider sectors (mainly central government groups) as required. **Decision making process:** Everything LG Inform does is owned and led by the representatives of the local government sector. They have working groups that are open to any representatives that municipalities care to offer up to take part. They are usually Information Managers and Data specialists in the local government sector; one group that looks after semantics and definitions, one group that looks after accuracy. Different experts are involved for a few hours a month to consult on the suggested changes.
Context	**Organisational dimension:** • **Inter-organisational.** **Administrative dimension:** • **National:** used nation-wide in local government. Looking to

	extend and link to central government too (in terms what additional information might be required that is shared by government). Also trying to move out and be applied to European countries (Belgium, Norway, The Netherlands) on some intellectual projects e.g. smart citizen.
Policy domain	Not restricted.
Enforcement policy	**Sharing:** sharing is not enforced, but encouraged on **voluntary** basis. The ESD vocabularies are available on Joinup.eu[131] under semantic assets section. **Reuse:** voluntary, reuse is encouraged via information campaigns. The esd-standards are acknowledged in England as the authoritative source of semantics for local governments and there is a drive to encourage all organisations to use them. They are identified as such on the data.gov.uk web site.
Authoritative source	LG Inform[132] is an authoritative source for hosting metadata and the raw data delivered to the metadata standards is coming from many different sources.
Licensing framework	Delivered under the UK Open Government Licence (OGL) – free to use and distribute. Vocabularies can be freely re-used under the OGL.
Quality controls	Quality controls rely on metadata experts who periodically review possible changes (via user compiled wish lists), e.g. of metadata standards.
Metadata schema	**Vocabulary:** well described common vocabulary which is available in different formats. The Integrated Public Sector Vocabulary (IPSV) is an 'encoding scheme' for populating the e-GMS[133] Subject element of metadata. It is fully compliant with ISO 2788[134] and BS 8723[135], the International and British Standards for monolingual thesauri vocabularies should be published according to SKOS and publicly available for free re-use. **Identifiers scheme:** Councils, authorities create their own identifiers (as new information is added, they create new unique identifiers). **Schema documentation:** there are no specific documents for schema documentation but all schemas are documented internally offering comments with minimal explanation. **Multilingualism:** English only although the European extensions are available in other source languages.
Management	
Update	**Update frequency:** different aspects of the model have different frequency of update (most changing from every 3 months and least changing, reviewed once a year). **Change management process:** experts (mainly people from local authorities) review candidates (suggested areas) for change, then the draft of the change is published for users' comments and decisions are made. The process takes around 6 months. **Version control:** every version itself is documented and is backwards compatible wherever possible.

[131] ESD: https://joinup.ec.europa.eu/catalogue/repository/esd-standards.

[132] LG Inform: http://lginform.local.gov.uk/

[133] E-GMS: http://www.esd.org.uk/standards/egms/

[134] ISO 2788: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=7776

[135] BS 8723: http://www.iskouk.org/presentations/DextreClarke_ISKOUKseminar1.pdf

Harmonization	**Creating mappings between related metadata sets:** common vocabularies are used to identify and map similar datasets. **Assessing alternative metadata sets:** it is not being done.
Documentation	**Publication:** publication is done manually. **Retrieval:** retrieval can be done by both humans (xml, doc, txt formats could be used) and machines (linked open data format could be used). A SPARQL[136] end point is also available[137]. **Supported formats of publication tool(s):** Formats are both human and machine readable.
Standards	Vocabularies constitute UK semantic standards for Local Government Vocabularies[138] conform to the SKOS standard.
Functionality of the tool(s)	Most of available tools are used to exploit data, also to gather and publish available government data openly.
Costs	There are 4 FTEs. There are also about 30-40 volunteers, who spend few hours per month on metadata management.
Benefits	Benefits offset the costs. Some of them: • Constant observation of country's (councils) progress in performance; • Data de-centralization but able to be aggregated for regional and national analyses.
Tool support	
Name of the tool(s) and documentation	In-house developed tools are built on a variety of languages, including C# / .NET, Java and PHP. Web services are Java connecting to a MySQL database. CMS is provided via Drupal.
Reusability	Vocabularies can be freely re-used under the OGL; • Application source code is generally not freely available but can be used by government organisations contributing to the esd-toolkit programme. • web services are accessible via an API[139]
Key functionality	• Develop reporting tools that compare councils with one another; • Functionality to identify councils in each region of the country and how each council area can be broken down for local communities; • Summarise and compare metrics across (pre-defined and user-defined) "comparison groups" of areas; • Reports can be written for sharing as HTML, PDF or MS Word documents or for embedding in web sites. • Report writers allow municipalities and areas within them to be compared via data on 1,800+ metrics.
Costs	Access to the standards and their metadata is free. Access to various applications and tools that are developed within esd-toolkit costs from £750 to £2,500 per annum (for municipalities) for access by all officers and use of the API. These costs are for access to various applications and tools that they have developed within esd-toolkit to deploy the standards. Private sector organisations pay according to data volumes they used. Costs for services of "Porism Ltd" for LG Inform amount to about £20.000 per quarter.

[136] SPARQL: http://www.w3.org/TR/rdf-sparql-query/

[137] Link to the end point: http://sparql.esd.org.uk/ds/query

[138] UK semantic standards for Local Government Vocabularies: standards.esd.org.uk

[139] API description: api.esd.org.uk

ANNEX III – SURVEY QUESTIONNAIRE

Survey on metadata management and governance in EU institutions and Member States

Introduction:

This survey is conducted in the context of the semantic methodologies Action (Action 1.1) of the Interoperability Solutions for European Public Administrations (ISA) Programme. Its purpose is to identify the current practice and existing solutions in the area of **inter-organisational metadata governance and management** in European public administrations. It will help the ISA Programme better understand the current situation and needs and more effectively target its initiatives. The results from the survey will afterwards be made publicly available in the semic.eu community on Joinup, the on-line collaborative platform of the ISA Programme.

We invite you to participate in this survey, particularly if your public administration has solutions for **inter-organisational metadata management and governance** that you would like to tell us about.

Join our SEMIC group on Joinup

Join our SEMIC group on LinkedIn

Follow @SEMICeu on Twitter

Glossary:

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)

Structural metadata. Data model or reference data.

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 management and governance:

1. Please provide the following personal data[140].

Last name	
First name	
E-mail address	
Telephone number	
Organisation	
Organisation URL	
Country	

2. Does your organisation apply metadata management and/or governance? At which level?

- **None** – my organisation does not apply any form of metadata management or governance;
- **Intra-organisational** – my organisation applies this internally;
- **Inter-organisational, national level** – my organisation applies this with other organisations at national level;
- **Inter-organisational, EU level** – my organisation applies this with other organisations at EU level;
- **Inter-organisational, international level** – my organisation applies this with other organisations at international level.

Please further elaborate on your choice(s).

Click here to enter text.

3. Please provide links or describe the *solutions* for metadata management and governance that are currently used by your organisation. These could include:
   a. **Methodologies and policies** for metadata governance and management (e.g.DAMA DM-BOK, Inter-institutional Metadata Management Committee - IMMC),
   b. **Standards** (e.g. ISO/IEC 11179 Metadata Registry standard, ISO 25964 Thesauri and interoperability, the Simple Knowledge Organisation System - SKOS, the Asset Description Metadata Schema - ADMS),
   c. **Tools** (e.g. VocBench).

Click here to enter text.

4. Does your organisation share and reuse structural metadata with organisations with whom you exchange information?

[140] Legal notice: As this survey collects and further processes personal data, Regulation (EC) 45/2001 of the European Parliament and the Council of 18 December 2000, on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data, is applicable. The following personal data is collected: first name, last name, e-mail address, and telephone number. This information will be made available to employees and contractors on behalf of the DG for Informatics, Unit B.2 (ISA programme) of the European Commission, who might use this information to contact you and to follow up on the questionnaire. No personal data will be published or shared with others without your permission.

a. Yes
b. No

Please briefly motivate your choice.

Click here to enter text.

5. Please indicate the most important benefit that your organisation receives by applying metadata management and governance.

   - **Improved data quality;**
   - **More efficient administrative processes;**
   - **Reduced system development and/or maintenance costs;**
   - **Other.**

Please briefly motivate your choice. When selecting the option 'Other' please provide your alternative here.

Click here to enter text.

6. What is the most important roadblock to inter-organisational metadata governance and management from the perspective of your organisation?

   - **Lack of buy-in from senior management;**
   - **Lack of methodologies;**
   - **Lack of standards;**
   - **Inadequate tool support;**
   - **Other.**

Please briefly motivate your choice. When selecting the option 'Other' please provide your alternative here.

Click here to enter text.