INTERNET-DRAFT
Intended Status: Informational
Expires: October 3, 2011

P. Spinosa (ITTIG/CNR)
E. Francesconi (ITTIG/CNR)
C. Lupo (CNIPA)

April 1, 2011

A Uniform Resource Name (URN) Namespace for Sources of Law (LEX) draft-spinosa-urn-lex-03.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/1id-abstracts.html

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on October 3, 2011.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

This document describes a Uniform Resource Name (URN) Namespace Identification (NID) convention as prescribed by the World Wide Web Consortium (W3C) for identifying, naming, assigning, and managing persistent resources in the legal domain.

Table of Contents

1 Introduction 5
1.1 The Purpose of Namespace "lex" 5
1.2 Entities Supporting this Standard 5
1.3 The Context 5
1.4 General Characteristics of the System 7
1.5 Linking a "lex" Name to a Document 8
1.6 Use of "lex" Names in References 8
1.7 Definitions 9
1.8 Terminology 9
2 Specification Template 9
2.1 Namespace ID 9
2.2 Registration Information 9
2.3 Syntax Used in this Document 10
2.4 Identifier structure 10
3 General Syntax of the "lex" Identifier 12
3.1 Allowed and Not Allowed Characters 12
3.2 Reserved Characters 12
3.3 Case sensitivity 13
3.4 National Characters and Diacritic Signs 13
3.5 Replacement of Spaces, Connectives and Punctuation Marks 13
3.6 Abbreviation Expansion 13
3.7 Acronyms 13
3.8 Date Format 14
3.9 Ordinal Numbers 14
4 Creation of the Source of Law "lex" Identifier 14
4.1 Basic Principles 14
4.2 Model of Sources of Law Representation 14
4.3 The Structure of the Local Name 15
4.4 Structure of the Document Identifier at Work Level . . . 16
4.5 Aliases 17
4.6 Structure of the Document Identifier at Expression Level 17
4.7 Structure of the Document Identifier at Manifestation Level 19
4.8 Sources of Law References 20
5 The Procedure of Uniform Names Assignment 21
5.1 Specifying the <jurisdiction> element of the URN "lex" . 21
5.2 Jurisdictional Registrar for Names Assignment 21
5.3 Identifier Uniqueness 22
5.4 Identifier persistence considerations 22
6 Principles of the Resolution Service 23
6.1 The General Architecture of the System 23
6.2 Catalogues for Resolution 24
6.3 Suggested resolver behaviour 24
7 Considerations 25
7.1 Conformance with URN Syntax 25
7.2 Validation mechanism 25

7.3 Scope 25
7.4 Namespace Considerations 26
7.5 Community Considerations 26
7.6 IANA Considerations 26
7.7 Security Considerations 26
8 References 27
8.1 Normative References 27
8.2 Informative References 27
9 Acknowledgments 28
10 Author's Addresses 28
Attachment A -- Summary of the syntax of the uniform names of the "lex" namespace 30
Attachment B -- Specific Syntax of the Identifier at Work Level . 34
B1 The <authority> element 34
B1.1 Indication of the Authority 34
B1.2 Multiple Issuers 34
B1.3 Indication of the Issuer 34
B1.4 Indication of the Body 34
B1.5 Indication of the Function 35
B1.6 Conventions for the Authority 35
B2 The <measure> element 35
B2.1 Criteria for the Indication of the Type of Measure . . . 35
B2.2 Further Specification to the Type of Measure 36
B2.3 Aliases for Sources of Law with Different Normative References 36
B2.4 Relations between Measure and Authority in the Aliases . 36
B3 The <details> element 37
B3.1 Indication of the Details 37
B3.2 Multiple Dates 37
B3.3 Unnumbered Measures 38
B3.4 Multiple Numbers 38
B4 The <annex> element 39
B4.1 Formal Annexes 39
B4.2 Annexes of Annexes 39
Attachment C -- Specific Syntax of the <version> Element of the Expression 40
C1 The <version> element 40
C1.1 Different Versions of a Legislative Document 40
C1.2 Identification of the Version 40

1 Introduction

1.1 The Purpose of Namespace "lex"

The purpose of the "lex" namespace is to assign an unequivocal identifier, in standard format, to documents that are sources of law. To the extent of this namespace, "sources of law" include any legal document within the domain of legislation (including bills), case law and administrative acts or regulations; therefore "legal doctrine" is explicitly not covered.

The identifier is conceived so that its construction depends only on the characteristics of the document itself and is, therefore, independent from the document's on-line availability, its physical location, and access mode.

This identifier will be used as a way to represent the references (and more generally, any type of relation) among the various sources of law. In an on-line environment with resources distributed among different Web publishers, uniform resource names allow simplified global interconnection of legal documents by means of automated hypertext linking.

1.2 Entities Supporting this Standard

The following entities support this proposal:

- ITTIG/CNR (Institute of Legal Information Theory and Techniques of the Italian National Research Council) - Italy;
- CNIPA (National Centre for ICT in Public Administration) - Italy;
- PRODASEN - IT Department of the Federal Senate - Brazil;
- LII (Legal Information Institute), Cornell Law School - USA

1.3 The Context

In the last few years a number of initiatives have arisen in the field of legal document management.

Since 2001 the Italian Government, through the CNIPA (National Authority for Information Technology in the Public Administration), the Ministry of Justice and ITTIG-CNR (the Institute of Legal Information Theory and Techniques of the Italian National Research Council) promoted the NormeInRete project. It was aimed at introducing standards for sources of law description and identification using XML and URN techniques.

Other national initiatives in Europe introduced standards for the description of legal sources [FRAN]: for example the Metalex project,

promoted by the University of Amsterdam and adopted by the Dutch Tax and Customs Administration, the Belgian Public Centers for Welfare and others; LexDania project in Denmark supported by the Danish Ministry of Justice; CHLexML in Switzerland developed by COPIUR, the Coordination Office for the Electronic Publication of Legal Data Federal Office of Justice; eLaw in Austria mainly coordinated by the Austrian Parliament.

Such initiatives, based in synergies between government, national research institutes, and universities, have defined national XML standards for legal document management, as well as schemes for legal document identification.

Outside Europe, similar initiatives have faced similar problems. For example, the Brazilian Senate carried out a feasibility study to provide unique and transparent identifiers to sources of law on the basis of the IFLA-FRBR model.
Similarly, the Akoma Ntoso (Architecture for Knowledge-Oriented Management of African Normative Texts using Open Standards and Ontologies) project provides a set of guidelines for e-Parliament services in a Pan-African context by proposing an XML document schema providing sophisticated description possibilities for several Parliamentary document types (including bills, acts and parliamentary records, etc.).
Finally, the Tasmanian Government provided advanced legislative information services through the EnAct project. It gave searchable consolidated Tasmanian legislation by automating much of the legislative drafting and consolidation process, as well as using SGML document representation. Numerous less-visible efforts in the United States and elsewhere have struggled with similar issues.

Several of these identifiers are based on a URN schema. The first national standard was defined in Italy within the NormeInRete project; to this the Danish LexDania and the Brazilian Lexml standard followed. Hungary, Slovenia and Switzerland expressed their interest in URN identifier for legislation as well. All these standards have a common internal structure, regarding both the hierarchy and the elements content.

In today's information society the processes of political, social and economic integration of European Union member states as well as the increasing integration of the world-wide legal and economic processes are causing a growing interest in exchanging legal information knowledge at national and trans-national levels.
The growing desire for improved quality and accessibility of legal information amplifies the need for interoperability among legal information systems across national boundaries. A common open standard used to identify sources of law at international level is an

essential prerequisite for interoperability.

Interest groups within several countries have already expressed their intention to adopt a shared solution based on a URN technique. The need for a unequivocal identifier of sources of law in different EU Member States, based on open standards and able to provide advanced modalities of document hyper-linking, has been expressed in several conferences by representatives of the Office for Official Publications of the European Communities (OPOCE), with the aim of promoting interoperability among national and European institution information systems. Similar concerns have been raised by international groups concerned with free access to legal information, and the Permanent Bureau of the Hague Conference on Private International Law is considering a resolution that would encourage member states to "adopt neutral methods of citation of their legal materials, including methods that are medium-neutral, provider-neutral and internationally consistent". In a similar direction the CEN Metalex initiative is moving, at European level, towards the definition of a standard interchange format for sources of law, including recommendations for defining naming conventions to them.

The "urn:lex" naming convention has interpreted all these recommendations, proposing an original solution for sources of law identification.

1.4 General Characteristics of the System

Registrants wish now to promote interoperability among legal information systems by the definition of a namespace convention and structure that will create and manage identifiers for legal documents. The identifiers will be:
- globally unique
- transparent
- persistent
- location-independent, and
- language-neutral.

These qualities will facilitate legal document management as well as provide a mechanism of stable cross-collections and cross-country references.
Language-neutrality is an especially important feature that will promote adoption of the standard by organizations that must adhere to official-language requirements. The proposed standard will provide useful guidance to both public and private groups that create, promulgate, and publish legal documents. Registrants wish to minimize the potential for creating conflicting proprietary schemes, while preserving sufficient flexibility to allow for diverse document types and to respect the need for local control of collections by an equally diverse assortment of administrative entities.

As usual, the problem is to provide the right amount guidance at the core of the standard while providing sufficient flexibility to cover a wide variety of needs. The proposed "lex" standard does this by splitting the identifier into parts. The first part uses a predetermined standard ("country/jurisdiction name standard") to specify the country (or more generally the jurisdiction) of origin for the legal document being identified; the remainder ("local name") is intended for local use in identifying documents issued in that country or jurisdiction. This second part depends only on sources of law identification system operating in that nation and it is mainly composed by a formalized information related to the enacting authority, the type of measure, the details and possibly the annex.

The identification system based on uniform names MUST include:
- a schema for assigning names capable of representing unambiguously any addressed source of law, namely legislation, case law and administrative acts, issued by any authority (intergovernmental, supranational, national, regional and local) at any time (past, present and future);
- a resolution mechanism - in a distributed environment - that ties a uniform name to the on-line location of the corresponding resources.

This document only considers the first of these requirements. It also contains a few references to the architecture of the resolution service and to the corresponding software.

1.5 Linking a "lex" Name to a Document

The "lex" name is linked to the document through meta-information which may be specified:
- internally to the document itself through a specific element within an XML schema or by an HTML META tag;
- externally by means of an RDF triple, a specific attribute in a database, etc.

One of these modalities is necessary for enabling automated construction and updating of catalogues (distributed and centralized) and the implementation of resolvers that associate the uniform name of a document with its physical location(s). The standard assumes no particular relationship between the originator of the document, its publisher, and the implementer of catalogues or resolution services. They may be the same entity, or not.

1.6 Use of "lex" Names in References

"lex" names will be used on a large scale in references as a HREF attribute value of the hypertext link to the referred document. This link can be created in two ways:
- by manually inserting, in the referring document, the link with the

uniform name: this is a burdensome procedure especially for documents that are already on-line;
- by automatically constructing (either permanently or temporarily) the link with the uniform name, through reference parsers of a text: this is a more time-saving procedure even if subject to a certain percentage of errors, since references are not always accurate or complete. This solution could nevertheless be acceptable for already published documents.

In any case, whatever the method adopted is, new documents produced in XML format compliant with the relative DTD/XMLSchema, SHOULD express references through the uniform name of the document referred to.

1.7 Definitions

According to this document, the following terms are used in the following meaning:
- Source of Law:
  is a general concept, and is used to refer to legislation, case law, regulations and administrative acts. In its broadest sense, the source of law is anything that can be conceived of as the originator of 'erga omnes' legal rules;
- Registrar:
  is an organization which shares and defines in any country or jurisdiction the assignment of the main components of the resource identifier through which its uniqueness is guaranteed. This task includes the definition of possible jurisdiction unit and the primary elements (issuing authority and type of legal measure) of uniform name, according to the characteristics of its own state or institution organization.

1.8 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2 Specification Template

2.1 Namespace ID

"lex" requested according to [RFC3406].

2.2 Registration Information

Version Number: 1.0
Date: 2011-04-01

Declared registrant of the namespace:

Institute of Legal Information Theory and Techniques (ITTIG)
Italian National Research Council (CNR)
Via de' Barucci, 20
50127 Florence
Italy

e-mail: lex@ittig.cnr.it

2.3 Syntax Used in this Document

This document uses the syntax common to many Internet RFCs, which is based on the BNF (Backus-Naur Form) meta-language. In particular:

- elements are included between angle brackets ("<" and ">");
- an element is separated from its specification by the string "::=";
- alternative elements are separated from each other by a vertical slash ("|");
- character strings are enclosed in quotes (" and ");
- optional parts are enclosed by square brackets ("[" and "]");
- a group of elements is enclosed by round brackets ("(" and ")");
- a symbol or an expression following an element or a group of elements indicates a factor of repetition, and, as in the regular expressions, takes the following formats:
  - ?   : 0 or 1 time;
  - +   : 1 or more times;
  - *   : 0 or more times;
  - {n} : <n> times;
  - {n,m}: from <n> to <m> times.

2.4 Identifier structure

The identifier has a hierarchical structure as follows:

"urn:lex:"<NSS>

where NSS is the Namespace Specific String composed as follows:

<NSS>::=<jurisdiction>":"<local-name>

where:

<jurisdiction> is the part providing the identification of the jurisdiction, generally corresponding to the country, where the source of law is issued. It is also possible to represent international organizations (either states or public administrations or private entities);

<local-name> is the uniform name of the source of law in the country or jurisdiction where it is issued; its internal structure is common to the already adopted schemas. It is able to represent all the aspects of an intellectual production, as it is a legal document, from its initial idea, through its evolution during the time, to its realisation by different means (paper, digital, etc.).

The <jurisdiction> element is composed of two specific fields:

<jurisdiction>::=<jurisdiction-code>[";"<jurisdiction-unit>]*

where:

<jurisdiction-code> is usually the identification code of the country where the source of law is issued; this code follows the standard [ISO3166] Alpha-2 (it=Italy, fr=France, dk=Denmark, etc.). In case of multi-national (e.g., European Union) or international (e.g., United Nations, Free Software Foundation) organizations the Top Level Domain Name (e.g., "eu") or the Domain Name (e.g., "un.org", "wto.int") is used instead of ISO 3166 code. In case such multi-national or international organization does not have a registered domain, in order to avoid ambiguities or collisions with actual domains, a domain name (according to the english acronym of the organization name) under the virtual domain "lex" is used. For example, the jurisdiction code of the European Economic Community is "eec.lex".

<jurisdiction-unit> are the possible administrative hierarchical sub-structures defined by each country or organisation according to its own legal system. This additional information can be used where two or more levels of legislative or judicial production exist (e.g., federal, state and municipality level) and the same bodies may be present in each jurisdiction. Then acts of the same type issued by similar authorities in different areas differ for the jurisdiction-unit specification. An example can be the following: "br:governo:decreto" (decree of federal government), "br;sao.paulo:governo:decreto" (decree of SU+00E3o Paulo state) and "br;sao.paulo;campinas:governo: decreto" (decree of Campinas municipality).

Examples of law sources identifiers are:

urn:lex:it:stato:legge:2003-09-21;456 (Italian act)
urn:lex:fr:etat:loi:2004-12-06;321 (French act)
urn:lex:es:estado:ley:2002-07-12;123 (Spanish act)
urn:lex:ch;glarus:regiere:erlass:2007-10-15;963 (Glarus Swiss Canton decree)
urn:lex:eu:commission:directive:2010-03-09;2010-19-EU (EU Commission Directive)

urn:lex:us:federal.supreme.court:decision:1963-03-18;372.us.335 (US FSC decision)
urn:lex:be:conseil.etat:decision:2008-07-09;185.273 (Decision of the Belgian Council of State)

3 General Syntax of the "lex" Identifier

3.1 Allowed and Not Allowed Characters

These characters are defined in accordance with the [RFC2141] "URN Syntax". For various reasons, later explained, in the "lex" <NSS> only a sub-set of characters is allowed. All other characters are either eliminated or converted.

For the full syntax of the uniform names in the "lex" space, please see Attachment A.

3.2 Reserved Characters

These characters MUST always and uniquely be used for the assigned purpose.
The first category includes those characters bearing a specific meaning in the general creation of the URI (Uniform Resource Identifier)[RFC3986]:

"%" "/" "?" "#"

The following characters instead are reserved in the specific "lex" namespace:

- "@" separator of the expression, that contains information on version and language;
- "$" separator of the manifestation, that contains information on format, editor, etc.;
- ":" separator of the main elements of the name at any entity;
- ";" separator of level. It identifies the introduction of an element at a hierarchically lower level, or the introduction of a specification;
- "+" separator of the repetitions of an entire main element (e.g., multiple authorities);
- "," separator of the repetitions of individual components in the main elements, each bearing the same level of specificity (e.g., multiple numbers);
- "~" separator of the partition identifier in references (e.g., paragraph of an article);

- "*" and "!" are reserved for future expansions.

3.3 Case sensitivity

The specific name <NSS> of the URN, as with URLs, is case-sensitive. Since the case does not change the logical identification of the source of law, the names belonging to the "lex" namespace are considered functionally equivalent independently from the case. To take advantage of memory caching, the specific name is always created in lower case.
(e.g., "Ministry" will be recorded as "ministry")

3.4 National Characters and Diacritic Signs

In order to keep editing and communication more simple and to avoid character percent-encoding, it is strongly recommended that national characters and diacritic signs are turned into base characters (e.g., the Italian term "sanitU+00E0" converted into "sanita", the French term "ministU+00E8re" converted into "ministere"). Otherwise, the characters have to be percent-encoded according to the UTF-8 character encoding [STD63] (e.g., "sanitU+00E0" encoded into "sanit%C3%A1").
Anyway each country or jurisdiction decides the uniform names encoding modality of all the sources of law issued within its territory.

3.5 Replacement of Spaces, Connectives and Punctuation Marks

All the language connectives (e.g., articles, prepositions, etc.), the punctuation marks and all the special characters (as apostrophes, dashes, etc.) are eliminated. The words left are connected each other by a dot (".") which substitutes the "space".
(e.g., "Ministry of Finances, Budget and of Economic Planning" becomes "ministry.finances.budget.economic.planning")

3.6 Abbreviation Expansion

All abbreviations indicating institutions (e.g., Min.), structures (e.g., Dept.), or legal measures (e.g., reg.), MUST be expanded.
(e.g., "Min." must be reported as "ministry")

3.7 Acronyms

The use of acronyms might be confusing and encourage ambiguity in uniform names (the same acronym may indicate two different institutions or structures), therefore their expansion is strongly recommended.
(e.g., "FAO" is to be expanded as "food.agriculture.organization")

3.8 Date Format

Dates are expressed by numbers in the ISO-8601 format:

yyyy-mm-dd

(e.g., "September 2, 99" will be written as "1999-09-02")

3.9 Ordinal Numbers

Any ordinal number included in a component of a document name (e.g., in the description of an institution body) MUST be indicated in Arabic numerals, regardless to the original expression: whether in Roman numerals, or with an adjective, or in Arabic numeral with apex, etc. (IV, third, 1U+00B0, 2^, etc.).
(e.g., "Department IV" becomes "department.4")

4 Creation of the Source of Law "lex" Identifier

4.1 Basic Principles

The uniform name must identify one and only one document (more precisely a "bibliographic entity") and is created in such a way that it is:
- self-explanatory ;
- identifiable through simple and clear rules;
- compatible with the practice commonly used for references;
- able to be created by references in the text, automatically (by parser) or manually;
- representative of both the formal and the substantive aspects of the document.

4.2 Model of Sources of Law Representation

According to FRBR (Functional Requirements for Bibliographic Records) model developed by IFLA (International Federation of Library Associations and Institutions), in a source of law, as in any intellectual production, 4 fundamental entities (or aspects) can be specified.

The first 2 entities reflect its contents:
- work: identifies a distinct intellectual creation; in our case, it identifies a source of law both in its being (as it has been issued or proposed) and in its becoming (as it is modified over time);
- expression: identifies a specific intellectual realisation of a work; in our case it identifies every different (original or up-to-date) version of the source of law over time and/or language in

which the text is expressed;
while the other 2 entities relate to its form:
- manifestation: identifies a concrete realisation of an expression; in our case it identifies realizations in different media (printing, digital, etc.), encoding formats (XML, PDF, etc.), or other publishing characteristics;
- item: identifies a specific copy of a manifestation; in our case it identifies individual physical copies as they are found in particular physical locations.

4.3 The Structure of the Local Name

The <local-name> of "lex" namespace MUST contain all the necessary pieces of information enabling the unequivocal identification of a legal document.
In the legal domain, at the "work" level, they are essentially four: the enacting authority, the type of measure, the details and possibly the annex.
It is often necessary to differentiate various expressions, that is:
- the original version and all the amended versions of the same document;
- the versions of the text expressed in the different official languages of the state or organization.

Finally the uniform name allows a distinction among diverse manifestations, which may be produced in multiple locations using different means and formats.
In every case, the basic identifier of the source of law (work) remains the same, but information is added regarding the specific version under consideration (expression); similarly a suffix is added to the expression for representing the characteristics of the publication (manifestation).
The information which forms a source of law uniform name at each level (work, expression, manifestation) is expressed in the official language of the related jurisdiction; in case of more official languages (as in Switzerland) or more involved jurisdictions (as in international treaties), more language-dependent names (aliases) are created.

Therefore, the more general structure of the national name appears as follows:

<local-name>::=<work>["@"<expression>]?["$"<manifestation>]?

However, consistent with legislative practice, the uniform name of the main original provision (work) becomes the identifier of an entire class of documents which includes: the original main document, the annexes, and all their versions, languages and formats subsequently generated.

4.4 Structure of the Document Identifier at Work Level

The structure of the document identifier is made of the four fundamental elements mentioned above, clearly distinguished one from the other in accordance with an order identifying increasingly narrow domains and competences:

<work>::=<authority>":"<measure>":"<details>[":"<annex>]*

where:

<authority> is the issuing or proposing authority of the measure (e.g., State, Ministry, Municipality, Court, etc.);

<measure> is the type of the measure, both public nature (e.g., constitution, act, treaty, regulation, decree, decision, etc.) as well as private one (e.g., license, agreement, etc);

<details> are the terms associated to the measure, typically the date (usually the signature date) and the number included in the heading of the act;

<annex> is the identifier of the annex, if any (e.g., Annex 1).

In case of annexes, both the main document and its annexes have their own uniform name so that they can individually be referenced; the identifier of the annex adds a suffix to that of the main document. In similar way the identifier of an annex of an annex adds an ending to that of the annex which it is attached to.

The main elements of the national name are generally divided into several elementary components, and, for each, specific rules of representation are established (criteria, modalities, syntax and order).
For the details regarding each element, please see the Attachment B.

Examples of <work> identifiers are:

urn:lex:it:stato:legge:2006-05-14;22
urn:lex:uk:ministry.justice:decree:1999-10-07;45
urn:lex:ch;glarus:regiere:erlass:2007-10-15;963
urn:lex:es:tribunal.supremo:decision:2001-09-28;68
urn:lex:fr:assemblee.nationale:proposition.loi:13.legislature;1762
urn:lex:br:estado:constituicao:1988-10-05;lex-1
urn:lex:fsf.org:free.software.foundation:general.public.license:2007-06-29;lex-1
urn:lex:nl:hoge.raad:besluit:2008-04-01;bc8581

It is worth to note that the type of measure is important to identify case law, as well as legislation, especially within the legal systems where cases, by tradition, are identified only through the year of release and a number. Since the aim of the "urn:lex" schema is to identify specific materials, the type of measure or the full date are able to provide discrimination between materials belonging to a specific case.

Here below is an example where the type of measure or the full date are essential for identify specific materials of a case:
- 4/59 Judgment of the EU Court of Justice 04/04/1960, Mannesmann AG and others / ECSC High Authority
  urn:lex:eec.lex:court.justice:judgment:1960-04-04;4-59
- 4/59 Order of the EU Court of Justice 18/05/1960, Mannesmann AG and others / ECSC High Authority
  urn:lex:eu:court.justice:order:1960-05-18;4-59

4.5 Aliases

International treaties involve more jurisdictions (the signing ones) so they are represented through more identifiers, each of them related to an involved jurisdiction. For example, a bilateral France and Germany treaty is identified through two URNs belonging to either "fr" or "de" jurisdiction
(e.g., "urn:lex:fr:etat:traite:..." and
"urn:lex:de:staat:vertrag:...")
since it pertains to both the French and the German jurisdiction.

In the states or organisations that have more than one official language, a document has more identifiers, each of them expressed in a different official language, basically a set of equivalent aliases. This system permits manual or automated construction of the uniform name of the referred source of law in the same language used in the document itself.
(e.g., "urn:lex:eu:council:directive:2004-12-07;31",
"urn:lex:eu:consiglio:direttiva:2004-12-07;31", etc.)

Moreover, a document can be assigned more than one uniform name in order to facilitate its linking to other documents. This option can be used for documents that, although unique, are commonly referenced from different perspectives. For example, the form of a document's promulgation and its specific content (e.g., a Regulation promulgated through a Decree of the President of the Republic).

4.6 Structure of the Document Identifier at Expression Level

There may be several expressions of a legal text, connected to specific versions or languages.

Each version is characterized by the period of time during which that text is to be considered as the valid text (in force or effective). The lifetime of a version ends with the issuing of the subsequent version.
New versions of a text may be brought into existence by:
- changes in the text (amendments) due to the issuing of other legal acts and to the subsequent production of updated or consolidated texts;
- correction of publication errors (rectification or errata corrige);
- entry into or departure from a particular time span, depending on the specific date in which different partitions of a text come into force.

Each of such versions may be expressed in more than one language, with each language-version having its own specific identifier.
The identifier of a source of law expression adds such information to the work identifier, using the following main structure:

<expression>::="@"<version>[":"<language>]?

where:

<version> is the identifier of the version of the (original or amended) source of law. In general it is expressed by the promulgation date of the amending act; anyway other specific information can be used for particular documents. If necessary, the original version is specified by the string "original" (for the details regarding this element, please see the Attachment C);

<language> is the identification code of the language in which the document is expressed, according to [ISO639-1] (it=Italian, fr=French, de=German, etc.); in case the code of a language is not included in this standard, the [ISO639-2] (3 letters) is used. This information is not necessary when the text is expressed in the unique official language of the country or jurisdiction.

Examples of document identifiers for expressions are:

urn:lex:ch:etat:loi:2006-05-14;22@originel:fr (original version in French)
urn:lex:ch:staat:gesetz:2006-05-14;22@original:de (original version in German)
urn:lex:ch:etat:loi:2006-05-14;22@2008-03-12:fr (amended version in French)
urn:lex:ch:staat:gesetz:2006-05-14;22@2008-03-12:de (amended version in German)
urn:lex:be:conseil.etat:decision:2008-07-09;185.273@originel:fr (original version in French of a Belgian decision)

4.7 Structure of the Document Identifier at Manifestation Level

To identify a specific manifestation, the uniform name of the expression is followed by a suitable suffix describing the:
- digital format (e.g., XML, HTML, PDF, etc.) expressed according to the MIME Content-Type standard [RFC2045], where the "/" character is to be substituted by the "-" sign;
- editorial staff who produced it, expressed according to its Internet domain name;
- possible components of the expressions contained in the manifestation. Such components are expressed by language-dependent labels representing the whole document (ex: "all") or the main part of the document (ex: "body") or the caption label of the component itself (e.g. Table 1, Figure 2, etc.);
- other features of the document (e.g., anonymized decision text).

The <manifestation> suffix will thus read:

```
<manifestation>::=<format>[";"<specification>"]*
                   ":"<editor>[";"<specification>]*
                  [":"<component>[";"<specification>]*]?
                  [":"<feature>[";"<specification>]*]?
```

To indicate possible features or peculiarities, each main element of the manifestation MAY be followed by further specifications, for example as regards <format> the version, for <editor> the archive name and the electronic publisher, etc.

(examples:
the original version the Italian act 3 April 2000, n. 56 might have the following manifestations with their relative uniform names:
- PDF format (vers. 1.7) of the whole act edited by the Italian Parliament:
  "urn:lex:it:stato:legge:2000-04-03;56$application-pdf;1.7:parlamento.it"
- XML format (version 2.2 DTD NIR) of the text of the act and PDF format (version 1.7) of the "Figura 1" (figure 1) contained in the body, edited by the Italian Senate:
  "urn:lex:it:stato:legge:2000-04-03;56$text-xml;dtd-nir-2.2:senato.it:testo"
  "urn:lex:it:stato:legge:2000-04-03;56$application-pdf;1.7:senato.it:figura.1"

the Spanish URN of the html format of the whole Judgement of the European Court of Justice n. 33/08 of 11/06/2009, in Spanish version, published in the Jurifast data base:
"urn:lex:eu:tibunal.justicia:sentencia:2009-06-11;33-08@original:es$text-html:juradmin.eu;jurifast:all")

Furthermore, it is useful to be able to assign a uniform name to a manifestation (or to a part of it) in case non-textual objects are involved. These may be multimedia objects that are non-textual in their own right (e.g. geographic maps, photographs, etc.), or texts recorded in non-textual formats, such as image scans of documents.

In these ways, a "lex" name permits:
- exploitation of all the advantages of an unequivocal identifier that is independent of physical location;
- a means to provide choice among different existing manifestations (e.g. XML or PDF formats, resolution degree of an image etc.) of the same expression.

4.8 Sources of Law References

References to sources of law often refer to specific partitions of the act (article, paragraph, etc.) and not to the entire document. An act partition is a logical subdivision of the text, that, in a structured format (as XML) fitting the document logical structure, is represented by an element with its own ID; this ID aims to identify the element and to locate it. In a mark-up that does not fit the logical structure of the text (as HTML), generally only the starting point of the partition, and not the element, is identified through a label (a <a name> tag).
Therefore, for allowing browsers to point to a specific partition, it is necessary that such partition is endowed with an unequivocal label or ID within the including document and its value is the same independently from the document format.
For enabling the construction of the partition identifier between different collections of documents, specific construction rules for IDs or labels SHOULD be defined and shared, within each country or jurisdiction, for any document type (e.g., for legislation, the paragraph 2 of the article 3 might have as label or ID the value "art3-par2", similarly for case-law, paragraph 22 of the judgment in Case 46/76 Bauhuis v Netherlands, might have as label or ID the value "par22").
Furthermore, it is useful to foresee the compatibility with applications able to manage this information (e.g., returning the proper element); these procedures are particularly useful in the case of rather long acts, such as codes, constitutions, regulations, etc.
For this purpose it is necessary that the partition identifier is transmitted to the servers (resolution and application) and therefore it cannot be separated by the typical "#" character of URI fragment, which is not transmitted to the server.

According to these requirements, the syntax of a reference is:

<URN-reference> ::= <URN-document> ["~" <partition-id>]?

(e.g., to refer to the paragraph 3 of the article 15 of the French Act of 15 may 2004, n. 106, the reference is written "urn:lex:fr:etat:loi:2004-05-15;106~art15-par3").

Using a different separator ("~") from the document name, the partition ID is not withheld by the browser but it is transmitted to the resolution process. This enables the resolver to retrieve (for example, out of a database), if it is possible, only the referred partition, otherwise to return the whole act.
Anyway, to make it effective pointing to the indicated partition, the resolver SHOULD transform the partition ID of each returned URL in a URI fragment; this is obtained appending to URL the "#" character followed by the partition ID (in the example above, the returned URL will be <URL-document>#art15-par3).

Anyway it is possible to use the general syntax (with "#"); in this case only the URN document component of the reference is transmitted to the resolver, therefore the whole document will be always retrieved.

5 The Procedure of Uniform Names Assignment

5.1 Specifying the <jurisdiction> element of the URN "lex"

Under the "lex" namespace, each country or international organization is assigned with a jurisdiction code, which characterizes the URNs of the source of law of that country or jurisdiction. This code is assigned according to the ISO 3166 Alpha-2 (as well as TLDN or DN for the organizations) representation and it is the value of the <jurisdiction-code> element, which preserves cross-country uniqueness of the identifiers.

5.2 Jurisdictional Registrar for Names Assignment

Any country or jurisdiction, who intends to adopt this schema, identifies a Jurisdictional Registrar, an organization which shares and defines the structure of the optional part (<jurisdiction-unit>) of the name, according to the organization of the state or institution. For example, in a federal state a <jurisdiction-unit> corresponding to the name of each member state (e.g. "br;sao.paolo", "br;minas.gerais", etc.) may be defined.

The process of assigning the <local-name> will be managed by each specific country or jurisdiction under the related <jurisdiction> element.
In any country the Jurisdictional Registrar shares and defines the assignment of the primary elements (issuing authority and type of legal measure) of the local names considering the characteristics of

its own state or institution organization.
Such a Registrar SHOULD establish, according to the guidelines indicated in the current document, a uniform procedure within the country or organization to define <local-name> elements, to take decisions upon normalizations and finally to solve and avoid possible name collisions as well as to maintain authoritative registries of various kinds (e.g., for authorities, types of measures, etc.). In particular, accurate point-in-time representations of the structure and naming of government entities are important to semantically-aware applications in this domain.
Moreover, the Registrar shares and defines the rules to construct partition IDs for each document type.
Finally, the Registrar will develop and publish the rules and the guidelines for the <local-name> construction as well as the predefined values and codes.

5.3 Identifier Uniqueness

Identifiers in the "lex" namespace are defined through a <jurisdiction> element assigned to the sources of law of a specific country or organization, and a <local-name> assigned by the issuing authority. The main elements (authority and type of measure) of the <local-name> are defined by the Jurisdictional Registrar, so that it is ensured that the constructed URNs are unique. The Jurisdictional Registrar SHOULD provide clear documentation of rules by which names are to be constructed, and SHOULD update and make accessible its registries.

Any issuing authority is responsible to define formal parameters to guarantee local name uniqueness by attributing, if necessary, a conventional internal number, which, combined with the other <local-name> components (authority, measure and date), builds an unequivocal identifier. Uniqueness is achieved by checking against the catalogue of previously assigned names.

5.4 Identifier persistence considerations

The persistence of identifiers depends on the durability of the institutions that assign and administer them. The goal of the "lex" namespace schema is to maintain uniqueness and persistence of all resources identified by the assigned URNs.

In particular, CNIPA and ITTIG-CNR, as proposers, are responsible of maintaining the uniqueness of the <jurisdiction> element; given that the <jurisdiction> is assigned on the basis of the long-held ISO 3166 Alpha-2 representation of the country (or the TLD name of the organization) and that the country or organization associated code is expected to continue indefinitely, the URN also persists

indefinitely.

The rules for the construction of the name are conceived to delegate the responsibility of their uniqueness to a set of authorities which is identified within each country or organization.

Therefore, each authority is responsible for assigning URNs which have a very long life expectancy and can be expected to remain unique for the foreseeable future. Practical and political considerations, as well as diverse local forms of government organization, will result in different methods of assigning responsibility for different levels of the name.
Where this cannot be accomplished by the implementation of an authoritative hierarchy, it can and SHOULD be done by creating consensus around a series of published rules for the creation and administration of names by institutions and bodies that operate by means of collaboration rather than compulsion.

Issuing authorities that operate in more localized scopes, ranging from the national down to the very local, MUST equally take responsibility for the persistence of identifiers within their scope.

6 Principles of the Resolution Service

6.1 The General Architecture of the System

The task of the resolution service is that of associating a "lex" identifier with a specific document address on the network. By contrast with systems that can be constructed around rigorous and enforceable engineering premises, such as DNS, the "lex" resolver will be expected to cope with a wide variety of "dirty" inputs, particularly those created by the automated extraction of references from incomplete or inaccurate texts. In this document, the result is a particular emphasis on a flexible and robust resolver design.

The system has a distributed architecture based on two fundamental components: a chain of information in DNS (Domain Name System) and a series of resolution services from URNs to URLs, each competent within a specific domain of the namespace.
Through the NAPTR records of the DNS (described in [RFC3403]), the client identifies the characteristics (protocol, port, site) of the service (e.g. according to [RFC2169]) capable of associating the relative URLs with the URN in question, thereby allowing access to the document.

A resolution service can delegate the resolution and management of hierarchically-dependent portions of the name.

Delegation of this responsibility will not be unreasonably withheld provided that the processes for their resolution and management are robust and are followed.

For the "lex" namespace, CNIPA and ITTIG-CNR will maintain the root zone "lex.urn.arpa" and, in correspondence with the adhesion of a new country (e.g., "br") or organization, will update the DNS information with a new record to delegate the relative resolution. This may be obtained by a regular expression that matches the initial part of the URN (e.g., "urn:lex:br") and redirects towards the proper zone (e.g., "lex.senado.gov.br").

Likewise the institution responsible for the jurisdiction uniform names (e.g., "urn:lex:br") has the task of managing the relative root in the DNS system (e.g., "lex.senado.gov.br" zone) and routing the resolution towards its resolvers on the basis of parts of the uniform names. In similar way it can delegate the resolution of country/organization sub-levels (e.g., "urn:lex:br;sao.paolo") towards the relative zone (e.g., "lex.sao-paolo.gov.br").

The resolution service is made up of two elements: a knowledge base (consisting in a catalogue or a set of transformation rules) and a software to query the knowledge base itself.

6.2 Catalogues for Resolution

Incompleteness and inaccuracy are rather frequent in legal citations, and incomplete or inaccurate uniform names of the referred document are thus likely to be built from textual references (this is even more frequent if they are created automatically through a specific parser). For this reason, the implementation of a catalogue, based on a relational-database, is suggested, as it will lead to a more higher flexibility in the resolution process.
In addition the catalogue must manage the aliases, the various versions and languages of the same source of law as well as the related manifestations.

It is suggested that each enacting authority implements its own catalogue, assigning a corresponding unambiguous uniform name to each resource.

6.3 Suggested resolver behaviour

First of all the resolver should separate the part corresponding to the partition ID, through the "~" separator, from the document name.

So, the resolution process SHOULD implement a normalization of the uniform name to be resolved. This may involve transforming some

components to the canonical form (e.g., filling out the acronyms, expanding the abbreviations, unifying the institution names, standardizing the type of measures, etc.). For this function authorities and types of measure registers are useful.

The resolver SHOULD then query the catalogue searching for the URN which corresponds exactly to the given one (normalized if necessary). Since the names coming from the references may be inaccurate or incomplete, an iterative, heuristic approach (based on partial matches) is indicated. It is worth remarking that incomplete references (not including all the elements to create the canonical uniform name) are normal and natural; for a human reader, the reference would be "completed" by contextual understanding of the reference in the document in which it occurs.

In this phase, the resolver should use the partition ID information to retrieve, if it is possible, only the referred partition, otherwise to return of the entire document.

Lacking more specific indications, the resolver SHOULD select the best (most recent) version of the requested source of law, and provide all the manifestations with their related items.
A more specific indication in the uniform name to be resolved will, of course, result in a more selective retrieval, based on any suggested expression and/or manifestations components (e.g. date, language, format, etc.).

Finally, the resolver SHOULD append to URLs the "#" character followed by partition ID, transforming it in a URI fragment for browser pointing.

7 Considerations

7.1 Conformance with URN Syntax

No special considerations.

7.2 Validation mechanism

The national Authority (or those it delegates) of each adhering country or organization is responsible of the definition or acceptance of the uniform name's primary elements (issuing authority and type of legal measure).

7.3 Scope

Global interest.

7.4 Namespace Considerations

In collaboration with the legislative XML community, registrants carried out a preliminary study of the URI alternatives to satisfy the key requirements.
The options analysed were: a private URI scheme, URL, PURL and URN. URN was considered the most appropriate URI given the requirements analysis.
Advantages we would emphasize are:
- greater flexibility in building the identifier;
- the capacity to represent name components that are not strictly hierarchical;
- the potential for clear division of the identifier into macro parts, main elements and components, using different separators;
- ease of managing optional parts of a name.

7.5 Community Considerations

The use of the "lex" namespace facilitates the interoperability of information systems used in the Public Administration at the national and international level. Moreover it allows the distribution of the legal information towards a federated architecture. In such an architecture, documents are directly managed by the issuing authorities, with resulting benefits in information authenticity, quality and currency. A shared identification mechanism resources guarantees that a distributed system will be as efficient and effective as a comparable centralized system.

Creators of Internet content that references legal materials - including publishers operating well outside the traditional arenas of legal publishing - benefit by the registration of the namespace because facilitates the linking of legal documents, whether by manual or automated means, and reduces the cost of maintaining documents that contain such references.

Any citizen or organisation with Internet web browser capability will be entitled to access the namespace and its associated application, registers, and resolution services, to facilitate document access.

7.6 IANA Considerations

This document includes a URN NID registration for "lex" for entry in the IANA registry of URN NIDs (see [RFC5226] for more information).

7.7 Security Considerations

This document introduces no additional security considerations beyond those associated with the use and resolution of URNs in general.

8 References

8.1 Normative References

[STD63] F. Yergeau, "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, November 2003.

[RFC2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3406] D Daigle, L., van Gulik, D., Iannella, R., and P. Faltstrom, "Uniform Resource Names (URN) Namespace Definition Mechanisms", BCP 66, RFC 3406, October 2002.

[RFC2141] R. Moats, K. R. Sollins, "URN Syntax", RFC 2141, May 1997.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax", STD 66, RFC 3986, January 2005.

[RFC3403] M. Mealling, Dynamic Delegation Discovery System (DDDS), Part Three: The Domain Name System (DNS) Database, RFC 3403, October 2002.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008

[ISO3166] ISO 3166, "Country name codes", ISO 3166-1:1997.

[ISO639-1] ISO 639-1, "Codes for the representation of names of languages" - Part 1: alpha-2 code, 2003.

[ISO639-2] ISO 639-2, "Codes for the representation of names of languages" - Part 2: alpha-3 code, 1999.

[RFC2169] R. Daniel, "A Trivial Convention for using HTTP in URN", RFC 2169, June 1997

[RFC2045] N. Freed, N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

8.2 Informative References

[SPIN] P.L. Spinosa, "The Assignment of Uniform Names to Italian Legal Documents", May, 2006

http://www.nir.it/sito_area3-ap_stan_assegnazione_nomi.htm

[FRAN] E. Francesconi, "Technologies for European Integration. Standards-based Interoperability of Legal Information Systems", ISBN 978-88-8398-050-3, European Press Academic Publishing, 2007.

9 Acknowledgments

The authors of this document wish to thank all the registrants for giving suggestions and comments.
They are also grateful to the Legislative XML community for the interesting discussions on this topic and to the Working Group "Identification of the legal resources through URNs" of Italian NormeInRete project for the provided guidance [SPIN].
The authors owe a debt of gratitude to Tom Bruce, director of the Legal Information Institute of the Cornell University Law School, for his contribution in revising this document and sharing fruitful discussions which greatly improved the final result. A special thank goes also to Joao Alberto de Oliveira Lima, legislative system analyst of the Brazilian Federal Senate, and to Attila Torcsvari, information management consultant, for their detailed comments on the first drafts of this document, which provided significant hints to the final version of the standard. Thanks also go to Robert Richards of the Legal Information Institute (Cornell University Law School), promoter and maintainer of the Legal Informatics Research social network, as well as to the members of this network, for their valuable comments on this proposal.
Finally, many thanks go to Loriana Serrotti who significantly contributed to the drafting of this document.

10 Author's Addresses

PierLuigi Spinosa, Enrico Francesconi
Istituto di Teoria e Tecniche dell'Informazione Giuridica (ITTIG)
Consiglio Nazionale delle Ricerche (CNR)
Via de' Barucci, 20
50127 Firenze
Italy
Telephone: +39 055 43995
e-mail: {pierluigi.spinosa, enrico.francesconi}@ittig.cnr.it

Caterina Lupo
Centro Nazionale per l'Informatica nella Pubblica Amministrazione

(CNIPA)
Viale Carlo Marx, 31/49
00137 Roma
Italy
Telephone: +39 06 85264262
e-mail: caterina.lupo@cnipa.it

Attachment A -- Summary of the syntax of the uniform names of the "lex" namespace

```
*-------------------------------------------------------------------
* General Structure of a Uniform Resource Name (URN)
* NID = namespace
* NSS = specific name
*-------------------------------------------------------------------
<URN> ::= "urn:"<NID>":"<NSS>

*-------------------------------------------------------------------
* Structure of a Uniform Resource Name (URN) of the "lex" namespace
*-------------------------------------------------------------------
<NID> ::= "lex"

<URN> ::= "urn:lex:"<NSS-lex>

*-------------------------------------------------------------------
* Structure of a "lex" specific name
*-------------------------------------------------------------------
<NSS-lex> ::= <jurisdiction>":"<local-name>

*-------------------------------------------------------------------
* Structure of the <jurisdiction> element
*-------------------------------------------------------------------
<jurisdiction> ::= <jurisdiction-code>[";"<jurisdiction-unit>]*

   <jurisdiction-code> ::= <lowercase>{2,4}

   <jurisdiction-unit> ::= <alfanum>[<normal>]*

*-------------------------------------------------------------------
* Structure of the <national-name> element
*-------------------------------------------------------------------
<national-name> ::= <work>["@"<expression>]?["$"<manifestation>]?

*-------------------------------------------------------------------
* Structure of the <work> element
*-------------------------------------------------------------------
<work> ::= <authority>":"<measure>":"<details>[":"<annex>]*

*-------------------------------------------------------------------
* Structure of the <authority> element
*-------------------------------------------------------------------
<authority> ::= <issuer>["+"<issuer>]*

   <issuer> ::= (<institution>[";"<body>]*[";"<function>]*) | <office>
```

```
      <institution> ::= <alfanum>[<normal>]*

      <body> ::= <alfanum>[<normal>]*

      <function> ::= <alfanum>[<normal>]*

      <office> ::= <alfanum>[<normal>]*

*-------------------------------------------------------------------
* Structure of the <measure> element
*-------------------------------------------------------------------
<measure> ::= <measure-type>[";"<specification>]*

   <measure-type> ::= <alfanum>[<normal>]*

   <specification> ::= <alfanum>[<normal>]*

*-------------------------------------------------------------------
* Structure of the <details> element
*-------------------------------------------------------------------
<details>  ::=  (<dates>|<period>)";"<numbers>

   <dates>  ::=  <date>[","<date>]*
"lex"
   <period>  ::= <alfanum>[<normal>]*

   <numbers> ::= (<document-id>[","<document-id>]*)|<number-lex>

      <document-id> ::= <alfanum>[<normal>|<other>]*

      <number-lex> ::= "lex-"<digit>+

*-------------------------------------------------------------------
* Structure of the <annex> element
*-------------------------------------------------------------------
<annex> ::= <annex-id>[";"<specification>]*

   <annex-id> ::= <alfanum>[<normal>]*

*-------------------------------------------------------------------
* Structure of the <expression> element
*-------------------------------------------------------------------

<expression> ::= <version>[":"<language>]?

*-------------------------------------------------------------------
* Structure of the <version> element
*-------------------------------------------------------------------
```

```
<version> ::= (<amendment-date>|<specification>)
              [";"(<event-date>|<event>)]*

   <amendment-date> ::= <date>

   <event-date> ::= <date>

   <event> ::= <alfanum>[<normal>]*

*-------------------------------------------------------------------
* Structure of the <language> element
*-------------------------------------------------------------------

<language> ::= <lowercase>{2,3}

*-------------------------------------------------------------------
* Structure of the <manifestation> element
*-------------------------------------------------------------------
<manifestation> ::= <format>[";"<specification>"]*
                    ":"<editor>[";"<specification>"]*
                    [":"<component>[";"<specification>]*]?
                    [":"<feature>[";"<specification>]*]?

<format> ::= <alfanum>[<normal>|"-"]*

<editor> ::= <alfanum>[<normal>|"-"]*

<component> ::= <alfanum>[<normal>|"-"]*

<feature> ::= <alfanum>[<normal>|"-"]*

*-------------------------------------------------------------------
* Structure of the date
*-------------------------------------------------------------------
<date> ::= <year>"-"<month>"-"<day>

   <year>  ::= <digit>{4}
   <month> ::= <digit>{2}
   <day>   ::= <digit>{2}

*-------------------------------------------------------------------
* Allowed characters
*-------------------------------------------------------------------
<allowed-lex> ::= <normal> | <other> | <reserved> | <future>

   <normal> ::= <alfanum> | "."

      <alfanum> ::= <lowercase> | <digit> | <encoded>
```

```
      <lowercase> ::= "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"

      <digit>     ::= "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" |
                      "8" | "9"

      <encoded>   ::= ("%" (<digit> | <hex-let>)){1,6}

      <hex-let>   ::= "a" | "b" | "c" | "d" | "e" | "f"


   <other>    ::= "-" | "_" | "'" | "=" | "(" | ")"

   <reserved> ::= ":" | "@" | "$" | "+" | ";" | "," | "~"

   <future>   ::= "*" |  "!"
```

Attachment B -- Specific Syntax of the Identifier at Work Level

B1 The <authority> element

B1.1 Indication of the Authority

The <authority> element of a uniform name may indicate, in the various cases:

- the actual authority issuing the legal provision. More specifically, the authority adopting the provision or enacting it;
- the institution where the provision is registered, known and referenced to, even if produced by others (e.g., the bills identified through the reference to the Chamber where they are presented);
- the institution regulated (and referred to in citations) by the legal provision even when this is issued by another authority (e.g., the statute of a Body);
- the entity that proposed the legal material not yet included in the institutional process (e.g. a proposed bill written by a a political party).

B1.2 Multiple Issuers

Some sources of law are enacted by a number of issuing parties (e.g., inter-ministerial decrees, agreements, etc.). In this case, the <authority> element contains all the issuing parties (properly separated), as follows:

<authority> ::= <issuer>["+"<issuer>]*

(e.g., "ministry.justice+ministry.finances")

B1.3 Indication of the Issuer

Each issuing authority is essentially represented by either an institutional office (e.g., Prime Minister) or an institution (e.g., Ministry); in the last case, the authority is indicated in accordance with the institution's hierarchical structure, from the more general to more specific (Council, Department, etc.), ending with the relative office (President, Director, etc.).
Therefore, the structure of the issuer is as follows:

<issuer> ::= (<institution>[";"<body>]*[";"<function>]*) | <office>

(e.g., "ministry.finances;department.revenues;manager")

B1.4 Indication of the Body

Depending on the kind of measure, the body within the issuing authority is unambiguously determined (e.g., the Council for Regional Acts) and normally it is not indicated in the references.
Just like in practice, the indication of the enacting authority is limited to the minimum in relation to the type of measure.
(e.g., "region.tuscany:act" and not "region.tuscany;council:act")

B1.5 Indication of the Function

Generally, the component <function> is indicated, sometimes instead of the body itself:
- in case of political, representative or elective offices (e.g., "university.oxford;rector:decree" instead of "university.oxford;rectorship:decree");
- when it refers to a top officer in the institution (e.g., general manager, general secretary, etc.) which is not always possible to associate a specific internal institutional structure to (e.g., "national.council.research;general.manager").

It is not indicated when it clearly corresponds to the person in charge of an institution (typically, a general director); in this case, only the structure and not the person in charge is indicated (e.g., "ministry.justice;department.penitentiary.administration").

The function MUST be indicated when:
- it is not the same of the director or the person in charge of the structure (for example, in case of an undersecretary, a deputy director, etc.);
- the type of measure may be both monocratic or collegial: the indication of the office eliminates the ambiguity.

B1.6 Conventions for the Authority

Acts and measures bearing the same relevance as an act, issued or enacted since the foundation of the State, have conventionally indicated "state" (expressed in each country official language) as authority; the same convention is used for constitutions, codes (civil, criminal, civil procedure, criminal procedure, etc) and international treaties.

B2 The <measure> element

B2.1 Criteria for the Indication of the Type of Measure

In uniform names the issuing authority of a document is mandatory.
This makes unnecessary to indicate any further qualification of the

measure (e.g., ministerial decree, directorial ordinance, etc.), even if it is widely used.
When the authority-measure combination clearly identifies a specific document, the type of measure is not defined through attributes referring to the enacting authority.
(e.g., "region.tuscany:act" and not "region.tuscany:regional.act")

B2.2 Further Specification to the Type of Measure

In the <measure> element, it is usually sufficient to indicate the type of a measure. As usual, references to sources of law, rather than through the formal details (date and number), may be made through some of their characteristics such as the subject-matter covered (e.g., accounting regulations), nicknames referring to the promoter (e.g., Bassanini Act) or to the topic of the act (e.g., Bankruptcy Law), etc..
In these cases, the type of measure MAY be followed by further specifications useful in referencing even if the details are lacking:

<measure>::=<measure-type>[";"<specification>]*

(e.g., "regulations;accounting" or "act;bankruptcy")

B2.3 Aliases for Sources of Law with Different Normative References

There are legislative measures that, although unique, are usually cited in different ways, for example through the legislative act introducing them into the legal order (President's decree, legislative decree, etc.) or through their legislative category (regulations, consolidation, etc.).
In order to ensure, in all the cases, the validity of the references, an alias that takes into account the measure category is associated to the uniform name, representing the legislative form.
(e.g., "state:decree.legislative:1992-07-24;358" and "state:consolidation;public.contracts:1992-07-24;358").

B2.4 Relations between Measure and Authority in the Aliases

The sources of law including different normative references are usually introduced in legislation through the adoption or the issuing of an act, which they are either included or attached to. It is, therefore, necessary to create an alias linking the two aspects of the same document. Specifically, the different measures can be:
- adopted/issued by an authority different from the one regulated by the provision (e.g., the statute of a Body); in this case, the correlation is established between two uniform names each featuring a completely different <authority> element
  (e.g., "italian.society.authors.publishers:statute" and

"ministry.cultural.activities+ministry.finances.budget.economic.
planning:decree");
- issued by the institution itself either because it has issuing authority or by virtue of a proxy (e.g., a provision that refers to the functioning of the Body itself); in this case, the two aliases share the first part of the authority;
(e.g., "municipality.firenze:statute" and "municipality.firenze;council:deliberation");
- issued by the same Body to regulate a particular sector of its own competence; in this case the <authority> element is the same (e.g., "ministry.justice:regulation;use.information.tools.telematic.process" and "ministry.justice:decree").

B3 The <details> element

B3.1 Indication of the Details

The details of a source of law usually include the date of the enactment and the identification number (inclusion in the body of laws, register, protocol, etc.).
Some measures can have multiple dates; there are also cases in which the number of the measure does not exist (unnumbered measures) or a measure has multiple numbers (e.g., unified cases). For these reasons, the set up of both elements (date and number) includes multiple values.

Some institutions (e.g., the Parliaments) usually identify documents through their period of reference (e.g., the legislature number) rather than through a date, which would be much less meaningful and never used in references (e.g., Senate bill S.2544 of the XIV legislature). In these cases, the component <period> is used in substitution of the component <dates>.

Usually details of a measure are not reported according to a specific sequence; in accordance with the global structure of the uniform name, which goes from the general to the specific, the sequence date-number has the following form:

<details>::=(<dates>|<period>)";"<numbers>

(e.g., "2000-12-06;126", "14.legislature;s.2544")

B3.2 Multiple Dates

Some sources of law, even if unique, are identified by more than one date; in this case, in the field <dates> all the given dates are to be reported and indicated as follows:

<dates>::=<date>[","<date>]*

(e.g., the measure of the Data Protection Authority of December 30, 1999- January 13, 2000, No. 1/P/2000 has the following uniform name: "personal.data.protection.authority:measure:1999-12-30,2000-01-13; 1-p-2000").

B3.3 Unnumbered Measures

Measures not officially numbered in the publications may have a non-unequivolcal identifier, because several measures of the same type can exist, issued on the same day by the same authority.
To ensure that the uniform name is unambiguous, the <numbers> field MUST, in any case, contain a discriminating element, which can be any identifier used internally, and not published, by the authority (e.g., protocol).
If the authority does not have its own identifier, one identifier MUST be created for the name system. In order to easily differentiate it, such number is preceded by the string "lex-":

<number-lex>::="lex-"[<digit>]+

(e.g., "ministry.finances:decree:1999-12-20;lex-3")

It is responsibility of the authority issuing a document to assign a discriminating specification to it; in case of multiple authorities, only one of them is responsible for the assignment of the number to the document (e.g., the proponent).
The unnumbered measures published on an official publication (e.g., the Official Gazette), instead of by a progressive number are recognized by the univocal identifying label printed on the paper. Such an identifier, even if unofficial but assigned to a document in an official publication, is to be preferred because it has the clear advantage to be public and therefore easier to be found.

B3.4 Multiple Numbers

Some legal documents (e.g., bills), even if unique, are identified by a set of numbers (e.g., the unification of cases or bills).
In this case, in the <numbers> field, all the identifiers are reported, according to the following structure:

<numbers>::=<document-id>[","<document-id>]*

(e.g., "2000-06-12;c-10-97,c-11-97,c-12-97")
The characters which are not allowed (e.g., "/") or reserved (e.g., ":"), including the comma, cannot exist inside the <document-id>, and therefore MUST be turned into "-".

This conversion may imply that the uniform name of the document is no more unique (e.g., removal 123-BIS and return 123/BIS of the bill 123 both are identified as "123-bis"); in this case, it is necessary to add a specific distinctive ending (e.g., "123-bis-removal" and "123-bis-return").

B4 The <annex> element

B4.1 Formal Annexes

Although annexes are an integral part of the legal document, they may be referred to and undergo amendments separately from the act to which they are annexed. It is, therefore, necessary that both the main document as well as each formal individual annex is univocally identified.

Formal annexes may be registered as separate parts or together with a legal provision; they may also be autonomous in nature or not. In any case, they MUST be given a uniform name, which includes the uniform name of the source of law to which they are attached, and a suffix which identifies the annex itself.

The suffix of formal annexes includes the official heading of the annex and, possibly, further specifications (e.g., the title) which will facilitate the retrieval of the annex in case the identifier is missing:

<annex>::=<annex-id>[";"<specification>]*

(e.g., "region.sicily;council:deliberation:1998-02-12;14:annex.a; borders.park")

The characters which are not allowed (e.g. "/") or which are reserved (e.g. ":") must not be featured in the <annex-id> and therefore MUST be turned into ".".

B4.2 Annexes of Annexes

When there are annexes to an annex, their corresponding identifiers are created by adding to the identifier of the original annex those of the annexes that are connected with it (that is, attached to it).

(e.g., Table 1 attached to Attachment A of the preceding legal act has the following uniform name:
"region.sicily;council:deliberation:1998-02-12;14:annex.a; borders.park:table.1;municipality.territories").

Attachment C -- Specific Syntax of the <version> Element of the Expression

C1 The <version> element

C1.1 Different Versions of a Legislative Document

The creation of an updated text of a document may have one of the following forms:

- "multi-version": when specific mark-ups which identify the modified parts of a document (added, substituted or deleted parts) and their related periods of effectiveness are indicated inside one single object (e.g., an xml file). Such a document will be able, in a dynamic way, to appear in different forms according to the requested date of effectiveness;
- "single-version": when, on the contrary, a new and distinct object is created for each amendment to the text at a given time. Each object is, therefore, characterized by its own period of validity.

In any case all the versions SHOULD be linked one another and immediately navigable.

C1.2 Identification of the Version

In order to identify the different time versions of the same act, to the uniform name of the original document has to be added a specific suffix.
Such a suffix identifies each version of a legal provision and includes, first and foremost, one of the following elements:

- the issuing date of the last amending measure taken into account;
- the date in which the communication of the rectification or of the errata corrige, is published;
- a specification which must identify the reason concerning the amendment (e.g., the specific phase of the legislative process), for the cases in which the date is not usually used (e.g., bills).

Anyway it is possible to add further specifications that will distinguish each of the different versions of the text to guarantee identifier unequivocalness. For example with regard to changes of the in-force or effectiveness of any partition or portion of the text itself (e.g., when the amendments introduced by an act are applied at different times) or different events occurring in the same date.

```
<version>::=(<amendment-date>|<specification>)
                [";"(<event-date>|<event>)]*
```

where:

- <amendment-date> contains the issuing date of the last considered

amendment or of the last communication of amendment. In case the original text introduces differentiated periods in which an act is effective and the information system produces one version for each of them, such element contains the string "original";
- <specification> any information useful to identify unambiguously and univocally the version;
- <event-date> contains the date in which a version is put into force, is effective or is published;
- <event> is a name assigned to the event producing a further version (e.g., amendment, decision, etc.).

The issuing date of an amending act was chosen as identifier of a version because it can be obtained from the heading (formal data).

(e.g., the name "state:royal.decree:1941-01-30;12@1998-02-19" identifies the updated text of the "Royal Decree of 30/1/1941, No. 12" with the amendments introduced by the "Law Decree of 19/2/1998, No. 51", without any indication of its actual entry into force. The same uniform name with the additional ending ";1999-01-01" indicates the in-force or effective version starting in a different date (from 1/1/99).

For a full compatibility, every updating of a text or of the effectiveness of a "multi-version" document implies the creation of a new uniform name, even if the object remains only one, containing the identifier of the virtually generated version, exactly as in the case of a "single-version" document. A specific meta-data will associate every uniform name with the period of time during which such a name together with its corresponding text is to be considered valid.

(e.g., the multi-version document containing the "R.D. of 01/30/1941, no. 12", updated by the amendments introduced by the "D.Lgs. of 02/19/1998, no. 51", contains the name of the original "state:royal.decree:1941-01-30;12" as well as the name of the updated version "state:royal.decree:1941-01-30;12@1998-02-19").

Please note that in case of attachments or annexes, the creation of a new version (even in the case of only one component) would imply the creation of a new uniform name for all the connected objects in order to guarantee their alignment (i.e., the main document, the attachments and annexes).