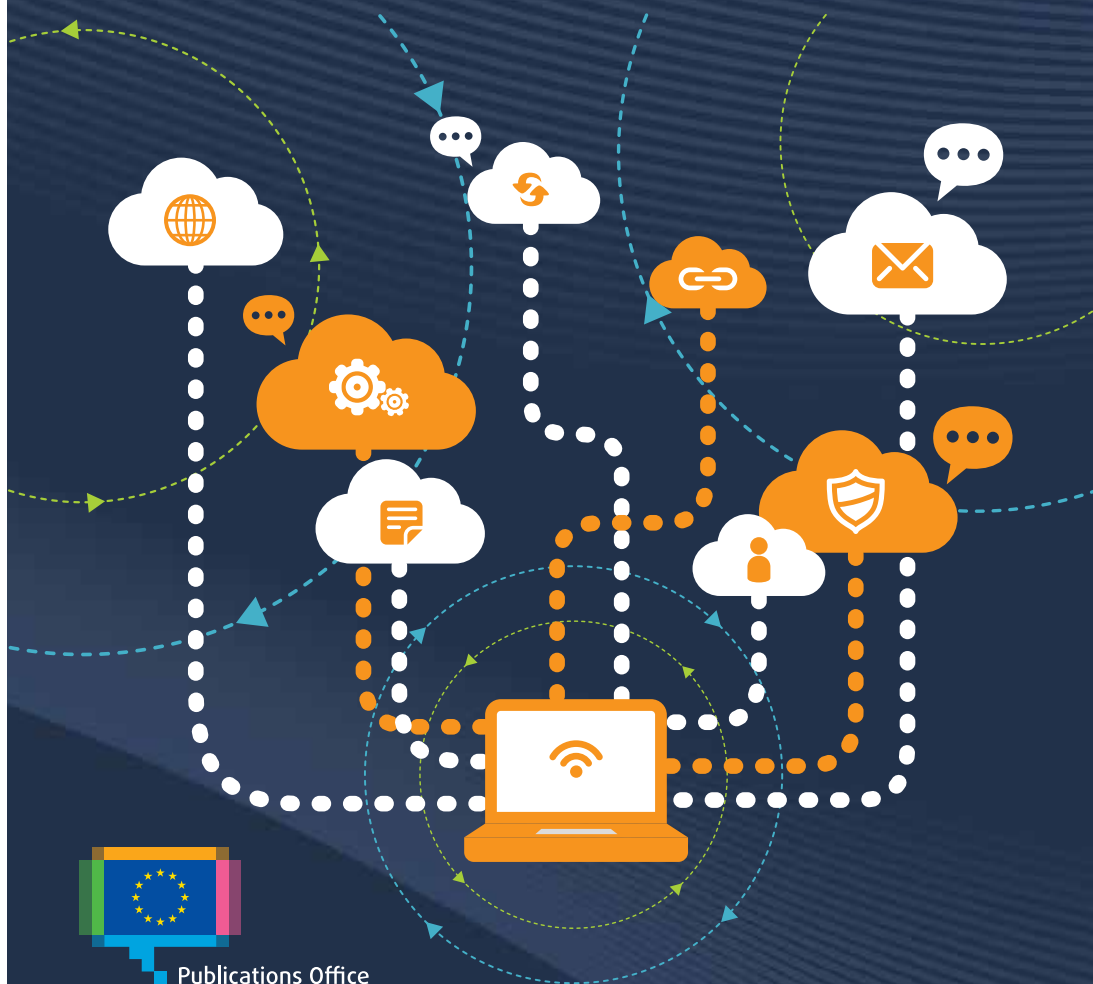




ELI implementation methodology

Good practices and guidelines

Author, ELI Task Force



More information on the European Union is available on the internet (<http://europa.eu>).

Luxembourg: Publications Office of the European Union, 2015

Print	ISBN 978-92-78-41354-5	doi:10.2830/990350	OA-01-15-858-EN-C
PDF	ISBN 978-92-78-41355-2	doi:10.2830/813167	OA-01-15-858-EN-N
ePUB	ISBN 978-92-78-41353-8	doi:10.2830/353863	OA-01-15-858-EN-E

© European Union, 2015
Reproduction is authorised provided the source is acknowledged.

Printed in Luxembourg

PRINTED ON ELEMENTAL CHLORINE-FREE BLEACHED PAPER (ECF)

ELI implementation methodology

Good practices and guidelines

Author, **ELI Task Force**



Contents

Foreword..... 5

Introduction..... 6

 Purpose 6

 Structure 6

 Target audience..... 6

 Scope 7

**PART 1: An introduction
to the European Legislation Identifier..... 9**

Background.....10

Origins 11

Objectives 11

Governance..... 14

ELI benefits 14

Work done so far..... 17

PART 2: ELI implementation methodology 19

Overview of the ELI implementation methodology.....20

Organisation and policy22

 STEP 1: Identify the business case22

 STEP 2: Estimate resources.....25

 STEP 3: Set up a governance structure30

 STEP 4: Set up an ELI implementation project.....33

 STEP 5: Formulate statements about URI
 and metadata services provided.....36

Pillar I: How to put in place legislation identifiers based
on ELI HTTP URIs38

 STEP 1: Capture requirements for HTTP URIs.....39

 STEP 2: Design the URI template.....42

 STEP 3: Configure web server for ELI URI resolution.....50

Pillar II: How to design an ELI metadata schema.....57

 STEP 1: Capture requirements for metadata.....58

 STEP 2: Design and publish metadata schema.....61

 STEP 3: Develop metadata extraction routines66



STEP 4: Deploy legislation metadata.....69

Pillar III: How to render ELI metadata machine-reusable72

STEP 1: Add microdata or RDFa tags in web pages
to facilitate parsing and extraction73

STEP 2: Provide an application programming interface (API)76

ANNEXES 79

ANNEX I: Glossary80

ANNEX II: ELI ontology.....86

ANNEX III: Self-assessment check list.....87



Foreword

This guide was created by the ELI Task Force and is a set of good practices based on feedback gathered from pioneering Member States that have implemented ELI.

The ELI Task Force was set up in December 2012, under the auspices of the Council of the European Union Working Party on E-law, to study the future development of the ELI standard.

At the time of publication, the task force was made up of Denmark, Ireland, France, Luxembourg (chair), the United Kingdom and the Publications Office of the European Union.

The task force aims to help Member States wishing to adopt ELI by sharing knowledge and expertise.



Introduction

Purpose

The purpose of this document is to provide publishers of legal information with a step-by-step methodology on how to implement the European Legislation Identifier (ELI).

Structure

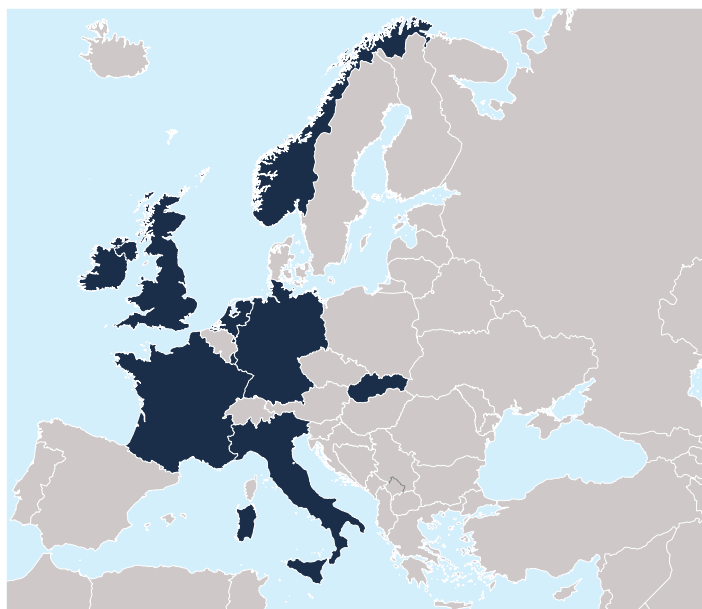
This document consist of two parts:

- **Part 1** provides an overview of ELI together with the main benefits derived from the implementation of ELI.
- **Part 2** provides a set of steps and good practices that publishers of legal information can apply in order to implement ELI. This part also provides an estimation of the resources needed to implement each of the three pillars of ELI.

Target audience

The target audience of this document are:

- **Decision-makers:** individuals that make decisions concerning the implementation of ELI in their countries. This group typically includes countries that have not yet implemented ELI, but that are interested to learn more about the potential associated benefits and experiences of those who have already implemented ELI.
- **Technology experts:** individuals that want to learn how to implement ELI.
- **ELI implementers:** those that have already implemented ELI in its totality or partially and that are interested in the progress of other ELI implementers.
- **Legal professionals:** individuals with a legal background that could be interested to learn more about URI templates, in order to know how to better search for legal information.



Scope

The guidelines in this document have been developed based on the findings gathered in the course of a study carried out in the first trimester of 2015 with 11 Member States/EFTA countries (i.e. France, Germany, Ireland, Italy, Luxembourg, Malta, the Netherlands, Norway, Slovakia, Switzerland and the United Kingdom) as well as the Publications Office of the European Union (OP).

The participating countries were selected based on a set of criteria aiming at taking into account different stages of ELI implementation, as described:

- impact analysis on existing systems, resources to be foreseen (i.e. Malta).
- analysis to identify URI templates, align metadata to the ELI scheme and define publication formats (i.e. Germany, Norway and Switzerland).
- development of the system according to the specifications and requirements defined in the 'design' stage and test it (i.e. Slovakia).
- implementation of the developed system previously tested (i.e. Ireland, Italy, France, Luxembourg, the United Kingdom and the Publications Office).



Furthermore, one country that was not planning to implement ELI at the time of the study (i.e. the Netherlands) was surveyed as well.

Further desk research complemented the findings from the interviews.

The results of the interviews and desk research were grouped according to the three 'pillars' of ELI specification — identification of legislation, metadata properties describing each legal resource and rendering the ELI metadata machine-reusable. Opinions and suggestions from the people interviewed were gathered and analysed to determine the main benefits, challenges and problems with ELI implementation, as they perceived it. These opinions and suggestions have been used as the basis for the implementation methodology and good practices in the remainder of this document.

PART 1:

An introduction to the European Legislation Identifier



Background

The conclusions of the Council of the European Union inviting the introduction of ELI explain that: ‘... a European area of freedom, security and justice in which judicial cooperation can take place requires not only knowledge of European law, but in particular mutual knowledge of the legal systems of other Member States, including national legislation’ ⁽¹⁾.

The exchange of legal information is key in this regard, but despite the increased availability of documents in electronic format, the Council conclusions report that the exchange of legal information originating from regional and national authorities at the European level is far from optimal, due to existing differences between national legal systems, including the technical systems used to store and display legislation through their websites.

To overcome such obstacles and improve interoperability between legal systems, the ELI Council conclusions invite the Member States to use identification of legislation and metadata properties describing each legal resource, and render the ELI metadata machine-reusable to reference national legislation in official journals and legal gazettes, so as to enable an effective, user-friendly and faster search, and exchange of legal information.

This is in line with the European Union commitment to open up legislation as part of the implementation of the G8 Open Data Charter ⁽²⁾ which aims to promote, amongst other things, transparency and government accountability. It also meets the recommendations included in the European Interoperability Framework (EIF) ⁽³⁾ on how to improve interoperability within the EU and across Member State borders and sectors.

By making legislation available on the web in a structured way, it will be easier to find, share and reuse legislation, as prescribed by the public sector information (PSI) directive ⁽⁴⁾. Finally, ELI aims to promote the access and exchange of legal information within and across borders, thereby contributing to the development of the common area of freedom, security and justice.

⁽¹⁾ OJ C 325/02, 26.10.2012., p. 3.

⁽²⁾ EU implementation of the G8 Open Data Charter (<https://ec.europa.eu/digital-agenda/en/news/eu-implementation-g8-open-data-charter>).

⁽³⁾ COM (2010) 744 final, Annex 2, Brussels, 16.12.2010.

⁽⁴⁾ Public sector information (PSI) directive 2003/98/EC (<http://eur-lex.europa.eu/legal-content/en/TXT?uri=CELEX:32003L0098>) and an amending directive 2013/37/EU (<http://eur-lex.europa.eu/legal-content/EN/TXT?uri=celex:32013L0037>).



The ELI initiative is supported in the context of Action 1.21 on facilitating the exchange of legislation data in Europe ⁽¹⁾ of the 'Interoperability solutions for European public administrations' (ISA) programme.

Origins

First established in the context of the European Forum of Official Gazettes ⁽²⁾, ELI has been further supported by the sub-group mandated by the Council of the European Union in the framework of the Working Party on E-law ⁽³⁾.

ELI stems from the acknowledgment that the World Wide Web defines a new paradigm for legal information access, sharing and enrichment.

In this context, consumers of legal information and publishers of legal information are faced with opportunities and challenges. On the one hand, there are new opportunities to access legal information in a more interconnected and interoperable way across legal information systems; on the other hand, there is a need to maintain the flexibility of each national legal system. Keeping this in mind, ELI is based on a progressive approach entailing the implementation of the following elements:

- identification of legislation via HTTP URIs;
- description of legal resources through metadata;
- rendering the ELI metadata machine-reusable through the serialisation of the metadata in compliance with the ELI ontology.

These three elements are commonly referred to as the three pillars of ELI.

Objectives

The purpose of ELI is to facilitate the access, sharing and interconnection of legal information published through national, European and global legal information systems. ELI aims primarily to facilitate the search, exchange and interconnection of legal information through national and European IT systems.

⁽¹⁾ http://ec.europa.eu/isa/documents/actions/more-about-action-1.21_en.pdf

⁽²⁾ <https://circabc.europa.eu/webdav/CircaBC/OPOCE/ojf/Information/prod/html/index.htm>

⁽³⁾ <http://www.consilium.europa.eu/en/council-eu/preparatory-bodies/working-party-e-law>



The European Legislation Identifier is used to:

1. identify legislation with a unique identifier which is recognisable, readable and understandable by both humans and computers, and which is compatible with existing technological standards;
2. describe legislation with a set of machine-readable metadata elements in compliance with a recommended ontology.

According to the Council of the European Union conclusions:

‘ELI should guarantee a cost-effective public access to reliable and up-to-date legislation. Benefiting from the emerging architecture of the semantic web, which enables information to be directly processed by computers and humans alike, ELI would allow a greater and faster exchange of data by enabling an automatic and efficient exchange of information.

Furthermore, ELI should give the Member States and the European Union a flexible, self-documenting, consistent and unique way to reference legislation across different legal systems. ELI URIs uniquely identify in a stable way each legislative act across the European Union, while at the same time taking into account the specificities of national legal systems.

ELI takes into account not only the complexity and specificity of regional, national and European legislative systems, but also changes in legal resources (e.g. consolidations, repealed acts, etc.). It is designed to work seamlessly on top of existing systems using structured data and can be taken forward by Member States at their own pace.’

While legal information has been available online via national legal information systems for a long time, ELI enables the putting in place of a genuine network of legal information, available as open data and reusable by all.

To achieve these objectives ELI relies on the infrastructure of the World Wide Web and the technologies of the semantic web. ELI is based on an approach which is characterised by the progressive adoption by the publishers of legal information of the following elements:

1. **Pillar I on the identification of legislation:** URI templates at the European, national and regional levels based on a defined set of components;
2. **Pillar II on metadata properties describing each legal resource:** definition of a set of metadata and its expression in a formal ontology;



3. **Pillar III to render the ELI metadata machine-reusable:** Integration of metadata into the legislative websites; in Annex 2 to the Council conclusions, it is suggested that the ELI metadata elements may be serialised in compliance with the World Wide Web Consortium (W3C) recommendation 'RDFa in XHTML: Syntax and Processing'.

The ELI Council conclusions do not specify as a technical requirement that in order to implement ELI the order of the pillars should be followed. However, the most common approach is to progressively implement ELI in line with the numbering of pillars.

ELI is based on the well-established model of 'Functional requirements for bibliographic records' (FRBR). FRBR distinguishes between the concepts of 'work' (intellectual or artistic creation), 'expression' (intellectual or artistic realisation of a work) and 'manifestation' (materialisation of one of the expressions of a work). Each time a 'work' is realised it takes the form of an 'expression'. The physical embodiment of an expression is a 'manifestation'. Figure 1 below illustrates how these concepts map to the ELI Council conclusions⁽¹⁾ as published in EUR-Lex. More specifically:

- The **LegalResource** is the conceptual entity that represents the intellectual content of the ELI Council conclusions, which is what the various language versions share. The intellectual content of the Council conclusions is realised by the text which is expressed in various language versions.
- The **LegalExpressions** are the language versions that are presented horizontally and express the intellectual content in actual text. The language versions are embodied by physical files.
- The **Formats** are the physical files presented vertically under each language.

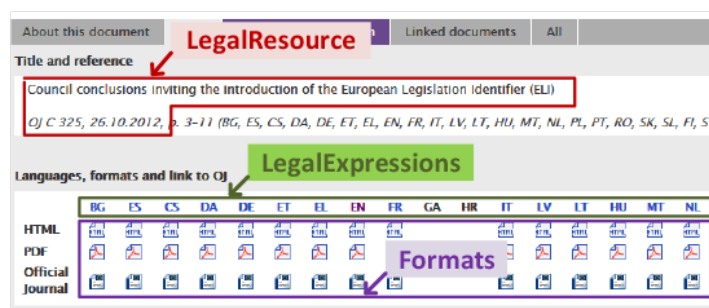


Figure 1 — Screenshot from EUR-Lex website depicting the three conceptual entities of the ELI model

⁽¹⁾ OJ C 325, 26.10.2012.



Governance

At the meeting of the Working Party on e-Law on 20 December 2012, the Chair of the Working Party proposed that an ELI Task Force should be set up to further study and define the development of a unique identifier for national and European legislation.

The ELI Task Force was initially composed of France, Luxembourg (chair), the United Kingdom and the Publications Office of the European Union; Denmark and Ireland joined the task force subsequently.

In order to help Member States, candidate countries, EFTA countries and others to adopt ELI, the task force aims to share knowledge/expertise, without imposing a strict schema on how ELI should be implemented and by taking into account national specificities. For this purpose a change and release management process is in place and users can contact the Publications Office for any information they would like to receive as well as to provide feedback and comments on ELI.

ELI benefits

The implementation of ELI brings a number of benefits for public administrations and publishers of legal information as well as citizens at large.

Benefit 1: access to legislation

The large amounts of legislation, as well as its diversity, can make access to it cumbersome. Thus, it is important that legislation is published in a way that allows people to use it easily and readily.

First of all the assignment of persistent identifiers as specified by ELI provides for a homogenous mechanism to identify legal information on the web, making it easier for users to look for legal information within and across legal systems.

The predictable nature of the ELI identifiers, which are designed in such a way as to stay as close as possible to the way legislation is cited, allows users to easily compose URLs to refer to published legislation. What is more, the design patterns of the ELI-compliant HTTP URLs are conceived to assist users looking for legal information when they do not have all the necessary elements to identify the specific piece of legislation. In this case, the user will be directed towards the right information through a partial input, based on the available information.

ELI also facilitates access to legal information by citizens as well as potential reusers of legal information. It allows referencing legislation at different levels of granularity, thus enabling different usages of legal in-



formation, such as social media discussions about particular elements of legislation, references to legislation in research papers, or developments of applications by businesses.

Furthermore, a common structure for metadata to describe legislation across national and European legislation services, as proposed by ELI, makes searching across the distributed collections easier. It also creates opportunities for aggregation services to harvest metadata from various sources to build indexes across collections.

Benefit 2: development of new services

The ever-evolving social, economic and technological environment in which publication services are operating enables smart reuse of data. The persistence of identifiers, the common specification of structured metadata and the publication format proposed by ELI give businesses the opportunity to develop innovative and sustainable value-added services on the basis of the published legal information.

Furthermore, the increased semantic interoperability in the context of legal information achieved through the adoption of ELI contributes to the development of the digital single market, which will lead to the modernisation of traditional industry, and which has estimated important gains in terms of additional annual growth, as well as a substantial boost for job creation in Europe.

Benefit 3: cost savings

Failure to link or share legislation within and between different public administrations results in inefficient public services and high information collection costs, increasing the burden on public administrations and citizens.

ELI promotes the publication of legislation in a structured and machine-readable way to promote linking and reuse of legal data. This also leads to cost savings for publishers, in terms of improved effectiveness of information flows and shorter time to publish legislation.

The ELI approach aims to construct common building blocks for naming and citing legal information in a flexible way so as to respect each country's unique legislative and legal tradition. This flexibility, together with the fact that ELI can be implemented in a progressive way on top of existing IT solutions and databases and that it is designed to work seamlessly on top of existing systems, makes it a cost-effective solution for publishers that want to make legal information available for reuse purposes.



Benefit 4: quality and reliability of data

The increased reuse of legislation triggers a growing demand to improve the quality of the published information. The structured publication of legal information based on the ELI specifications enables review and feedback from users, which can help increase the quality and reliability of information.

Furthermore, ELI allows the better structuring of legal information in terms of internal links within a text and external links to other documents that are mentioned within a text (for example laws, decrees or EU legislation).

Quality and reliability of data is key in order to establish more effective, simplified and user-friendly services which are crucial for fostering the trust of businesses and citizens in digital services.

Benefit 5: transparency

There is an ever-growing demand for openness and transparency in modern societies. Increased transparency is an important objective across public administrations as demonstrated in the context of the European Union, where the principles of openness and transparency feature in EU primary law.

By making legislation more accessible, ELI contributes to making it easier for citizens and watchdog organisations to be better informed and monitor the work done by governments. Transparency guarantees greater legitimacy and accountability of the administration in a democratic system because citizens have the opportunity to better understand the considerations underpinning the law in order to exercise their democratic rights.

Benefit 6: interoperability

ELI is based on the emerging architecture of the semantic web and linked open data, enabling greater and faster exchange of data through the automatic and efficient exchange of information. This is achieved through semantic interoperability. Semantic interoperability is a means to achieving greater efficiencies where there is a clear need and where there is clear demand from relevant users.

The ELI approach is based on the idea that in order to generate greater efficiencies, it is necessary to reach a minimum level of agreement on the way legislation is identified and described. The ELI approach thus allows structuring and identifying legislation in a sufficiently uniform way and at the same time takes into account the specificities of each national legal system by providing a flexible and progressive approach to its implementation.



ELI also provides a standard way to publish legislation as machine-readable open data, creating opportunities to automatically discover and reuse published legislation, as well as to ensure legal reference certainty.

By allowing humans and machines to have improved access to legal information, ELI sets the basis for a web of legal information. Making legal information part of the web of data allows users to easily discover legal information across countries and reference to legal information more easily, as well as improve their understanding of legal information thanks to the interconnections with different sources of information across domains.

Work done so far

There are a number of activities that have been carried out in the context of the ELI initiative. These activities are listed below:

- **Refine and extend the ELI ontology for all EU legislation.** The ontology, drafted by the Publications Office, can be found here: <http://publications.europa.eu/mdr/eli/index.html>
- **Develop a pilot project to demonstrate how ELI can be used to enable the exchange of information between systems (i.e. national legal information systems and EUR-Lex).** This pilot was carried out by the Publications Office and legislation.gov.uk (National Archives) between September and December 2014.
- **Conduct a feasibility study into the use and integration of ELI in the European Commission's National Implementing Measures (NIM) database and national third parties' applications to monitor the transposition of European legal acts.** This study was carried out by the Publications Office between September and December 2014.
- **Carry out a discovery phase to identify user needs and technical constraints which the IT systems of a selected number of Member States/EFTA countries must adhere to for them to be successful in the implementation of ELI.** This activity was carried out by the Publications Office in the context of the ISA Programme Action 1.21. It was carried out between December 2014 and October 2015.
- **Provide an assistance package in the form of guidelines.** This document is the result of this activity and was developed by the Publications Office between January and October 2015.
- **Develop a technical ELI implementation manual.**
- **Align ELI with other standards by engaging with groups and communities working in areas that are related to ELI.**



- **Document ELI success stories and provide news about progress made.** This activity is carried out by the Publications Office and aims to share information on how ELI is used across Europe. The success stories are available on the ELI website.
- **Develop a platform to give access to EU legislation via ELI URIs:** <http://data.europa.eu/eli>
- **Organise ELI workshops to assist Member States in implementing ELI.**

PART 2:

**ELI implementation
methodology**



Overview of the ELI implementation methodology

This section introduces the ELI implementation methodology, providing a step-by-step approach that publishers of legal information can follow to implement one or more ELI pillars. This is accompanied by a set of good practices that have been identified based on the experience of stakeholders that have already implemented one or more of the ELI pillars. The ELI implementation methodology consists of two parts, as depicted in Figure 2 below.

1. **Organisation and policy:** this part mainly addresses decision-makers and is composed of five main steps.
2. **Technical implementation:** this part addresses a more technical audience that is involved with the actual ELI implementation. It is divided according to the three ELI pillars and for each pillar a number of steps have been outlined.



Organisation and policy



Technical implementation

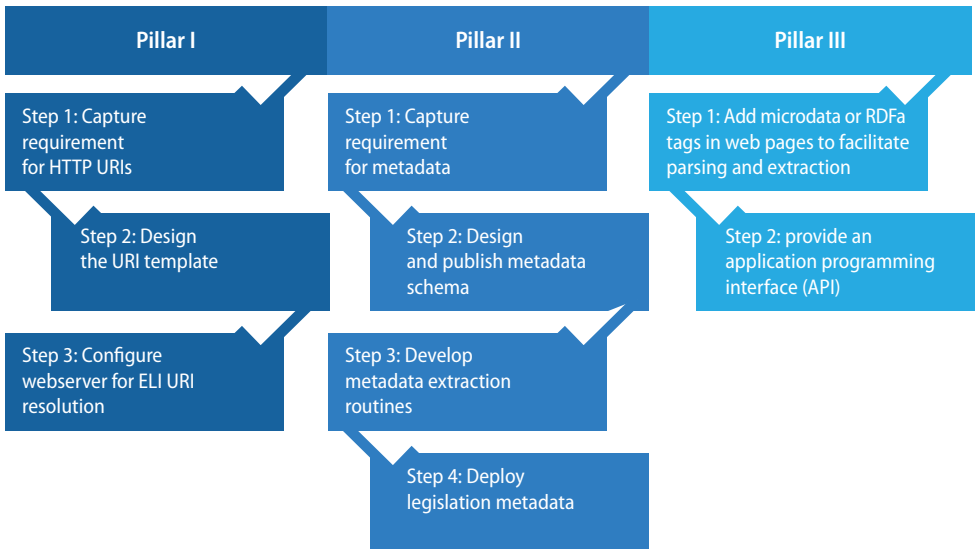


Figure 2 — Overview of the ELI implementation methodology



Organisation and policy

STEP 1: Identify the business case

Publishers of legal information should develop a business case to help decisionmakers decide whether to implement ELI and the allocation of resources that are necessary to implement it. The business case should include the following elements:

- **An executive summary** which outlines the ELI project and contains the key considerations that will be discussed in the business case, including the timeline for business case implementation and completion, as well as the projected benefits and costs.
- **Scope**, defining the coverage of ELI implementation both from a chronological and type(s) of collection perspective. With regard to the chronological aspect, publishers of legal information need to decide whether ELI will be implemented in relation to all legislation that has ever existed or to a sub-part, such as for example only newly adopted legislation. This decision depends significantly on what legislation is available online or is planned to be made available online. For example, it might be that legal information that goes back to the 19th century is not available electronically. This is the case in Italy, where legal acts adopted from 1933 onwards are available in electronic format, but not those from 1861 to 1932. Conversely, Luxembourg has a complete and accurate inventory of current legislation and national regulations published in the *Répertoire Analytique du Droit Luxembourgeois* (<http://www.legilux.public.lu>) with a classification made via keywords (one per item). More specifically, all issues of the *Mémorial* (national official journal) are available since the first issue in 1617. Most of the material from the years 1617-1933 is only available as scanned pages in PDF. All texts in force are also presented in HTML (created from an XML source) with a link to the page of the *Mémorial* where they were published. All acts in force have also been converted to XML, but not all of the historic material will be made available in a structured format.

With regard to the type of collection(s), publishers of legal information might decide that ELI will be applied only to legislation or to a broader range of legal information. For example, on the one hand, France decided to include summaries of legislation within the scope of ELI because this collection is regularly consulted by users to get a better understanding of legislation. On the other hand, the Publications Office decided to limit the scope of ELI implementation to the EU legislation published in the Official Journal L series.



- **Problem statement** highlighting the issue or goal that the business case for the implementation of ELI aims to solve. The 'ELI benefits' presented in Part 1 of this guide can provide guidance on common issues and objectives that ELI aims to address. An example of a concrete problem statement entails the need to establish clear links between an EU directive and related national implementing measures. The issue related to the fact that it is not always easy to establish which national provision transposes which EU directive and even if the relationship is identified, there is a need to ensure that such information is kept up to date. In the past public administrations used static tables of correspondence that were manually maintained. Such an exercise not only was extremely cumbersome, but also led to information that was not always accurate and up-dated.
- **A proposed solution** which describes the project and options available to solve the business problem, explaining how the implementation of ELI addresses and resolves the problem statement. For example, ELI provides a useful means to improve the links between EU directives and national implementing measures thanks to the fact that it promotes semantic interoperability between legal information systems across countries. To demonstrate its feasibility, a pilot was developed in the context of the project: 'Feasibility study on providing access to national implementing measures via EUR-Lex' led by the Publications Office together with legislation.gov.uk (National Archives). The pilot demonstrates how the ELI URI template, metadata schema and serialisation enable semantic interoperability between national legal systems publishing national implementing measures and EUR-Lex, where the EU directives are published.
- **An estimation** of resources which indicates what is needed to implement ELI, including allocation of resources, timeline for implementation and resources needed. For guidance on how to define the resources necessary for ELI implementation the reader is invited to move to the next step, Step 2 'Estimate resources'.



Good practice 1:

Building on the knowledge and experience of others (support by the ELI Task Force)

Problem statement

When approaching ELI implementation, publishers of legal information might be inclined to work in isolation, without looking for existing experience and know-how. This leads to inefficiencies as time and money might be spent in investigating aspects which are already known inside other organisations that have already implemented ELI.

Recommendation

Publishers of legal information can learn from the experience of those that have already implemented ELI and can ask for help and assistance from the ELI Task Force. Furthermore, a list of references of ELI coordinators as well as ELI Task Force members is made available on the EUR-Lex website.

Building on shared knowledge allow publishers of legal information to anticipate risks and avoid mistakes made by others in the past, saving time and resources. Learning from the findings of legal analysis already carried out by other publishers as well as from technical implementation helps in avoiding reinventing the wheel and optimising the use of resources.

For instance, some countries greatly benefited from the support provided by the ELI Task Force, both in terms of technical support as well as more general background knowledge, in order to acquire a better understanding of how ELI works and the benefits derived from its implementation.

DOs

- Learn from the experience of those who have already implemented any of the pillars of ELI
- Share lessons learned while implementing ELI

DON'Ts

- Work in isolation
- Reinvent the wheel



BENEFITS



Quality and reliability

Avoids common pitfalls



Development of new services

Optimisation of resources and better focus on aspects to promote the development of added-value services



STEP 2: Estimate resources

When estimating the resources to implement ELI, publishers of legal information should take into account the fact that ELI implementation effort varies depending on:

- each ELI pillar, and
- the digitalisation degree of the existing publication process for legal information.

To help publishers carrying out such analysis, this document provides below an overview of the resources that should be allocated depending on whether:

1. a publisher of legal information needs to start implementation from the very beginning (**Scenario 1**); or
2. a publisher of legal information is already well advanced (**Scenario 2**).

For each scenario, it is possible to identify:

- the resources estimation associated with each of the implementation steps, which are further elaborated in the remainder of this document; and
- relevant professional profiles. Two key profiles are considered relevant:
 - **Knowledge information officer**, who should have a good understanding of information technologies in general, semantic web in particular, and expert knowledge of legal information. Experience in working with multiple metadata standards, frameworks and controlled vocabularies, familiarity with XML and XML schema languages together with knowledge of XPath, XSLT or XQuery are recommended; furthermore, experience in creating and implementing metadata schemas and taxonomies and controlled vocabularies together with gathering user requirements is a plus.
 - **Semantic web programmer/developer**, who should be proficient in software development in general, and have extensive experience with semantic web technologies. Important skills necessary for a semantic web programmer/developer are: experience in developing HTTP URI specifications; familiarity with data modelling, semantic web design principles, RDF, RDF/XML, Turtle, OWL, RDFa and triple stores; experience with languages for querying RDF (e.g. SPARQL); experience with web services technologies such as SOAP/WSDL, and RESTful web service architecture; expert knowledge of XML, XSLT, HTML5, CSS, JavaScript and related front-end web technologies; experience with



one or more programming languages such as: Java, Perl, Python, Ruby or PHP; and experience with data modelling and data extraction, harvesting techniques and tools as well as configuration and administration of web servers.

Figure 3, Figure 4 and Figure 5 below provide an overview of the estimated resources for the implementation of each ELI pillar. Implementation efforts are expressed in full-time equivalent (FTE), which is a unit to indicate the workload of an employed person.



Figure 3 — Resources allocation Pillar I



Figure 4 — Resources allocation Pillar 2

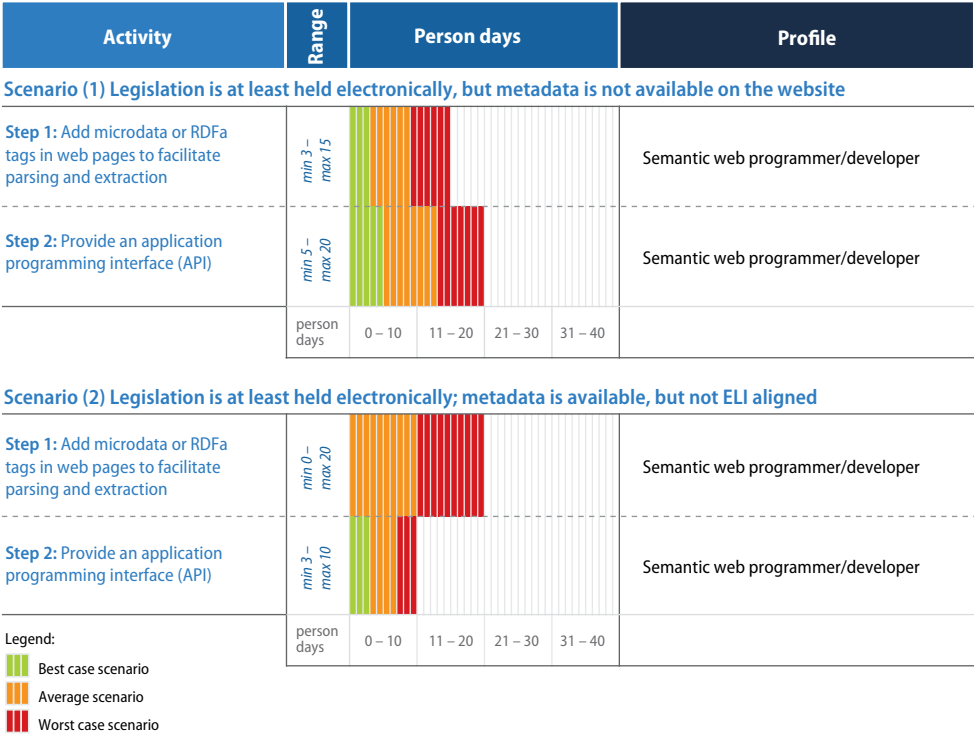


Figure 5 — Resources allocation Pillar 3



Good practice 2: Estimating implementation costs

Problem statement

In order to decide whether to implement ELI, it is fundamental to provide an accurate estimation of implementation costs. The development of semantic web solutions such as ELI necessitates publishers of legal information to take into account not only the technical developments required to implement ELI but also other associated costs.

Recommendation

To secure appropriate funds allocation, publishers of legal information should take into account three main cost drivers:

- product, including hardware and software;
- process, including organisational aspects;
- personnel, taking into account the need for different profiles.

To ensure accurate estimation, publishers of legal information should identify constraints so as to ensure that the estimates are meaningful.

The costs estimation should take into account not only the costs encountered during the development of each ELI pillar, but also the maintenance of the service provided. Publishers of legal information can rely on a variety of approaches to carry out their costs estimation including:

1. consulting the estimation of resources provided in this document;
2. comparing costs encountered by other publishers of legal information when implementing ELI or similar projects;
3. involving technical experts to estimate the necessary effort.



BENEFITS



Quality and reliability

Support to the effective implementation of ELI with appropriate allocation of resources

DOs

- Estimate effort, for example using full time-equivalent (FTE), which is a unit that indicates the workload of an employed person in a way that makes workloads across various contexts comparable
- Estimate cost for hardware and software
- Estimate the cost to manage risks

DON'Ts

- Estimate implementation effort without consulting technical experts
- Underestimate coordination efforts



STEP 3: Set up a governance structure

Once the decision to implement ELI has been taken, publishers of legal information should ensure that all stakeholders involved in the implementation project have a clear understanding of the decision-making process. For this purpose it is recommended to set up a governance structure which, for example, could be composed of: a Steering Committee, providing direction for implementation and review of the progress of ELI implementation; a Governance Committee, in charge of managing the operational team(s); and an Operational Team(s), mainly composed of technical staff, responsible for the implementation and maintenance of ELI.

When setting up the governance structure, publishers of legal information should take into account the ELI Council conclusions, which establish, in Article 15(c) and Annex 3, that Member States should appoint a central organisation as the national coordination point for ELI. More specifically, Annex 3 states that:

‘3. On national implementation

1.1. The national ELI coordinator

1. Each Member State using the ELI must appoint a national ELI coordinator. One country must not have more than one ELI coordinator.
2. The national ELI coordinator is responsible for:
 - (a) Reporting on the progress of the ELI implementation;
 - (b) Defining the applicable URI template(s) and communicating them to the Publications Office;
 - (c) Documenting available metadata and its relationship to the ELI metadata schema (if applicable);
 - (d) Sharing and disseminating information on ELI.
1. The national ELI coordinator should provide information to be published on the ELI website, as defined in paragraph 4, information describing the way the ordinal number is composed.’



Setting up a national coordination point facilitates the adoption of common policies relevant in the context of the ELI implementation, such as policies on persistent identifiers and metadata. Establishing a common policy, while not obligatory, can help in ensuring that relevant technical and organisational aspects are identified and tackled in a coordinated manner before any project is started. The references of all national ELI coordinators are available on the ELI website (EUR-Lex).

Once the organisational structure is established, publishers of legal information can then define change and release management procedures as well as quality control procedures. The ELI Task Force can provide guidance in this regard, if their assistance is requested.



Good practice 3:

Setting up a central organisation as a national coordination point for ELI implementation

Problem statement

ELI implementation requires carrying out a number of tasks which, depending on the existing administrative structure, might be assigned to different entities. Such a decentralised organisation, however, might lead to a number of drawbacks, most notably:

- inconsistencies in formulating coherent URI and metadata policies;
- increased costs for ELI implementation and maintenance.

Recommendation

In line with Article 15(c) of and Annex 3 to the ELI Council conclusions, publishers of legal information should appoint a national ELI coordinator responsible for:

- reporting on the progress of the ELI implementation;
- defining the applicable URI template(s) and communicating them to the Publications Office;
- documenting available metadata and its relationship to the ELI metadata schema (if applicable);
- sharing and disseminating information on ELI.

The national coordination points for ELI implementation are available on the ELI website.

DOs

- Follow the ELI Council conclusions and appoint a national coordinator for ELI implementation

DON'Ts

- Allow a decentralised ELI implementation at national level



BENEFITS



Interoperability

Coherent URI and metadata policies implemented at national level



Quality and reliability

Increased opportunities to align efforts for ELI implementation at national level



STEP 4: Set up an ELI implementation project

ELI can be implemented progressively. This means that publishers do not have to go through the full implementation of the three pillars at once. Instead, it is possible to take a step-by-step approach. For example, at the time of drafting this implementation methodology, Italy had decided to limit the ELI implementation to pillar I, while Ireland, France, Luxembourg, the United Kingdom and the Publications Office had decided to implement the first two ELI pillars.

The fact that publishers of legal information decide at a certain point to limit the ELI implementation to one or two pillars does not exclude the possibility of implementing the third one later on. For example, at the time of writing, Norway had a clear indication of the benefits derived from the implementation of the first two pillars, while the business case for the implementation of the third pillar was not clear at this stage and therefore its implementation had not been investigated yet. Keeping this flexibility in mind, the key steps that publishers of legal information need to take into account when setting up the ELI implementation project are:

1. **Define priorities** based on user requirements (cf. Organisation and policy, Step 1) and policy agendas. Since ELI can be implemented following a step-by-step approach, publishers of legal information can prioritise deliverables between 'must have' and 'nice to have'.
2. **Set up a team** including a project manager, representative of users, experts in information management and the semantic web, and additional stakeholders that are impacted by the ELI implementation.
3. Choose an **implementation approach** which emphasises the use of iterative developments whereby regular releases are foreseen (for example every 3 months) which allow building and refining the ELI implementation based on user feedback and use cases.
4. **Build the project timeline**, taking into account the time required to design, build and deploy the ELI pillars. By following the iterative implementation approach, the timeline will consist of short iterations of the design, construction and deployment phases for each pillar.
5. **Implement** the ELI pillars in line with the developed timeline and following the implementation steps provided in this document (cf. Pillar I 'How to put in place legislation identifiers based on ELI HTTP URIs', Pillar II 'How to design an ELI metadata schema' and Pillar III 'How to render ELI metadata machine-reusable').
6. Encourage use and **raise awareness** by providing relevant documentation and presenting the service via workshops, webinars and conferences.



Setting up the ELI implementation project will help to:

- identify high-level implementation options: typically this would entail the analysis of implementing only Pillar I or the first two pillars or all three pillars;
- understand whether to go ahead with the implementation and with what pillars;
- define efforts required for the ELI implementation and deadlines to be met;
- target the main functionalities to be implemented while adopting ELI.



Good practice 4: Structuring your implementation project

Problem statement

Each ELI pillar can be implemented progressively. Therefore, publishers of legal information have flexibility in deciding how to go about the ELI implementation. However, the great flexibility offered by ELI might be challenged by the absence of a well-structured project, as it can be difficult to identify the right scope and activities.

Recommendation

In order to ensure a successful implementation of ELI, publishers of legal information need to have a structured methodology for all the execution phases of the project. The methodology described in this document can be used to structure the work plan. Overall, key steps to be taken into account include:

- gathering requirements;
- analysis of existing publication processes;
- designing the solution;
- tests against real data;
- implementation.

In line with the implementation methodology presented in this document, publishers of legal information can break down the implementation of the three ELI pillars as follows:

- **Pillar I (HTTP URIs):** 1. Capture requirements for ELI HTTP URIs; 2. Design the URI template; 3. Configure the web server for ELI URI resolution;
- **Pillar II (metadata):** 1. Capture requirements for metadata; 2. Design and publish the metadata schema; 3. Develop metadata extraction routines; 4. Deploy legislation metadata.
- **Pillar III (rendering ELI metadata machine-reusable):** 1. Add microdata or Rdfa tags in web pages to facilitate parsing and extraction; 2. Provide an application programming interface (API).

DOs

- Build a proof of concept to demonstrate ELI implementation feasibility
- Break down the large-scale implementation into phases
- Take into account risks

DON'Ts

- Initiate your project before carrying out assessment and planning
- Postpone scope definition



BENEFITS



Quality and reliability

Project management in a control environment helps in delivering within agreed scope, budget, time and quality



STEP 5: Formulate statements about URI and metadata services provided

Once publishers of legal information decide to implement one or more ELI pillars, it is advisable to formulate a statement indicating the key aspects offered by the service provided. Important elements that the statement should contain are as follows:

- **Governance:** The national coordination point for ELI should be described as being the authoritative source of information for the implementation of ELI.
- **ELI URI template(s):** The statement should provide information about the URI template(s) and provide an indication of the URI persistence. The statement could be published on the ELI website.
- **Metadata:** The statement should also refer to the metadata schema, in case the ELI implementation also includes the second pillar. Such information should also provide information on existing relationships supported by the ELI metadata as well as mappings with other standards (cf. Pillar II, Steps 2, 3 and 4). Furthermore information about the quality of metadata could be included, as this is also important in terms of liability for the publishers of legal information.

In general, publishers of legal information are encouraged to say what level of commitment they want to set up, without having to necessarily quantify it.

Finally, the statement should equally address technical and non-technical profiles and should be published on the ELI website to raise awareness and promote reuse of legal information.

Examples of statements about ELI URI can be found following the links below:

PARTICIPANT	REFERENCE TO DOCUMENTATION
France	http://www.eli.fr
Ireland	http://www.irishstatutebook.ie/pdf/ELI_URI_schema.pdf
Publications Office	http://publications.europa.eu/mdr/eli/documentation/uri_templates.html
United Kingdom	http://www.legislation.gov.uk/developer/uris



Good practice 5:

Writing a policy document providing guarantees for long-term persistence

Problem statement

The absence of sufficient guarantees for the persistence of URIs and metadata when implementing ELI can have a negative impact on the level of trust among potential consumers of legislation. Reduced trust may, in turn, result in a limited use of the legal information published, therefore limiting the potential for the creation of value-added services.

Recommendation

By drafting a policy document that provides enough guarantees with regard to persistence and quality of the information provided, publishers increase the confidence to rely on both HTTP URIs as well as metadata, to access and reuse legal information. Such a policy document should address: 1. Needs: for example, the need for stable URIs, or the need to avoid duplication of identifiers for the same resource; 2. Governance structure: there should be a short description of how the management of URIs and metadata is performed, and who is responsible for a proper maintenance of the publication service; 3. Technical aspects: the policy document should make a reference to the technical aspects of the URI policy and metadata policy. The policy document should be published on the ELI website and promoted by the publishers of legal information.



BENEFITS



Access to legislation

Clear understanding of access and reuse requirements



Interoperability

Promotion of greater interoperability thanks to clear documentation



Quality and reliability

Increased level of commitment to provide a quality data service

DOs

- Write and publish a policy document for persistent ELI URIs and metadata on the ELI website
- Promote the ELI website
- Describe how persistence will be ensured if the national coordination point changes
- Consult policy documents of those that have already implemented ELI

DON'Ts

- Ignore the needs of consumers of legislation to have guarantees of persistence
- Take for granted the trust of consumers of legislation in your service levels



Pillar I: How to put in place legislation identifiers based on ELI HTTP URIs



STEP 1: Capture requirements for HTTP URIs

The first step that publishers of legal information should take when approaching the implementation of the first ELI pillar (i.e. HTTP URIs) is to get a detailed understanding of user needs and — if needed — update the business case (cf. Organisation and policy, Step 1). Many techniques are available for gathering user requirements. In order to gain a complete overview of the requirements that are relevant for the implementation of ELI HTTP URIs, publishers of legal information might consider the following approaches:

- **Interviews:** Publishers of legal information should gather detailed information on what type of tasks users carry out with legal information and understand how they cite legal information and what type of information they expect to find online. Interviews are very useful in this regard because they allow the carrying out of an in-depth analysis and uncover new requirements.
- **Questionnaires:** Questionnaires are useful when publishers of legal information need to gather a great amount of requirements from a wide audience, including users in remote locations that cannot be interviewed individually. It is advisable to complement the results gathered via questionnaires with more in-depth approaches.
- **Prototyping:** This technique allows the gathering of requirements that will help publishers of legal information to build the initial version of the HTTP URI. The development of a prototype on HTTP URIs will help users provide feedback which will be used to refine the HTTP URI implementation. This process should continue until the HTTP URIs template meets the critical mass of user needs.
- **Use cases:** Use cases are useful to describe how and for what purpose HTTP URIs are used. Examples of use cases for HTTP URIs might include: consultation of legislation as it stood at a specific point in time; access to the consolidated version of specific piece of legislation; reference to legal information in social media.

On top of gathering user requirements, publishers of legal information should take into account needs that are more related to the characteristics of legal information as well as the environment where the HTTP URIs will be developed, such as:

- Whether there is a **single authority** that attributes the identifiers or whether the attribution of identifiers is done in a decentralised manner. For example, in a federal system like Germany, since there are 16 *Bundesländer* that promulgate legislation, the attribution of identifiers is decentralised;
- The fact that not all legal information might have an identifier to be used to **uniquely identify** legal information, which means that



alternative options should be considered. This is the case for Luxembourg where, for example the following regulation:

Règlement grand-ducal du 2 février 2015 portant organisation de la Conférence nationale des élèves

is identified uniquely with an ELI-compliant HTTP URI by adding a sequential number (cf. 'n7') as shown below:

<http://eli.legilux.public.lu/eli/etat/leg/rgd/2015/02/02/n7> ⁽¹⁾

Another example can be seen in Norway. Before Lovdata ⁽²⁾ existed not all regulations had numbers. As a result, when Lovdata came into existence, it introduced numbers to legislation in a systematic way. Where numbers did not exist (i.e. for regulations and statutes published before 1982), Lovdata retroactively assigned them and from 1982 all regulations got an official number published in the Legal Gazette.

- The fact that certain publishers of legal information will have **two distinguished identifiers**: one for production purposes and the second one used for dissemination purposes.

For example, France uses ELI as well as the 'NOR' which is used for production purposes. The NOR is composed of 12 alphanumerical characters.

For example, the *Décret n° 2007-435 du 25 mars 2007 relatif aux actes et aux conditions d'exercice de l'ostéopathie* has the following **NOR**: SANH0721330D and the following two **ELI HTTP URIs**:

1. ELI: <http://www.legifrance.gouv.fr/eli/decret/2007/3/25/SANH0721330D/jo/texte>
 2. Alias: <http://www.legifrance.gouv.fr/eli/decret/2007/3/25/2007-435/jo/texte>
- The need to identify whether legislation is published as an **individual resource**, or in the context of an official journal, or both. For example, Italy provides ELI for both:
 - **Official journal**: www.gazzettaufficiale.it/eli/id/1989-03-02/089A0882/serie_generale, where:
 - 1989-03-02: date legal gazette;
 - 089A0882: identifier;

⁽¹⁾ The elements of the ELI templates are explained here:

etat = jurisdiction/agent, country-wide

leg = type, legislation

rgd = sub-type, règlement grand ducal

2015/02/02 = date of signature

n7 = sequence number, distinguishing legislation signed on the same day

⁽²⁾ Lovdata is a foundation which publishes judicial information on Norway (link: https://lovdata.no/info/information_in_english)



- serie_generale: OJ series;
 - /eli/{jurisdiction}/{agent}/{sub-agent}/{year}/{month} /{day}/
 {type}/{natural identifier}/{level 1...}/{point in time}/{version}/
 {language}.
 - **Individual resource:** <http://www.gazzettaufficiale.it/eli/Decreto/2013/65/art4>, where:
 - 'Decreto': type of act;
 - '2013': year; '65': number of act; 'art4': sub-part.
-
- Take into account when **consolidation** takes place (when applicable). For example, France consolidates legislation immediately after the amending act has been adopted and has put in place a specific database for it, which is called 'LEGI'. Ireland instead consolidates legislation on a case-by-case basis and mostly for large acts.
 - The existence of **more than one calendar**. This is the case for Ireland and the UK where prior to 1922 and 1963 respectively the regnal years were used to identify the year that legislation was published. To overcome such a challenge, these countries decided to redirect URLs with a calendar year before these dates to a URI based on the Gregorian calendar. Although such an approach worked for the majority of the cases, there was a small number of items of legislation held on the UK website where this redirection could not have been done unambiguously. For such cases, a list of candidates is given in response to the user, who then has to choose the right legislation.
 - Existence of **legacy identifiers**. It might be that when publishers of legal information implement ELI, they already assign identifiers to legal information. When this is the case, these identifiers can continue to be supported. For example, in the past Ireland went through two iterations of URI patterns. The first implemented pattern was not user friendly and it had to be replaced around 2004. The second URI pattern was a user-friendly pattern that had some resemblance to the current ELI URI schema. Due to the fact that there are still a high number of websites that link in to the eISB (the electronic Irish Statute Book) using the old pattern, redirection is planned from the old to the new HTTP ELI patterns.



STEP 2: Design the URI template

The ELI Council conclusions provide the ELI URI template that publishers of legal information can use to design the HTTP URIs for legal information.

The ELI HTTP URI template has been conceived in such a way as to enable the exchange of legal information across information systems and, to do so, it prescribes the design of the URI template on the basis of any of the following components:

/eli/{jurisdiction}/{agent}/{sub-agent}/{year}/{month}/{day}/{type}/{natural identifier}/{level 1...}/{point in time}/{version}/{language}

Each publisher can build its own URIs using these components, which are all optional, and do not have a predefined order. The use of a limited set of components provides a consistent way to identify legal information while keeping the necessary flexibility to take into account the specificities of each country's legal system. This in turn makes it easier to access legal information not only within a legal system but also across borders.

In order to be useful, identifiers of legal information should be easy to use, thus the closer identifiers are to the way the legislation is cited the easier it is to understand them, remember them and use them. For illustrative purposes, two examples are provided below; the first identifier is far from the way legislation is cited, the second one (ELI HTTP URI) is close to the way legislation is cited.

Examples of identifiers close to the way legislation is cited

LEGAL INSTRUMENT	IDENTIFIER FAR FROM THE WAY LEGISLATION IS CITED	IDENTIFIER CLOSE TO THE WAY LEGISLATION IS CITED (ELI HTTP URI)
EU: Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the reuse of public sector information	32003L0098	http://eur-lex.europa.eu/eli/dir/2003/98/oj
Italy: DECRETO 27 febbraio 2013, n. 65 Regolamento, di cui all'articolo 16, comma 1 del decreto legislativo 1° giugno 2011, n. 93, per la redazione del Piano decennale di sviluppo delle reti di trasporto del gas naturale	13G00106	www.gazzettaufficiale.it/eli/Decreto/2013/65



As the examples above demonstrate, it is easier to remember and re-use an identifier that is close to the way legislation is cited compared to a number with a more complex structure that is less intuitive.

When designing the ELI URI template(s) publishers of legal information should follow the following steps:

- identify how legal information is cited (Good practice 6);
- ensure that each legal resource is identified uniquely (Good practice 7);
- take into account the concepts of the ELI ontology: work, interpretation and manifestation (Good practice 8);
- identify sub-parts of legal information (Good practice 9);
- follow good practices for HTTP URIs (Good practice 10); and
- test URI templates (Good practice 11).

Publishers of legal information can find a number of examples of ELI HTTP URIs on the ELI website. An example is the HTTP URI template implemented in Luxembourg shown together with a concrete example in the table below.

COUNTRY/LEGAL INSTRUMENT	ELI TEMPLATE/EXAMPLE
Luxembourg	<code>/eli/etat/leg/{type}/{année}/{mois}/{jour}/{id}</code>
Règlement grand-ducal du 2 février 2015 portant organisation de la Conférence nationale des élèves	http://eli.legilux.public.lu/eli/etat/leg/rgd/2015/02/02/n7

A second example is the HTTP URI template used in Ireland.

COUNTRY/LEGAL INSTRUMENT	ELI TEMPLATE/EXAMPLE
Ireland	<code>/eli/{year}/{type}/{Natural identifier}/{Level 1...}/{Point in Time}/{Version}/{language}</code>
Motor Vehicle (Duties and Licences) Act 2012 (10/2012)	http://www.irishstatutebook.ie/eli/2012/act/10/enacted/en

In addition, examples of ELI HTTP URIs templates used in France can be found here: <http://www.eli.fr/> ⁽¹⁾.

Finally, examples of ELI HTTP URIs templates developed by the Publications Office can be found in the specific section on URI templates on the [Metadata Registry](http://publications.europa.eu/mdr/eli/documentation/uri_templates.html) ⁽²⁾ website.

⁽¹⁾ <http://www.eli.fr/en/constructionURI.html>

⁽²⁾ http://publications.europa.eu/mdr/eli/documentation/uri_templates.html



Good practice 6:

Designing URI template(s) to stay as close as possible to existing citation practice

Problem statement

Legislation is cited in different ways depending on the country, type of legal instrument and user. For example, legislation may be cited according to: the title (short or long); the number of the legislation; the date of adoption. Citation practices are used to identify legal information and therefore the design of URI templates should comply with existing citation practices. If publishers of legal information do not follow this approach, they greatly reduce the potential to share, reuse, cite and link legal information.

Recommendation

When designing URI patterns, publishers should stay as close as possible to existing citation practices. This can be achieved by using any of the available elements of the ELI URI template in any order. It is important that URIs are flexible and user friendly. Elements found in ELI URI templates that are commonly used in citation practices are: year, type of legislation, identifier and version (⁽¹⁾). See the following French example:

Vue le décret n° 2014-1169 du 10 octobre 2014 modifiant diverses dispositions réglementaires du code de la défense

ELI HTTP URI: <http://www.legifrance.gouv.fr/eli/decret/2014/10/10/DEFD1415169D/jo/texte>

A relevant standard for ELI implementation is the specification of the URI described in the Internet Engineering Task Force (IETF) Request for Comment (RFC) 3986, 'Uniform Resource Identifier (URI): Generic Syntax' (⁽²⁾). Further guidance on how to effectively use URIs in the context of the semantic web is given by W3C in the document 'Cool URIs for the Semantic Web' (⁽³⁾). Additional principles that should be applied in designing URI templates are described in the document '10 Rules for Persistent URIs' (⁽⁴⁾).

DOs

- Design ELI URI template(s) close to citation practice
- Ensure that URIs are human readable/reader friendly

DON'Ts

- Add elements in the URI template that are not necessary to identify a piece of legislation in a unique way



BENEFITS



Access to legislation

Improved discovery of legislation



Interoperability

Greater interoperability by means of common URI design patterns



Quality and reliability

Ensures availability of quality URIs

(⁽¹⁾) Examples of URI patterns can be found on the ELI website.

(⁽²⁾) <https://tools.ietf.org/rfc/rfc3986.txt>

(⁽³⁾) <http://www.w3.org/TR/cooluris>

(⁽⁴⁾) <https://joinup.ec.europa.eu/community/semic/document/10-rules-persistent-uris>



Good practice 7: Modelling HTTP URIs by treating each piece of legislation as a unique resource

Problem statement

If one HTTP URI resolves to more than one piece of legislation, a request for that HTTP URI will receive in response all the pieces of legislation resolved by that URI. This is called the URI homonym problem, and can lead to confusion and interoperability conflicts, as it might not be clear what resource is associated with a URI.

Recommendation

When identifying legislation by means of HTTP URIs, it is important to first ensure that each HTTP URI identifies in a unique way each piece of legislation, i.e. publishers should ensure that a given HTTP URI will identify one and only one piece of legislation. Difficulties might occur typically when:

- the legislation does not have a number;
- the same number is attributed to more than one piece of legislation.

If the legislation does not have a number, it is recommended to attribute one in the URIs. This is the approach followed by Luxembourg, where legislation is cited without number as the example shows:

Règlement grand-ducal du 2 février 2015 portant organisation de la Conférence nationale des élèves.

The URI identifier for this act includes 'n7' in order to uniquely identify it

<http://eli.legilux.public.lu/eli/etat/leg/rgd/2015/02/02/n7>

If the same number is attributed to more than one piece of legislation, it is important that the URI is designed in such a way that it includes enough elements to distinguish the two pieces of legislation, for example by including dates of adoption, the author, jurisdiction, etc. The choice of which elements to select is left up to the publishers of legal information, since each legal system has its own specificities.



BENEFITS



Access to legislation

Unambiguous assignment of URIs to legislation resources makes the legislation more discoverable



Interoperability

Reduced risk of interoperability conflicts

DOs

- Design a mechanism to assign additional numbers or character strings to disambiguate URIs if they lead to the same texts
- Use aliases in order to strike a balance between consistency of URI scheme and reality

DON'Ts

- Try to model an official journal; treat legislation as resources by themselves



Good practice 8:

Taking into account the concepts of the ELI ontology: work, interpretation and manifestation

Problem statement

ELI describes legal information based on the well-established model of 'Functional requirements for bibliographic records' (FRBR) and distinguishes between the concepts of:

- 'work' (intellectual or artistic creation);
- 'interpretation' (also called 'expression', which represents the intellectual or artistic realisation of a work); and
- 'manifestation' (materialisation of one of the expressions of a work).

Each time a 'work' is realised it takes the form of an 'interpretation'. The physical embodiment of an interpretation is a 'manifestation'.

In line with the commonly known best practices for linked data, such a structure enables the retrieval of various resources representations via content negotiation (cf. Good practice 12), but to do so it is necessary to ensure that the HTTP URLs that identify the legal information also follow this structure.

Recommendation

While the use of the FRBR structure should be transparent to the user when navigating through the URLs, publishers of legal information should follow the FRBR structure when designing the HTTP URLs.

By following the FRBR structure, publishers of legal information can provide the set of metadata at each level of the FRBR hierarchy, contributing to the development of the 'web of data'.

For example for Directive 2013/37/EU the ELI HTTP URLs could look like:

- eli/dir/2013/37/oj - work
- eli/dir/2013/37/oj/eng - interpretation (expression)
- eli/dir/2013/37/oj/eng.rdf manifestation.

DOs

- Take into account the ELI ontology structure when designing the ELI URIs

DON'Ts

- Structure ELI HTTP URIs following another representation than the FRBR model



BENEFITS



Access to legislation

Allows retrieval of various legal resources representations



Interoperability

Contributes to the creation of the 'web of data' since it allows provision of the set of metadata at each level of the FRBR hierarchy



Good practice 9: Considering the right level of granularity when identifying legislation

Problem statement

Citations to legal information can include references to sub-parts of a legal act. However, the production workflows are not always designed to handle identification of legal information beyond the identification of the act as whole. Limiting the identification of legal information to a general level, without identifying its sub-parts, hampers the potential reuse of legal information. For example, users might want to cite directly a specific provision of a piece of legislation when using social media or other online tools, instead of having to refer to the legislation as a whole.

Recommendation

When implementing ELI, it is important to identify the right level of granularity of the HTTP URIs. For this purpose, the Council conclusions define the ELI template components 'Level 1', 'Level 2', 'Level 3' and 'Level n' as references to subdivisions of an act, such as articles or even smaller subdivisions. For example, the UK has developed a URI model that considers the 'section' parameter. This parameter refers to particular sections, articles and regulations within a piece of legislation. The template of the URI including this parameter is the following:

- <http://www.legislation.gov.uk/id/{type}/{year}/{number}/{section}>

The identifier URI for Act of the Scottish Parliament 2014, Number 2 is:

- <http://www.legislation.gov.uk/id/asp/2014/2>

The identifier URI for Section 3 of the same document is:

- www.legislation.gov.uk/id/asp/2014/2/section/3

In general, if the production workflow is based on XML structures, it may be relatively easy to determine the more granular levels by processing the XML records.



BENEFITS



Transparency

Improved understanding of the context in which legal information exists



Access to legislation

Increases the type of information that is searchable and indexable



Interoperability

Allows better interaction between information systems



Savings

Increased interoperability between IT systems can lead to reuse of information, and thus creates opportunities for cost optimisations

DOs

- Design ELI URIs that take into account sub-parts of legal acts
- Keep in mind the way external users will want to use the URI, for example in social networks or in value-added services

DON'Ts

- Assign ELIs only at the level of acts and ignore the benefits of more granular identification and description
- Use anchors in the HTML code



Good practice 10: Following common good practices for HTTP URIs

DOs

- Use or create an integrated workflow for publication that includes the assignment of identifiers
- Use elements in the URI template(s) to indicate sub-parts of legislation (e.g. an article of a piece of legislation you want to identify)
 - Create pleasing and recognisable identifiers so that people will understand what they can expect to find when they click on it
- Build the ELI HTTP URIs out of existing web infrastructure, because this is what people are familiar with

DON'Ts

- Use more elements than necessary in the URI template
- Include elements in the ELI that might be prone to change
 - Use different URIs to identify the same piece of legislation.

Problem statement

Designing URI templates in isolation, without learning from the experiences of those who have already done it, increases the risk of repeating the same mistakes, duplicating analysis and wasting resources by solving problems that have already been addressed by others.

Recommendation

While it is not possible to provide a fit-for-all URI template, there are a number of good practices that publishers can follow in order to foster the access to and reuse of legal information. Key good practices include:

- avoiding using more elements than necessary to identify in a unique way legal information;
- staying as close as possible to existing citation practice (cf. Good practice 6);
- avoiding using elements that tend to change in HTTP URIs.

These elements are important to guarantee the persistence of the URIs. Furthermore, when URIs are deprecated, meaning that they should no longer be used, they should nevertheless be maintained both in terms of the actual URI and the resource they identify. Once minted, a URI should not be deleted and a redirection mechanism should be put in place (cf. Good practice 15). Where URIs identify resources that are versioned, new URIs should be minted for each new version of the resource as well as a URI for the latest version.



BENEFITS



Quality and reliability

High-quality URIs which meet users' needs



Good practice 11: Testing URIs against real data

Problem statement

When designing URI templates, publishers of legal information might face a number of challenges such as being able to correctly identify sub-parts of legislation, to identify uniquely legislation that does not have a unique identifier or to handle dates where more than one type of calendar year applies. These difficulties may not be immediately apparent during the design phase of ELI HTTP URIs and it may not be possible to foresee all possible scenarios.

Recommendation

To reduce the risk of developing URI design templates that do not take into account all possible scenarios, it is recommended that tests be performed as early as possible. The tests should reveal issues related to the correspondence between URIs and the way legislation is cited; identification of consolidated acts (when applicable); and resolution mechanisms. Tests should also include simulations of probable scenarios that can happen during the publication process, such as: 1. create a new legal act, 2. update the content of an act by altering its structure, 3. replace an act, 4. amend an act, 5. consolidate an act, etc. The tests should be designed to match the context of each implementer. For example, the Publications Office had to find a way to assign unique and persistent URIs to legislation that did not have an identifier, such as corrigenda which led to the following URI template for corrigenda:

http://data.europa.eu/eli/{typedoc}/{year}/{natural_number}/corrigendum/{pub_date}/oj where 'typedoc' is the resource type of the corrected act and 'pubdate' is the date of publication of the corrigendum in the Official Journal.



BENEFITS



Access to legislation

Ensures that all possible scenarios are taken into account



Interoperability

Ensures quality of URI templates, which in turn increases trust and potential for others to rely on them



Quality and reliability

Reduced risks of wrongly identifying legal information

DOs

- Choose diverse datasets for testing purposes
- For each tested dataset, compare the obtained ELI URIs to the way people cite legislation
- Test all the activities related to your URI management processes
- Test as many scenarios as possible and check results with users from the business side

DON'Ts

- Start large-scale implementation without testing your ELI URI template(s)
- Test only main scenarios related to management activities of your URIs



STEP 3: Configure web server for ELI URI resolution

Publishers of legal information should design HTTP URIs in such a way that they resolve to legislation. Resolution means that if an HTTP request is issued (either by a human using a web browser, or by a machine/an application), the system returns either:

- a page with a description of a legal resource;
- the legal resource itself (a document).

Legislation resources must be associated to content information so that resolution of HTTP URIs leads to a response with data.

ELI does not impose a specific solution. Depending on different scenarios, publishers of legal information may decide whether to return metadata and content or one of the two. For example, if a publisher of legal information does not have metadata, but only the full text of the legislation in PDF format, ELI does not oblige the publisher to implement metadata. In this case, it is possible that the publisher of legal information decides to only return the actual legal resource.

It is useful to consider the following components:

- **HTTP client:** a browser (used by humans) or an application (used by machines) that issues HTTP requests;
- **web server:** an application that receives HTTP requests and sends HTTP responses back. These responses can either be redirections or web documents; and
- **legislation registry:** The legislation resources are contained in a register; the registry is a software application used to manage them.

The resolution mechanism refers to the service that handles URI requests of HTTP clients and attempts to obtain a representation of the resource that the URIs identify.

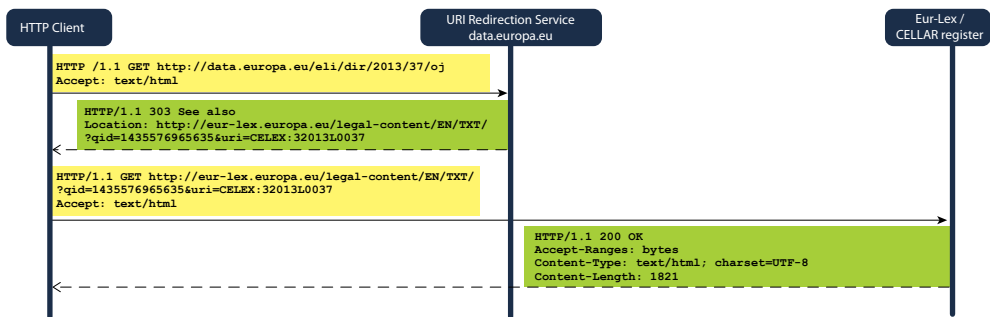


Figure 6 — Example of URI Redirection (HTTP 303 redirection)

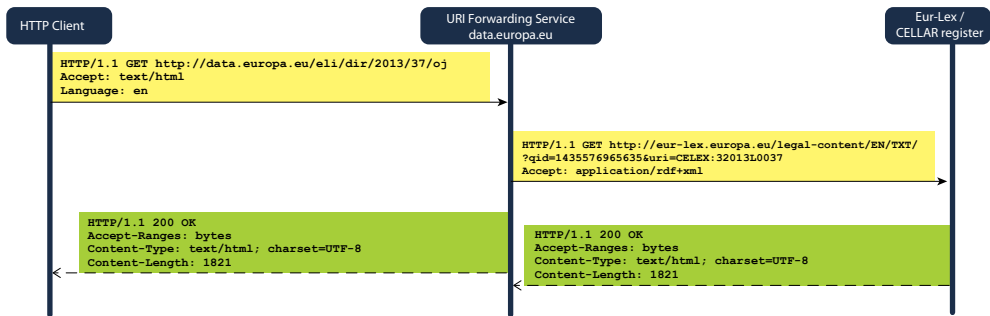


Figure 7 — URI forwarding

It is recommended that tests be carried out to uncover and fix any potential issue related to the design of URIs or the resolution mechanism.

Finally, it is important to always schedule production deployments and publicly announce such events well in advance. It is also useful for the publishers of legal information to offer a short description of the changes that will occur after deployment. By making such information publicly available, consumers of legislation are aware of upcoming changes and can be prepared.



Good practice 12:

Using content negotiation and/or identifiers for language and format variants

Problem statement

If legislation — or the metadata describing legislation — is available in multiple languages and formats, publishers of legal information need to decide how to identify these different languages and formats.

Recommendation

To identify language and format variants, different persistent URIs can be used possibly in combination with content negotiation. As described in Good practice 8, different language versions of legislation can be considered to be different 'interpretations', whereas different formats can be considered to be different 'manifestations' of the same 'work' following the ELI ontology. Minting different URIs for these language and format variants of the same 'work' is a good practice, whereby only meaningful language and format indicators should be used. For example, the use of technology-specific extensions like .asp, .aspx, .jsp in URIs should be avoided as this is prone to change; meaningful format indicators like .html, .rdf, .pdf should be allowed. For example, the following URI could be used to distinguish an English-language variant in HTML format:

<http://data.europa.eu/eli/dir/2013/37/oj/eng.html>.

Additionally, it might be useful to apply content negotiation, although this is not explicitly required by the ELI Council conclusions. Content negotiation is the process that takes place between a web server and a web agent to select and provide the best representation for a given response, when there are more representations available. Such a mechanism allows sending a request to a web server and specifying the desired specific representation in the header of the request. For example, requesting the resource: <http://data.europa.eu/eli/dir/2013/37/oj> and providing in the request's header the 'Accept' parameter with the value: 'text/html', and 'Accept-Language' parameter with the value 'en', would allow the server to choose the appropriate response, which in this example could be: <http://data.europa.eu/eli/dir/2013/37/oj/eng.html>.

DOs

- Include meaningful format indicators in ELI URIs such as .html, .rdf, etc.
- Consider using content negotiation

DON'Ts

- Use technology-specific extensions in the ELI URIs like .asp, .aspx, .jsp, etc.



BENEFITS



Access to legislation

Better access to legislation by allowing consuming applications to request the format that is most appropriate in a particular situation



Interoperability

Increases the flexibility of adding or removing representation formats of resources while preserving persistence of URIs



Good practice 13: Providing service-level guarantees

Problem statement

Potential consumers of legal information might be reluctant to rely on URIs if:

- they do not have enough guarantees that URIs are resolvable and persistent; and
- they do not have a reasonable request–response time and an acceptable level of data quality.

Recommendation

It is recommended that service-level guarantees be offered as part of the URI policy. The service-level guarantee should be a commitment to offer:

- resolvable URIs;
- long-term persistence of minted URIs;
- a reasonable request–response time;
- reasonable quality of legislation. The service-level guarantee should also provide information about how issues encountered by the legislation consumers will be reported, who will address reported issues and how the resolution will be communicated back to those who reported.

DOs

- Plan to provide service-level guarantees
- Plan to monitor performance of the resolution mechanism and the provision of service-level guarantees
- Ensure consumers of legislation can access service-level guarantees
- Review the service-level guarantees on a regular basis and adapt whenever necessary
-

DON'Ts

- Allow poor performance of the resolution mechanism, as this will reduce the interoperability potential of your solution
- Ignore issues reported by consumers of legislation



BENEFITS



Access to legislation

Increased consumption of published legislation



Interoperability

Improved trust and thus reuse potential



Quality and reliability

Improved reliability of legal information published



Good practice 14: Using HTTP 303 redirection ⁽¹⁾

Problem statement

Publishers of legal information that use HTTP URIs can identify either the legal resource itself (i.e. the document) or a web document describing that piece of legislation. For example Directive 2013/37/EU on the reuse of public sector information is described on one of the web pages of EUR-Lex ⁽²⁾, but a user can also access the document itself ⁽³⁾. The user may not be interested in the web page, but be interested in the document itself. So two URIs are needed, one for the piece of legislation and a second one for the web page describing the piece of legislation.

Recommendation

When publishing URIs that are meant to identify entities which are not a web document, such as a piece of legislation, publishers of legal information should mint URIs for the piece of legislation and also allow users to get a description of the piece of legislation (if available) using standard web technologies. One possibility is to implement a redirect mechanism, which is a World Wide Web technique for making a web page available under more than one URL address. In line with the approach proposed by the W3C interest group Note Cool URIs for the Semantic Web ⁽⁴⁾, one possibility is to respond to a GET for the URI representing the conceptual resource with an HTTP status code 303 (See other) and a location HTTP header with the URL of the actual file. The HTTP client will then issue a new GET for the URL received. The URI resolution mechanism should implement a redirection mechanism that works as follows: when an HTTP client issues an HTTP request for a piece of legislation identified by a certain URI, the response will be an HTTP 303 'See also' redirection response code with a new URI. The HTTP client will issue a request with the newly received URI which will be forwarded to the register holding the document describing the piece of legislation which will be sent back to the client.

DOs

- Plan to implement redirection mechanism

DON'Ts

- Limit URIs to resource-level identification



BENEFITS



Quality and reliability

Clear distinction between the legal resource and the web page describing it

⁽¹⁾ https://en.wikipedia.org/wiki/HTTP_303

⁽²⁾ <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1440509555192&uri=CELEX:32013L0037>

⁽³⁾ <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:EN:PDF>

⁽⁴⁾ <http://www.w3.org/TR/cooluris/#r303uri>



Good practice 15:

Monitoring website usage and links to legislation on the web through statistics

Problem statement

Monitoring the website traffic might not always be a common practice among publishers of legal information. As a consequence, they have a very limited insight about usages of legal information and any problems encountered by the users. Without having a good overview of the users' experience, publishers of legal information might miss opportunities to correct errors and ultimately improve the service according to the users' needs.

Recommendation

It is recommended that tools be implemented to regularly analyse the website traffic to optimise and improve access to published legislation. Key indicators that should be monitored include: 1. Visits: the number of visits to the website during a given time period; 2. Unique visitors: the number of different people who visited your website during a given time period; 3. Visitor demographics: the profile of the website's audience such as the geographic location, language, gender, age, etc.; 4. Bounce rate and time on site: whether visitors find what they are looking for in the website or if they leave it immediately; 5. Type of sources: it segments the traffic by specific sources and mediums, such as search engines, referring sites, email or custom campaigns; 6. Access by redirections of URLs vs. access by redirections by a search engine: users that look for legislation according to the URI templates; 7. Social media interactions: amount of visitors that interact with social media profiles.



BENEFITS



Quality and reliability

A careful watch of key traffic metrics will allow the provision of the stated service-level guarantees, and constantly optimise and improve the quality of your service

DOs

- Implement website traffic analytics tools
- Watch key performance indicators to constantly look for improvements

DON'Ts

- Ignore surveying the website traffic, as this can lead to a lack of persistence of your services



Good practice 16: Monitoring service levels and persistence of URIs

Problem statement

The added value that ELI brings in terms of increased persistence and common use of metadata to describe legal information could be greatly hampered by the absence of a reliable online service. In fact, if users are not able to access legal information in a consistent and timely fashion due to poor service performance, they will decrease the level of trust in the service and will tend not to rely on it in the future.

Recommendation

It is recommended that tools be deployed that allow the monitoring of the website traffic. Such monitoring will allow the discovery of issues such as broken links or long request–response times. By finding such issues on time, the appropriate actions can be taken in a timely manner to remedy the situation — thus providing the guarantee of a quality service.

The analysis of weblog files can provide valuable insights on the web server's performance. GoAccess ⁽¹⁾ is a popular choice for this task. It can run in a terminal window, and is useful to quickly analyse and view web server statistics in real time. It can be used to generate HTML, JSON and CSV reports from the weblogs.

Besides monitoring performance of the web traffic generated by visitors, it is recommended that load and stress tests be performed. Such tests allow understanding of the behaviour of a system under a specific expected load and its limits. Some of the most popular tools for load and stress testing are: curl-loader ⁽²⁾, ApacheBench ⁽³⁾ and Httpperf ⁽⁴⁾.

It is recommended that the response time to issues reported by the legislation be monitored and statistics gathered about the resolution of issues. This will allow the optimisation of the issue reporting–resolving mechanism.

DOs

- Monitor website service levels. Fix any observed or reported issue in a timely manner
- Review and optimise your issue-resolution mechanism

DON'Ts

- Ignore low request–response times and broken links
- Ignore issues reported by consumers of legislation



BENEFITS



Quality and reliability

Better quality of the service thanks to prompt replies to any occurring issues

⁽¹⁾ <http://goaccess.io>

⁽²⁾ <http://curl-loader.sourceforge.net>

⁽³⁾ <http://httpd.apache.org/docs/2.2/programs/ab.html>

⁽⁴⁾ <https://github.com/httpperf/httpperf>



Pillar II: How to design an ELI metadata schema



STEP 1: Capture requirements for metadata

The second ELI pillar is described in the annex to the Council of the European Union conclusions ⁽¹⁾ which explains that:

‘While a structured URI can already identify acts using a set of defined components, the attribution of additional metadata established in the framework of a shared syntax will set the basis to promote interchange and enhance interoperability between legal information systems. By identifying the metadata describing the essential characteristics of a resource, Member States will be able to reuse relevant information processed by others for their own needs, without having to put into place additional information systems. Therefore, while Member States are free to use their own metadata schema, they are encouraged to follow and use the ELI metadata standards with shared but extensible authority tables, which permit to meet specific requirements. The ELI metadata schema is intended to be used in combination with customised metadata schemas.’

Metadata provides the means for identifying, describing, classifying and finding legal information and enabling users to access and reuse it and in the context of legal information it includes for example: title, date of adoption, date of signature, author of a given piece of legislation. Having the ability to search and access legal information through that metadata makes it much easier for someone to locate a specific piece of legal information and use the information. In order to ensure that the metadata provided meet the user needs, publishers of legal information should gather user requirements using the techniques highlighted in Pillar I, Step 1 and, if necessary update the business case (cf. Organisation and policy, Step 1).

Relevant aspects that publishers of legal information should investigate are:

- **the reasons** users access legal information, such as personal or professional reasons;
- the **way(s) users access** legal information, for example whether they type key words into general search engines or reference numbers directly on legal information websites;
- **what information users are looking** for and what obstacles they encounter, for example with regard to the identification of relationships between various pieces of legislation or consultation of specific articles;

⁽¹⁾ OJ C 325, 26.10.2012.



- **profiles of users**, whether they belong to the private or public sector, whether they are mainly national or international or internal users (users studies, website statistics on usage and links into your legislation on the web);
- **usages of legal information**, for example whether users cite a piece of legislation, establish cross-references between different types of data, reference legal information via social media or develop mobile applications;
- **the expectations** users have in terms of availability and quality of legal information.

Thanks to the requirements collected, publishers of legal information will be able to identify what legal information should be described according to ELI-compliant metadata. For example, if users are mainly interested in consulting in-force legislation, an option could be to focus on applying ELI-compliant metadata to this collection first. Another option could be to apply ELI-compliant metadata not only to legislation, but also to relevant information that helps users to understand what the legislation is about, such as summaries of legislation.



Good practice 17: Following a user-centered design taking into account type of users

Problem statement

Publishers of legal information might implement the ELI second pillar without knowing who will use the service and for what purposes. Discarding such information increases the risk of implementing a solution that does not answer to the needs of those that will consume legal information, therefore reducing its potential to be used.

Recommendation

It is recommended that the metadata be defined taking into account both user profiles (such as legal professionals, journalists, academics) and their requirements by looking at: 1. What metadata helps users find what they are looking for; 2. For what purposes users want to consult legal information; 3. What format can help consumers of legal information to further use it; 4. How often legal information is consulted; 5. What information users need to understand legal information.

Luxembourg, for example, gathered a list of queries that users would like to make such as: all acts that implement a directive (or an article thereof) on a specific date, after deadline or partially; all acts adopted by a ministry in a given year, with a specific subject (such as 'water'); acts based on parliamentary number or date of signature; all acts adopted that are based on a particular article of the constitution.

DOs

- Carry out an in-depth investigation of user requirements
- Document user requirements in the business case
- Identify key metrics to measure whether the needs are met

Collecting user requirements greatly helps publishers of legal information to better understand what ELI metadata should be implemented. For example, if users look for legal information by the legislation number, it is a good idea to include the metadata 'eli:id_local'. Similarly, if users are interested in the date of entry into force, the following metadata can help: 'eli:in_force'. To then ensure that the metadata actually meets the identified requirements, publishers of legal information should monitor the website usages based on predefined metrics (cf. Good practices 15 and 16).

DON'Ts

- Engage in development without performing an appropriate analysis of user requirements



BENEFITS



Quality and reliability

Provides services that meet user requirements



STEP 2: Design and publish metadata schema

Designing a metadata schema in line with ELI specifications is a fundamental step to enable publishers of legal information to ‘... identify and exchange legal information originating from regional and national authorities at the European Level’ as explained by the ELI Council conclusions ⁽¹⁾.

There are three approaches that publishers of legal information can follow when designing a metadata schema that is ELI compliant and these are:

1. **Use the ELI metadata schema (ELI ontology)** ⁽²⁾: This option entails designing a metadata schema using the ELI ontology, including the full list or a subset of ELI-compliant metadata elements.
2. **Customise the ELI metadata schema**: This second option entails defining an application profile of the ELI ontology that is a subset of the ELI classes and properties — with specific controlled vocabularies — and possibly supplemented with additional metadata elements. This option is in line with the ELI Council conclusions which state that:

‘... while Member States are free to use their own metadata schema, they are encouraged to follow and use the ELI metadata standards with shared but extensible authority tables, which permit to meet specific requirements. The ELI metadata schema is intended to be used in combination with customised metadata schemas’ ⁽³⁾.

3. **Provide a mapping to the ELI metadata schema**: The third option entails creating mappings between the ELI ontology and other existing ontologies used for a similar purposes. This option is most useful for those publishers of legal information that already use a metadata schema and that want to align it with the ELI metadata schema at conceptual level only, i.e. using another syntax binding. Examples of mapping tools that publishers of legal information can use for this purpose are:

- Mapping.semic.eu ⁽⁴⁾;
- Silk: a tutorial on the use of silk to align controlled vocabularies is available on Joinup ⁽⁵⁾.

⁽¹⁾ OJ C 325, 26.10.2012, p. 1.

⁽²⁾ http://publications.europa.eu/mdr/eli/documentation/ELI_Ontology-v1.0.htm

⁽³⁾ OJ C 325, 26.10.2012, p. 6.

⁽⁴⁾ <http://mapping.semic.eu>

⁽⁵⁾ <https://joinup.ec.europa.eu/community/semantic/document/tutorial-use-silk-aligning-controlled-vocabularies>



With regard to the third approach, Switzerland has already matched ELI metadata with fields of two other standards: Akoma Ntoso and URN:LEX. This resulted in the OPC Naming Convention ‘Namenskonvention chLEX und KAV (URI-Schema) P2 final’.

Similarly, Norway has developed a first draft of ELI URI templates and mapping of metadata to the ELI ontology. The mapping was done by identifying which metadata from the ELI ontology are present in the national specifications, either directly fetched from their own metadata, translated from it or derived from other data. Norway also identified which metadata cannot be mapped, e.g. the metadata ‘agent/authority’ has been mapped, while metadata ‘sub-agent/sub-authority’ has not been mapped due to lack of correspondence.

Additionally, using structured metadata allows the transformation of documents from native formats (such as XML) to a different one. For example, the National Archives (UK) is working to support HTML5 and has been experimenting with a near one-to-one serialisation of Akoma Ntoso in HTML5. The goal has been to follow the structure of Akoma Ntoso as closely as possible, while using all of the native semantics of HTML5. The native XML schema governing UK legislation is called the Crown Legislative Markup Language (CLML). A second example is represented by Italy that uses the URN:LEX standard, which can be serialised to ELI-compliant specifications.

Independently from the approach used, publishers of legal information should store the metadata schema in an **authoritative source**, make it available to the public on a public website or by download, define a release schedule and put in place a change and release management process.

This information should be made available in both human and machine-readable format so that people can:

- understand the metadata; and
- reuse the metadata to develop value-added services.

Further information on the metadata specifications can be found in the [Documentation](#) ⁽¹⁾ section on the [Metadata Registry](#) ⁽²⁾ website.

⁽¹⁾ <http://publications.europa.eu/mdr/eli/documentation/index.html>

⁽²⁾ <http://publications.europa.eu/mdr/eli/index.html>



Good practice 19: Considering the right level of granularity when creating metadata

Problem statement

ELI metadata can describe legal acts at different levels of granularity, and publishers of legal information can choose to implement very little metadata, for example by identifying the type of legal information using 'eli:type_document' to very granular descriptions, using for example 'eli:relevant_for', 'eli:passed_by', 'eli:publisher', etc. The first approach is easier to follow, because it requires less metadata to be generated. However, in certain circumstances generating more granular metadata meets user requirements better.

Recommendation

It is recommended that the appropriate level of granularity be chosen when creating metadata that describes legal information, in line with user requirements and needs. More granular metadata offers a greater variety of information that can be reused, enabling greater levels of technical manipulation. It is important to strike the right balance between a level of granularity that allows publishers of legal information to organise and maintain content in the most appropriate way and helping consumers of legal information to find and reuse what they are looking for. Publishers of legal information should also look for other relevant initiatives that classify sub-parts of a piece of legislation and that can be used when implementing ELI. Examples of such initiatives include:

- the Interinstitutional Formats Committee (IFC) Common Vocabulary on subdivisions of acts (currently under development); and
- the Publications Office Authority Table on Subdivisions (also under development) ⁽¹⁾.



BENEFITS



Access to legislation

More functionalities available to access legislation



Interoperability

Further exchange of information allowed



Quality and reliability

More detailed information available



Development of new services

More information is provided and more functionalities can be built on top of it

DOs

- Create sufficient metadata at various levels of granularity
- Create sustainable metadata for future purposes balancing the cost of creation and maintenance

DON'Ts

- Create metadata with insufficient details

⁽¹⁾ <http://publications.europa.eu/mdr/authority/index.html>



Good practice 20: Structuring information by linking it to other documents

Problem statement

Legislation does not exist in isolation. For example, there is national legislation that transposes EU directives or national legislation that amends or repeals other national legislation. The ability to establish the relationship between different pieces of information, such as for example, establishing the link between basic acts and related amendments, is key for supporting a good understanding of the evolution of a given law. If publishers of legal information do not take the necessary steps to ensure that legal information is linked to relevant data, it is more difficult for users to understand, share and reuse legal information.

Recommendation

Publishers of legal information should design the ELI metadata schema making explicit relationships between acts. Examples of relationships that can be expressed using the ELI metadata schema include: textual references, amendments, consolidations, transposition of EU law into national laws. ELI metadata helps publishers of legal information to make explicit a variety of relationships between legal resources, boosting the analysis that can be carried out on legal information. Examples of available ELI metadata ⁽¹⁾ include:

- 'related to', defined as 'Reference to draft bills, judgments, press releases, etc.';
- 'consolidates', defined as 'Reference to the consolidated version of the resource';
- 'transposes', defined as 'References to other legal resources that allow Member States to adopt relevant legislation'.

Linking information to other documents can also help in developing innovative applications, for example, to create visualisations that enable a quick overview of relationships between documents ⁽²⁾.

DOs

- Reuse vocabularies whenever possible, as identified in the ELI metadata specifications

DON'Ts

- Miss the opportunity to make the links between legal resources explicit



BENEFITS



Access to legislation

Easier identification of the wider context to which legal information relates



Interoperability

Improved links and exchange of information

⁽¹⁾ OJ C 325, 26.10.2012, p. 6.

⁽²⁾ Cf. ELI success story about the *Service central de législation (SCL)* in Luxembourg.



Good practice 21: Mapping ELI metadata with other standards

Problem statement

Publishers of legal information might have implemented metadata standards that predate ELI and they might be inclined to think that implementing ELI would entail making a choice between one standard over another one. However, such an approach would increase the barriers to the management of digital resources and would go against a key objective of ELI, which is increasing interoperability between information systems.

Recommendation

If publishers of legal information already use metadata standards to describe legal information, they can map them to ELI-compliant metadata thanks to the clear and precise definition of the elements that can be found on the ELI website. In such a way, it will be possible to benefit from the increased interoperability offered by the use of ELI, while maintaining existing metadata standards. It is recommended to document the mappings using the relationships of the SKOS vocabulary:

- exact match: two concepts are equivalent;
- close match: two concepts are very close;
- related match: two concepts are related;
- broad match: one concept is more general;
- narrow match: one concept is more specific.

An example, of possible mapping between ELI ontology ⁽¹⁾ elements and Akoma Ntoso ⁽²⁾ is provided here:

- eli:type_document > akn:documentType (exact match)
- eli:cites > akn:citation (exact match).



BENEFITS



Access to legislation

Alignments of metadata standards foster access



Interoperability

Mapping metadata facilitates the exchange of information



Savings

Simplification of the information exchange process and reuse of information

DOs

- Document and publish your mappings
- Use controlled vocabularies to implement and validate metadata

DON'Ts

- Keep your metadata schema isolated from other metadata schema standards

⁽¹⁾ <http://publications.europa.eu/mdr/eli>

⁽²⁾ <http://docs.oasis-open.org/legaldocml/akn-core/v1.0/csprd01/part2-specs/akn-core-v1.0-csprd01-part2-specs.html>



STEP 3: Develop metadata extraction routines

Once the metadata schema is ready and the metadata sources identified, publishers of legal information should develop routines for on-going ELI metadata extraction, transformation and validation. To do so publishers of legal information should have a clear overview of the production and dissemination of legislation taking into account:

- How publication of legal information is **organised**, including how amendments to legislation are managed as well as how consolidation is carried out (when applicable). For example, with regard to consolidation there are entities that consolidate legislation as soon as the amending act is adopted (this is for example the case for France, Norway, the Publications Office, Slovakia), while others carry out consolidation at less regular intervals (such as Ireland);
- **Available resources** (i.e. legal information) from which metadata is extracted; and
- **Available tools** to generate metadata (if any). An example of an on-going project about tools that can support the process of editing legislation is LEOS — ‘Legislation Editing Open Software’ project ⁽¹⁾.

To facilitate the ongoing extraction of ELI metadata, publishers of legal information can automatically extract it from the content of the legal text. For example, this is the case with regard to the United Kingdom where metadata is created automatically as part of the workflow which also includes validation services. A different approach is followed by Luxembourg, where metadata is automatically created after publication when the XML is generated. Metadata can also be added during the production system stage. This is the case with regard to Ireland, where a basic generic set of metadata is added to aid general searching. To be noted is the fact that a more structured approach is planned to be implemented together with the implementation of the ELI metadata. Metadata will be added automatically where possible but there will be some manual entry of metadata especially in relation to legislation transposing EU directives. A similar approach is followed by the Publications Office where metadata is created during the publication process by the institutions responsible for the legal information to be published online. The Publications Office stores the metadata in the Cellar and it is enriched through a manual legal analysis. To support this enrichment the Publications Office has developed a tool called ‘RDFedit’ ⁽²⁾. This is a similar approach to that followed by Slovakia, where some metadata is filled by the drafters of legislation (e.g. author, type of act, etc.), other metadata is added to the document during the legislative process and some more during the promulgation of the

⁽¹⁾ http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-13action_en.htm

⁽²⁾ <https://joinup.ec.europa.eu/software/rdfedit/description>



document in the official legal gazette. Germany follows yet another approach, whereby metadata is added during the consolidation process by the Federal Office of Justice. The basis for the intellectual analysis carried out is the data registration scheme for standards (DES-N) that has been coordinated between the documentation agencies at the federal level.

If possible, publishers of legal information should develop automated transformation from existing metadata to the new ELI-compliant metadata schema, for example using SQL, XSLT or SPARQL scripting languages.

Furthermore, to ensure high-quality metadata, publishers of legal information should include a validation mechanism. A popular tool that supports the execution of a quality control process of data and metadata is Open Refine ⁽¹⁾. Started at Google and, at the time of writing of the present guide, being an open source project, Open Refine provides powerful functionalities to manually verify the quality of sample metadata and fix data inconsistencies.

Finally, publishers of legal information should also take into account the fact that there might be a need to add a new metadata element into the metadata schema, update an existing one or remove an existing element. For this purpose it is recommended to put in place a formal process to manage all these changes. A common change management process covers the following steps.

1. Record and justify any change request.
2. Assess the impact of a requested change and make a recommendation.
3. Implement the change and test, or reject the change and provide an explanation.
4. Schedule, and release the implementation into the production environment.
5. When metadata become deprecated, which means that it is considered obsolete and thus its use is discouraged, it is important to carry out an impact assessment on the systems using the deprecated element. Users should be informed of such a change and backward compatibility should be ensured.

⁽¹⁾ <http://openrefine.org>



Good practice 22: Creating metadata as part of the workflow, including validation services

Problem statement

When metadata that describes legal information is created manually after legislation is published online, publishers of legal information run the risk of causing errors and incoherencies, due to the fact that manually created metadata tends to be less accurate, especially in the absence of validation services.

Recommendation

Now that legal information is increasingly available in digital form, publishers of legal information have the opportunity to apply metadata at the time of creation as part of the workflow and in a semi-automated way. This approach ensures that metadata can be used throughout the publishing process, enabling publishers of legal information to concentrate human intervention where it matters. The machine-generated metadata will typically include information that can be gleaned from the document structure or process stage (e.g. title, identifiers, dates and author). The human involvement can then be focused on adding information that is not possible to extract automatically from the document (e.g. subject classifications, links to other legal information) and on ensuring quality controls. Validation can also be incorporated in the automated tools, including spell checks, numbering, consistency of numbering and link checking.

One example of automatic creation of metadata as part of the workflow comes from the UK, where Legislation Services, in conjunction with its subcontractor, manages the whole production process for the legislation at the central level for the UK and for Scotland, Wales and Northern Ireland.

DOs

- Deploy workflow tools that capture metadata along the way
- Automate validation as much as possible
- Ensure that content specialists focus on carrying out intellectual analysis and quality assurance

DON'Ts

- Keep metadata for internal use only
- Generate metadata at the end of the publishing process
- Attribute all metadata in a manual way



BENEFITS



Quality and reliability

Increased reliability of the workflow by minimising human intervention to a few selected tasks



Development of new services

Ability to reuse the process for different information workflows



Savings

Balanced approach between automatic and manual intervention when attributing metadata which leads to: effectiveness; decreased error rate; increased reliability of the process as a whole



STEP 4: Deploy legislation metadata

Publishers of legal information should ensure that the descriptive metadata, such as the title, dates, author, etc., is made available in a **public queryable repository**.

Descriptive metadata allows users to search, find and select legal information in their repositories and, by making metadata accessible, publishers greatly increase the discoverability of legal information as well as its further use.

In order to avoid reinventing the wheel, it is recommended to reuse as much as possible existing reference data that other publishers of legal information use. For example, the Netherlands uses the following reference data:

- government organisations and public bodies
- audiences
- information types
- municipalities
- thematic classification.

The UK uses reference data mainly for internal purposes and includes:

- controlled vocabularies (e.g. legislation types)
- authority files (e.g. UK government agencies and legislation-making bodies).

Norway uses the following predefined reference data:

- countries
- geographical area (ocean) collaboration with Norwegian mapping authority
- ministries.

The Publications Office uses authority tables that are available via the Metadata Registry ⁽¹⁾ as well as translation tables and Eurovoc.

To ensure that metadata is ready for deployment, publishers of legal information should ensure its:

- **completeness:** the metadata elements should describe the legal information in line with the expected level of granularity;
- **accuracy:** the properties and values should be correctly defined in line with the ELI metadata specifications;

(1) <http://publications.europa.eu/mdr>



- **conformance:** the metadata should describe what it claims to;
- **consistency:** the metadata model should be technically correct and well structured according to the ELI ontology;
- **accessibility:** the metadata and access methods should be well documented.



Good practice 23: Generating standard metadata formats

Problem statement

Metadata is key to ensuring that legal information can be accessed and reused across information systems. Describing legal information with metadata allows it to be understood by both humans and machines in ways that promote interoperability. However, such potential for interoperability may be hampered in the absence of standard metadata, because only a few will know the meaning of the customised metadata and others will find it difficult to understand and process it.

Recommendation

It is recommended that ELI-compliant metadata schema be used to ensure greater accessibility and interoperability across legal information systems as well as greater effectiveness of the publication process.

The creation of formal metadata through adherence to ELI metadata is essential for managing and sharing legal information. ELI metadata enables interoperability among legal information systems that use the same metadata standards, greatly enhancing the exchange of legal information. To ensure interoperability, publishers of legal information should use consistent ontologies throughout legal information and metadata should use the same terms for describing legal information.

The ELI metadata can be accessed via the Publications Office's Metadata Registry website: <http://publications.europa.eu/mdr/eli>



BENEFITS



Access to legislation

Facilitates search criteria for search engines



Interoperability

Allows the exchange of information in a standard way



Development of new services

Facilitates the development of added-value services



Savings

Increased degree of automation for querying, accessing and exchanging information

DOs

- Use ELI-compliant metadata scheme to describe legal information
- Make available metadata for reuse

DON'Ts

- Use custom-made metadata



Pillar III: How to render ELI metadata machine- reusable



STEP 1: Add microdata or RDFa tags in web pages to facilitate parsing and extraction

If publishers of legal information already publish legal information on web pages using HTML or create new web pages, there is a preliminary step that they are recommended to take before adding machine-readable mark-ups.

This step entails ensuring that the web pages are compliant with standard syntax conventions so as to facilitate metadata parsing and extraction. In fact, if publishers of legal information add microdata or RDFa without ensuring that web pages are compliant with standard syntax convention, they make it more difficult for consumers of legal information to extract and parse the metadata.

The standard syntax conventions that publishers of legal information are invited to use for their web pages are XHTML (eXtensible HyperText Markup Language) ⁽¹⁾ or HTML5 ⁽²⁾.

Once the web pages are compliant with these standard syntaxes, the next step entails serialising the metadata so as to make it machine-readable.

The Council conclusions ⁽³⁾ explain that 'ELI metadata elements may be serialised in compliance with the W3C Recommendation "RDFa in XHTML: Syntax and Processing"'.

Depending on the context, different metadata serialisations are possible. For example, the Publications Office has published its metadata using RDF/XML ⁽⁴⁾.

In order to create RDF documents it is possible to use the textual syntax called Turtle ⁽⁵⁾. Another syntax that can be used is RDF/XML ⁽⁶⁾, which has widespread support in tools for consuming linked data.

Finally, when legal information publication is limited to HTML documents, the metadata may be embedded in an HTML page using any of three alternative formats: microdata ⁽⁷⁾, RDFa ⁽⁸⁾ and JSON-LD ⁽⁹⁾.

If publishers of legal information decide to add RDFa markups to the web pages, they are recommended to follow the RDFa specifications for 'Syntax and processing rules for embedding RDF through attributes' ⁽¹⁰⁾.

⁽¹⁾ <http://www.w3.org/TR/xhtml1>

⁽²⁾ <http://www.w3.org/TR/html5>

⁽³⁾ OJ C 325, 26.10.2012, p. 4.

⁽⁴⁾ <http://data.europa.eu/eli/dir/2013/37/oj/eng.rdf>

⁽⁵⁾ <http://www.w3.org/TeamSubmission/turtle>

⁽⁶⁾ <http://www.w3.org/TR/rdf-syntax-grammar>

⁽⁷⁾ http://schema.org/docs/gs.html#microdata_how

⁽⁸⁾ <http://www.w3.org/TR/rdfa-lite>

⁽⁹⁾ <http://www.w3.org/TR/json-ld>

⁽¹⁰⁾ <http://www.w3.org/TR/rdfa-syntax>



Good practice 24: Publishing metadata in non-proprietary formats

Problem statement

When selecting metadata formats to describe legal information, publishers of legal information have many options to choose from. How publishers choose to represent metadata is a primary factor in someone else's ability to use it in the future. Thus it is important to think carefully about what format will be best for managing, sharing and preserving the metadata.

Recommendation

In order to promote the reuse of metadata, publishers of legal information should make it available in a non-proprietary format. To enable further use, publication in XML or in one or more of the expression formats for RDF (RDF/XML, Turtle and RDFa) should be considered.

Some examples of the formats being used in different participants are the following:

- France: XHTML and RDFa.
- Ireland: The current metadata is embedded HTML. Thus it is available for access on a per-document basis.
- Italy: HTML and XML.
- Norway: Limited set of metadata can be screen scraped from the public website using HTML.
- Publications Office: Metadata can be downloaded from EUR-Lex in a notice view. Metadata in the Cellar is available in XML and RDF. The new system that is currently prototyped allows the retrieval of metadata in RDF/XML, Turtle and HTML+RDFa.
- Slovakia: Metadata could be made available publicly; API is available; for metadata, XML and JSON are used.
- Switzerland: it is planned to make metadata available in XML.
- UK: Metadata is provided in RDF. Examples are shown at: <http://www.legislation.gov.uk/developer/formats/rdf>. For the moment, there is no direct relationship with the ELI schema.



BENEFITS



Transparency

More legal information available in a standard way



Interoperability

Improved access and exchange of information through standard formats



Development of new services

The formats created to publish metadata can be reused for other purposes



Savings

Increased degree of automation for querying, accessing and exchanging information

DOs

- Use open machine-readable standard format like XML, RDF or Rdfa

DON'Ts

- Use proprietary formats



STEP 2: Provide an application programming interface (API)

Besides adding machine-readable metadata, publishers of legal information might consider providing an application programming interface (API).

There are lots of ways to use an API for legal information. Most notably, APIs can help publishers of legal information to build a website and also enable others to more easily incorporate legislation into their own services.

To do so, publishers of legal information should take the following steps:

1. **Decide on the protocols to provide.** If metadata is available, the ELI HTTP URLs for works and interpretations should resolve to metadata. In this respect, publishers of legal information should decide in which formats the metadata will be provided, preferably using content negotiation. Possible formats are: JSON-LD, RDF/XML, Turtle, CSV.

ELI HTTP URLs for manifestations could resolve to the actual legislation document or a metadata document with a link to the legislation document.

Optionally, publishers of legal information could provide a query interface — for example, based on RESTful⁽¹⁾ design principles or based on the SPARQL protocol.

2. **Implement an application programming interface.** Publishers of legal information should then develop a static or dynamic web application capable of rendering machine-readable metadata for legislation that conforms to the ELI-compliant metadata schema, the chosen syntax binding and the protocol.
3. **Publish API documentation.** Finally, publishers of legal information should publish documentation about the API. An example of such documentation is the one provided by the Publications Office which can be found following this link: <http://data.europa.eu/eli>

As far as the use of additional formats is concerned, the discovery phase identified the following formats per country/entity:

- **France:** HTML (xhtml), RTF, PDF and RDF.
- **Germany:** HTML, PDF, XML and EPUB.
- **Ireland:** HTML and PDF. XML format will be added following a proposed redesign of the website.

(1) https://en.wikipedia.org/wiki/Representational_state_transfer



- **Italy:** HTML, PDF, XML.
- **Luxembourg:** All of the legislation in force is published in HTML and PDF. The HTML is derived from the XML.
- **Malta:** The publication format used is PDF. Legislation from 2003 onwards is searchable whereas legislation before 2003 and up to 1984 is available as images.
- **Netherlands:** PDF, HTML, XML, ODT, RDF.
- **Norway:** Default format of Legal Gazette is HTML. Professional application enables the export of files in PDF, Word, etc. All formats are produced from *bukus* (markup language of Lovdata).
- **Publications Office:** Information is held natively in XML with the logical mark-up defined by the Formex format (⁽¹⁾). Legislation is published in both HTML and PDF/A-1a. Additionally, digital signatures are provided with the PDF files so that authenticity of the document can be verified.
- **Slovakia:** HTML, XML, PDF. These formats are produced automatically by the portal.
- **Switzerland:** It is planned to derive PDF, HTML and any other output format from the main XML source.
- **UK:** XML, HTML, RDF/XML, Atom.

(¹) <http://formex.publications.europa.eu/formex-4/formex-4.htm>



Good practice 25: Publishing legislation online as structured machine-readable data

Problem statement

When legal information is published online in a non-structured way (such as image scan), data cannot be easily retrieved or processed by machines. An additional consequence is that legal information is not accessible in a standard way, limiting the potential for interoperability as well as for parsing and extracting relevant information.

Recommendation

Publishers of legal information should publish legal information in a structured machine-readable way. While Microsoft Word or PDF could offer an excellent display format, they do not facilitate machine-processing. They should publish legal information using machine-readable formats such as CSV, JSON, XML or RDF, RDFa to facilitate the access to data and to create the opportunity to process this data programmatically. For example, the United Kingdom publishes legislation in HTML, RDF/XML and Atom and the Netherlands in HTML, XML, ODT and RDF.

In addition, publishers of legal information should consider making their data available through an API (application programming interface).

DOs

- Publish information in structured formats such as: CSV, JSON, XML, RDF, RDFa

- Scan legislation on paper and apply optical character recognition tools to create a text layer that can be used for indexing and searching

- Structure well your metadata, which will allow the serialisation of one or multiple standards

DON'Ts

- Publish data in PDF format only



BENEFITS



Transparency

Promotes transparency by making it easier to consume information by machines



Access to legislation

Increases discoverability of legal information



Development of new services

Increases reusability, because legislation is available in a structured machine-readable way



ANNEX I: Glossary

TERM	EXPLANATION
Act	A law enacted by a parliament or similar legislative body.
Akoma Ntoso	<p>Akoma Ntoso is an initiative managed by the United Nations Department of Economic and Social Affairs (Undesa), Division for Public Administration and Development Management (DPADM) and the University of Bologna.</p> <p>Akoma Ntoso defines a 'machine-readable' set of simple technology-neutral electronic representations (in XML format) of parliamentary, legislative and judiciary documents. The objective is to provide a framework for the exchange of 'machine-readable' parliamentary, legislative and judiciary documents.</p> <p>Akoma Ntoso XML schemas make 'visible' the structure and semantic components of relevant digital documents so as to support the creation of high-value information services to deliver the power of ICTs to increase efficiency and accountability in the parliamentary, legislative and judiciary contexts.</p> <p>(Source: http://www.akomantoso.org)</p>
Amendment	An effect that changes the text of legislation.
Availability	<p>Availability is the extent to which the data can be accessed. This also includes the long-term persistence of data.</p> <p>(Source: https://joinup.ec.europa.eu/community/ods/description)</p>
Chapter	A numbered level of division within an act or other legislation.
'Coming into force' date	The date on which a legislative provision or an effect comes into force.
Consistency	<p>Consistency is the extent to which the data does not contain contradictions that would make its use difficult or impossible.</p> <p>(Source: Open Data Support)</p>
Consolidation	<p>Consolidation consists of the integration in a legal act of its successive amendments and corrigenda.</p> <p>(Source: http://eur-lex.europa.eu/content/help/faq/intro.html#top)</p>
Content negotiation	Content negotiation is a mechanism of the HTTP protocol by which different documents (or 'representations') can be returned for the same URI. This is typically used to return a page in the language of the client for the same URL, or, in the context of the web of data, to return either a human-readable page or a machine-readable data file about the same URI.
Data integration	<p>Almost any interesting use of data will combine data from different sources. To do this it is necessary to ensure that the different datasets are compatible: they must use the same names for the same objects, the same units or co-ordinates, etc. If the data quality is good this process of data integration may be straightforward but if not it is likely to be arduous. A key aim of linked data is to make data integration fully or nearly fully automatic. Non-open data is a barrier to data integration, as obtaining the data and establishing the necessary permission to use it is time-consuming and must be done afresh for each dataset.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/data-integration)</p>



Database	<p>(i) Any organised collection of data may be considered a database. In this sense the word is synonymous with dataset.</p> <p>(ii) A software system for processing and managing data, including features to extend or update, transform and query the data. Examples are the open source PostgreSQL, and the proprietary Microsoft Access.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/database)</p>
Dataset	<p>Any organised collection of data. 'Dataset' is a flexible term and may refer to an entire database, a spreadsheet or other data file, or a related collection of data resources.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/dataset)</p>
Dereferencing	<p>The act of retrieving a representation of a resource identified by a URI</p> <p>(Source: http://www.w3.org/)</p>
Discoverability	<p>It is not enough for open data to be published if potential users cannot find it, or even do not know that it exists. Rather than simply publishing data haphazardly on websites, governments and other large data publishers can help make their datasets discoverable by indexing them in catalogues or data portals.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/discoverable)</p>
Division	<p>A term we use to denote any one of the hierarchical levels into which a piece of legislation may be divided.</p>
Dublin Core	<p>The Dublin Core Metadata Initiative, or 'DCMI', is an open organisation supporting innovation in metadata design and best practices across the metadata ecology. DCMI's activities include work on architecture and modelling, discussions and collaborative work in DCMI communities and DCMI task groups, global conferences, meetings and workshops, and educational efforts to promote widespread acceptance of metadata standards and best practices.</p> <p>DCMI maintains a number of formal and informal liaisons and relationships with standards bodies and other metadata organisations.</p> <p>Dublin Core is the common name for a generic set of metadata elements that were originally conceived as a simple mechanism to describe physical or digital resources, but developed into one of the main metadata standards for the semantic web.</p> <p>(Source: http://dublincore.org)</p>
Effect	<p>Any impact that one legislative provision may have on another. The most familiar type of effect is an amendment that changes the text of the affected legislation, but there are also types of effect that do not change the text, such as where a provision is said to be 'modified' or 'applied'.</p>
Extent	<p>The term 'extent' when used in legislation refers to the jurisdiction(s) for which it is law.</p>
Five stars of open data	<p>A rating system for open data proposed by Tim Berners-Lee, founder of the World Wide Web. To score the maximum five stars, data must (1) be available on the web under an open licence, (2) be in the form of structured data, (3) be in a non-proprietary file format, (4) use URIs as its identifiers (see also RDF), (5) include links to other data sources (see linked data). To score three stars, it must satisfy all of (1)-(3), etc. The open definition requires data to score three stars in order to qualify as open, not requiring RDF or linking. This permits data of a wider variety of types and sources to be open, without the work of creating linking information.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/five-stars-of-open-data)</p>



Hierarchy	<p>'Hierarchy' and 'hierarchical structure' are terms we use to denote the levels of division within a piece of legislation and the relationship between them. For example, the level of a cross-heading in an act comes below the part level in the hierarchy, but above the section level.</p>
Human-readable	<p>Data in a format that can be conveniently read by a human. Some human-readable formats, such as PDF, are not machine-readable as they are not structured data, i.e. the representation of the data on disk does not represent the actual relationships present in the data.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/human-readable)</p>
Identifier	<p>The name of an object or concept in a database. An identifier may be the object's actual name (e.g. 'London' or 'W1 1AA', a London postcode) or a word describing the concept ('population') or an arbitrary identifier such as 'XY123' that makes sense only in the context of the particular dataset. Careful choice of identifiers using relevant standards can facilitate data integration.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/identifier)</p>
Internet	<p>A structured collection of data presented in a form that people can understand and process. Information is converted into knowledge when it is contextualised with the rest of a person's knowledge and world model.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/internet)</p>
Interoperability	<p>Interoperability, within the context of European public service delivery, is 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</p> <p>(Source: http://ec.europa.eu/isa/documents/eif_brochure_2011.pdf)</p>
Legislation	<p>The generic term for laws of any type. The terms 'piece of legislation' and 'item of legislation' are used within this guide to mean a whole legislative document of any type, for example an act or statutory instrument.</p>
Linked data	<p>Linked data is a set of design principles for sharing machine-readable data on the web for use by public administrations, business and citizens.</p> <p>(Source: Open Data Support)</p>
Long title	<p>Acts and measures have two titles, the 'short title' and the 'long title'. The 'long title' sets out the purposes of the act, sometimes at great length, whereas the 'short title' is a more convenient short form by which the act will usually be known. For example, the Petroleum Act 1998 (short title) has a long title that reads:</p> <p>'An Act to consolidate certain enactments about petroleum, offshore installations and submarine pipelines.'</p>
Metadata	<p>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.</p> <p>(Sources: National information, standards organisation)</p>
Machine-readable	<p>Data in a data format that can be automatically read and processed by a computer, such as CSV, JSON, XML, etc. Machine-readable data must be structured data. Compare human-readable.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/machine-readable)</p>



OCR	Optical character recognition (OCR) is the mechanical or electronic conversion of images of typewritten or printed text into machine-encoded text. It is widely used as a form of data entry from printed paper data records, whether passport documents, invoices, bank statements, computerised receipts, business cards, mail, printouts of static data or any suitable documentation. It is a common method of digitising printed texts so that it can be electronically edited, searched, stored more compactly, displayed online and used in machine processes such as machine translation, text-to-speech, key data and text mining.
Open data	Open data is data that can be freely used, reused and redistributed by anyone — subject only, at most, to the requirement to attribute and share alike. (Source: Opendefinition.org)
Open format	A file format whose structure is set out in agreed standards, overseen and published by a non-commercial expert body. A file in an open format enjoys the guarantee that it can be correctly read by a range of different software programs or used to pass information between them. Compare proprietary. (Source: http://opendatahandbook.org/glossary/en/terms/open-format)
Original (Basic act or As enacted or As made):	The original version of the legislation as it stood when it was initially adopted. No changes have been applied to the text.
Paragraph	A provision, usually numbered, constituting the lowest level of division in a schedule. (But note that the term 'paragraph' may also be used in legislation to denote certain levels of sub-division within a provision.)
Part	A division of the main body or a schedule in an item of legislation, usually forming part of a numbered sequence of parts. A part may be further subdivided hierarchically into chapters, cross-headings and numbered sections (or paragraphs, if in a schedule).
Preamble	Words appearing near the beginning of an act after the long title, stating the reasons for passing the act. The use of preambles is optional and they are now rare. Any preamble would appear in the introductory text.
Primary legislation	General term used to describe the main laws passed by the legislative bodies. It is to be distinguished from secondary legislation.
Processability	The processability of data is the extent to which it can be understood and handled by automated processes. (Source: Open Data Support)
Proprietary	(i) Proprietary software is owned by a company which restricts the ways in which it can be used. Users normally need to pay to use the software, cannot read or modify the source code, and cannot copy the software or resell it as part of their own product. Common examples include Microsoft Excel and Adobe Acrobat. Non-proprietary software is usually open source. (ii) A proprietary file format is one that a company owns and controls. Data in this format may need proprietary software to be read reliably. Unlike an open format, the description of the format may be confidential or unpublished, and can be changed by the company at any time. Proprietary software usually reads and saves data in its own proprietary format. For example, different versions of Microsoft Excel use the proprietary XLS and XLSX formats. (Source: http://opendatahandbook.org/glossary/en/terms/proprietary)



Provision	<p>The term provision is used to describe a definable element in a piece of legislation that has legislative effect. Most commonly in the help documentation and messages on this site it will be used to refer to a section (or corresponding element such as a paragraph in a schedule or an article in an order) but it can also refer to higher-level divisions such as parts or chapters.</p>
Public sector information	<p>Information collected or controlled by the public sector.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/public-sector-information)</p>
RDF	<p>Resource description framework (RDF) is a model and a set of syntaxes for sharing data in the web.</p> <p>The basic concepts of RDF are described in the W3C recommendation 'RDF 1.1 Concepts and Abstract Syntax'.</p> <p>(Source: http://www.w3.org/TR/rdf11-concepts)</p> <p>A primer, aimed to convey basic knowledge required to effectively use RDF, is published by W3C under the title 'RDF 1.1 Primer'.</p> <p>(Source: http://www.w3.org/TR/rdf11-primer)</p> <p>Specific syntaxes for the expression of RDF are given in the documents 'RDF 1.1 XML Syntax' and 'RDF 1.1 Turtle', both published by W3C.</p> <p>While RDF is increasingly being used for the description of resources on the web, there are no easy ways to check validity of RDF data against specified patterns, the way that XML can be validated not just for valid syntax but also for conformity to application-specific rules. The W3C Working Group 'RDF Data Shapes' has been established with the mission 'to produce a W3C Recommendation for describing structural constraints and validate RDF instance data against those'.</p> <p>(Source: http://www.w3.org/2014/data-shapes)</p>
RDFa	<p>Resource Description Framework in Attributes (RDFa) is a way of expressing RDF-style relationships using simple attributes in existing markup languages such as HTML.</p> <p>RDF in Attributes provides a mechanism to include data structured in RDF in HTML pages, providing a set of markup attributes that enable machine-readable information to be embedded in web pages.</p> <p>The formal specification is contained in the document 'RDFa Core 1.1 — Third Edition', while the approach is explained in the 'RDFa 1.1 Primer — Third Edition'. A dedicated website that outlines the objectives and the benefits of RDFa and provides pointers to informational materials and an index of tools is available at rdfa.info.</p> <p>(Source: http://www.w3.org/TR/rdfa-syntax)</p>
RDF/XML	<p>The RDF XML serialisation is one of the possible ways of serialising RDF data, using an XML syntax (see also Turtle)</p>
Revised legislation	<p>We use the terms 'revise', 'revised' and 'revision' to refer to the editorial process of incorporating amendments and carrying through other effects into legislation.</p>
Section	<p>A provision, usually numbered, constituting the lowest level of division in the main body of an act or other primary legislation.</p>
Serialisation	<p>Serialisation is the process of translating data structures or object state into a format that can be stored (for example, in a file, or transmitted across a network connection link) and reconstructed later in the same or another computer environment.</p>



Short title	The title by which an act or measure is usually known. It is to be distinguished from the long title, which sets out the purposes of the legislation.
Standard	<p>A published specification for e.g. the structure of a particular file format, recommended nomenclature to use in a particular domain, a common set of metadata fields, etc. Conforming to relevant standards greatly increases the value of published data by improving machine readability and easing data integration.</p> <p>(Source: http://opendatahandbook.org/glossary/en/terms/standard)</p>
Statute	An item of primary legislation, such as an act or measure.
Statute book	A term we use to denote the totality of the statute law in force at any particular time.
Statutory instrument	A type of secondary legislation made under authority contained in acts of parliament.
SPARQL	<p>SPARQL is a language to query RDF data. It is to RDF data what SQL is to relational databases. An overview of the main aspect of SPARQL is contained in the document 'SPARQL 1.1 Overview'. The formal specification is defined in the document 'SPARQL 1.1 Query Language'.</p> <p>(Source: Open Data Support)</p>
Turtle	Turtle is one of the possible ways to serialise RDF data, using a text syntax. This is usually more compact and readable than RDF/XML.
URI	<p>A uniform resource identifier (URI) is a compact sequence of characters that identifies an abstract or physical resource. URI enables interaction with representations of the resource over a network, typically the World Wide Web, using specific protocols.</p> <p>A URI can be composed of:</p> <p>URL: uniform resource locator</p> <p>URN: uniform resource name.</p> <p>The most common form of URI is the URL, but it can also be found as an URN or as a combination of both. This last option is the one that provides URI with more consistent characteristics for the identification of resources.</p> <p>(Source: ISA's 10 rules for persistent identifiers)</p>
URL	Uniform resource locator. It can be defined as a URI that provides the means to access a web resource. URLs occur most commonly to reference web pages (http), but can also have a role in file transfer (ftp), email (mailto) or even database access (JDBC).
URN	Uniform resource name. It is the historical name for a uniform resource identifier (URI) that uses the urn scheme.
Version	<p>1. 'Version' may refer to the 'as enacted' version of the legislation or the 'latest available (revised)' version.</p> <p>2. 'Version' is also used in the context of revised legislation to refer to one of any number of versions of a provision (or higher level of division of legislation) that may exist in any number of different versions, usually created as a result of amendments made to it.</p>



ANNEX II: ELI ontology

The latest version of the ELI ontology can be found on the [Metadata Registry](#) website ⁽¹⁾.

⁽¹⁾ <http://publications.europa.eu/mdr/eli/index.html>



ANNEX III: Self-assessment check list

The following check list allows decision-makers to perform a self-assessment of their solution for publishing legislation, from an ELI perspective.

STEP	QUESTION	YES	NO	N/A
Organisation and policy				
1	Have you identified your business case? ⁽¹⁾	o	o	o
2	Have you estimated the resources for the implementation of the ELI pillars? ⁽²⁾	o	o	o
3	Have you set up a governance structure? ⁽³⁾	o	o	o
4	Have you set up an ELI implementation project? ⁽⁴⁾	o	o	o
5	Have you formulated a URI and metadata policy? ⁽⁵⁾	o	o	o
Pillar I — Identification of legislation				
1	Have you captured requirements relevant for HTTP URIs?	o	o	o
2	Have you designed URI templates? ⁽⁶⁾	o	o	o
3	Have you configured the web server for ELI UROI resolution? ⁽⁷⁾	o	o	o
Pillar II — ELI metadata				
1	Have you captured requirements relevant for metadata? ⁽⁸⁾	o	o	o
2	Have you designed the metadata schema? ⁽⁹⁾	o	o	o
3	Have you developed the metadata publishing routines? ⁽¹⁰⁾	o	o	o
4	Have you published the metadata? ⁽¹¹⁾	o	o	o
Pillar III — Render the ELI metadata machine-reusable				
1	Have you adapted the web pages to facilitate parsing and extraction? ⁽¹²⁾	o	o	o
2	Have you provided the metadata in machine readable format? ⁽¹³⁾	o	o	o

⁽¹⁾ Good practice 1: Building on the knowledge and experience of others (support by the ELI Task Force)

⁽²⁾ Good practice 2: Estimating implementation costs

⁽³⁾ Good practice 3: Setting up a central organisation as a national contact point for ELI implementation

⁽⁴⁾ Good practice 4: Structuring your implementation project

⁽⁵⁾ Good practice 5: Writing a policy document proving guarantees for long-term persistence

⁽⁶⁾ Good practices 6: Designing URI template(s) to stay as close as possible to existing citation practices; Good practice 7: Modelling HTTP URIs by treating each piece of legislation as a unique resource; Good practice 8: Taking into account the concepts of the ELI ontology: work, interpretation and manifestation; Good practice 9: Considering the right level of granularity when identifying legislation; Good practice 10: Following common good practices for HTTP URIs; Good practice 11: Testing URIs against real data

⁽⁷⁾ Good practice 12: Using content negotiation and/or identifiers for language and format variants; Good practice 13: Providing service-level guarantees; Good practice 14: Using HTTP 303 redirection; Good practice 15: Monitoring website usage and links to legislation on the web through statistics; Good practice 16: Monitoring service level and persistence of URIs

⁽⁸⁾ Good practice 17: Following a user-centred design taking into account type of user

⁽⁹⁾ Good practice 18: Considering the right level of granularity when creating metadata; Good practice 19: Structuring information by linking it to other documents; Good practice 20: Mapping ELI metadata with other standards

⁽¹⁰⁾ Good practice 21: Creating metadata as part of the workflow, including validation services

⁽¹¹⁾ Good practice 22: Generating standard metadata formats

⁽¹²⁾ Good practice 23: Publishing metadata in non-proprietary formats

⁽¹³⁾ Good practice 24: Publishing legislation online as structured machine-readable data



■ Publications Office

Print	ISBN 978-92-78-41354-5	doi:10.2830/990350	OA-01-15-858-EN-C
PDF	ISBN 978-92-78-41355-2	doi:10.2830/813167	OA-01-15-858-EN-N
ePUB	ISBN 978-92-78-41353-8	doi:10.2830/353863	OA-01-15-858-EN-E