European Forum of Official Gazettes

3rd meeting
Ljubljana, September 14 – 16, 2006

Use of XML for the production and distribution of the official gazettes

Final Report of the Forum Working group on XML

Authors:

Søren Broberg Nielsen
(Chairman)

Holger Bagola
(Secretary)

Copenhagen and Luxembourg, September 2006

Table of contents

0. Management report

This paper gives the final report on the work carried out by the working group 'XML: Common use of XML for the production and distribution of official gazettes' (short WG 'XML').

The working group was given the task to propose a vocabulary in form of a XML schema which simplifies the development and use of XML models across the different official gazettes which are specific for legislative documents.

The analyses of the working group show that a common vocabulary covering both structure and metadata would not be recommendable. A common vocabulary on structure would not be successful from two different reasons;

1. Important differences in legislative culture causing different structures in legislative documents.
2. The advanced status of XML based projects in the different countries leaving no possibility to introduce a new schema on structure.

On the other hand a common vocabulary regarding metadata is proposed by the working group, while the analyses showed that metadata across legislative systems were highly congruent, and none of the existing metadata standards are appropriate for the juridical description of documents.

The working group has created a metadata glossary with agreed definitions, and from that a common XML Schema and DTD were written.

Two obvious benefits are stressed in the report. The common metadata can be used as an inspiration or a check list to ensure that the most typical subjects of legislative metadata are covered, and implementing the common metadata will facilitate the creation and maintenance of cross system portals giving access from the same website to multiple legal information systems.

1. Introduction

The working group 'XML: Common use of XML for the production and distribution of official gazettes' (short WG 'XML') was founded in the context of the *European Forum of Official Gazettes*. The basic mission is described as follows:

> The XML project aims that the cooperation between national and European official gazette publishers in the IT sector should lead to a vocabulary which simplifies the development and use of XML models which are specific for legislative documents. This type of vocabulary — in the form of an XML schema — would contain models for common objects but would also allow for definitions which take account of particularities in any given national legislation.

Starting from this point of view, the task of the working group was the development of XML models in form of XML Schema element definitions. These models are both related to the description of the document structures in legislative acts as well as to the markup of metadata information.

The working group is composed of representatives from these countries: the Czech Republic, Denmark, Germany, Greece, Estonia, Former Yugoslav Republic of Macedonia, France, Hungary, Italy, Lithuania, the Netherlands, Portugal, Romania, Sweden and the European Union. Other countries expressed their interest of being informed about the result of the working group: Austria, Belgium, Ireland, Finland and the United Kingdom. In the beginning the working group was chaired by Nina Koch (Denmark). But because of other tasks in the context of the Forum she passed the chair to Søren Broberg Nielsen (Denmark).

In order to get to know the members of the working group, a questionnaire asking for the experiences and expectations was sent out to the designed members. The answers show that at least basic experiences are present concerning the XML based methodologies and technologies. But regarding the two techniques for designing XML based grammars — DTD (document type definitions) and XML Schema —, it becomes obvious that most experiences are still limited to the DTD variant. This image gets clearer when the expectations are taken into account:

— the design of legal documents, which concentrates on the granularity used for the markup of the different document components,

— the exchange of experiences, which expresses the wish to discuss different approaches and to learn from the success and problems met in different steps of the work,

— publishing technologies which are based on or profit from the use of XML based markup and

— technical and organisational work with XML which includes experiences in the use and configuration of tools.

Four meetings were organised (see 2.1.). The minutes of the are available in the 'Members area' (section 'Ongoing projects') of the Forum web site (http://forum.europa.eu.int/irc/opoce/ojf/info/data/prod/html/index.htm).

The report is structured in different parts. After some information on the use of XML in the production of official gazettes in the various countries a description of the undertaken works gives detailed information on the meetings, the discussed subjects as well as the interests. Contacts with other working groups will be summarized afterwards.

Chapter 3 will give a general overview of the discussions which lead to limitations of the initial mission and the definition of the work to be executed by the working group. The chapters 4 and 5 describe the results of the work, a glossary of common metadata as well as its technical implementation as a W3C XML Schema and a DTD. The last chapter will underline the benefits of the use of a set of common metadata.

The report is completed by the metadata vocabulary, the schema, the DTD, instructions on how to integrate the grammars and where to access them. The last annex contains a list of all members taking part in the working group meetings.

Already in the first meeting, representatives of the countries, when talking about their different experiences in the work with XML, outlined that they all could profit from the exchange of experiences. The lack of intuitive and user friendly interfaces to the XML authoring tools was identified as one of the most serious problems for XML projects.

Countries	XML
EU	from 1.5.2004, the OJ will be composed using XML mark-up (so-called "FORMEX v.4" - FORMEX: Formalised Exchange of Electronic Publications; before, the Office used SGML (since 1985: FORMEX V.1, V.2, V.3)
Italia	The NORMA editor and the NiR editor were developed in the context of the Norme-in-rete project. The two editors thus became the main editing tools in the Italian legislative process. The NORMA editor is an add-on to MS-Word, and the NiR editor is based on JAVA and XML technology. A good expertise on use of XML.
Sweden	Earlier adopter of XML. Their first project fails due to the reluctance of the users to deal with XML tag. Learning from their mistakes, the Swedish team presented a new prototype which proves to be much more user-friendly than the first one and the acceptance by the users could be foreseen soon.
Denmark	Denmark has considered using the Italian NiR editor in the LexDania project. But in the first phase of the project a server based transformation from .doc-files to Lex Dania XML will be used.
Germany	Bundesanzeiger print: partly XML-based; eBundesanzeiger: XML-based; Bundesgesetzblatt: partly XML-based; Use of DTD
France	Use of XML and schema. Rely on dematerialization or digitize document in their production process
Former Yugoslav Republic of Macedonia	A production system is under development. It is completely based on XML and also allows for the charging of information in an information system on legislation.
Netherlands	A good experience of XML. Develop an XML based tool to manage the consolidation process. Special interest was drawn to the consolidation facilities in that system which allows extracting a document in the particular form at any date in its life.
Romania	A production system based on XML is under development.

It is well known that in other countries there are important developments of production and information systems in this context; therefore it was rather regrettable that no reports on experiences were available.

2. Works undertaken

2.1. Meetings

Four meetings were organised in 2005 and 2006. They were all hosted by the Publication Office of the European Union in Luxembourg.

The first meeting took place in March 2005. The main items of the first meeting were the following:

— Presentation of members,

— The XML view on legal documents,

— Presentation of the Circa discussion forum,

— Structure of European and national legislative acts,

— Transformation of European directives into national law,

— Terminology and metadata.

The second meeting which was organised in June 2005 concentrated on the following items:

— Status of XML in the different countries,

— Report on the Swedish project,

— NiR and the NiR editor,

— Discussion of editorial tools,

— Presentation of the Dutch legal information system and the consolidation approach,

— Metadata.

The third meeting in March 2006 took into account the following items:

— Report on the Copenhagen meeting of the *Forum*,

— Presentation and discussion of the working plan 2006,

— Continuation of the metadata vocabulary,

— Discussion of a namespace or inclusion based approach for the integration of the metadata specifications in local applications,

— Presentation of the Eurovoc thesaurus,

— Reports on metadata indexation in the various countries.

The fourth meeting in June 2006 was supposed to be the last one; these items had been inscribed on the agenda:

— Report on news from the *Forum*,

— Reports on news from the various countries,

— Presentation of the n-Lex system which can possibly profit from the definition of common metadata,

— Presentation of the current status of the metadata vocabulary,

— Discussion of open issues on the metadata vocabulary, definition of cardinality of elements, naming conventions, versioning,

— Presentation and discussion of a draft table of contents for the final report.

In the last meeting it was also proposed to develop a prototype system which should illustrate the benefits of the use of a common metadata vocabulary.

2.2. Subjects

As already stated before, one of the common problems concerned the availability of tools which offer good interfaces between the human operators and the XML instances. As most of the human operators are not experts in XML issues, but specialists in drafting legislation, it is obvious that they should not be confronted by XML markup. The different presentations as well as the experiences of users showed that there is not a system which can cover all needs. In most cases editing tools will have to be configured for local needs or specific applications have to be developed.

The development of the metadata vocabulary was based on a synopsis of the metadata which are used in the different systems in the countries. It turned out that the items could be classified into two groups. The first one was related to the national or local legislation system. These metadata can hardly be re-used on a more generic level. The other group, however, contain elements which are present in more or less all of the mentioned systems.

The discussion also revealed that the development of a common metadata vocabulary has to proceed on two levels. A glossary has to be defined which does not only list the terms to be used, but give definitions on what is meant by a specific item. Only then the terms could be used to create an XML grammar for the concerned metadata.

A special item of the discussion concerned the description of the contents of legal acts within the metadata section. The use of thesauri and/or ontologies has to taken into account. The multilingual approach such as presented by Eurovoc is out of the scope of most of the national systems. But it was underlined that for the domains covered by Eurovoc the approach is useful. It has of course to be completed by additional information on those subjects which are not covered by European legislation and as a consequence not taken into account by Eurovoc.

Another subject concerns the re-use of metadata information by systems giving access to legislative documents. The n-Lex system which is still experimental could certainly be improved and simplified if documents foresee commonly readable description of their contents.

2.3. Interests

In the meetings of the working group issues more or less related with the mission of the WG have emerged. Members of the working group showed interest in many different aspects of the use of technology in all parts of the production of the official gazette. The interests expressed by the members was often rooted in challenges experienced in the different countries, and as often in a working group others have experienced the same challenges. By that the WG showed its usefulness by giving the opportunity to exchange common challenges and solutions.

Among the many divers interests expressed by the members two issues proved to be of general interest. The two issues were:

- Editorial tools for drafting legislation
- Automatic indexation of legislation

The WG chose different approaches to the issues. In the working group there were members with an extensive knowledge of developing and/or implementation of editors thus giving the opportunity to examine the issue in one of the meetings, and on the contrary expertise regarding automatic indexation was not sufficiently present among the WG members, which has led the WG to recommend the founding of a new working group.

As mentioned all the projects in the working group had considered the issue of an appropriate and user friendly editor for drafting legislation in a XML environment. One of the main motives to concern about editors is the more and more widespread wish for capturing the legislative text at the source. If drafters write the legislative text in XML the text is ready for integration into information systems, publishing and archiving procedures, and there is no need for additional mark up in the prepress production. An additional mark up that is both costly and opens opportunities for mistakes and errors.

The working group examined the case of editors on the second meeting with a presentation and demonstration of the NIREditor which was developed in the context of the Norme-in-rete project.

The discussion in the working group showed that there are at least three solutions to the editor question:

- a specific editor (such as the NIREditor),
- a generic XML editor such as XMetal which was used in the Swedish project and which can be configured in a way that all tags are hidden for the user, or
- a generic word processor (e. g. Microsoft Word) with additional processing into XML such as LegisWrite from the EU.

The discussion showed that there were various arguments for and against the different approaches, and therefore there could not be drawn any specific conclusions. Some of the arguments are:

"a specific editor will only allow valid XML with no use for an *a posteriori* validation to the annoyance of the drafter"

"it is easier to do additional processing with the legislative texts than introducing a new editor to the users"

"the cost to develop and support a specific editor is too high"

"cost of software license for a generic XML editor is too high, if the editor should be used by all drafters"

"a specific editor gives the possibilities to aid the drafter with legal techniques at a much higher level"

In the third meeting the WG discussed indexation vis-à-vis common metadata representing indexation. To introduce the aspects of indexation a presentation of the Eurovoc system as a multilingual thesaurus for the indexation of legislation was given by the Publication Office. During the discussion the usefulness was clearly underlined, but it was also stated that Eurovoc in its current version does not cover all domains of national legislation, but concentrates on the European needs. In the discussion it also surprised many WG members how time-consuming a proper and thorough indexation is both regarding the time spend on every document and the overhead covering training of staff and administration of the thesaurus.

The discussion inspired the delegation from the Czech Republic to propose a comprehensive examination of the issue of automatic indexation of legislation within the WG. It was however decided that the issue was out of scope in relation to the XML Working Group, but so interesting and useful that the Chair of the European Forum of Official Gazettes proposes that a new working group is founded concerned with automatic indexation of legislation.

2.4. Contact with other working groups

Two of the forums working groups cover issues that are of interest of the working group 'XML', and it was decide to contact the working groups concerning authentification and data capturing. The objective of these contacts was to give knowledge to the other working groups so that they could take the XML dimension into account during their work.

Regarding the working group “Data Capturing at the Source” we posed these two questions:

- Introducing XML in data capturing at the source, is it feasible? And what are the means?

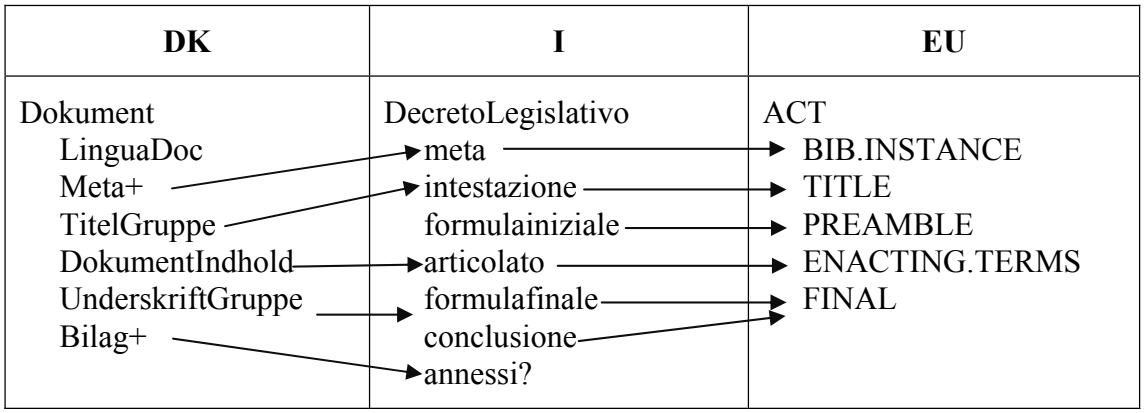
Regarding the working group “Electronic publication of legislation: Methods of authentification of the texts” the attention was drawn to two aspects of XML and authentification:

- If XML is chosen as the master document in the production of the electronic legal gazette, what is the legal status of the master document? Is the XML master document the authentic document or is it the XML-document combined with the presentation?
- What are the possibilities in authentification in the XML world?

3. Scope

The expected result from the initial mission of the working group was an XML vocabulary for the description of the structure of legal acts in the different countries. Behind this objective there was among others the idea to avoid the development of different XML grammars in the various countries.

The starting point of the discussion was a comparison of the markup in three different acts: from the European Union, from Italy and from Denmark. A first view on legal acts in different countries proves that there are obvious similarities, but also important differences. The comparison of the document structure at the highest level makes this situation evident. The following synopsis shows the markup as chosen in Italian, Danish and European laws:



The differences mainly mean that corresponding elements are absent in other legislations. The number of these differences is even bigger when details are taken into account. The following example illustrates this situation:

IT *articolo*	=	EU *article*
	≠	DA *artikel*

On the other side it is also clear that up to a certain extent, but only theoretically, the markup systems can be exchanged between countries if the chosen names had a meaning in the other country.

Having a look at the different approaches, it becomes obvious that the markup on the contents level is done on two levels, a semantic markup for the main structural document components and a more generic markup of elements on a deeper level. A common approach, however, could still be realised by using the following approach.

First of all, for the naming of elements semantic names have to be dropped and replaced by a system of recursive containers. Semantic information could be maintained on the attribute level. The advantage is that all validation rules are independent from a given language, so the control mechanisms could be re-used in all countries. National specific names could be integrated by means of predefined grammar fragments which are included in the moment when the DTD or schema is used.

Although this method could be a solution, it will certainly not meet all the needs which exist in the different countries. On the other side the problem of indicating semantic correspondences is not solved, it is only transferred to another level. As the values for the

attributes have to be defined outside the core grammar, the replacement of the DTD by schemas is nearly mandatory, especially because the conventions for creating a system of attribute values are rather limited for a DTD. Another fact concerns the readability for a human user. In general, it is said that XML instances are made for machines and the markup as shown in the following example only proves this opinion:

```
<document doc.type="act">
  <sem.obj sem.obj.type="title">
    <p type="doc.title">
      VERORDNUNG (EG) Nr. 362/2005 DER KOMMISSION
    </p>
    <p type="doc.title">
      vom 3. März 2005
    </p>
    <p type="doc.title">
      zur Ablehnung von Anträgen auf Erteilung von Ausfuhrlizenzen
      im Getreidesektor für Erzeugnisse des KN-Codes 1101 00 15
    </p>
  </sem.obj>
  <sem.obj sem.obj.type="preamble">
    <sem.obj sem.obj.type="preamble initialization">
      DIE KOMMISSION DER EUROPÄISCHEN GEMEINSCHAFTEN —
    </sem.obj>
    <sem.obj sem.obj.type="group of visa">
      <p type="visa">
        gestützt auf den Vertrag zur Gründung der Europäischen
        Gemeinschaft,
      </p>
      <p type="visa">
      gestützt auf die Verordnung (EG) Nr. 1784/2003 des Rates vom
      29. September 2003 über die gemeinsame Marktorganisation für
      Getreide (1),
    </p>
    <p type="visa">
      gestützt auf die Verordnung (EG) Nr. 1342/2003 der Kommission
      vom 27. Juli 2003 mit besonderen Durchführungsbestimmungen über
      Einfuhr- und Ausfuhrlizenzen für Getreide und Reis (2),
      insbesondere auf Artikel 8 Absatz 1,
    </p>
  </sem.obj>
  <sem.obj type="group of recitals">
      <p type="group of recitals initialization">
        in Erwägung nachstehenden Grundes:
      </p>
      <p type="recital">
        Die Anzahl der Anträge auf im Voraus festgesetzte
        Erstattungen für Erzeugnisse des KN-Codes 1101 00 15 ist
        bedeutend und von spekulativem Charakter. Es sollten deshalb
        alle Anträge abgelehnt werden, die am 1. März 2005
        eingereicht wurden —
```

```
      </p>
    </sem.obj>
    <sem.obj type="preamble final">
      HAT FOLGENDE VERORDNUNG ERLASSEN:
    </sem.obj>
  </sem.obj>
  <sem.obj sem.obj.type="enacting terms">
    <sem.obj sem.obj.type="article">
      <p type="article header">
        Artikel 1
      </p>
      <p type="alinea">
        Gemäß Artikel 8 Absatz 1 der Verordnung (EG) Nr. 1342/2003
        wird die am 1. März 2005 beantragte Erteilung von Lizenzen
        für die Ausfuhr von Erzeugnissen des KN-Codes 1101 00 15
        abgelehnt.
      </p>
    </sem.obj>
    <sem.obj sem.obj.type="article">
      <p type="article header">
        Artikel 2
      </p>
      <p type="alinea">
        Diese Verordnung tritt am 4. März 2005 in Kraft.
      </p>
    </sem.obj>
  </sem.obj>

<sem.obj type="final">
    <p type="applicability">
      Diese Verordnung ist in allen ihren Teilen verbindlich und gilt
      unmittelbar in jedem Mitgliedstaat.
    </p>
    <p type="place date">
      Brüssel, den 3. März 2005
    </p>
    <sem.obj sem.obj.type="signature">
      <p type="affiliation">
        Für die Kommission
      </p>
      <p type="signatory">
        J. M. SILVA RODRÍGUEZ
      </p>
      <p type="function">
        Generaldirektor für Landwirtschaft und Entwicklung des
        ländlichen Raumes
      </p>
    </sem.obj>
  </sem.obj>
</document>
```

Furthermore, the information on XML based projects in the different countries show that XML is used nearly in all countries. The development of grammars is finished or at a very advanced level. So for the development of a common vocabulary, problems might arise from the fact that resources already have been appointed, so replacing the home made solution which corresponds exactly to the needs will hardly be welcome.

So the development of a vocabulary for common structural markup does not seem to be feasible because of important differences in legislative culture on the one hand side and because of the advanced status of XML based projects in the different countries.

The discussions, however, also showed that there is an important conformity on the metadata level. A synoptic inventory listed the various metadata elements which are in use in the various countries. On this basis a list of about thirty elements could be identified, which are of interest in the various legal information systems.

The list also made evident some terminological difficulties. Therefore it was decided to create a glossary with definitions for each term which is part of the metadata element collection. The glossary was completed by some terms which describe general components of an act, but do not belong to the metadata elements.

On the XML level the grammar was built on the basis of W3C XML Schema technology. For each of the identified elements an XML model was defined. A supplementary container was created to keep the metadata information together and to add cardinality to the elements. It turned finally out that most of the elements are optional, either because the information is not available in all documents or because it is not used in some countries.

As already mentioned in the Introduction, some countries have not yet taken the decision to migrate their XML based systems to schemas; they maintain a DTD oriented approach. This is why the schema is accompanied by a corresponding DTD. It has to be kept in mind that the DTD was derived from the schema and cannot cover all the features offered by a schema grammar.

For the implementation of the schema three methods were presented:

1. copying the contents of the metadata schema into the local schema;
2. including the metadata schema by means of the xs:include element or
3. defining a namespace and call the metadata schema by means of the namespace location address (xs:import element).

The first approach needs an important management in the case the metadata schema evolves. So it was no longer followed. Whether including or importing the metadata schema depends on the organisation of the local schema. Therefore it was decided to maintain both solutions. As it is not possible to refer to the same schema for inclusion and import, a container schema was created which in fact includes the original schema. The namespace reference in the local schema has to be linked to this container. For the use of the schema within the instances this method remains transparent.

According to the XML standard the decision to use a DTD is a decision against namespaces, as they are not supported within a DTD. So the metadata DTD can only be included by means of a public or system reference call.

The import or inclusion of a schema or a DTD is only executed when the grammar is opened by a coherent tool, p. ex. a parser. A *conditio sine qua non* for this approach is the online availability of the metadata schema and DTD. Therefore both grammars will be accessible at the same address as the Formex schema which is used for the documents published in the *Official Journal of the European Union*.

In order to demonstrate how useful the common metadata vocabulary can be, a prototype will be developed. Documents from different countries on the same contents domain will be searched and presented to the user. He will then not have to know the local methodologies applied for the integration of metadata information.

4. Metadata vocabulary

The need of metadata for the description of document contents is in fact not new. International standards for the exchange of bibliographic information and/or library catalogues have been created and are used regularly.

One of these standards, perhaps the most important one, is *MARC21*, *Ma*chine *R*eadable *C*ataloguing, which was developed and is maintained by the American Library of Congress. It offers a record based system with rather a high variety of metadata fields. The XML adaptation of this standard is not very satisfying as it resembles the abovementioned approach for a common markup as can be seen in the following illustration:

```
<collection xmlns="http://www.loc.gov/MARC21/slim">
  <record>
    <leader>01142cam 2200301 a 4500</leader>
    <controlfield tag="001">92005291</controlfield>
    <controlfield tag="003">DLC</controlfield>
    <controlfield tag="005">19930521155141.9</controlfield>
    <controlfield tag="008">920219s1993 caua j 000 0
    eng</controlfield>
    <datafield tag="010" ind1="" ind2="">
      <subfield code="a">92005291</subfield>
    </datafield>
    <datafield tag="020" ind1="" ind2="">
      <subfield code="a">0152038655 :</subfield>
      <subfield code="c">$15.95</subfield>
    </datafield>
    <datafield tag="040" ind1="" ind2="">
      <subfield code="a">DLC</subfield>
      <subfield code="c">DLC</subfield>
      <subfield code="d">DLC</subfield>
    </datafield>
    <datafield tag="042" ind1="" ind2="">
      <subfield code="a">lcac</subfield>
    </datafield>
    …
    <datafield tag="520" ind1="" ind2="">
      <subfield code="a">A poem about numbers and their
        characteristics. Features anamorphic, or distorted, drawings
        which can be restored to normal by viewing from a particular
        angle or by viewing the image's reflection in the provided
        Mylar cone.</subfield>
    </datafield>
    <datafield tag="650" ind1="" ind2="0">
      <subfield code="a">Arithmetic</subfield>
      <subfield code="x">Juvenile poetry.</subfield>
    </datafield>
```

```
    <datafield tag="650" ind1="" ind2="0">
      <subfield code="a">Children's poetry, American.</subfield>
    </datafield>
    <datafield tag="650" ind1="" ind2="1">
      <subfield code="a">Arithmetic</subfield>
      <subfield code="x">Poetry.</subfield>
    </datafield>
    <datafield tag="650" ind1="" ind2="1">
      <subfield code="a">American poetry.</subfield>
    </datafield>
    <datafield tag="650" ind1="" ind2="1">
      <subfield code="a">Visual perception.</subfield>
    </datafield>
    <datafield tag="700" ind1="1" ind2="">
      <subfield code="a">Rand, Ted,</subfield>
      <subfield code="e">ill.</subfield>
    </datafield>
  </record>
</collection>
```

Another standard is derived from MARC. METS, *M*etadata *E*ncoding & *T*ransmission *S*tandard, was also developed by the Library of Congress. It consists of 34 fields; its objective is as follows:

> The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library expressed using the XML schema language of the Word Wide Web Consortium. The standard is maintained in the Network Development and MARC Standards Office of the Library of Congress, and is being developed as an initiative of the Digital Library Federation.

An interesting fact is that it is based on an XML grammar.

Another standard is frequently used in the context of XML based document markup, although it is originally not based on or limited to XML. DCMI, *D*ublin *C*ore *M*etadata *I*nitiative, offers 15 fields for the collection of metadata. They are often used as markup within the container system RDF (Resource Description Framework):

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description
    rdf:about="http://media.example.com/audio/guide.ra">
    <dc:creator>Rose Bush</dc:creator>
    <dc:title>A Guide to Growing Roses</dc:title>
    <dc:description>Describes process for planting and nurturing
      different kinds of rose bushes.</dc:description>
    <dc:date>2001-01-20</dc:date>
  </rdf:Description>
</rdf:RDF>
```

It is obvious for this approach that 15 fields might not be sufficient for certain projects. This is why DCMI offers the possibility of extending the fields by the definition of subfields. One project which proceeded on this way is MIReG, *M*anag*i*ng *R*esources for *e-G*overnment. MIReG is the result of an IDA project between the United Kingdom and the EU.

All these approaches are not appropriate for the juridical description of documents. On the first meeting of the working group a list of metadata fields was identified and completed meanwhile (see the synopsis in Annex 1). This list unveils the following circumstances:

— some of the mentioned fields are used in various countries under more or less the same name,

— other fields are specific for a certain country, the concept seems not to be taken into account elsewhere,

— some of the definitions are overlapping and need to be clarified.

In order to analyse how far existing standards could be used for the metadata fields collected in the synopsis, DCMI was applied leading to the following results:

Metadata fields collected in the synopsis	DCMI
identifier on metadata level	identifier
information (issue number, issue name, approval date, printing date)	identifier
reference to OJ	relation
Subject	subject
Act	./.
additionally used abbreviation	./.
authors service (ministry)	creator
code number	./.
consolidated or not	type
Contents	description
date of adoption/signature	date
date of application	date
date of decision	date
date of enactment	date
date of entering into force	date
date of expiration	date
date of publishing/publication	date
date of repeal	date
date of the law	date
document type	./.
keywords (subject of the document)	subject
language of the document	language
nickname of the legal document (subject, friendly name)	./.
number used in the ministry archives	./.

Metadata fields collected in the synopsis	DCMI
official abbreviation	./.
original vs. amendment	./.
Preamble	./.
reference to basic legislation	relation
reference to changing text (if consolidated)	relation
reference to directive (if applicable)	relation
reference to draft	source
reference to modified acts	relation
reference to next version (if consolidated)	relation
reference to previous version (if consolidated)	relation
reference to source (if consolidated)	relation
reference to translation	relation
References	relation
region code	./.
re-print or not	type
Signature	creator
status (in force or not)	type
Summary	./.
Title	title

Some of the fields are indeed covered by corresponding descriptors in DCMI. However, most of these descriptors have to be used more than once, so the real meaning of a metadata field is no longer evident. Other fields are not at all covered. As a consequence the need of a specific markup system for metadata in legal documents became obvious.

So it was decided to identify relevant metadata fields and to define their roll within the juridical analysis of the object.

5. Technical solution

The metadata glossary has been transformed into an XML schema (see annex 2A) and a parallel DTD (see annex 2C). Both grammars are mutually interchangeable. The parsing, however, with the schema gives more details on the correctness of the instance. This becomes evident when comparing the models for the Celex number: in the DTD it is only a sequence of characters, in the schema a pattern is defined which helps to control the syntax of the codes while parsing.

The names of the defined elements all start with the prefix ‘efog.’. This guarantees that the elements may always be identified within the local grammar or in the instances. The elements are defined in alphabetical order, just to simplify the access to the elements for a human user.

The various models are always completed by annotations or comments which explain the semantic role of a given element. These explanations are based on the definitions in the metadata glossary.

Although it is possible to use the various elements at any place within an instance, it is recommended to integrate them within a single block. This should simplify researches in the documents. Furthermore a control of the cardinality is only possible within the predefined block. Another possibility is the creation of an element in the local schema which is related to the efog-type. From a technical point of view this approach is as efficient as the use of the efog-block element; it only has the disadvantage that for researching the instances the name of the block should be known.

The grammars will be available from the Formex website which is managed by the Publications Office. The URL to the directory is the following:

```
http://formex.publications.europa.eu/schema/
```

It is completed by the name of the grammar. The URL is also valid for the DTD. The directory will always contain all versions of the grammars. So in case of improvements there is not necessarily the need to review and update all instances. The syntax of the names of the grammars has been defined as follows:

```
[ 1]   name         ::=   name.schema | name.DTD

[ 2]   name.schema  ::=   'efog_' 'cnt_'? date '.xsd'

[ 3]   name.DTD     ::=   'efog_' date '.dtd'

[ 4]   date         ::=   d{8}

[ 5]   d            ::=   '0'|'1'|'2'|'3'|'4'|'5'|'6'|'7|'8'|'9'
```

The date reflects the date when the version of the grammar was adopted. The constant part '`cnt_`' is only to be used when the schema is referred to by means of the namespace technology (the schema is to be found in annex 2B).

As discussed earlier in this document, the implementation can be achieved on different ways: a copy-paste method, the inclusion of the schema, a namespace reference. For the DTD the copy-paste method and the inclusion are feasible. The copy-paste method is highly deprecated; details on inclusion and namespace references can be found in annex 2D.

6. Benefits of common metadata

In this chapter the working group will exemplify the benefits of common metadata by different use cases that hopefully will promote the dissemination of the common metadata. It is important to understand that the implementation of the common metadata does not exclude the use of system specific metadata, with the proposed naming convention there should not be any conflicts.

The main benefits of the common metadata could be:

- Improved access to a set of standard definitions of metadata terms through the Internet.
- Improved standardisation of metadata for dissemination and international comparisons.
- Support to XML structures for searching and exchanging metadata on legislation.

6.1 Check list for metadata design

When creating or redesigning a legal information system, the common metadata can be used as an inspiration or a check list to ensure that the most typical subjects of legislative metadata are covered. This advantage covers also the proposed grammar, where metadata designers can use the XML-schema definitions when defining local metadata.

6.2 Facilitating access to legislation across different legal information systems

By implementing the common metadata it will facilitate the creation and maintenance of cross system portals giving access from the same website to multiple legal information systems. This is made possible when the portal can be designed with a uniform interface to metadata in the different systems.

The opposite case is also facilitated when new legal information systems are created or existing systems are renewed. The new or redesigned sites can more easily be integrated in the portal; if the portal supports the common metadata the access to a metadata search is more or less prefabricated.

Besides the more technical benefits for cross system portals the users will supposedly find it easier and more intuitive to query the different systems, when metadata is common the query in different systems will not be hinder by misunderstanding of “where-to-search-for-what”.

Annexes

1. Metadata vocabulary

Metadata Glossary

Version 1.00

Abbreviation, →Official abbreviation, →Other used abbreviation

Act, [basic element], an official legislative document, decree or law made by the legislative body (OALD, s.v. act $_4$)

Act number, [metadata element] official number of an →act, identification

Alias, →Nickname

Applicability, [metadata element] specification of the start and end date of the applicability of an act; in some cases the applicability may depend on conditions other than dates

Authority, [metadata element], competent body

Bibliographic citation, [metadata element], reference to another →act

Celex, [metadata element] identification of an EU law by means of its Celex number; the Celex number is composed as follows:

```
I.      Complex version

[ 1]    CelexNumber  ::=  Sector.1|Sector.2|Sector.3|Sector.4|
                          Sector.6|Sector.7|Sector.E

[ 2]    Sector.1     ::=  '1' Year DocType.1 Qualifier.1+ NumCurr.1
[ 3]    Sector.2     ::=  '2' Year DocType.2 NumCurr NumSeq?
                          Correction*
[ 4]    Sector.3     ::=  '3' Year DocType.3 NumCurr NumSeq?
                          Correction*
[ 5]    Sector.4     ::=  '4' Year DocType.4 NumCurr NumSeq?
                          Correction*
[ 6]    Sector.6     ::=  '6' Year DocType.6 NumCurr    NumSeq?
[ 7]    Sector.7     ::=  '7' Year DocType.7 numCurr NumSeq?
                          Correction*
[ 8]    Sector.E     ::=  'E' Year DocType.E NumCurr NumSeq?

[ 9]    Year         ::=  d{4}

[10]    DocType.1    ::=  'A'|'B'|'C'|'D'|'E'|'F'|'G'|'H'|'I'|'K'|
                          'M'|'N'|'R'|'T'|'U'
[11]    DocType.2    ::=  'A'|'D'|'P'|'X'
[12]    DocType.3    ::=  'A'|'B'|'C'|'D'|'E'|'F'|'G'|'H'|'J'|'K'|
                          'L'|'M'|'O'|'Q'|'R'|'X'
[13]    DocType.4    ::=  'A'|'D'|'X'
[14]    DocType.6    ::=  'A'|'B'|'C'|'D'|'F'|'H'|'J'|'O'|'P'|'S'|
                          'T'|'V'|'X'
```

```
[15]   DocTYpe.7    ::=  'L'
[ 1]   DocType.E    ::=  'A'|'C'|'G'|'J'|'P'|'X'

[16]   Qualifier.1  ::=  '/' c{3}

[17]   NumCurr      ::=  d{4}
[18]   NumCurr.1    ::=  '/' d{2}

[19]   NumSeq       ::=  '(' d{2} ')'

[20]   Correction   ::=  'R(' d{2} ')'

[21]   c            ::=  'A'|'B'|'C'|'D'|'E'|'F'|'G'|'H'|'I'|'J'|
                         'K'|'L'|'M'|'N'|'O'|'P'|'Q'|'R'|'S'|'T'|
                         'U'|'V'|'W'|'X'|'Y'|'Z'
[22]   d            ::=  '0'|'1'|'2'|'3'|'4'|'5'|'6'|'7'|'8'|'9'
```

II. Simplified version

```
[ 1]   CelexNumber  ::=  Sector Year Doctype ((Qualifier.1+
                         NumCurr.1)|(NumCurr NumSeq? Correction*))
[ 2]   Sector       ::=  '1'|'2'|'3'|'3'|'4'|'6'|'7'|'E'
[ 3]   Year         ::=  d{4}
[ 4]   DocType      ::=  'A'|'B'|'C'|'D'|'E'|'F'|'G'|'H'|'I'|'J'|
                         'K'|'L'|'M'|'N'|'O'|'P'|'Q'|'R'|'S'|'T'|
                         'U'|'V'|'X'|

[ 5]   Qualifier.1  ::=  '/' c{3}

[ 6]   NumCurr      ::=  d{4}
[ 7]   NumCurr.1    ::=  '/' d{2}

[ 8]   NumSeq       ::=  '(' d{2} ')'

[ 9]   Correction   ::=  'R(' d{2} ')'

[10]   c            ::=  'A'|'B'|'C'|'D'|'E'|'F'|'G'|'H'|'I'|'J'|
                         'K'|'L'|'M'|'N'|'O'|'P'|'Q'|'R'|'S'|'T'|
                         'U'|'V'|'W'|'X'|'Y'|'Z'
[11]   d            ::=  '0'|'1'|'2'|'3'|'4'|'5'|'6'|'7'|'8'|'9'

Example: 32005L0084
```

Corrigenda, [metadata element], definition of a bibliographic reference to a correcting document as well as to the corrected version if available

Date of abrogation, [metadata element] specification of the date when the law is repealed; it has to be remembered that in particular cases, only fragments of a law are repealed; the authority which launches the repeal has to be mentioned

Date of entering into force, [metadata element, date related to →act and →partition], date of entering into the legal systems

Date of passing, [metadata element, date related to →act], final acceptance by the Parliament

Date of publication, [metadata element, date related to →act], date of publication in the official medium

Date of signature, [metadata element, date related to →act], signature of the →act by the authority foreseen in the Constitution or similar

Document type, [metadata element] specification of the legal value of a given document (e.g. 'directive', 'regulation', 'decision' ...)

EU reference, [metadata element], definition of a reference to EU legislation both in form of the →Celex identifier and the reference to the Official Journal (→OJ reference)

Eurovoc, [metadata element for indexing] container for Eurovoc descriptors; it consists of keywords each of them containing a descriptor

Index, [metadata element] indexation according to local needs and constraints, if possible with reference to the local thesaurus, ontology or taxonomy; the element is composed of keywords each of them containing a single descriptor

Legal hierarchy, [metadata element] specification of the legal hierarchy in which a document is created (e.g. 'primary law', 'secondary law', 'tertiary law')

Nickname (var. Alias), [metadata element], title currently used in the spoken language

Norm, [basic element], interpretation of the →provision

Official abbreviation, [metadata element], abbreviation prescribed in the →act

Official title, [metadata element], name given to the resource, usually by the Creator or Publisher (RFC 2413, p. 3), indicates the object of the →act, informs about the subject of the →act as precisely as possible and necessary

Other used abbreviation, [metadata element], abbreviation currently used in the spoken language

Partition, [basic element], well identified part within the enacting terms of an →act

Provision [basic element], conceptual and textual entities, condition or stipulation in a legal document (OALD, s.v. provision $_4$)

Reference to the Official Journal, [metadata element], reference to the official publication of a legal →act

Registration, [metadata element], reference to the author's internal registration or archiving system; the element is supposed to be useful for citizens if they require additional information on a given law

Short Title, [metadata element], citation title prescribed in the →act_

Signatory, [metadata element], name(s) of person(s) empowered by a competent body to sign an →act

References

OALD : *Oxford Advanced Learner's Dictionary of Current English.* A. S. Hornby. Fourth edition: chief editor: A. P. Cowie. Eighth impression with corrections. Oxford: University Press 1992. ISBN 0 19 431110 4.

RFC 2413 : Weibel, S.; Kunze, J.; Lagoze, C.; Wolf. M. 1998. *Dublin Core Metadata for Resource Discovery.* http://www.ietf.org/rfc/rfc2413.txt [last visited: 4 July 2005]

2. Technical solution

A. The schema solution

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
   xmlns:xs="http://www.w3.org/2001/XMLSchema"
   elementFormDefault="qualified"
   attributeFormDefault="unqualified">
   <xs:annotation>
      <xs:documentation>
         Version: 1.00
      </xs:documentation>
      <xs:documentation>
         All defined elements start with the prefix efog which stands
         for European Forum of Official Gazettes.
      </xs:documentation>
      <xs:documentation>
         Naming convention for element names
      </xs:documentation>
      <xs:documentation>
         elementName  ::= "efog." localName
      </xs:documentation>
      <xs:documentation>
         localName    ::= ((a-z)|"-")+
      </xs:documentation>
   </xs:annotation>
<!--
T y p e   d e f i n i t i o n s
-->
   <xs:complexType name="t_efog">
      <xs:annotation>
         <xs:documentation>
            Definition of complex type for integration into local schema
         </xs:documentation>
      </xs:annotation>
      <xs:sequence>
         <xs:element ref="efog.abbreviation"
            minOccurs="0"
            maxOccurs="unbounded"/>
         <xs:element ref="efog.abrogation"
            minOccurs="0"
            maxOccurs="unbounded"/>
         <xs:element ref="efog.act-number"
            minOccurs="0"/>
         <xs:element ref="efog.applicability"
            minOccurs="0"/>
         <xs:element ref="efog.authority"
            minOccurs="0"/>
         <xs:element ref="efog.bibliographic-citation"
            minOccurs="0"
            maxOccurs="unbounded"/>
         <xs:element ref="efog.date-enter-into-force"
            minOccurs="0"/>
```

```
        <xs:element ref="efog.date-passing"
          minOccurs="0"/>
        <xs:element ref="efog.date-publication"/>
        <xs:element ref="efog.date-signature"
          minOccurs="0" maxOccurs="unbounded"/>
        <xs:element ref="efog.document-type"/>
        <xs:element ref="efog.eurovoc"
          minOccurs="0"/>
        <xs:element ref="efog.index"
          minOccurs="0"/>
        <xs:element ref="efog.nickname"
          minOccurs="0"
          maxOccurs="unbounded"/>
        <xs:element ref="efog.reference-corrigendum"
          minOccurs="0"
          maxOccurs="unbounded"/>
        <xs:element ref="efog.reference-eu"
          minOccurs="0"/>
        <xs:element ref="efog.reference-oj"
          minOccurs="0"/>
        <xs:element ref="efog.registration"
          minOccurs="0"/>
        <xs:element ref="efog.signatory"
          minOccurs="0"
          maxOccurs="unbounded"/>
        <xs:element ref="efog.title"
          maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
<!--
.....................................................................
-->
    <xs:simpleType name="t_date">
      <xs:annotation>
        <xs:documentation>
          A date is supposed to be encoded according to the model
          yyyymmdd (y: year, m: month, d: day)
        </xs:documentation>
      </xs:annotation>
      <xs:restriction base="xs:string">
        <xs:pattern value="\d{8}"/>
      </xs:restriction>
    </xs:simpleType>
```

```
<!--
E l e m e n t   d e f i n i t i o n s
-->
<!--
.. super-container for metadata information
-->
  <xs:element name="efog" type="t_efog">
    <xs:annotation>
      <xs:documentation>
        Definition of generic container for metadata
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.. Metadata information
-->
  <xs:element name="efog.abbreviation">
    <xs:annotation>
      <xs:documentation>
        official: abbreviation prescribed in the act;
        other: abbreviation currently used in the spoken language
      </xs:documentation>
    </xs:annotation>
    <xs:complexType mixed="true">
      <xs:attribute name="type" use="required">
        <xs:simpleType>
          <xs:restriction base="xs:string">
            <xs:enumeration value="official"/>
            <xs:enumeration value="other"/>
          </xs:restriction>
        </xs:simpleType>
      </xs:attribute>
    </xs:complexType>
  </xs:element>
<!--
.............................................................
-->
  <xs:element name="efog.abrogation">
    <xs:annotation>
      <xs:documentation>
        information on abrogation of an act or a fragment
      </xs:documentation>
      <xs:documentation>
        specification of the date (mandatory), the target (optional)
        and the originator (law, jurisdiction, etc.) (optional)
      </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="efog.date-abrogation"/>
        <xs:element ref="efog.target" minOccurs="0"/>
        <xs:element ref="efog.originator" minOccurs="0"/>
      </xs:sequence>
```

```
    </xs:complexType>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.act-number" type="xs:string">
    <xs:annotation>
      <xs:documentation>
        official number of an act, identification
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.applicability">
    <xs:annotation>
      <xs:documentation>
        specification of the applicability of an act, optional
        specification of reasons or conditions
      </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="efog.date-start"/>
        <xs:element ref="efog.date-end" minOccurs="0"/>
        <xs:element ref="efog.condition" minOccurs="0"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.authority" type="xs:string">
    <xs:annotation>
      <xs:documentation>
        competent body
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.bibliographic-citation">
    <xs:annotation>
      <xs:documentation>
        reference to another act
      </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:choice>
        <xs:element ref="efog.uri"/>
        <xs:sequence>
          <xs:element ref="efog.series" minOccurs="0"/>
```

```
            <xs:element ref="efog.ojno" minOccurs="0"/>
            <xs:element ref="efog.date-publication"/>
            <xs:element ref="efog.page" minOccurs="0"/>
          </xs:sequence>
        </xs:choice>
      </xs:complexType>
    </xs:element>
<!--
.....................................................................
-->
    <xs:element name="efog.celex">
      <xs:annotation>
        <xs:documentation>
          Celex number; attention: the pattern only covers directives
          and their corrigenda
        </xs:documentation>
      </xs:annotation>
      <xs:simpleType>
        <xs:restriction base="xs:string">
          <xs:pattern
            value="
              (\d|E)
              \d{4}
              (A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|P|Q|R|S|T|U|V|X)
              (
                (
                  (/(A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|
                  P|Q|R|S|T|U|V|W|X|Y|Z){3}
                )+
                (/\d{2})
              )
              |
              (\d{4}
                (\(\d{2}\))?
                (R\(\d{2}\))*)
              )
          "/>
        </xs:restriction>
      </xs:simpleType>
    </xs:element>
<!--
.....................................................................
-->
    <xs:element name="efog.condition">
      <xs:annotation>
        <xs:documentation>
          expression of reason or condition in the context of the
          applicability of an act
        </xs:documentation>
      </xs:annotation>
      <xs:complexType>
        <xs:sequence>
          <xs:element ref="efog.p" maxOccurs="unbounded"/>
        </xs:sequence>
```

```
    </xs:complexType>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.date-abrogation" type="t_date">
    <xs:annotation>
      <xs:documentation>
        date of abrogation of an act or a part of it
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.date-end" type="t_date">
    <xs:annotation>
      <xs:documentation>
        date of the end of applicability
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.date-enter-into-force" type="t_date">
    <xs:annotation>
      <xs:documentation>
        date of entering into the legal systems
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.date-passing" type="t_date">
    <xs:annotation>
      <xs:documentation>
        final acceptance by the Parliament
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.date-publication" type="t_date">
    <xs:annotation>
      <xs:documentation>
        date of publication in the official medium
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
```

```
-->
  <xs:element name="efog.date-signature">
    <xs:annotation>
      <xs:documentation>
        signature of the act
      </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:simpleContent>
        <xs:extension base="t_date">
          <xs:attribute name="quality"
            type="xs:string"
            use="optional"/>
        </xs:extension>
      </xs:simpleContent>
    </xs:complexType>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.date-start" type="t_date">
    <xs:annotation>
      <xs:documentation>
        date of start of applicability
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.document-type" type="xs:string">
    <xs:annotation>
      <xs:documentation>
        specification of type of document such as regulation,
        directive, decision etc.
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.eurovoc">
    <xs:annotation>
      <xs:documentation>
        container for Eurovoc descriptors
      </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="efog.keyword" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
<!--
```

```
.....................................................................
-->
   <xs:element name="efog.index">
     <xs:annotation>
       <xs:documentation>
         container for descriptors of a local indexation system
       </xs:documentation>
     </xs:annotation>
     <xs:complexType>
       <xs:sequence>
         <xs:element ref="efog.keyword" maxOccurs="unbounded"/>
       </xs:sequence>
     </xs:complexType>
   </xs:element>
<!--
.....................................................................
-->
   <xs:element name="efog.keyword" type="xs:string">
     <xs:annotation>
       <xs:documentation>
         descriptor from Eurovoc or a local indexation system
       </xs:documentation>
     </xs:annotation>
   </xs:element>
<!--
.....................................................................
-->
   <xs:element name="efog.nickname" type="xs:string">
     <xs:annotation>
       <xs:documentation>
         title currently used in the spoken language
       </xs:documentation>
     </xs:annotation>
   </xs:element>
<!--
.....................................................................
-->
   <xs:element name="efog.ojno" type="xs:string">
     <xs:annotation>
       <xs:documentation>number of the official journal/gazette referred
to</xs:documentation>
     </xs:annotation>
   </xs:element>
<!--
.....................................................................
-->
   <xs:element name="efog.originator" type="xs:string">
     <xs:annotation>
       <xs:documentation>
         originator of the limitation of applicability of an act
       </xs:documentation>
     </xs:annotation>
   </xs:element>
<!--
```

```
.....................................................................
-->
  <xs:element name="efog.p" type="xs:string">
    <xs:annotation>
      <xs:documentation>
        container for any information which has to be kept together
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.page" type="xs:string">
    <xs:annotation>
      <xs:documentation>
        first page of a document within a oj reference
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.reference-corrigendum">
    <xs:annotation>
      <xs:documentation>
        reference to a corrigendum
      </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:choice>
        <xs:element ref="efog.uri"/>
        <xs:sequence>
          <xs:element ref="efog.series" minOccurs="0"/>
          <xs:element ref="efog.ojno" minOccurs="0"/>
          <xs:element ref="efog.date-publication"/>
          <xs:element ref="efog.page" minOccurs="0"/>
        </xs:sequence>
      </xs:choice>
      <xs:attribute name="roll" use="required">
        <xs:simpleType>
          <xs:restriction base="xs:string">
            <xs:enumeration value="active"/>
            <xs:enumeration value="passive"/>
          </xs:restriction>
        </xs:simpleType>
      </xs:attribute>
    </xs:complexType>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.reference-eu">
    <xs:annotation>
      <xs:documentation>
```

```
          container for the reference to a European directive with the
          Celex number as identifier and the bibliographic location
        </xs:documentation>
      </xs:annotation>
      <xs:complexType>
        <xs:sequence>
          <xs:element ref="efog.celex" minOccurs="0"/>
          <xs:element ref="efog.reference-oj"/>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
<!--
.....................................................................
-->
    <xs:element name="efog.reference-oj">
      <xs:annotation>
        <xs:documentation>
          reference to the official publication of a legal act
        </xs:documentation>
      </xs:annotation>
      <xs:complexType>
        <xs:choice>
        <xs:choice>
          <xs:element ref="efog.uri"/>
          <xs:sequence>
            <xs:element ref="efog.series" minOccurs="0"/>
            <xs:element ref="efog.ojno" minOccurs="0"/>
            <xs:element ref="efog.date-publication"/>
            <xs:element ref="efog.page" minOccurs="0"/>
          </xs:sequence>
        </xs:choice>
        </xs:choice>
      </xs:complexType>
    </xs:element>
<!--
.....................................................................
-->
    <xs:element name="efog.registration" type="xs:string">
      <xs:annotation>
        <xs:documentation>
          information on the local registration
        </xs:documentation>
      </xs:annotation>
    </xs:element>
<!--
.....................................................................
-->
    <xs:element name="efog.series" type="xs:string">
      <xs:annotation>
        <xs:documentation>
          series of official journal/gazette
        </xs:documentation>
      </xs:annotation>
    </xs:element>
```

```
<!--
.....................................................................
-->
  <xs:element name="efog.signatory" type="xs:string">
    <xs:annotation>
      <xs:documentation>
        name(s) of person(s) empowered by a competent body to sign an
        act
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.target" type="xs:string">
    <xs:annotation>
      <xs:documentation>
        definition of the target of the limitation of applicability
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.title">
    <xs:annotation>
      <xs:documentation>
        official: name given to the resource; short: citation title
        prescribed in the act
      </xs:documentation>
    </xs:annotation>
    <xs:complexType mixed="true">
      <xs:attribute name="type">
        <xs:simpleType>
          <xs:restriction base="xs:string">
            <xs:enumeration value="official"/>
            <xs:enumeration value="short"/>
          </xs:restriction>
        </xs:simpleType>
      </xs:attribute>
    </xs:complexType>
  </xs:element>
<!--
.....................................................................
-->
  <xs:element name="efog.uri" type="xs:anyURI">
    <xs:annotation>
      <xs:documentation>
        electronic address
      </xs:documentation>
    </xs:annotation>
  </xs:element>
<!--
.....................................................................
-->
</xs:schema>
```

B. The schema container

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
   xmlns:xs="http://www.w3.org/2001/XMLSchema"
   elementFormDefault="qualified"
   attributeFormDefault="unqualified"
   xmlns="http://xxx">
   <xs:annotation>
     <xs:documentation>
       Version 1.00
     </xs:documentation>
     <xs:documentation>
       Container which includes the original schema and has to be
       referred to in the case of namespace import
     </xs:documentation>
   </xs:annotation>
   <xs:include schemaLocation="efog_20060703.xsd"/>
</xs:schema>
```

C. The DTD solution

```
<?xml version="1.0" encoding="UTF-8"?>
<!ENTITY    % t_date           "#PCDATA"                          >
<!--
   Element container
-->
<!ELEMENT   efog
              (efog.abbreviation*,efog.abrogation*,
              efog.act-number?,efog.applicability?,
              efog.authority?,efog.bibliographic-citation*,
              efog.date-enter-into-force?,efog.date-passing?,
              efog.date-publication,efog.date-signature*,
              efog.document-type,efog.eurovoc?,efog.index?,
              efog.nickname*,efog.reference-corrigendum*,
              efog.reference-eu?,efog.reference-oj?,
              efog.registration?,efog.signatory*,efog.title+)     >
<!--
.....................................................................
-->
<!ELEMENT   efog.abbreviation
                   (#PCDATA) >
<!ATTLIST   efog.abbreviation
                   type            (official|other)    #REQUIRED >
<!--
   official: abbreviation prescribed in the act;
   other: abbreviation currently used in the spoken language
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.abrogation
              (efog.date-abrogation, efog.target?,
              efog.originator?)                                   >
```

```
<!--
  information on abrogation of an act or a fragment
  specification of the date (mandatory), the target (optional)
  and the originator (law, jurisdiction, etc.) (optional)
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.act-number
              (#PCDATA)                                              >
<!--
  official number of an act, identification
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.applicability
              (efog.date-start, efog.date.end?, efog.condition?)  >
<!--
  specification of the applicability of an act, optional
  specification of reasons or conditions
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.authority
              (#PCDATA)                                              >
<!--
  competent body
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.bibliographic-citation
              (efog.uri | (efog.series?, efog.ojno?,
              efog.date-publication, efog.page?))                    >
<!--
  reference to another act
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.celex
              (#PCDATA)                                              >
<!--
  Celex number; attention: the pattern only covers directives
  and their corrigenda
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.condition
              (efog.p+)                                              >
```

```
<!--
   expression of reason or condition in the context of the
   applicability of an act
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.date-abrogation
               (%t_date;)                                            >
<!--
   date of abrogation of an act or a part of it
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.date-end
               (%t_date;)                                            >
<!--
   date of the end of applicability
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.date-enter-into-force
               (%t_date;)                                            >
<!--
   date of entering into the legal systems
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.date-passing
               (%t_date;)                                            >
<!--
   final acceptance by the Parliament
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.date-publication
               (%t_date;)                                            >
<!--
   date of publication in the official medium
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.date-signature
               (%t_date;)                                            >
<!ATTLIST   efog.date-signature
               quality     CDATA       #IMPLIED                      >
<!--
   signature of the act
-->
```

```
<!--
.....................................................................
-->
<!ELEMENT   efog.date-start
              (%t_date;)                                             >
<!--
  date of start of applicability
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.document-type
              (#PCDATA)                                              >
<!--
  specification of type of document such as regulation, directive,
  decision etc.
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.eurovoc
              (efog.keyword+)                                        >
<!--
  container for Eurovoc descriptors
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.index
              (efog.keyword+)                                        >
<!--
  container for descriptors of a local indexation system
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.keyword
              (#PCDATA)                                              >
<!--
  descriptor from Eurovoc or a local indexation system
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.nickname
              (#PCDATA)                                              >
<!--
  title currently used in the spoken language
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.ojno
              (#PCDATA)                                              >
```

```
<!--
  number of the official journal/gazette referred to
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.originator
              (#PCDATA)                                              >
<!--
  originator of the limitation of applicability of an act
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.p
              (#PCDATA)                                              >
<!--
  container for any information which has to be kept together
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.page
              (#PCDATA)                                              >
<!--
  first page of a document within a oj reference
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.reference-corrigendum
              ((efog.uri)|(efog.series?,efog.ojno?,
              efog.date-publication, efog.page?))                    >
<!ATTLIST   efog.reference-corrigendum
              roll      (active|passive)      #REQUIRED              >
<!--
  reference to a corrigendum
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.reference-eu
              (efog.celex?, efog.reference-oj)                       >
<!--
  container for the reference to a European directive with the
  Celex number as identifier and the bibliographic location
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.reference-oj
              (efog.uri|(efog.series?, efog.ojno?,
              efog.date-publication, efog.page?))                    >
<!--
  reference to the official publication of a legal act
-->
```

```
<!--
.....................................................................
-->
<!ELEMENT   efog.registration
              (#PCDATA)                                              >
<!--
  information on the local registration
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.series
              (#PCDATA)                                              >
<!--
  series of official journal/gazette
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.signatory
              (#PCDATA)                                              >
<!--
  name(s) of person(s) empowered by a competent body to sign an
  act
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.target
              (#PCDATA)                                              >
<!--
  definition of the target of the limitation of applicability
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.title
              (#PCDATA)                                              >
<!ATTLIST   efog.title
              type      (official| short)      #REQUIRED             >
<!--
  official: name given to the resource; short: citation title
  prescribed in the act
-->
<!--
.....................................................................
-->
<!ELEMENT   efog.uri
              (#PCDATA)                                              >
<!--
  electronic address
-->
<!--
.....................................................................
-->
```

D. Implementation

a) Schema

— Inclusion

The metadata schema can be included by means of a top level declaration in the local schema, such as:

```
<xs:include
schemaLocation= "http://formex.publications.europa.eu/schema/efog_
20060703.xsd"/>
```

Afterwards all elements of the included fragment can be used as if they were elements proper to the local schema. The local schema has to be updated if a new version of the metadata schema exists. The inclusion has no consequence for the instances.

— Namespace reference

The namespace solution is more complicated, but is recommended as the most conformant methodology. In order to avoid any problems with the default namespace in the metadata schema and the local schema, the local schema has to refer to the metadata schema container. The reference is managed by importing the container:

```
<xs:import
namespace="efog"
schemaLocation= "http://formex.publications.europa.eu/schema/efog_
cnt_20060703.xsd"/>
```

Additionally a namespace has to be defined in the schema declaration which refers to the same as the abovementioned one:

```
<xs:schema
xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified"
attributeFormDefault="unqualified"
xmlns:md="efog">
```

In this example the namespace prefix is declared as `md`; any other value which corresponds to the XML naming conventions will work as well. The following extract from a local schema shows how to use the referenced schema:

```
<xs:element name="document">
     <xs:complexType>
          <xs:sequence>
               <xs:element ref="md:efog"/>
               <xs:element ref="title"/>
               <xs:element ref="preamble"/>
               <xs:element ref="enacting.terms"/>
               <xs:element ref="final"/>
          </xs:sequence>
     </xs:complexType>
</xs:element>
```

The highlighted line demonstrates how the efog-block element is referred to.

The instances have to be prepared for the use of the namespace methodology as well. The following fragment shows how this can be managed, just by temporarily redefining the default namespace:

```
<document
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
     xsi:noNamespaceSchemaLocation="document-v2.xsd">
     <efog xmlns="http://xxx">
          <efog.act-number>337/2006</efog.act-number>
          <efog.authority>Commission</efog.authority>
          <efog.date-enter-into-force>20060225
          </efog.date-enter-into-force>
          <efog.date-signature>20060224</efog.date-signature>
          <efog.reference-oj>L 055, 20060225, 0001-0002
          </efog.reference-oj>
     </efog>
     <title>
          <p>COMMISSION REGULATION (EC) No 337/2006</p>
          <p>of 24 February 2006</p>
          <p>establishing the standard import values for
          determining the entry price of certain fruit and
          vegetables</p>
     </title>
...
```

b) DTD

The DTD approach does not allow namespace references. Although the W3C namespace recommendation is nearly as old as the XML standard itself, it was not integrated for reasons of compatibility with the SGML standard. The metadata DTD may be included in the local DTD by means of a parameter entity. The following example illustrates this approach:

```
<!ENTITY % efog SYSTEM
"http://formex.publications.europa.eu/schema/efog_20060703.dtd"
>
%efog;
```

Thus the metadata DTD becomes a component of the local DTD. Afterwards it may be reused without any further interventions:

```
<!ELEMENT document    (efog, title, preamble, enacting.terms,
                      final)
>
```

Nothing special has to be foreseen for the instances.

3. Members of the Working Group

Czech Republic

Pavel Blümel; DTP; Ministry of the Interior of the Czech Republic; Bartůňkova 4/1159; pošt. schr. 10; CZ-149 01 Praha 47 pošta 415; Phone: +420 974 887 380; E-mail: pavel.blumel@tmv.cz;

Emil Horčička; Ministerstvo vnitra České republiky; Director, Community Law Department; Náměstí Hrdinů 3; CZ-140 21 Praha; Phone: +420 974 817 270; E-mail: horcicka@mvcr.cz;

Zdeněk Jirouš; Network system manager; Ministry of the Interior of the Czech Republic; Bartůňkova 4/1159; pošt. schr. 10; CZ-149 01 Praha 47 Chodov; Phone: +420 974 887 369; E-mail: jirous@tmv.cz;

Milan Snížek; Ministry of the Interior of the Czech Republic; Community Law DepartmentNad Štolou 3; CZ-173 34 Praha 7; Phone: +420 974 832 968; E-mail: snizek@mvcr.cz

Denmark

Søren Broberg Nielsen; Head of Division; Civil Affairs Division; Legal Information Division; Borgergade 28,3.; DK-1300 København K; Phone: +45 3332 5222; E-mail: soren.nielsen@ civilstyrelsen.dk; *Chairman of the working group*

Estonia

Tanel Rõigas; IT juht; Riigi Teataja Kirjastus; Lossi plats 1a; EE-15163 Tallinn; Phone: +372 6935872; E-mail: tanel.roigas@rtk.ee

European Union

Holger Bagola; Principal Administrator; Office for Official Publications of the European Communities; Directorate A "Official Journal and Access to legislation" — "Formats"; 2, rue Mercier; L-2985 Luxembourg; Phone: +352 2929 42354; E-mail: Holger.Bagola@ cec.eu.int; *Secretary of the working group*

France

Emmanuel Cortes; Responsable de la section prépresse; Direction des Journaux Officiels; Service Informatique; 26, rue Desaix; F-75727 Paris Cedex 15; Phone: +33 1 40 58 80 01; E-mail: emmanuel.cortes@journal-officiel.gouv.fr;

Jean-François Imokrane; Chef du service informatique; Direction des Journaux officiels; 26, rue Desaix; F-75015 Paris Cedex 15; Phone: +33 1 40 58 77 21; E-mail: jean-francois. imokrane@ journal-officiel.gouv.fr

Germany

Uwe Mähren;l Amsterdamer-Straße 192; D-50735 Köln; Phone: +49 221 97668 311; E-mail: uwe. maehren@bundesanzeiger.de

Ilona Natke; Planung und Entwicklung elektronischer Publikationsplattformen; Amsterdamer Straße 192; D-50735 Köln; Phone: +49 221 97668 157; E-mail: ilona.natke@ bundesanzeiger.de;

Helmut Klein; Diplom-Informatiker; Teamleiter Systementwicklung Backoffice ; juris GmbH; Gutenbergstraße 23; D-66117 Saarbrücken; Phone: +49 681 5866-199; E-mail: helmut. klein@ juris.de;

Greece

Victoria Salavou; Kapodistriou 34; GR-104 32 Athinai; Phone: +30 210 5279156; E-mail: vsalavou@ et.gr;

Hungary

Béla Pusztai; IT Director; Magyar Hivatalos Közlönykiadó; Somogyi Béla 6; H-1085 Budapest; Phone: +36 1 266 5095; E-mail: pusztai.bela@mhk.hu

Italy

Catarina Lupo; Ufficio Standardizzazione e rappresentazione delle informazioni; Centro Nazionale per l'Informatica nella Pubblica Amministrazione; Via Isonzo, 21/B; I-00198 Roma; E-mail: lupo@cnipa.it

Lithuania

Birute Leonaniciene; Gedimino pr. 53; LT-01109 Vilnius; E-mail: Bileonrs@lrs.lt

Kastytis Trapinskas; Gedimino pr. 53; LT-01109 Vilnius; Phone: +370 52396301; E-mail: katrap@lrs.lt;

Former Yugoslav Republic of Macedonia

Biljana Cvetanovska; Head of Legal affairs department; Partizanski Odredi 29; MK-1000 Skopje; Phone: +389-2-329-88-60; E-mail: biljanac@slvesnik.com.mk;

Andon Stefanovsk; Partizanski Odredi 29; MK-1000 Skopjei; Phone: +389-2-329-88-60; E-mail: andon@slvesnik.com.mk

The Netherlands

Eric Stevens; Manager Infromatiediensten; SDU; Binckhorstlaan 403; Postbus 16262; NL-2500 BG Den Haag; Phone: +31 70 378 95 36; E-mail: E.Stevens@sdu.nl

Marc Woltering; Ministerie van Binnenlandse Zaken en Koninkrijksrelaties — Diensten en Agentschappen; Postbus 84011; NL-2508 AA Den Haag;; Phone: +31 70 8887877; E-mail: Marc.Woltering@ictu.nl

Portugal

Miguel Palma; Phone: +351 213 945700; E-mail: miguel.palma@dr.incm.pt

Romania

Claudia Acsinte; Str. Parcului nr 65, sector 1; RO-Bucaresti; Phone: +4072 3316925; E-mail: claudia.acsite@ramo.ro

Mihai Mocanu; Sos. Panduri, nr , bl P 33, sector 5; RO-Bucarest; Phone: +4072 4388366; Phone: mihai.mocanu@ramo.ro

Andrei Moldovanu; Head of Service; Monitorul Oficial; Str. Parcului nr. 65, sectorul 1; RO-Bucureşti; Phone: +04021 224 09 71 110; E-mail: amoldovanu@ramo.ro

Adriana Niculescu; Monitorul Oficial; Str. Parcului nr. 65, sectorul 1; RO-Bucureşti; Phone: +04021 224 09 71 110; E-mail: adriana_niculescu@ramo.ro;

Gheorghe Toma; Str. Parcului nr. 65, sector 1; RO-Bucaresti; Phone: +4072 4388342; E-mail: gheorghe.toma@ramo.ro

Sweden

Niklas Sojde; Phone: +46 8 405 4406; E-mail: niklas.sojde@adm.ministry.se

Per-Anders Svensson; Head of the Office of the Swedish Code of Statutes ; Ministry of Justice; Directorate-General of Administrative Affairs; Division for Legal and Linguistic Draft Revision; Rosenbad 4; S-103 33 Stockholm; Phone: +46 8405 1000; E-mail: per-anders.svensson@justice.ministry.se