DELIVERABLE

Project Acronym: EUCases
Grant Agreement number: 611760
Project Title: European and National Legislation and Case Law Linked in Open Data Stack

D2.2 Legal XML-schema (XSD)

Revision: final

Authors:
Guido Boella University of Torino
Loredana Cupi University of Torino
Luigi di Caro University of Torino
Monica Palmirani University of Bologna, CIRSFID
Livio Robaldo University of Torino
Andrea Violato Nomotika

Project co-funded by the European Commission within FP7-ICT-2013-SME-DCA

Dissemination Level

PU Public
PP Restricted to other programme participants (including the Commission Services)
RE Restricted to a group specified by the consortium (including the Commission Services)
CO Confidential, only for members of the consortium (including the Commission Services)
## Revision History, Status, Abstract, Keywords, Statement of Originality

### Revision History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Author</th>
<th>Organisation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>25.5.2014</td>
<td>Monica Palmirani</td>
<td>University of Bologna</td>
<td>First release from the sub-contract with University of Bologna</td>
</tr>
<tr>
<td>2.0</td>
<td>5.6.2014</td>
<td>Andrea Violato</td>
<td>Nomotika</td>
<td>Integration</td>
</tr>
<tr>
<td>3.0</td>
<td>10.6.2014</td>
<td>Loredana Cupi</td>
<td>University of Torino</td>
<td>Corrections</td>
</tr>
<tr>
<td>4.0</td>
<td>15.6.2014</td>
<td>Monica Palmirani</td>
<td>University of Bologna</td>
<td>Final version</td>
</tr>
<tr>
<td>5.0</td>
<td>22.6.2014</td>
<td>Guido Boella</td>
<td>University of Bologna</td>
<td>Final check</td>
</tr>
</tbody>
</table>

### Date of delivery

<table>
<thead>
<tr>
<th>Contractual:</th>
<th>Actual:</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.05.2014</td>
<td>23.06.2014</td>
</tr>
</tbody>
</table>

### Status

- final ✗ /draft ☐

### Abstract (for dissemination)

This document presents the XML Schema following the Akoma Ntoso legislative XML standard to be used in EUCases. The schema has been carefully selected to fit the needs of the projects, exploiting the modular structure of the standard. The XML Schema will be used to store legal documents (regulations, case law and doctrine) in a structured format to improve search and publication as linked open data.

### Keywords

- Legal XML, Akoma Ntoso, XML Schema
Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.
Table of Content

Introduction ................................................................................................................................................... 6

1  State of the Art ........................................................................................................................................ 7

2  Legal XML Standards ................................................................................................................................ 9

3  EUCases Needs & Requirements ............................................................................................................ 13
   3.1  Akoma Ntoso ....................................................................................................................................... 14
   3.2  Multi-layer Legal Document Modelling ............................................................................................... 15
   3.3  Akoma Ntoso Design Principles ........................................................................................................... 15

4  Akoma Ntoso Document Structure ........................................................................................................ 18
   4.1  Main structure ..................................................................................................................................... 18
   4.2  Example of Act Analysis ....................................................................................................................... 19
   4.3  meta Block ........................................................................................................................................... 23

5  Acts in Akoma Ntoso .............................................................................................................................. 25

6  Judgments in Akoma Ntoso ...................................................................................................................... 36
   6.1  Main element of the header ................................................................................................................ 36
   6.2  Structure of main part of the judgments ............................................................................................. 36
   6.3  Metadata of judgments ....................................................................................................................... 37
   6.4  Citations in judgments ......................................................................................................................... 38
   6.5  Rule of law, stare decisis and ratio decidendi, obiter dicta ................................................................. 39

7  Judgment Analysis Example ................................................................................................................... 40
   7.1  Akoma Ntoso Judgment Body .............................................................................................................. 42
   7.2  Akoma Ntoso Judgment Example ........................................................................................................ 44

8  Docs in Akoma Ntoso ............................................................................................................................. 60

9  EUCases Requirements .......................................................................................................................... 62

10 Akoma Ntoso for EUCases ....................................................................................................................... 63

11 Akoma Ntoso Compliance Levels ........................................................................................................... 65

12 Compliance with ELI, ECLI standards ........................................................................................................ 66
   12.1  Naming convention for citations ......................................................................................................... 66
   12.2  ELI/ECLI ........................................................................................................................................... 66
   12.3  URN:LEX ........................................................................................................................................... 68
   12.4  Akoma Ntoso: URIs and Citations .................................................................................................. 69
   12.4.1  Example of a EU Directive with Annexes .................................................................................... 70
   12.4.2  Virtual citations ........................................................................................................................... 71
   12.4.3  Example of a EU Directive with Annexes .................................................................................... 72
12.5 Comparison and Comments ................................................................................................................. 72
12.5.1 Akoma Ntoso resolver ................................................................................................................ 73

13 Final XML-schema (XSD) for EUCases ..................................................................................................... 74
13.1 Element: ref ......................................................................................................................................... 74
13.2 Element: judgement ........................................................................................................................... 74
13.2.1 Element: judgementBody ........................................................................................................... 75
13.2.2 Element: conclusions ................................................................................................................. 75
13.2.3 Element: introduction ............................................................................................................... 76
13.2.4 Element: background ............................................................................................................... 76
13.2.5 Element: motivation ................................................................................................................. 77
13.2.6 Element: decision ...................................................................................................................... 78
13.3 Element: judicial ............................................................................................................................... 78
13.4 Element: act ...................................................................................................................................... 79
13.4.1 Element: coverPage ................................................................................................................... 79
13.4.2 Element: preface ...................................................................................................................... 80
13.4.3 Element: preamble ................................................................................................................... 81
13.4.4 Element: body .......................................................................................................................... 81
13.4.5 Element: conclusions .............................................................................................................. 82

14 Next steps .............................................................................................................................................. 84

15 References ............................................................................................................................................. 89

List of Figures

Fig. 1 – Akoma Ntoso main skeleton ......................................................................................................... 18
Fig. 2 – Act schema ................................................................................................................................. 19
Fig. 3 – Metadata schema ....................................................................................................................... 23
Fig. 4 – Judgment main structure in Akoma Ntoso ................................................................................ 36
Fig. 5 – Judgment Body sub-elements in Akoma Ntoso ........................................................................ 37
Fig. 6 – Actor qualification ..................................................................................................................... 38
Fig. 7 – Shepard's method qualification ................................................................................................ 38
Fig. 8 – Scheme of an FRBR structure .................................................................................................. 59
Fig. 9 – URI composition ....................................................................................................................... 59
Introduction

A legal XML schema based on the legal XML standards (like NormeinRete, Akoma Ntoso or CEN Metalex) is required to structure the documents, represent metadata and (cross-)links, and format the legislative and judicial documents on different devices.

We selected as standard Akoma Ntoso since it as discussed in the next section is the new generation of standard and it is increasingly adopted by parliaments, among which the European Parliament.

This document will be the basis for the text to XML transformation tool, for the linking tools and for the database schema where the documents will be collected and metadata organized. It will allow also the indexing of documents with ELI/ECLI standards adopted by EU. Moreover, it will allow access to single parts of norms for previewing in the applications deriving from the project, linking of documents with each other, and linking the documents with ontologies and classifications.

This deliverable is the result of the Task 2.2 with the contribution of the subcontractor CIRSFID University of Bologna, the excellence center in legal informatics where the Akoma Ntoso standard has been promoted.
1 State of the Art

There are many XML standards applied in law, but few have all the characteristics for representing the legal knowledge and the legal documents in the proper way. Many have been created only for Web publishing and make no distinction between representation and presentation of content (Formex is one example). Others (such as LexDania and NormeInRete) are specific to one country and are difficult to adapt to other legal systems. Others (the US House of Representatives XML standard) do not distinguish the applicative-procedural part from mere representation of the document.

Several legal document standards were introduced in the last twenty years. EnAct, written by Timothy Arnold-Moore for the government of Tasmania, was the first system in 1995 producing the point-in-time\(^1\) legislative database in SGML. In 1992, the LII (Legal Information Institute) of Cornell Law School, launched by Peter Martin and Tom Bruce, has been provided on the web (HTML) the consolidated United States Code\(^2\). AustLII, Australasian Legal Information Institute, co-funded by Graham Greenleaf in 1995, makes today accessible on the web more of 400 legal database using HTML. Eur-Lex began to consolidate database of European Legislation in 1999 using Formex, an SGML data standard now translated in XML (Formex v4)\(^3\). Norway activates on 1st January 2001 a web service by Lovdata\(^4\) and provides consolidated legislation. France transformed in 2002 the commercial service Jurifrance into a public web portal called Legifrance\(^5\), including consolidated text in mixed format (HTML, XML, PDF). Austria with eLaw project (2004) transformed its previous database RIS (1983) into a web collection of authentic documents, dematerializing completely the legal Official Gazette publication The Emilia-Romagna Region (Italy) started the consolidation of regulations in 2003 using the NormeInRete XML schema\(^6\), and the Italian High Court of Cassation started the same mark-up in 2005 and now it is approaching to consolidate the set of document. Senate of Italy uses till 2013 now Akoma Ntoso for the bill for producing the side-by-side comparison text for underlining the differences between two versions (mostly coming from other institutions). Senate of Brazil from June 30th, 2009 launched the parliamentary consolidated database (LexMLBrazil\(^7\) with the point-in-time function based on a customisation of XML Akoma Ntoso schema. The Library of Congress of Chile\(^8\) provides the actualized legislation using the national XML schema for legal resources, and from the 2009 offers, by the LeyChile service, all the versions over the time of the legal documents starting from the 1998. Kenya Law Report\(^9\) is now converting their database of laws in XML documents marked-up in Akoma Ntoso standard.

---

\(^1\) Point-in-time is the function that permits to manage all the versions of the document over the time and not only the original document and the last version.

\(^2\) http://www.law.cornell.edu/uscode/

\(^3\) http://formex.publications.europa.eu/index.html

\(^4\) http://www.lovdata.no/info/lawdata.html

\(^5\) http://www.legifrance.gouv.fr/

\(^6\) http://demetra.regione.emilia-romagna.it/

\(^7\) http://projeto.lexml.gov.br/documentacao/resumo-em-ingles

\(^8\) http://www.leychile.cl/Consulta

\(^9\) http://www.kenyalaw.org/update/index.php
Akoma Ntoso is now used by the following institutions:

1. Senate of Brazil (act, bill, consolidation, point-in-time);
2. European Parliament (bill and amendments);
3. Library of Congress of Chile (bill and debates);
4. Senate of Italy (bill publication in open data);
5. Parliament of Uruguay (bill workflow);
6. US Code Consolidation service (code management);
7. State of California (xml standard for document management);
8. Hong Kong City State (xml standard for document management);
10. Federal Chancellery of Switzerland (publication in gazette);
11. High Court of Cassation of Italy (xml standard for document management);
12. Chamber of Deputies of Italy (ongoing).

Several other institutions are now evaluating to use Akoma Ntoso standard especially in the European Forum Gazette.

However not all these standard are oriented to preserve the legal principles, the semantic web design and the pattern oriented approach. Four classes can be identified as follows:

- The first-generation legal-document XML standard\(^{10}\) was mainly aimed at describing legal text and its structure with a similar approach to the database model or the typography-based word-processing models.
- The second generation\(^{11}\) paid more attention to modeling and describing texts, their structure, and metadata. Yet, the elements were described without prior abstract analysis of the data classes, and the result was a long list of tags, with complex inclusions of DTDs or XML schemas and much overlapping between metadata and text definition, along with weak instruments for linking text to other layers.
- The third generation\(^{12}\) is based on patterns. Patterns define the properties of a class and its grammar—providing a content model and specifying the class’s behavior and hierarchy in relation to other classes—in such a way that any additional tag will belong to an existing abstract class, which allows consistency to be maintained over time. A basic principle is that we must be very careful to divide the text, its structure, its metadata, and the ontology, so that we can reliably track any new layer placed on top of the bare text. A pattern ensures clarity of design by defining general rules that no longer impose real constraints in markup, but on the downside, this means that the standard lacks prescriptive constraints.
- The fourth generation (examples are RELEX NG, Schematron, and DSD) uses the pattern jointly with co-constraint grammars to resolve the aforementioned problem of lack of prescriptiveness.

Akoma Ntoso belongs to the third generation and is in the process of becoming a fourth generation Legal XML standard. For this reason we take it as an example of success in the introduction of Legislative XML markup.

\(^{10}\) Such as EnAct or Formex.
\(^{11}\) Such as NiR or LexDania.
\(^{12}\) Such as CEN/Metalex and Akoma Ntoso.
## Legal XML Standards

There are different legal XML standards nowadays used in the different Parliaments, Official Gazette bodies, entities, public administrations. For an exhaustive description see:


This is an updated table of comparison in order to understand the main differences:

<table>
<thead>
<tr>
<th>Date</th>
<th>Legal XML Schema</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>EnAct - Tasmania, Australia, New Zealand e Canada</td>
<td>Dedicated to the legislative documents form common law tradition (especially from Commonwealth) It includes point-in-time metadata. It is an evolution of an original schema provided in SGML. For this reason there isn't distinction between the content representation and the metadata part. It is mostly dedicated to model the presentation of the legal document on the web without a distinction between rendering and content. First generation</td>
</tr>
<tr>
<td>1999</td>
<td>FORMEX data model - EUR-LEX</td>
<td>Dedicated to the EU legislative documents. It includes point-in-time metadata for managing the consolidation of the directives and regulations. It is an evolution of an original schema provided in SGML. For this reason there isn't distinction between the content representation and the metadata part. It is mostly dedicated to model the presentation of the legal document on the web without a distinction between rendering and content. The schema manages also the multilingual features. First generation</td>
</tr>
<tr>
<td>2001</td>
<td>NormeinRete – Italy</td>
<td>Dedicated to the Italian legislative documents. It includes point-in-time metadata for managing the consolidation. There are</td>
</tr>
<tr>
<td>Year</td>
<td>Description</td>
<td>Details</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>2001</td>
<td>House of Representatives DTD</td>
<td>Dedicated to the US parliamentary documents (bill, act, committee reports, records, etc.). It is mostly dedicated to model the typographic elements for managing the publications. Recently they have added intensively metadata, but not distinguished by the content of the legal text. The point in time is not managed. The schema doesn't manage the multilingual features.</td>
</tr>
<tr>
<td>2002</td>
<td>MetaLex and SDU BWB - NL</td>
<td>Dedicated to the Netherland legislative documents. It is mostly dedicated to model the structure of the legal document. The point in time is managed but only for the Netherland legal tradition (no retroactive modifications). The schema doesn’t manage the multilingual features.</td>
</tr>
<tr>
<td>2003</td>
<td>LexDania – Denmark</td>
<td>Dedicated to the Danish legislative documents. It includes point-in-time metadata for managing the consolidation. There are dedicated tags for modelling metadata, even if they are mixed inside of the content. There is a clear distinction between the presentation and the content layers. There are 53 XML-schemas for managing the documents, oriented to the local national legal drafting tools.</td>
</tr>
<tr>
<td>Year</td>
<td>Organization</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>2004</td>
<td>eLaw - Austria</td>
<td>Dedicated to the Austrian legislative documents. It includes point-in-time metadata for managing the consolidation. There are dedicated tags for modelling metadata, even if they are mixed inside of the content. There is a clear distinction between the presentation and the content layers. However the XML schema is too much oriented to the local national legal drafting tools and publication purpose. The schema doesn’t manage the multilingual features.</td>
</tr>
<tr>
<td>2006</td>
<td>CHLexML - Switzerland</td>
<td>Dedicated to the Switzerland legislative documents. It is mostly dedicated to model the structure of the legal document for the presentation on the Web and for typographical reasons. The schema manages the multilingual features and the basic elements for the point-in-time.</td>
</tr>
<tr>
<td>2008</td>
<td>AKOMA NTOSO - United Nations for Pan-African Parliaments</td>
<td>This standard manages with a unique XSD schema all the parliamentary documents, judiciary documents, legislation documents, legal document. It clearly separates metadata to the content, presentation to the content, rendering to the official text. It is based on patterns. It is not designed for a particular legal system or legal drafting tool. It is able to cover different scenarios, in different languages,</td>
</tr>
<tr>
<td>Year</td>
<td>Description</td>
<td>Details</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>2009</td>
<td>Crown XML Schema for Legislation - United Kingdom</td>
<td>It is the UK XML standard for legislative documents. It is oriented to the common law system. It manages the metadata in a separate way from the content. Some information are clearly oriented to the publication but in meantime with particular attention to the semantic web. They use a sophisticated mechanism for tracking the jurisdiction of the different kingdoms. It manages the lifecycle of the document and in some part also the normative qualification (e.g. saving clauses). It is not based on pattern. Second generation.</td>
</tr>
</tbody>
</table>

Akoma Ntoso is the only one standard at third generation, based on pattern, that provides one single XSD schema for managing different type of legal documents from legislative act to case law.

Akoma Ntoso is an XML schema for modelling parliamentary, legislative, and judiciary documents. The AKOMA NTOSO XML schema makes the structural and semantic components of digital legislative documents fully accessible to machine-driven processes, thereby supporting the creation of high-quality legislative information services and greatly improving efficiency and accountability in parliamentary, legislative, and judicial contexts.

The unique XSD is able to mark-up different documents:
- acts, bills, constitutions, regulations, decrees, treaties, directives, ordinances, etc.;
- debate reports (e.g. Hansard), assembly transcripts, etc.;
- documents, annexes, committee reports;
- judgments;
- amendments, lists of amendments;
- gazettes, codes, books, etc.
3 EUCases Needs & Requirements

The primary goal of EUCases is to collect multiple legal resources coming from different open data portals or public web sites with the aim of providing a harmonized service - based on semantic web techniques - to help legal experts retrieve relevant legal information pertaining to particular thematic sectors.

The importance of having a standard legal database is to be able to take the data and bring them into a single scheme to provide the independence from domain, language, format etc.

Akoma Ntoso results to be the right choice for EUCases since it is becoming the most widely used in international scenarios, both in institutional and private environments.

For example, currently, Akoma Ntoso schemas are implemented in different scenarios:

- EU Parliament uses Akoma Ntoso for modelling amendments, amendment list, bills, proposals, consolidated version of those documents in 24 languages;
- LexML Brazil Project of the Senate. The project is using a customization of Akoma Ntoso for the document management (and URN:LEX as naming convention);
- Senate of Italy uses Akoma Ntoso for publishing bills in open data;
- Library Congress of Chile uses Akoma Ntoso for managing debates and recently also bills and acts;
- State of California uses Akoma Ntoso as back-office document standard;

There are several projects "under construction":

- Uruguay Parliament will use Akoma Ntoso for modelling all the law-making process of the bills;
- Federal Chancellery of Switzerland will use Akoma Ntoso for the publication of bills, acts, consolidated code in the Official Journals;
- European Commission aims to adopt Akoma Ntoso for the document management in order to favour the interoperability with EU Parliament;
- Nicaragua Assembly will use Akoma Ntoso for managing all the life-cycle of the law-making system;
- Still, Akoma Ntoso seems to fully fit the role of "standard" for its features and growing usage worldwide. Then, it has the ability to be applied to legislation and case law, and to facilitate formal ontological citations (URI, by means of EL/ECLI).

Another important fact is that it is maintained and so always up-to-date with respect to recent issues, and the evolution of legal document analysis. Indeed, since its born in 2006, Akoma Ntoso is at its third version. The governance now is under OASIS standardization body. This guarantees a stable and robust management process over time.

Finally, it has the great advantage to be carved out on specific situations through the use of a tool that is able to generate sub-schemes according to some expressivity features in real-time.

The main functionalities of EUCases of relevance to the document standard are:

- the use of one XML document standard as a harmonized form for a variety of documents representation. In fact, one of the requirements for
EUCases is to manage the complexity of the forms in which legal information are expressed in different European scenarios, both from the structural point of view (how do they relate) and from the technical point of view (txt, pdf, html). In the choice of the XML standard we searched for a standard able to manage the information in any legal and regulatory environment touched by EUCases.

- **the use of an XML document standard to capture the semantics of relevant parts of the document.** The characteristic of any XML standard is that of conferring "semantics", as well as the structure, to a document. This allows one to process with higher accuracy the information found in pieces of texts, also from the point of view of meaning. This aspect is fundamental to EUCases, as it aims to link together semantically convergent legal information belonging to different sources, conceptual plans, and language. This is one feature of Akoma Ntoso.

- **the use of an XML document standard to include metadata, semantic annotation and content inside the same document.** EUCases will feature the ability to uniformly manage large amounts of legal information, regardless of the scenarios and contexts of origin. So it needed an XML standard that guarantees interoperability between information and ontological annotations. Characteristics that we found in Akoma Ntoso.

- **the semantic annotation and classification of text to help query EUCases.** Ontological annotations and metadata are critical to facilitate the construction of a system capable to manage the complexity of the language, and making it manageable disambiguating the classification of texts, concepts, and sources. Consequently, it was deemed important to highlight this aspect in the choice of Akoma Ntoso.

- **multilingual features.** EUCases born in a multilingual context. As a result, the XML standard must allow the harmonious management of the different language versions of the same text, as Akoma Ntoso does.

- **links to ontologies, thesauri and vocabularies of classification.** The added value of EUCases will be the possibility to semantically search for different types of information, namely legislation, court decisions and legal concepts. Therefore, we identified the XML standard to provide the best management of these aspects. Akoma Ntoso, in particular, presents the ideal tools for these needs.

### 3.1 Akoma Ntoso

Akoma Ntoso is an XML schema for modelling parliamentary, legislative, and judiciary documents. The AKOMA NTOSO XML schema makes the structural and semantic components of digital legislative documents fully accessible to machine-driven processes, thereby supporting the creation of high-quality legislative information services and greatly improving efficiency and accountability in parliamentary, legislative, and judicial contexts.

The unique XSD is able to mark-up different documents:

- acts, bills, constitutions, regulations, decrees, treaties, directives, ordinances, etc.;
- debate reports (e.g. Hansard), assembly transcripts, etc.;
- documents, annexes, committee reports;
- judgments;
• amendments, lists of amendments;
• gazettes, codes, books, etc.

3.2 Multi-layer Legal Document Modelling

Over the last ten years, the e-law scientific community has invested much effort in modelling and representing legal resources using XML standards (Metalex/CEN for legislative documents, NormInRete for laws, Akoma Ntoso for parliamentary activities and judiciary documents, LegalXML OASIS initiative for electronic court filing, and the Australian Judgment XML standard). Some projects have even used XML markup to structure precedents and different interpretations. But in order to capture and represent legal knowledge embedded in case-law – including judicial reasoning – a supplementary semantic layer is required based on the Tim Berners-Lee semantic cake, and ontologies can make a significant contribution to this layer. Our project is focused on the representation of the contents of judicial decisions and on their connection to the text, based on a shared ontology to capture the semantics of legal concepts.

The approach is based on a multi-layer paradigm, where the legal resource is managed at separate levels, linked to one other but organized in a way that allows redundancy of representation with multiple annotations, multiple interpretations, and multiple ontologies. The syntactical approach is based on the following schema:

• Text annotation in XML: using the Akoma Ntoso standard we are able to properly mark-up the structure of judgements and citations;
• Metadata annotation: using the Akoma Ntoso metadata block we capture not only metadata concerning the lifecycle of the document (e.g. workflow of the trial, formalities, jurisdiction, level of judgments), but also different parts of the decision, such as the minority report or dissenting opinion;
• Ontology annotation: this is achieved using external OWL definitions and a special mechanism to link to the XML text;
• Rules: even with the functionalities of version 2.0, OWL is unable to represent complex and defeasible legal arguments. It is therefore necessary to extend the model with rule modelling, using argumentation theory.

3.3 Akoma Ntoso Design Principles

We believe that the Akoma Ntoso design principles form a good basis for any technical format used to create, manipulate and archive legal documents:

LEGAL AND LEGISLATIVE ORIENTATION. It is important to provide a representation of the main structures of legal documents using a principled approach that provides the best combination of technological competence and legal knowledge. The XML schema has to capture the relevant legal metadata of a single act (law, decree, etc.), along with metadata for the whole collection of documents per se (e.g. Official Gazette n. 70, 26 March 2011 is a separate work). Secondly, legislative documents express and embed cultural values, legal principles, a territory’s sovereignty, historical traditions as well as the intellectual preoccupations of the political class. It is incumbent upon us to represent all this knowledge in the best possible way to preserve the legal values for the next generation. For instance, written form is required ad substantiam or ad probationem.

DESCRIPTIVENESS. The document’s original descriptions (e.g. terms used to identify hierarchical elements) must be preserved and excessive generalization of elements' names
must be avoided. The names of the mark-up elements must be as self-explanatory as possible (e.g., while article elements clearly represent the basic normative unit; an element such as block is too general and ambiguous). Descriptiveness is useful for evaluating whether tags have been applied correctly and in conformance with agreed rules or guidelines. It also allows future users to understand the mark-up applied to a document even if the documentation or the document schema is lost.

**PRESCRIPTIVENESS.** The technical rules should be based on rules and constraints taken directly from the legal domain. These rules can be used to increase the quality of available legal information, for example translating legal-drafting rules into local technical rules. The use of XML schemas provides a guide for authors, editors and documentalists, allowing them to achieve better harmonization over time. However, it is also important not to impose formats that are too prescriptive and difficult to extend and adapt to real-life scenarios.

**SELF-CONTAINMENT.** Each document should have at least one place to store all additional information needed to access, use, and understand the content and the metadata of the document itself. This reduces the need to access external documents or resources (e.g. ontological specifications or traditional databases) to “understand” a document. Self-containment means treating access to external resources as an optional step rather than a fundamental mechanism. Making document collections independent of architectural choices and technological evolutions is a key aspect in achieving long-term preservation.

**STRUCTURE MODELLING.** The Legal XML standards fully describe the document’s original structure when expressed in XML. Proper attention is given to the textual content and the associated metadata associated, so as to respect the author’s subdivision of the document (e.g. Parliament). The hierarchical order of the parts of the documents is respected, as it expresses a precise legal message of priority.

**SEPARATION OF LEVELS.** It is important to identify with precision all the layers of information contained inside a document. Metadata must never be confused with content. Akoma Ntoso features a layered structure that splits the legal content in various layers of information, from the most objective (the text approved by an authority) to the more subjective (e.g. scholarly annotations added by a publisher). This strict separation allows different level of trust to be assigned to each level, enabling information to be added on top of other information without tainting the most trusted pieces of information.

**NAMING POLICY.** URIs, combining both instant usability and future extendibility, should be used wherever there is a need to address resources or concepts. This makes it possible to correctly and specifically support a large class of references, including static legal references, dynamic legal references, data-level object inclusion, and ontological references.

**TEMPORAL MODEL.** It is essential to have a correct temporal model to represent a legal document’s dynamicity over time. The temporal model should allow the representation of versions, variants and documents containing a plurality of versions, as well as static and dynamic references.

**WORKFLOW MODEL.** It is possible to express workflow events and how these events have influenced the document. Workflow events capture the procedural steps required to maintain the chain of validity of the legal document. Sometimes the workflow event modifies the text of the document, and in such cases there is a strict relationship with the lifecycle events. Sometimes the workflow step is only a formal phase in the legal process (e.g. voting), but is nevertheless important for the enforceability of the legal document. It is possible to link these events to concepts and instances of the legal domain ontologies used in a particular environment.

**INTERCHANGE DOCUMENT FORMAT.** The Legal XML standards provide a data-interchange format between different standards, legacy systems, application layers, and data formats. It is both an interchange and interoperability standard compliant with CEN Metalex specifications.
**Homogeneous Publishing Format.** Legal documents should contain enough information to allow publishing systems to operate without any human intervention. Formats that do not allow enough details to be expressed in a native way, result in a longer, more complex and less automatable publishing processes. It should also be easy to identify the kind of document and the legal system in which a document has been produced, to further simplify the publishing phase.

**Neutral and Open Document Format.** It is paramount for long-preservation that any format adopted for legal documents is public, open and well documented. The adoption of any closed or proprietary format would hinder the use of the produced documents as part of any public data-sharing initiative such as the Open Gov Initiative.

**Reliance on Existing Standards.** Instead of introducing new ad-hoc technologies, a good format should rely on current and widespread standards such as XML, URIs, XML namespaces, XML schema languages (e.g. XML Schema, Relax NG, Schematron). It is also important to offer compatibility with emerging formats like RDF and OWL.

**Metadata Modelling.** It must be possible to associate documents with a rich set of metadata, allowing complex information to be expressed. It should be possible to provide, where available, rich semantic annotations, so that marked-up documents can be used for advanced reasoning tasks such as comparing abstracts.

**Minimal Required Metadata Set Format for Queries.** A minimal set of metadata should be guaranteed to allow interoperable queries between heterogeneous collections of legal documents expressed in different local and national standards. Initiatives such as IPEX, N-Lex or Eur-Lex could take advantage of the definition of this minimal set of metadata to manage, at least at a basic level, all the legal documents coming from different countries with a single data software architecture without the need to converge the actual data formats.

**Pattern Design Oriented.** The Legal XML standards are designed to respect classes of patterns, thereby allows future extensions to be carried out in a simple and backward-compatible way. Pattern-based designs define a small number of element classes. Each class interacts with another in a very precise way: for example, elements in the block class can include elements of the inline class but cannot include other block elements; paragraphs (block) can include links (inline) or stylistic variations like bold (inline) but cannot include other paragraphs (block). When a new element is introduced for a concept not previously covered, it is placed in one of these classes. The mechanism provides a strong benchmark to understand whether this new concept has been carefully and correctly introduced in the XML schema. This helps keep the Akoma Ntoso design simple and easily extensible.

To summarize, Akoma Ntoso is a third generation XML vocabulary for legal documents whose design is driven by current good practices for XML formats in addition to the particular requirements for legal documents.

---

13 The Open Government Directive has been issued by the Executive Office of the President of the United States on December 8, 2009.
14 http://www.ipex.eu/ipex/
15 http://eur-lex.europa.eu/n-lex/
16 http://eur-lex.europa.eu/
4 Akoma Ntoso Document Structure

4.1 Main structure

Akoma Ntoso is composed of 308 elements and defines a regular structure for any type of document managed with this legal XML standard:

Fig. 1 – Akoma Ntoso main skeleton

**Metadata**

<table>
<thead>
<tr>
<th>Preliminary pages information</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first part provides information about the type of document. We typically find the title, number, the issuing authority, and the document’s legal basis.</td>
</tr>
<tr>
<td>We also include a cover page that with a specific layout for providing information about the relevant authority (e.g. the cover page of the EU Parliament includes the EU Parliament logo, and this logo is mandatory in any official EU Parliament document.)</td>
</tr>
<tr>
<td>This part also includes introductory text (e.g. preamble, citations, recitals).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Main text</th>
</tr>
</thead>
<tbody>
<tr>
<td>The main legal part of the document, where the content is represented, can have different structures depending on the type of document and its purposes.</td>
</tr>
</tbody>
</table>
The only mandatory parts in Akoma Ntoso are the identification metadata and the main text block (body):

Fig. 2 – Act schema

The main skeleton of Akoma Ntoso is composed of:

1. metadata
2. cover page
3. introductory part
4. preamble
5. hierarchy of the body
6. end part of the document
7. annexes
8. normative references inside the text
9. other metadata in the text (e.g. names of persons, locations, data, quantities, etc.).

4.2 Example of Act Analysis

We will now analyse an act from the UK:
High Speed Rail (Preparation) Act 2013

CHAPTER 51

Explanatory Notes have been produced to assist in the understanding of this Act and are available separately.
3. introductory part

High Speed Rail (Preparation) Act 2013
CHAPTER 31
CONTENTS
1 Preparatory expenditure
2 Financial reports
3 Short title, commencement and short title

High Speed Rail (Preparation) Act 2013

An Act to make provision authorising expenditure in preparation for a high speed railway transport network. [21st November 2013]

4. preamble

B E IT ENACTED by the Queen's most Excellent Majesty, by and with the advice and consent of the Lords Spiritual and Temporal, and Commons, in this present Parliament assembled, and by the authority of the same, as follows:—
5. hierarchy of the body

Section 2

2. Financial reports

1. The Secretary of State must prepare a report on expenditure under section 1 in relation to each financial year.

2. Each report must contain details of—
   (a) expenditure incurred under section 1 during the financial year to which the report relates (with capital and revenue expenditure specified separately in respect of each head of expenditure referred to in section 1(4)(a) to (c));
   (b) the extent to which expenditure incurred under section 1 during that year represents an overspend or underspend as against the budget for such expenditure for the year;
   (c) the likely effect of any such overspend or underspend on a total budget of £50.1 billion in 2011 prices (which includes construction and the cost of rolling stock);
   (d) total expenditure incurred under section 1 up to the end of that year;
   (e) revenue or assets received in that year in connection with expenditure incurred under section 1.

3. Each report must also contain an account of the vocational qualifications gained during the financial year by individuals employed by persons appointed under an enactment to carry out activities in connection with preparing for, and constructing, the network referred to in section 1(1).

4. In this section, “financial year” mean—
   (a) the period beginning with the day on which this Act is passed and ending with 31 March 2015; and
   (b) each subsequent period of 12 months.

5. The Secretary of State must lay each report under this section before Parliament as soon as is reasonably practicable after the end of the financial year to which it relates.

6. No report is required in relation to a financial year in which there is nothing to record under subsection (5)(a) and (e).

6. end part of the document

None

7. annexes

None

8. normative references inside of the text

separately in respect of each head of expenditure referred to in section 1(4)(a) to (c));

(c) the likely effect of any such overspend or underspend on a total budget of £50.1 billion in 2011 prices (which includes construction and the cost of rolling stock).

(a) involves the construction of railway lines connecting at least—
   London,
   Birmingham,
   the East Midlands,
   Sheffield,
   Leeds, and
   Manchester, and

(1) This Act extends to England and Wales and to Scotland.

(3) This Act may be cited as the High Speed Rail (Preparation) Act

9. other metadata in the text (e.g. name of persons, locations, data, quantities, etc.).

quantity (50.1 billion)
4.3 meta Block

The metadata block is divided into 11 other sub-blocks:

![Metadata schema diagram]

Fig. 3 – Metadata schema
<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>identification</td>
<td>Identification of the legal document following the FRBR method. There are specific metadata for identifying the language of the document and translation from official legal document. This is the only mandatory part.</td>
</tr>
<tr>
<td>publication</td>
<td>Publication metadata</td>
</tr>
<tr>
<td>classification</td>
<td>Classification metadata</td>
</tr>
<tr>
<td>lifecycle</td>
<td>List of events that have affected the document</td>
</tr>
<tr>
<td>workflow</td>
<td>List of workflow events where the document is involved</td>
</tr>
<tr>
<td>analysis</td>
<td>Legal analysis of norms or of the text. This part is highly valuable but resource-intensive, as it requires the interpretation of a legal expert. For legislative acts. It includes all modifications proposed or included in the document. For debates (e.g. Hansard), it includes the quorum and the voting result. For judgments, it includes the classification of each case-law referred to in the decision.</td>
</tr>
<tr>
<td>temporalData</td>
<td>Temporal metadata</td>
</tr>
<tr>
<td>references</td>
<td>References to other documents, entities, persons, organizations, roles, status, concepts, etc. used inside the document.</td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:choice&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCPerson&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCOrganization&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCConcept&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCOBJECT&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCEVENT&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCLOCATION&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCPROCESS&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCROLE&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCTERM&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;xsd:element ref=&quot;TLCREFERENCE&quot;/&gt;</code></td>
</tr>
<tr>
<td></td>
<td><code>&lt;/xsd:choice&gt;</code></td>
</tr>
<tr>
<td>notes</td>
<td>Notes from agents different to the author of the document. e.g. consolidation, interpretation, or library notes.</td>
</tr>
<tr>
<td>proprietary</td>
<td>Block dedicated to proprietary tags</td>
</tr>
<tr>
<td>presentation</td>
<td>Block dedicated to the presentation specification of a particular layout</td>
</tr>
</tbody>
</table>
5 Acts in Akoma Ntoso

Here is an example of the previous UK act marked up in Akoma Ntoso:

```xml
<?xml version="1.0" encoding="UTF-8"?>
  <!-- http://www.legislation.gov.uk/ukpga/2013/31/contents -->

  <act name="act">
    <meta>
      <FRBRWork>
        <FRBRthis value="/uk/act/2013-11-21/2013/31/main"/>
        <FRBRuri value="/uk/act/2013-11-21/2013/31"/>
        <FRBRdate date="2013-11-21" name="Enactment"/>
        <FRBRauthor href="#ukParliament" as="#author"/>
        <FRBRcountry value="ul"/>
      </FRBRWork>
      <FRBRExpression>
        <FRBRthis value="/uk/doc/2013-11-21/2013/31/eng@/main"/>
        <FRBRuri value="/uk/doc/2013-11-21/2013/31/eng@"/>
        <FRBRdate date="2013-11-21" name="Expression"/>
        <FRBRauthor href="#ukParliament" as="#editor"/>
        <FRBRlanguage language="eng"/>
      </FRBRExpression>
      <FRBRManifestation>
        <FRBRthis value="/uk/doc/2013-11-21/2013/31/eng@/main.xml"/>
        <FRBRuri value="/uk/doc/2013-11-21/2013/31/eng@/.akn"/>
        <FRBRdate date="2014-01-08" name="XMLConversion"/>
        <FRBRauthor href="#palmirani" as="#editor"/>
      </FRBRManifestation>
    </meta>

    <publication date="2013-11-21" name="Official Gazette" showAs="Official Gazette" number="999"/>

    <classification source="#ukParliament">
      <keyword dictionary="#dictionary" showAs="c. 31" value="2013 CHAPTER 31"/>
    </classification>
  </act>
</akomaNtoso>
```
Lifecycle of the events

<analysis source="#palmirani"/>

<restrictions source="#sec_3__subsec_1__cnt_1">
  <restriction type="jurisdiction" eEId="rest1" href="#sec_1"/>
  <restriction type="jurisdiction" eEId="rest2" href="#sec_2"/>
  <restriction type="jurisdiction" eEId="rest3" href="#sec_3"/>
</restrictions>

Analysis of the jurisdiction limited to England, Wales and Scotland

<references source="#palmirani"/>

<original eEId="ro1" href="/uk/doc/2013-11-21/2013/31/eng@" showAs=""/>

<TLCOrganization eEId="ukParliament" href="/ontology/organizations/uk/ukParliament" showAs="ukParliament"/>

<TLCRole eEId="author" href="/ontology/roles/uk/author" showAs="Author of Document"/>

<TLCRole eEId="editor" href="/ontology/roles/uk/editor" showAs="Editor of Document"/>

<TLCConcept href="/ontology/concept/uk/ews" showAs="England and Wales"/>

<TLCLocation href="/ontology/location/uk/london" showAs="London"/>

<TLCLocation href="/ontology/location/uk/birmingham" showAs="Birmingham"/>

<TLCLocation href="/ontology/location/uk/eastMidlands" showAs="the East Midlands"/>

<TLCLocation href="/ontology/location/uk/sheffield" showAs="Sheffield"/>

<TLCLocation href="/ontology/location/uk/leeds" showAs="Leeds"/>

<TLCLocation href="/ontology/location/uk/manchester" showAs="Manchester"/>

<TLCConcept href="/ontology/concept/uk/c. 31" showAs="c. 31"/>

<TLCConcept href="/ontology/concept/uk/price" showAs="Price"/>

<TLCConcept href="/ontology/concept/uk/price2011" showAs="Price"/>

Reference to the external ontology used to refer to entities used in the XML file
ELIZABETH II

High Speed Rail (Preparation) Act 2013

2013 CHAPTER 31

Explanatory Notes have been produced to assist in the understanding of this Act and are available separately.

High Speed Rail (Preparation) Act 2013

2013 CHAPTER 31
An Act to make provision authorising expenditure in preparation for a high speed railway transport network.

BE IT ENACTED by the Queen's most Excellent Majesty, by and with the advice and consent of the Lords Spiritual and Temporal, and Commons, in this present Parliament assembled, and by the authority of the same, as follows:—

The Secretary of State may, with the approval of the Treasury, incur expenditure in preparation for a high speed railway transport network.

The Secretary of State may, with the approval of the Treasury, incur expenditure in preparation for a high speed railway transport network.

The Secretary of State may, with the approval of the Treasury, incur expenditure in preparation for a high speed railway transport network.
The network referred to in subsection (1) is a network which—

(a) involves the construction of railway lines connecting at least—

<table>
<thead>
<tr>
<th>Location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>London</td>
<td>#london</td>
</tr>
<tr>
<td>Birmingham</td>
<td>#birmingham</td>
</tr>
<tr>
<td>the East Midlands</td>
<td>#eastMidlands</td>
</tr>
<tr>
<td>Sheffield</td>
<td>#sheffield</td>
</tr>
<tr>
<td>Leeds</td>
<td>#leeds</td>
</tr>
<tr>
<td>Manchester</td>
<td>#manchester</td>
</tr>
</tbody>
</table>
connects with the existing railway transport network as well as with such other parts of the transport network (including roads, footpaths, cycleways, airports and light railways) as the Secretary of State considers appropriate.

The expenditure which may be incurred under subsection (1) includes expenditure on—

preparation for the construction of any railway line and any other infrastructure proposed to be included at any time in the network referred to in subsection (1), and preparation for the provision of services as part of that network.
The expenditure mentioned in subsection (3) includes expenditure incurred—

(a) on pre-construction activity (such as surveying and design),

(b) in acquiring property, and

(c) in providing compensation in respect of property likely to be affected.

In this Act, references to incurring expenditure include incurring liabilities.

Expenditure incurred under this section is to be defrayed out of money
The Secretary of State must prepare a report on expenditure under section 1 in relation to each financial year. Each report must contain details of—

- expenditure incurred under section 1 during the financial year to which the report relates (with capital and resource expenditure specified separately in respect of each head of expenditure referred to in section 1(4)(a) to (c));
- the extent to which expenditure incurred under section 1 during that year represents an overspend or underspend as against the budget for such expenditure for the year;
(c) the likely effect of any such overspend or underspend on a total budget of £50.1 billion in 2011 prices (which includes construction and the cost of rolling stock);

(d) total expenditure incurred under section 1 up to the end of that year;

(e) sums or assets received in that year in connection with expenditure incurred under section 1.
This Act comes into force on the day on which it is passed.

This Act may be cited as the High Speed Rail (Preparation) Act 2013.

This Act extends to England and Wales and to Scotland.

The Secretary of State, with the approval of the Treasury, may, with the approval of the Secretary of State, incur expenditure in preparation for a high speed railway transport network.

In this section, "financial year" means—
6 Judgments in Akoma Ntoso

The judgment in Akoma Ntoso has a particular structure: the header captures the main information concerning the parties, Coram, neutral citation, document numbers and identification information; the body represents the main part of the judgment including the decision; the conclusion captures the signatures.

6.1 Main element of the header

The header is the part of the judgment where all the main data concerning the administrative information of the case-law should be recorded: parties, Coram composition, number of the case-law, hearing dates, etc.

6.2 Structure of main part of the judgments

The body is divided into four main blocks:

- the introduction, where usually (especially in common law decisions) the history of the trial is introduced;
- the background, describing the facts of the case;
- the motivation, where the judge presents arguments supporting his/her decision;
- the decision, where the final outcome is stated by the judge.

This division is essential for detecting facts and factors from the background: in the motivation we detect arguments and counterarguments and in the decision the conclusion of the legal argumentation process. Such text should be annotated by legal experts with the help of a special editor tool (e.g. Norma-Editor) that allows links to be created from the text to the metadata and the ontology classes.
6.3 Metadata of judgments

The metadata of judgments is divided into four main blocks:

**Descriptive metadata**: that objectively track judgment metadata such as date of publication, the ID number of the case, natural citation, names of judges, the relevant jurisdiction, level of the judgment, nature of the case, type of court, parties, lawyers, and so on.

**Classification metadata**: include the subject-matter of the case (using a thesaurus), together with the reportable or not reportable case-base. This metadata provides a way to filter reportable judgments, following the common law tradition that highlights cases that produce a new rule of law.

**Lifecycle**: history of the evolution of the document, with versioning.

**Workflow metadata**: tracking each step of the document's production process. Each actor in the workflow chain can annotate the document with his/her specific metadata, since multi-annotation of the same fragment of text is allowed.

**References**: all documents cited in the judgment and links to all other logically connected documents.

**Semantic annotation**: the classification of the text from a legal point of view, especially in the motivation part.

**Ontology**: definition of mechanisms for linking fragments of text with macro classes such as People, Organization, Role, Action, Event, Term, Location, etc.

It is possible to annotate very specific knowledge using the above metadata. In the following fragment of text, we capture the role of each person involved in the trial: Mr. Du Plessis is a lawyer, with the role of advocate for the appellant, instructed by the Kruger Inc. Annotating this information with XML allows complex queries to be carried out such as: "select all judgments where Du Plessis plays the role of instructor for the appellant on behalf of a third Inc. company".
6.4 Citations in judgments

Each judgment citation could be qualified using Shepard's method for representing which references are in favour of the current judgment argumentation. The list of qualifications include: support, meaning that the cited judgment supports the current decision; isAnalogTo, meaning that the current case-law is analogous with a cited precedent; distinguished, meaning that the current precedent is distinguished from the cited case-law. Of particular importance is overrules, detecting the case-law cited by the judgment's motivation whose rule of law the judge intends to overrule. This qualification mechanism helps to reinforce the main arguments used by the judge to provide evidence and parameters (e.g. list of the cited case-law with the role played in the argumentation).
<xsd:choice maxOccurs="unbounded">
   <xsd:element ref="supports"/>
   <xsd:element ref="isAnalogTo"/>
   <xsd:element ref="applies"/>
   <xsd:element ref="extends"/>
   <xsd:element ref="restricts"/>
   <xsd:element ref="derogates"/>
   <xsd:element ref="contrasts"/>
   <xsd:element ref="overrules"/>
   <xsd:element ref="dissentsFrom"/>
   <xsd:element ref="putsInQuestion"/>
   <xsd:element ref="distinguishes"/>
</xsd:choice>

6.5 Rule of law, stare decisis and ratio decidendi, obiter dicta

One of the main principles in common law judgments is to define a rule of law fixing the pattern for the similar future cases. This monotonic mechanism is called stare decisis and is meant to guarantee equal application of justice in comparable cases. Stare decisis is applied only to a particular and relevant part of the decision called ratio decidendi, excluding accessory arguments called obiter dicta. Research conducted on the ontology framework (see section 4) shows the importance of marking-up relevant and meaningful parts of the text. By marking up the text of the ratio decidendi and using this information in combination with Shepard's qualification method for cited case-law, in richer information is provided to the argumentation engine devoted to the legal reasoning. For these reasons, we will define new metadata in the analysis block of Akoma Ntoso for qualifying the ratio decidendi and the obiter dicta.
## 7 Judgment Analysis Example

### Header
- Name of the court
- Judgment Type
- Seat of court (Pretoria)
- Jurisdiction
- Location of the court
- Case number
- Name of case (Haupt vs. Brewers)
- Subject of the case
- Parties
- Names of Judges (Coram)
- Dates: delivery, hearing, publication, registration, etc.
- Neutral citation: Media Neutral Citation – (e.g. [2006] SCA 39 (RSA))
- Summary/Abstract

### Inline elements for marking up:
- judgmentType
- judgmentTitle
- judgmentNumer
- courtType
- neutralCitation(a sort of URI)
- parties
- coram
- judgmentDates

### Parties
Party is a person or organization
Types of party are:
- plaintiff(s) - the party initiating
proceedings usually by issuing a summons (or writ) in an action

- defendant(s) – the party sued in criminal proceedings or an action usually commenced by issue of summons (or writ)
- accused – the defendant in a criminal case
- applicant(s) – the party initiating proceedings using the application procedure
- respondent(s) – the party sued in an application procedure
- appellant(s) – a party who approaches the higher court challenging a decision of a lower court
- amicus curiae – a friend of the court; the name given to a member of the bar, or other bystander, who advises the court regarding a point of law or fact upon which information is required
- intervening party - a third party who, to protect his/her own interest, interposes and becomes a party to legal proceedings pending between other parties
- witness

**Coram**

Coram is a set of judge: a panel of judges

Judge is a person

**NOTE:** the address and any other information about the parties should be stored outside the XML document. We suggest using a local ontology or a DB.
7.1 Akoma Ntoso Judgment Body

ORDER

Appeal from: High Court, Pretoria (Rabie J sitting as court of first instance)

The appeal is dismissed with costs

JUDGMENT

BRAND JA (Scott, Farlam, Lewis et Jaftha JA concurring)

[1] The appellant (‘Fourway’) is a long distance haulier. The respondent (‘the Agency’) owes its existence to the South African National Roads Agency Limited and National Roads Act 7 of 1998 (‘the Act’). The dispute between them originates from an accident which occurred in the early evening of 28 September 2003 on the N1 national road between Polokwane and Mokopane in the Limpopo province. The two vehicles involved were an articulated truck and a light delivery van. The articulated truck was driven at the time by an employee of Fourway who was acting in the course and scope of his employment.

[2] The articulated truck was on its way from an asbestos mine in Zimbabwe to Durban harbour carrying about 34 tonnes of chrysotile asbestos, destined for export. As a result of the collision, the truck overturned and spilled its cargo onto practically the entire surface of a portion of the national road and its surroundings. Because of the hazardous nature of asbestos powder, the spillage required an extensive cleaning up and decontamination operation.

[3] To facilitate the cleaning up and decontamination process, the traffic authorities closed the section of the national road involved and diverted the traffic in both directions onto an alternative road. This lasted for about 24 hours. The section of the national road which was closed forms part of a toll
Body

The body is composed of four structural elements

- Introduction
- Background
- Motivation
- Decision

Each block is organised hierarchically or in a sequence of blocks.

Each block contains:

- Quoted text/Citations
- Notes

Introduction: the summary of the case
Background: the description of the fact
Motivation: judicial argumentation
Decision: judges’ decision and final order

Conclusions

- Signatures
- Date
- Place

Voting part

- majority opinion
- dissenting opinion
- plurality opinion

F R MALAN
Acting Judge of Appeal
• concurring opinion
• memorandum opinion

CONCUR:
HOWIE P
NAVSA JA
NUSENT JA
COMBRINCK AJA

7.2 Akoma Ntoso Judgment Example
LIFECYCLE OF THE DOCUMENT

WORKFLOW STEPS

ANALYSIS OF THE CASE-LAW CITATIONS

REFERENCES TO CONCEPTS and DOCUMENTS
Sections starting at [16] and ending at [35] were deleted in order to simplify the example.
## EVENTS

### DATE OF THE DOC

HEARD: 5 November 2008

DELIVERED: 26 November 2008

CORRECTED:

### SUMMARY

**SUMMARY:** Delict – pure economic loss – meaning of policy considerations relevant in determining remoteness of damage – application of flexible test

**ORDER**

On appeal from: High Court, Pretoria (Rabie J sitting as court of first instance) The appeal is dismissed with costs

**JUDGMENT**

BRAND JA (Scott, Farlam, Lewis et Jafta JJA concurring)

---

**BACKGROUND**

The appellant ('Fourway') is a long distance haulier. The respondent commercial in confidence
### INTRODUCTION

('the Agency') owes its existence to the South African National Roads Agency Limited and National Roads Act 7 of 1998 ('the Act').

The dispute between them originates from an accident which occurred in the early evening of 26 September 2003 on the N1 national road between Polokwane and Mokopane in the Limpopo province. The two vehicles involved were an articulated truck and a light delivery van. The articulated truck was driven at the time by an employee of Fourway who was acting in the course and scope of his employment.

The articulated truck was on its way from an asbestos mine in Zimbabwe to Durban harbour carrying about 34 tonnes of chrysolite asbestos, destined for export. As a result of the collision, the truck overturned and spilled its cargo onto practically the entire surface of a portion of the national road and its surroundings. Because of the hazardous nature of asbestos powder, the spillage required an extensive cleaning-up and decontamination operation.

To facilitate the cleaning-up and decontamination process, the traffic authorities closed the section of the national road involved and diverted the traffic in both directions onto an alternative road. This lasted for about 24 hours. The section of the national road closed forms part of a toll road. The alternative route was not subject to toll. As a result of the closure, two toll plazas – as defined in the Act – could not collect toll fees. Based on these facts, the agency as the entity authorised by s 27 of the commercial in confidence
Act to levy and collect toll fees on toll roads, instituted an
action in delict against Fourway for the damages it allegedly suffered
in the form of loss of toll revenue in an amount of R105 996.67.

At the commencement of the trial, the parties asked the court a quo
(Rabie J) to order a separation of issues. In terms of the separation order, the issues relating to the liability of Fourway were to be decided first, while the Agency's alleged damages stood over for later determination. The preliminary issues were decided in favour of the agency. Hence the court declared Fourway liable for such damages as the Agency may prove in respect of the lost revenue it would have collected at the two toll plazas involved, but for the closure of the road. It also ordered Fourway to pay the costs of the preliminary proceedings.

Fourway's appeal against that judgment is with the leave of the court a quo. Part of the controversy on appeal was brought about by a shift in the focus of the defence advanced by Fourway and the resulting mutation of the issues involved. A convenient starting point for an account of this is the opening address by counsel for the Agency, at the beginning of the trial. With reference to the pleadings, at that stage defined the issues between the parties as...
follows:

- Whether or not the respondent had the necessary authority to collect toll fees on that portion of the toll road which was closed as a result of the collision.
- Whether the collision occurred as a result of the negligence of the driver employed by Fourway.
- Whether the occurrence of the collision necessitated the decontamination operation and the closure of the road.

Counsel for Fourway did not react to this definition of the issues. During the trial, Fourway formally conceded the issue referred to in (a) and the evidence led by the parties therefore dealt exclusively with the issues in (b) and (c).

But in argument at the end of the trial, Fourway's counsel, for the first time, raised two further contentions. First he submitted that the Agency's claim was for the recovery of pure economic loss which required the existence of a legal duty on the part of Fourway and that the Agency had failed to plead or establish the existence...
of such a legal duty. Secondly he submitted that the Agency had failed to establish the requirement of legal causation with reference to the loss which formed the basis of its claim.

As we know from the result, the court a quo dismissed all defences relied on by Fourway, including those originally raised under what I categorised as

- As to \[8\]

On appeal, it was conceded on behalf of Fourway that the court a quo was correct in deciding the issues under \[b\] and against it. In consequence, the only issues on appeal turned on the contentions that were raised...
for the first time in argument at the end of the trial. They can be summarised thus:

1. Whether the court a quo correctly came to the conclusion that the Agency’s claim is not a claim for pure economic loss.

2. If not, how the issue of wrongfulness should have been dealt with in the light of the fact that it was not pertinently raised in the pleadings.

3. Whether the court a quo correctly came to the conclusion that the damages claim by the Agency cannot be regarded as too remote.

The court a quo’s finding that the damages claimed did not result from pure economic loss clearly emanated from its understanding of that concept. That understanding appears from the following statements in the judgment:

‘The economic loss in this sense comprises patrimonial loss that does not result from a direct invasion of a subjective right of the person who commercial in confidence
suffered the loss.'
</span>

And that <eol/>

' the aforesaid rights of the plaintiff...  

[i.e the Agency's statutory rights to operate a toll road and  

fees] were clearly subjective rights worthy of protection  

plaintiff could enforce against other people.'
</span>

And that <eol/>

'[c]onsequently, the loss suffered by the plaintiff is not a  

pure economic loss, but the direct result of a direct  

subjective rights which was as such unlawful.'
</span>

I do not share the court a quo's understanding of what is  

'meant by  

'pure economic loss' in the present context.  

I believe its meaning to be far less metaphysical. As  

explained by Harms JA in  

para 1, it means simply this:
</span>

"Pure economic loss" in this context connotes loss  

that does not arise directly from damage to the plaintiff's  

person or property  

but rather in consequence of the negligent act itself, such  

being put to extra expenses or the diminution in the value
of property.'

(See also <i>Lillicrap, Wassenaar and Partners v Pilkington Brothers (SA) Pty Ltd</i> 1985 (1) SA 475 (A) 497I-498H; Trustees, <i>Two Oceans Aquarium Trust v Kantey &amp; Templer (Pty) Ltd</i> 2006 (3) SA 138 (SCA) para 14; <i>Wille's Principles of South African Law</i> 9 ed, (General editor: Francois du Bois) sv 'Delict' by Daniel Visser, 1105; Neethling, Potgieter &amp; Visser, <i>Law of Delict</i>, 5 ed 268 et seq).

Thus understood, the Agency's claim, in my view, falls squarely within the ambit of pure economic loss. As formulated, its claim was for loss of revenue in the form of toll fees resulting from the closure of the road. The Agency did not allege, nor did it set out to prove in evidence, that it was the owner of the road; that the road was physically damaged by the collision; or that the closure of the road resulted from any physical damage to the road. The Agency's argument on appeal, that in terms of <ref>accessed via languages/1998-03-31/eng@/main.xml#sec27</ref> of the Act it was in fact the owner of the road on which the collision occurred, is of no consequence and misses the point. For present purposes the question is not whether the Agency is in fact the owner of the road. The point is that it did not rely on such ownership to support its claim.

Recognition that we are dealing with a claim for pure economic loss brings in its wake a different approach to the element of wrongfulness. This results from the principles which have been formulated by this court so many
times in the recent past that I believe they can by now be regarded as trite.

These principles proceed from the premise that negligent conduct which manifests itself in the form of a positive act causing physical damage to the property or person of another is prima facie wrongful. By contrast, negligent causation of pure economic loss is not regarded as prima facie wrongful. Its wrongfulness depends on the existence of a legal duty. The imposition of this legal duty is a matter for judicial determination involving criteria of public or legal policy consistent with constitutional norms.

In the result, conduct causing pure economic loss will only be regarded as wrongful and therefore actionable if public or legal policy considerations require that such conduct, if negligent, should attract legal liability for the resulting damages (see eg Minister of Safety and Security v Van Duivenboden 2002 (6) SA 431 (SCA) paras 12 and 22; Gouda Boerdery BK v Transnet 2005 (5) SA 490 (SCA) para 12; Telematrix (supra) paras 13-14; Trustees, Two Oceans Aquarium Trust (supra) paras 10-12).

Note of the dissenting:

In this light, so Fourway contended on appeal, the Agency was obliged to allege in its pleadings not only that the negligent conduct relied upon was wrongful, but that it also had to allege and prove the facts relied upon to substantiate the considerations of policy giving rise to a legal duty on the part of Fourway's employee. As a result of the Agency's failure to adhere to these rules of litigation, so the argument went, neither the policy considerations relevant to the question of wrongfulness, nor the factual basis...
underlying such policy considerations, were identified and investigated during the trial. In consequence, so the argument concluded, it would be prejudiced if the issue of wrongfulness were to be summarily disposed of at the appeal. 

Fourway therefore suggested that, unless this court upholds its contention that the damages claimed are too remote – to which I shall presently return – the issue of wrongfulness should be postponed and decided with the rest of the issues concerning the quantum of the Agency's damages, which are standing over in any event.

The proposition that a plaintiff claiming pure economic loss must allege wrongfulness, and plead the facts relied upon to support that essential allegation, is in principle well founded. In fact, the absence of such allegations may render the particulars of claim expiable on the basis that no cause of action had been disclosed (see eg Trope v SA Reserve Bank 1992 (3) SA 208 (T) at 214A-G; Indac Electronics (Pty) Ltd v Volkskas Bank Ltd 1992 (1) SA 783 (A) 797E; Telematrix (supra) para 2). But, as we know, Fourway did not file an exception. The trial proceeded without any objection on its part. In the circumstances it would be futile to investigate whether an exception, if properly and timeously taken, would have been successful. As I see it, the question is rather whether, despite the lack of necessary opportunity to produce the facts it would seek to rely on for the determination of the policy considerations pertaining to wrongfulness in its favour. Conversely stated, the question is whether Fourway has shown prejudice, in the sense that it would
have conducted its case in a materially different way if the Agency's claim for pure economic loss had been properly pleaded.

(See eg <i>Shill v Milner</i> 1937 AD 101 at 105;
<i>Robinson v Randfontein Estates GM Co Ltd</i> 1925 AD 173 at 198;
<i>Collen v Rietfontein Engineering Works</i> 1948 (1) SA 413 (A) at 433;
<i>Stead v Conradie</i> 1995 (2) SA 111 (A) at 122A-H.)

As I see it, the proposal by Fourway that the issue of wrongfulness be referred back for determination by the trial court therefore depends on the outcome of two discrete enquiries. First, can this court, on the basis of the facts available, decide that, as a matter of policy, Fourway should be held liable for the loss of revenue claimed by the Agency? If not, that would be the end of the matter. The Agency would have failed to make out a case. A decision on the other hand that the issue of wrongfulness should on the facts available be determined in favour of the Agency will lead to the next enquiry. The question is: can it be said that, if the issue of wrongfulness had been properly pleaded by the Agency, Fourway would have conducted its case any differently? If not, the Agency is entitled to succeed. It is therefore only a finding of potential prejudice on the part of Fourway that can justify a referral back to the trial court.
For these reasons the appeal is dismissed with costs, including the costs of two counsel.

JUDGE OF APPEAL

APPEARANCES:
FOR APPELLANT: H DREYER SC
A DU PLESSIS
INSTRUCTED BY: MACGREGOR STANFORD KRUGER INC, PRETORIA
CORRESPONDENTS: E G COOPER ATTORNEYS, BLOEMFONTEIN
FOR RESPONDENT: A C FERREIRA SC
I ELLIS
INSTRUCTED BY: FRIEDLAND HART INC, PRETORIA
CORRESPONDENTS: ISRAEL SACKSTEIN MATSEPE INC, BLOEMFONTEIN
8 Docs in Akoma Ntoso

For marking up a general document like a paper or an extract of the book we can use the document type `<doc>`.

The type doc permits to use all the general elements of Akoma Ntoso that are not linked to any specific semantic:

<table>
<thead>
<tr>
<th>Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hcontainer</td>
<td>The element hcontainer is a generic element for a hierarchical container. It can be placed in a hierarchy instead of any of the other hierarchical containers. The attribute name is required and gives a name to the element.</td>
</tr>
<tr>
<td>container</td>
<td>The element container is a generic element for a container. It includes elements belonging to the block pattern.</td>
</tr>
<tr>
<td>block</td>
<td>The element block is a generic element for a container. It can be placed in a container instead of any of the other blocks. The attribute name is required and gives a name to the element.</td>
</tr>
<tr>
<td>tblock</td>
<td>The element tblock (titled block) is used to specify a container for blocks introduced by heading elements, similarly to a hierarchical structure</td>
</tr>
<tr>
<td>inline</td>
<td>The element inline is a generic element for an</td>
</tr>
<tr>
<td>Element</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>inline</td>
<td>It can be placed inside a block instead of any of the other inlines. The attribute name is required and gives a name to the element.</td>
</tr>
<tr>
<td>marker</td>
<td>The element marker is a generic element for a marker. It can be placed in a block instead of any of the other markers. The attribute name is required and gives a name to the element.</td>
</tr>
</tbody>
</table>

These elements combined with the part `<references>` could be enriched with the proper semantic.
9 EUCases Requirements

The primary goal of EUCases is to collect multiple legal resources coming from different open data portals or public web sites with the aim of providing a harmonized service - based on semantic web techniques - to help legal experts retrieve relevant legal information pertaining to particular thematic sectors.

The main functionalities of EUCases of relevance to the document standard are many. First, sing one XML document standard as a harmonized form for a variety of documents representation. In particular, Akoma Ntoso provides a set of tools able to manage in a unified way the majority of legal documents. Furthermore, it’s possible to create subschemas to appropriately structure the individual types of legal information making them processable and uniforms.

Then, it is possible to use the XML document standard to capture the semantics of relevant parts of the document. More in detail, the purpose of Akoma Ntoso is the allocation of structure and semantics in a legal text, regardless of its origin, type, nationality and linguistic context to which they belong.

With Akoma Ntoso, it’s possible to include metadata, semantic annotation and content inside the same document, since it provides suitable tools to be able to allow the management of citations within texts, references between texts, references to concepts belonging to ontologies, thesauri or classifications. Akoma Ntoso handles these semantic annotations, in addition to metadata useful for classifiers and revisions of the text.

All such semantic information also help query EUCases. Metadata, semantic annotations, ontologies allow structured information in XML to be easily interoperable with each other, ensuring a high return in terms of performance, quality and completeness of information as well as value-added service.

The multilingual features of Akoma Ntoso fully match with EUCases requirements. In fact, EUCases born in a multilingual context, and the ability to handle this particular scenario by Akoma Ntoso is extremely important. The ability to make processable the linguistic “versioning” of the text using FRBR ontologies limits the problems of users coping with this particular aspect of the data.

As mentioned earlier, Akoma Ntoso also ensures interoperability through direct references to concepts, citations and references among legal texts, ontological concepts and so forth.
10 Akoma Ntoso for EUCases

For EUCases, we propose to use only some parts of Akoma Ntoso, and specifically:

**Metadata** (only the identification FRBR metadata, linguistic metadata, classification metadata, analysis of the case-law references):

It’s structured information about a document, data or other information content. Metadata has a wide range of benefits and it plays a very important function in helping users to find and retrieve content. The main purpose of these editorial additions is to facilitate in the discovery of relevant information, but a more sophisticated use of it helps to organize electronic resources, to track the usage of a document and to connect content to other content. Metadata values are labelled and collected according to a common ontology, i.e. an organized description of the metadata categories that describe the resources. A shared ontology is fundamental to provide a way for managing, organizing and comparing metadata. All metadata terms must be assigned a URI (a uniform resource identifier) that identifies the term. Akoma Ntoso uses location-independent and permanent identifiers to identify resources such as Parliamentary Documents and ancillary texts, thereby allowing long-term and reliable access to them via web-based technologies.

**Introductory part** (for understanding the identification URI):

The introductory part is the part of the document where are expressed the information for identifying the document in univocal way. The date, the type of document, the author of the document, the number of the document. Those information are fundamental for naming the document in persistent and univocal way.

All AKOMA NTOSO documents are identified by a unique name expressed as a specific URI. This Akoma Ntoso Naming Convention provides a simple mechanism needed to ensure that metadata belonging to the core set can be named and referred to in a consistent way. Being able to enhance clarity in cases of potential ambiguity and to refer to documents in a way that is both human and machine-readable is very useful.

**Header** (for judgments and for detecting parties, the Coram composition, persons and institutions involved in the case-law):

The header - located on top-level of the compression package - should be used both as a container for introductory content and a container of all prefacing material of judgements concerning parties, Coram composition, court type, neutral citation, party, judge, lawyer, document number, hearing dates, etc.

**Preamble** (for acts):

The act is the Akoma Ntoso type of document for modelling the approved law. It is organized in some main blocks (meta, preface, preamble, etc.). The element preamble is used as a container of an opening clause that introduces the main body of the document or for some document it includes also the motivations and the justifications statements (e.g. EU directive, decree, executive order.).

**Hierarchy** (of the structure of body in as much detail as possible):

This part is a container with a set of nested sections of elements that has a hierarchy organization (part, chapter, section, subsection, etc.). Each part of this hierarchy structure has particular sub-elements number of the part (<num>); title of the part (<heading>); subtitle (<subheading>). Each hierarchy part could contain <intro> and <wrap> elements for any introductory text or for the ending sentence. Each level of the nesting can contain either more nested sections or blocks. No text is allowed directly inside the hierarchy, but only within the appropriate block element (or, of course, titles and numbering).

**End Matter** (end part of the document embedded in a rough block):
This element is the final block in the document and it is used as a container of all concluding material and closing formulas (the date, the signature, the location, etc.).

**Annexes** (light mark-up):
Annex are called attachments in Akoma Ntoso. All attachments are external documents containing explicit references to the main current act. In the European legislation, the annex has not other heading that the word ANNEX with, eventually, an number in the case of multiple annexes (in fact, each annex may itself carry another annex, and so on indefinitely) So, the constraint to structure an independent document is an inconvenient. Akoma Ntoso allows a more simple representation of annexes by a structural element <annex>. 
11 Akoma Ntoso Compliance Levels

To help each body find the subset of Akoma Ntoso XML-schema suitable for its needs and requirements, Prof. Fabio Vitali of the University of Bologna has developed a sub-schema extractor web service. This web service is able to extract only the part of the XML-schema in which the end-user is interested, and can be accessed from http://akn.web.cs.unibo.it/aknssg/aknssg.html.

The OASIS TC LegalDocML, that are now managing the standard, have defined five levels of compliance to the Akoma Ntoso schema. Of course, technical validation is a pre-requisite for compliance.

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>structure of the document</td>
</tr>
<tr>
<td>2</td>
<td>structure and naming convention of URI/IRI (FRBR metadata)</td>
</tr>
<tr>
<td>3</td>
<td>structure, naming convention, basic metadata (e.g. normative references)</td>
</tr>
<tr>
<td>4</td>
<td>structure, naming convention, basic metadata, advanced metadata (e.g. events, modifications, qualification of the document, etc.)</td>
</tr>
<tr>
<td>5</td>
<td>structure, naming convention, basic and advanced metadata, enriched semantic elements (e.g. references, location, quantity, term, person, etc.)</td>
</tr>
</tbody>
</table>

For the EUCases project, we recommend adherence to level 3 with the addition of any automatic entity included in the normative text as persons, organizations, dates, quantities, locations, and terms from EuroVoc.
12 Compliance with ELI, ECLI standards

Akoma Ntoso permits to respect the compliance with ELI and ECLI standards as well as with URN:LEX. Moreover Akoma Ntoso is able to manage better versioning, linguistic variants, multiple interpretation and publication of the same legal source.

12.1 Naming convention for citations

The naming convention of the legal documents is a key point for implementing interoperability among documents, among institutions, among heterogeneous document collections. The document needs a unique identifier and it is important also to define a robust methodology for representing the citations. The citations are the fragment of legal text that point out to legal document. We need to have a naming convention also for the citation (fragment, range, list of articles, internal citations, static citations, dynamic citations). Several standard raised recently in the international scenario especially when the Open Government movement accelerated the accessibility and the reusability of the legal sources in the web following the semantic web paradigm. It is so true that currently the OASIS international standardization body has opened a Technical Committee for conciliating all those standard and so to provide a unique specification (LegalCiteM TC\(^{22}\)).

The members of the LegalCiteM TC belong to the main standard used in EU and at international sector. We can identify the following emerging standards for the URI: ELI/ECLI; URN:LEX; Akoma Ntoso.

12.2 ELI/ECLI

The European Commission and the Council deliberated at 26.10.2012, the “Council conclusions inviting the introduction of the European Legislation Identifier (ELI)” (2012/C 325/02) for introducing the ELI naming convention in order to manage the URI and the legal citation in harmonized way among the main UE institutions (Parliament, Council, Commission, Court of Justice) and also for favouring the interoperability among the Member States.

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

\(^{22}\) https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legalcitem
For the case-law the European Commission released the standard for the UR ECLI (European Case Law Identifier). The Court of Justice of the European Union is the 'national' ECLI co-ordinator for the EU and introduced ECLI on 24 March 2014.

12.3 URN:LEX

URN:Lex24 (A Uniform Resource Name (URN) Namespace for Sources of Law (LEX) started in 1999 with the NormeInRete project (Italian government project for URI). When the project ended an extended version, including international approach, was submitted to the Internet Engineering Task Force (IETF) back in 2010 and was the product of a number of groups. The Institute of Legal Information Theory and Techniques of the Italian National Research Council led the charge and the initiative also involved Cornell’s Legal Information Institute, Senate of Brazil, Senate of Italy, Chamber of Deputies of Italy, several Italian regions.

The URN mechanism was not originally based on HTTP protocol but on the URN approach, but recently, after the Akoma Ntoso naming convention, the URN:LEX specifications includes also a part for making the standard compliant with HTTP protocol. After the work done in the CEN Metalex URN:LEX included also the FRBR vision.

local-name = work ["@" expression] ["$" manifestation]

The WORK minimal elements in URN:LEX are25:

- country or jurisdiction enacting the document
- enacting authority
- document type
- relevant creation date or period
- relevant number or any disambiguating feature.

The EXPRESSION minimal elements in URN:LEX are:

specific version of the corresponding work are:

- event date associated to a specific version
- event creating such new version
- validity date that specifies the content validity
- language in which the document is written.

The MANIFESTATION minimal elements in URN:LEX are:

- data format used (e.g., xml, pdf, ...)
- publisher or editorial staff (responsible for content
- markup and metadata)
- components contained in the manifestation (the
- document itself, tables, maps, etc.)
- possible additional features (e.g., on charge, anonymized
- decision text, etc.).

An example is the following:

<table>
<thead>
<tr>
<th>art. 5 of the Ministry of Justice decree of 24 july 1992, n. 358</th>
</tr>
</thead>
</table>

24 http://datatracker.ietf.org/doc/draft-spinosa-urn-lex/
25 Francesconi E., Summer School LEX2013.
The URN: LEX provides a very detailed and complex standard and for this reason sometime it is difficult to implement an effective resolver able to includes all the possible rules of the grammar.

One of the main problems is with the authority filed as the type of document that are required. It needs an updated vocabulary for resolving the name of the abbreviation of the authority, owner of the document, as well as a repository of all the different type of document names (decree, legislative.decree, regional.act, etc.).

We take the following example:


it is the URI in URN:Lex of the Spanish version of the case-law 33-08. Because the name of the authority is in the beginning of the URI and it is a mandatory field, without a central repository of all the name of the authorities (tibunal.justicia) and their abbreviation, we can’t model the URI of the citation starting from the text (NLP).

### 12.4 Akoma Ntoso: URIs and Citations

Akoma Ntoso uses a naming convention for URIs and citations based on the FRBR ontology and it is an evolution of the CEN Metalex\(^{26}\) principles and to the URN:LEX outcomes. FRBR is a standard adopted by the IFLA Study Group on the Functional Requirements for Bibliographic Records. The naming convention is based on Work, Expression and Manifestation that distinguishes the abstract level, the logical level and the physical level of the legal resource.

The Akoma Ntoso naming convention for the URI is based also on HTTP protocol for permitting a rapid and robust redirection, derefering, navigation following the HTTP literature.

(From the release notes delivered with the XSD release by Fabio Vitali, University of Bologna)

“In keeping with FRBR\(^{27}\), each bibliographic document should be viewed from three perspectives:

- as a Work, that is, a distinct work of authorship, an author’s intellectual or artistic creation of the mind. In the arts this might be Shakespeare’s Hamlet, or in the legal domain, a specific Act;
- as an Expression, that is, a distinct intellectual or artistic creation of the mind through which a Work is expressed, interpreted, or otherwise represented. In the arts, this might be Zeffirelli’s 1990 movie of Hamlet; in the legal domain, a specific version of an Act; and
- as a Manifestation, that is, the physical embodiment of a given Expression. In the arts, this might be the DVD on which the movie is recorded; in the legal domain, a specific publication in the Gazette of a particular version of an Act.

In our case, the legal document is interpreted as a bibliographic resource, and is accordingly identified and named from three perspectives: as a work, an expression and a manifestation. As a Work, the document bears the name of the law it

---


\(^{27}\) Functional Requirements for Bibliographic Records.
represents (e.g., Legislation 34); the Work will in turn contain parts that can be associated with a specific Expression (e.g., an amended part of Legislation 34); and there will be a Manifestation (e.g., the XML file representing Legislation 34).”

It is therefore important to determine at this stage which parts of the entire document are to be converted into an XML Manifestation and which parts are best kept in other formats (such as PDF or TIFF). We must accordingly assess the main document, its annexes, and table of contents to identify the appropriate type of document required.

For each legislative resource, the expert should identify the URI, taking into account the three main aspects mentioned previously: Work, Expression, and Manifestation.

It is therefore important to define the type of the document, the country where the corresponding law was enacted, the main language, the date of the document, and its version if there is more than one available.

On the other hand, a document can be composed of different parts, such as annexes, explanatory statements, and tables, or it may contain a number of other official documents. It is therefore essential, when identifying an URI, that we also fully analyse the structure of the document’s components so as to separate the way the content has been logically organized by the author (i.e. parliament for legislative documents, a judge for judicial ones) from the way it has been physically organized, usually on the basis of technical criteria. In other words, the Work URI should reflect the author’s original logical structure, so as to preserve over time the original form and hierarchy of the annexes or that of other material that make up the full document. The physical organization can follow different criteria related to the purpose of the application or technical choices. Although we might find three components in the URI at the Work level, there may be a single, unique URI component at the Manifestation level, making the document easier to manage.

12.4.1 Example of a EU Directive with Annexes

(From the release notes delivered with the XSD release by Fabio Vitali, University of Bologna):

“Akoma Ntoso WORK naming convention is composed by the following main elements:

• Country (a two-letter code according to ISO 3166-1 alpha-2);
• Type of document – in case of an Akoma Ntoso XML representation, this value must correspond to the element immediately below the akomaNtoso root element (e.g., act, bill, debateReport, etc.);
• Any specification of document subtype, if appropriate – in case of an Akoma Ntoso XML representation, this value must correspond to the content of element FRBRsubtype in the metadata (optional);
• The emanating actor, unless implicitly deducible by the document type (e.g. acts and bills do not usually require actor, while ministerial decrees do) – in case of an Akoma Ntoso XML representation, this value must correspond to the content of element FRBRauthor in the FRBRWork section of the metadata (optional);
• Original creation date (expressed in YYYY-MM-DD format or just YYYY if the year is enough for identification purposes) – in case of an Akoma Ntoso XML representation, this value must correspond to the content of element FRBRdate in the FRBRExpression section of the metadata;
• Number or title or other disambiguating feature of the work (when appropriate, otherwise the string nn) – in case of an Akoma Ntoso XML representation, this value must correspond to the content of element FRBRnumber or FRBRname, respectively, of the metadata.

Akoma Ntoso EXPRESSION naming convention add at the end of the WORK:
• the human language code in which the expression is drafted (a three-letter code according to ISO 639-2 alpha-3);
• the date of the version;
• #character if we add the option parts;
• Any content authoring information to determine the authoritativeness of the text content. This is separate and independent of the authoring information relative to the metadata and mark-up, which are among the features of the manifestation (optional) – in case of an Akoma Ntoso XML representation, these values must correspond to the content of elements in the FRBRExpression section of the metadata.
• Any content-specification date (as opposed to validity dates) (optional).

Akoma Ntoso MANIFESTATION naming convention adds at the end of the EXPRESSION the format of the file (e.g. PDF, xml, doc, etc.).

12.4.2 Virtual citations
It is frequent to cite an act in force in a given time, but without to know the date of proper versioning. For this reason Akoma Ntoso includes a particular mechanism for permitting a dynamic reference:
It is a reference to the directive 2014/23/EU closest to the 2014-05-30 given date

The same for the language versions in case we have legal system that admits different legal binding variant (e.g. EU institutions):
/eu/act/2014-02-26/2014-23-EU/:deu/main
It is a reference to all the versions of the EU directive in German.
12.4.3 Example of a EU Directive with Annexes\textsuperscript{29}

| of 26 February 2014 |
| on the award of concession contracts |

This directive has a main document and from eleven annexes.

Here after the URI of the document:

work uri: /eu/act/2014-02-26/2014-23-EU
work uri of the main document: /eu/act/2014-02-26/2014-23-EU/main
work uri for the annex: /eu/act/2014-02-26/2014-23-EU/main/annex1

expression uri: /eu/act/2014-02-26/2014-23-EU/eng@ (original)
expression uri of the main document: /eu/act/2014-02-26/2014-23-EU/eng@/main (original)
expression uri of the annex: /eu/act/2014-02-26/2014-23-EU/eng@/main/annex1
expression uri of the main where we have specified the author of the publication: /eu/act/2014-02-26/2014-23-EU/eng@/#eur-lex/main

manifestation uri for all the package: work uri: /eu/act/2014-02-26/2014-23-EU/eng@/main.akn
manifestation uri for the main document: work uri: /eu/act/2014-02-26/2014-23-EU/eng@/main.xml

In the future, when this directive will be modified and a new version will be released by the EUR-LEX, we can add also the date of the version in the expression URI:

expression uri: /eu/act/2014-02-26/2014-23-EU/eng@2014-05-20
this means to have a new version of the same work in force at 2014-05-20.

12.5 Comparison and Comments

The ELI standard includes several syntactical problems in the management of the versioning, language variant, jurisdiction, authority and languages.

The URN:LEX standard is more expressive to Akoma Ntoso and this fact makes it more difficult to manage with different authorities, sub-authorities, type of documents, etc. Also the time model is slightly different. URN:LEX uses the date of the publication of the modificatory act as date of versioning. Akoma Ntoso uses the date of the enter in force of the modifications as the date of the new version of the act.

Akoma Ntoso permits to manage a light HTTP naming convention, that includes the main part of the URI for legal sources and it could be a interchange standard also for the others.

\textsuperscript{29} \url{http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0023&from=EN}
Here after an example of comparison between the three standards:

on the recognition of professional qualifications


|----------------------|-----|---------|-------------|

### 12.5.1 Akoma Ntoso resolver


This resolver, implemented by University of Bologna (Fabio Vitali, Luca Cervone, Monica Palmirani) and released in open source, permits to navigating, resolving, dereferenting the URI citations using Akoma Ntoso model and in the meantime to map the local naming convention coming from different naming conventions.

Using this resolver it is possible to map ELI and URN:LEX in Akoma Ntoso model and so to implement the persistency, the interoperability and the HTTP robustness principles.
13  **Final XML-schema (XSD) for EUCases**

As said above, for EUCases, we propose to use only some parts of Akoma Ntoso. We are interested in annotating case law and laws, for which we will respectively use the document type element `<judgement>` `<act>` `<doc>`. Furthermore, we will use the tag `<ref>` for annotating references to a document with legal status (other case law / laws, codes, etc.). In this section, we briefly describe the tags we are going to use in the project.

For using only some parts of Akoma Ntoso there are two methods following the official documentation of the schema:

a) to use the unique XSD schema akomantoso.xsd;

b) to use the service http://generator.akomantoso.org/ for extracting a sub-schema necessary for a particular purposes.

We do not provide a detailed description of all tags and subtags used in EUCases. It only briefly describes the main ones. The full specifications of the subschema(s) used in EUCases are the extracts of the sub-schema judgement, act and doc extracted by the official service http://generator.akomantoso.org/. This guarantees to have the last updated version any time.

See also the http://sinatra.cirsfid.unibo.it/XSDocViewer/ for the official documentation of the schema in visual way.

### 13.1 Element: ref

<table>
<thead>
<tr>
<th>Name</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Locally-defined complex type</td>
</tr>
<tr>
<td>Documentation</td>
<td>The element ref is an inline element containing a legal reference (i.e. a reference to a document with legal status and for which an Akoma Ntoso URI exists).</td>
</tr>
</tbody>
</table>

The inline `<ref>` is the text referring to the document, e.g. “article 2.1”, “penal code”, etc. The tag will specify an attribute `href` that will be filled with the Akoma Ntoso naming convention that is able to mapping on ELI and ECLI of the referred document, eventually augmented with a relative internal index to a subpart of the document using the naming convention of id mandatory for the compliance of Akoma Ntoso.

### 13.2 Element: judgement

<table>
<thead>
<tr>
<th>Name</th>
<th>Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>judgementStructure</td>
</tr>
<tr>
<td>Documentation</td>
<td>Element judgement is used for describing the structure and content of a judgement. The type judgementStructure specifies the overall content model of the document types that describe judgements.</td>
</tr>
</tbody>
</table>
XML Instance Representation

<judgement contains="singleVersion[0..1]">
    <judgementBody> ... </judgementBody> [1]
    <conclusions> ... </conclusions> [0..1]
</judgement>

13.2.1 Element: judgementBody

Name   judgementBody
Type   judgementBodyType
Documentation   The element judgementBody is the container of the main hierarchy of a judgement document. The type judgementBodyType specifies a content model of the main hierarchy of a judgement document.

XML Instance Representation:

<judgementBody class="xsd:string [0..1]" style="xsd:string [0..1]"
title="xsd:string [0..1]", period="xsd:anyURI [0..1]", status="statusType [0..1]"
id="xsd:ID [0..1]", evolvingId="xsd:NMTOKEN [0..1]", refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]", xml:space="[0..1]", alternativeTo="xsd:anyURI [0..1]">
    Start Sequence [1..*]
    Start Choice [1]
        <introduction> ... </introduction> [1]
        <background> ... </background> [1]
        <motivation> ... </motivation> [1]
        <decision> ... </decision> [1]
    End Choice
End Sequence
</judgementBody>

13.2.2 Element: conclusions

Name   Conclusions
Type   basicopt
Documentation   The element conclusion is used as a container of all concluding material (e.g. dates, signatures, formulas, etc.). Its type is basicopt, specifying the element is optional.
XML Instance Representation:

<conclusions class="xsd:string [0..1]" style="xsd:string [0..1]"
title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
eld="xsd:ID [0..1]" wld="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">
Start Sequence [1..*]
Start Choice [1]
Start Choice [1]
Start Choice [1]
<blockList> ... </blockList> [1]
<block> ... </block> [1]
End Choice
<p> ... </p> [1]
End Choice
End Sequence
</conclusions>

We omit the description of the subtags <blockList>, <block> and <div>, used to frame the text of the lists and paragraphs of the case law. We will use only <p>.

13.2.3 Element: introduction

Name: introduction
Type: maincontent

Documentation: This element is a structural container for the section of a judgement containing introductory material. The complex type maincontent is used by container elements that can contain basically any other Akoma Ntoso structure.

<introduction class="xsd:string [0..1]"
style="xsd:string [0..1]" title="xsd:string [0..1]" period="xsd:anyURI [0..1]"
status="statusType [0..1]" eld="xsd:ID [0..1]" wld="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">
Start Choice [1..*]
Start Choice [1]
Start Choice [1]
Start Choice [1]
<blockList> ... </blockList> [1]
<block> ... </block> [1]
End Choice
<p> ... </p> [1]
End Choice
</introduction>

13.2.4 Element: background

Name: Background
Type maincontent

Documentation This element is a structural container for the section of a judgement containing the background.

XML Instance Representation

```
<background class="xsd:string [0..1]" style="xsd:string [0..1]"
title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
eId="xsd:ID [0..1]" wId="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">

Start Choice [1..*]
  Start Choice [1]
    Start Choice [1]
      <blockList> ... </blockList> [1]
    <tblock> ... </tblock> [1]
  End Choice
  <p> ... </p> [1]
  End Choice
  <div> ... </div> [1]
End Choice
</background>
```

13.2.5 Element: motivation

Name Motivation

Type maincontent

Documentation This element is a structural container for the section of a judgement containing the motivation.

XML Instance Representation

```
<motivation class="xsd:string [0..1]" style="xsd:string [0..1]"
title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
eId="xsd:ID [0..1]" wId="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">

Start Choice [1..*]
  Start Choice [1]
    Start Choice [1]
      <blockList> ... </blockList> [1]
    <tblock> ... </tblock> [1]
  End Choice
  <p> ... </p> [1]
  End Choice
  <div> ... </div> [1]
End Choice
</motivation>
```
13.2.6 Element: decision

Name decision
Type maincontent
Documentation
<type> Element</type><name> decision</name><comment>this element is a structural container for the section of a
judgement containing the decision</comment>

XML Instance Representation
<decision class="xsd:string [0..1]" style="xsd:string [0..1]"
title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
eId="xsd:ID [0..1]" wId="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">
Start Choice [1..*]
Start Choice [1]
Start Choice [1]
<blockList> ... </blockList> [1]
<block> ... </block> [1]
End Choice
End Choice
<block> ... </block> [1]
End Choice
End Choice
End Choice
</decision>

13.3 Element: judicial

Each judgment citation could be qualified using Shepard's method for representing which
references are in favour of the current judgment argumentation. We omit the detailed
description of each subtag <judicial>, for which we address the reader to the official
documentation. Below, we report only a brief description of the subtags.

Name judicial
Type judicialArguments
Documentation The element judicial is a metadata container of the analysis of
the judicial arguments of a judgement

XML Instance Representation:
• <supports>: The element supports is a metadata element specifying a reference to a source
  supported by the argument being described
• <isAnalogTo>: The element isAnalogTo is a metadata element specifying a reference to a
  source analog to the argument being described
- `<applies>`: The element applies is a metadata element specifying a reference to a source applied by the argument being described.
- `<extends>`: The element extends is a metadata element specifying a reference to a source extended by the argument being described.
- `<restricts>`: The element restricts is a metadata element specifying a reference to a source restricted by the argument being described.
- `<derogates>`: The element derogates is a metadata element specifying a reference to a source derogated by the argument being described.
- `<contrasts>`: The element contrasts is a metadata element specifying a reference to a source contrasted by the argument being described.
- `<overrules>`: The element overrules is a metadata element specifying a reference to a source overruled by the argument being described.
- `<dissentsFrom>`: The element dissentsFrom is a metadata element specifying a reference to a source dissented from the argument being described.
- `<putsInQuestion>`: The element putsInQuestions is a metadata element specifying a reference to a source questioned by the argument being described.
- `<distinguishes>`: The element distinguishes is a metadata element specifying a reference to a source being distinguished by the argument being described.

(see Monica Palmirani papers).

### 13.4 Element: act

<table>
<thead>
<tr>
<th>Name</th>
<th>act</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>hierarchicalStructure</td>
</tr>
<tr>
<td>Documentation</td>
<td>Element act is used for describing the structure and content of an act. The type hierarchicalStructure specifies the overall content model of the document types that are hierarchical in nature, especially acts and bills</td>
</tr>
</tbody>
</table>

XML Instance Representation:

```
<act contains="versionType [0..1]" name="" [1]>
  <coverPage> ... </coverPage> [0..1]
  <preface> ... </preface> [0..1]
  <preamble> ... </preamble> [0..1]
  <body> ... </body> [1]
  <conclusions> ... </conclusions> [0..1]
</act>
```

### 13.4.1 Element: coverPage

<table>
<thead>
<tr>
<th>Name</th>
<th>coverPage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>basicopt</td>
</tr>
</tbody>
</table>
The element coverPage is used as a container of the text that acts as a cover page. The complex type basicopt defines the content model and attributes used by basic containers such as coverPage and conclusions. Here the id attribute is optional.

XML Instance Representation:

```
<coverPage class="xsd:string [0..1]" style="xsd:string [0..1]"
title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
eId="xsd:ID [0..1]" wId="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">
  Start Sequence [1..*]
    Start Choice [1]
      <blockList> ... </blockList> [1]
      <tblock> ... </tblock> [1]
    End Choice
    Start Choice [1]
      <table> ... </table> [1]
      <p> ... </p> [1]
    End Choice
  End Sequence
</coverPage>
```

### 13.4.2 Element: preface

<table>
<thead>
<tr>
<th>Name</th>
<th>preface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>prefaceopt</td>
</tr>
<tr>
<td>Documentation</td>
<td>The element preface is used as a container of all prefacing material (e.g. headers, formulas, etc.). The complex type prefaceopt defines the content model and attributes used by preface. Here the id attribute is optional</td>
</tr>
</tbody>
</table>

Inside of the preface there are important inline elements: `<docNumber>`, `<docTitle>`, `<docDate>`, `<docType>`, `<docAuthority>`

XML Instance Representation:

```
<preface class="xsd:string [0..1]" style="xsd:string [0..1]"
title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
eId="xsd:ID [0..1]" wId="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
refersTo="xsd:anyURI [0..1]"
xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">
  Start Sequence [1..*]
    Start Choice [1]
      <blockList> ... </blockList> [1]
      <tblock> ... </tblock> [1]
  End Sequence
</preface>
```
13.4.3 Element: preamble

<table>
<thead>
<tr>
<th>Name</th>
<th>preamble</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>preambleopt</td>
</tr>
<tr>
<td>Documentation</td>
<td>The element preamble is used as a container of the text that opens the main body of the document as a preamble. The complex type preambleopt defines the content model and attributes used by preambles. Here the id attribute is optional</td>
</tr>
</tbody>
</table>

XML Instance Representation:

```
<preamble class="xsd:string [0..1]" style="xsd:string [0..1]"
    title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
    eId="xsd:ID [0..1]" wid="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
    refersTo="xsd:anyURI [0..1]"
    xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">
    Start Sequence [1..*]
    Start Choice [1]
    Start <blockList> ... </blockList> [1]
    <tblock> ... </tblock> [1]
    End Choice
    Start Choice [1]
    <table> ... </table> [1]
    <p> ... </p> [1]
    End Choice
    End Sequence
</preamble>
```

13.4.4 Element: body

<table>
<thead>
<tr>
<th>Name</th>
<th>body</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>bodyType</td>
</tr>
<tr>
<td>Documentation</td>
<td>The element body is the container of the main hierarchy of a hierarchical document (e.g, an act or a bill). The type bodyType specifies a content model of the main hierarchy of a hierarchical</td>
</tr>
</tbody>
</table>
Within the tag <body>, we may include ten subtags. We omit the detailed description of the subtags and of their compositional rules, because they are rather intuitive and follow the basic structure of the laws (cf. Section 5 above).

The reader is addressed to the documentation, extracted by the official release of Akoma Ntoso, which is downloadable from the URL eucases.eu/fileadmin/EUCases/documents/subschemaAKN.7z

Basically, laws are usually drawn up according to a hierarchical structure in which the text is subdivided into sections or chapters. These are subdivided into clauses or articles, subparagraphs, etc. For each item in which the law is subdivided we have a homonym tag where the corresponding text is enclosed: <title>, <chapter>, <section>, <subsection>, <article>, <paragraph>, etc.

### 13.4.5 Element: conclusions

<table>
<thead>
<tr>
<th>Name</th>
<th>conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>basicopt</td>
</tr>
</tbody>
</table>
The element conclusion is used as a container of all concluding material (e.g. dates, signatures, formulas, etc.).

XML Instance Representation:

```xml
<conclusions class="xsd:string [0..1]" style="xsd:string [0..1]"
  title="xsd:string [0..1]" period="xsd:anyURI [0..1]" status="statusType [0..1]"
  eId="xsd:ID [0..1]" wId="xsd:NMTOKEN [0..1]" GUID="xsd:NMTOKEN [0..1]"
  ="xsd:anyURI [0..1]"
  xml:lang="[0..1]" xml:space="[0..1]" alternativeTo="xsd:anyURI [0..1]">
  Start Sequence [1..*]
    Start Choice [1]
      Start Choice [1]
        <blockList> ... </blockList> [1]
        <tblock> ... </tblock> [1]
      End Choice
    Start Choice [1]
      <table> ... </table> [1]
      <p> ... </p> [1]
    End Choice
  End Choice
  End Sequence
</conclusions>
```
14 Next steps

The next steps in the project will concern the implementation of software and interfaces that automatically build and use XML documents in Akoma Ntoso, starting from the legislative documents downloaded by the crawlers.

Building XML documents in Akoma Ntoso

Two tools will be implemented: the LinkingTool and the Text2XML tool.

LinkingTool

The goal of the LinkingTool is to identify in free text references to norms and case law or to subparts of them (paragraphs, subparagraphs, items in a numbered list, etc.). References are identified by enclosing the text within the tag `<ref href="..."></ref>`.

Documents downloaded by the crawlers are files in different formats (pdf, html, etc.). The text within these documents is extracted, eventually via special libraries freely available on the web, and processed via NLP tools like postaggers and parsers.

Starting from the result of the NLP tools, the LinkingTool builds XML documents where references are identified. For example, with respect to the following Italian text:

```italian
...In relazione alla commissione dei delitti di cui agli articoli 615-quater e 615-quinquies del codice penale, si applica all'ente la sanzione pecuniaria sino a trecento...
```

The LinkingTool will return the following XML document:

```xml
...In relazione alla commissione dei delitti di cui agli articoli 615-quater e 615-quinquies del codice penale, si applica all'ente la sanzione pecuniaria sino a trecento...
```

The strings “AKN1” and “AKN2” refer to the AKN reference to the law (cf. §11, “Compliance with AKN standard that is also compliance with ECLI standards”) while “#615-quarter” and “#615-quinquies” are internal indexes that refer to the mentioned articles within the law.

LegalText2XML

The LegalText2XML module will take as input the text extracted from the documents downloaded by the crawlers and it will build XML documents in the Akoma Ntoso sub-schema associated with the original documents downloaded by the crawlers.

The first sub-module of the LegalText2XML tool is a statistical classifier that inspects the text and chooses the proper Akoma Ntoso sub-schema that has to be associated with the input document.

Afterwards, the LegalText2XML will identify the several XML tags of the selected Akoma Ntoso sub-schema by processing the text via NLP tools.
DECRETO LEGISLATIVO 8 giugno 2001, n. 231

Disciplina della responsabilità amministrativa delle persone giuridiche, delle società e delle associazioni anche prive di personalità giuridica, a norma dell’articolo 11 della legge 29 settembre 2000, n. 300.

Capo I

RESPONSABILITÀ AMMINISTRATIVA DELL’ENTE

SEZIONE I

Principi generali e criteri di attribuzione della responsabilità amministrativa

Art. 1.

Soggetti

1. Il presente decreto legislativo disciplina la responsabilità degli enti per gli illeciti amministrativi dipendenti da reato.

2. Le disposizioni in esso previste si applicano agli enti forniti di personalità giuridica e alle società e associazioni anche prive di personalità giuridica.

3. Non si applicano allo Stato, agli enti pubblici territoriali, agli altri enti pubblici non economici nonché agli enti che svolgono funzioni di rilievo costituzionale

[…]

The output of the LegalText2XML will be the following:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<akomaNtoso xmlns = "http://docs.oasis-open.org/legaldocml/ns/akn/3.0/CSD10" xmlns:html = "http://www.w3.org/1999/xhtml"
>  
<act name="act">
  <meta>
  </meta>
  <preface>
    <p>
      <docType eld="dcTyp1">DECRETO LEGISLATIVO </docType>
      <docDate eld="dcTyp1-dcDt1"> 8 giugno 2001</docDate>,
      n.
      <docNumber eld="dcTyp1-dcNmbr1">
        231
      </docNumber>
    </p>
    </preface>
  </act>
</akomaNtoso>
```
Disciplina della responsabilità amministrativa delle persone giuridiche, delle società e delle associazioni anche prive di personalità giuridica, a norma dell'articolo 11 della legge 29 settembre 2000, n. 300.

CAPO I

RESPONSABILITÀ AMMINISTRATIVA DELL'ENTE

Principi generali e criteri di attribuzione della responsabilità amministrativa

Il presente decreto legislativo disciplina la responsabilità degli enti illeciti amministrativi dipendenti da reato.
Le disposizioni in esso previste si applicano agli enti forniti di personalità giuridica e alle società e associazioni anche prive di personalità giuridica.

Non si applicano allo Stato, agli enti pubblici territoriali, agli altri enti pubblici non economici nonché agli enti che svolgono funzioni di rilievo costituzionale.

Using Akoma Ntoso documents

The Akoma Ntoso XML documents constitute fine-grained semantic descriptions of the legal documents. These descriptions are a useful knowledge base for web interfaces that help legal professionals across Europe with navigating legislation.

Laws and case laws may be displayed in the portal together with hyperlinks connecting a legal document with the other documents mentioned therein, thanks to the references identified by the LinkingTool. A legal professional could drastically reduce his/her work time by following the hyperlinks rather than by searching the mentioned documents via standard search engines.

On the other hand, it would be possible to classify, via statistical machine-learning and clustering techniques, the singular sections and paragraphs of the documents, identified by the LegalText2XML tool. For instance, the sections of case law specifying the motivations of
the verdict can be used as input for a statistical classifier able to find other case law with similar motivational bases, that may be displayed in the web interfaces. Of course, the correctness and effectiveness of the links displayed depend on the precision of the Akoma Ntoso documents. However, the state-of-the-art NLP tools used to build them are unable to achieve highest precision. Therefore, the Akoma Ntoso documents will eventually require human tuning, once the tools have been built them automatically. To this end, we will consider the possibility of implementing further interfaces to help annotators with fixing errors and building new links that the NLP tools are unable to find automatically. It will be possible to develop such web interfaces both for developers and annotators and for the final users via web 2.0 techniques. For instance, the user interface could provide widgets to enter feedbacks about the quality of the displayed information. In case of errors, users may inform the system of errors and inaccurate annotations. These feedbacks are both useful to the developers for manually tuning the knowledge and to the statistical models that may be automatically filled with false positives in order to achieve better performances.

**DECRETO LEGISLATIVO 8 giugno 2001, n. 231**

*Disciplina della responsabilità amministrativa delle persone giuridiche, delle società e delle associazioni anche prive di personalità giuridica, a norma dell’articolo 11 della legge 29 settembre 2000, n. 300.*
15 References


[7]. International Working Group on FRBR and CIDOC CRM Harmonisation. FRBR Object-Oriented Definition and Mapping to FRBRER (v. 0.9 draft), ed. by C. Bekiari, M. Doerr, and P. Le Boeuf, 2008.


Further information:

Akoma Ntoso official website: http://www.akomantoso.org/

Akoma Ntoso Open Source Tools
AT4AM EU Parliament tool video: http://vimeo.com/48325937
AT4AM for all: http://www.at4am.org/
LIME UNIBO Web Editor: http://lime.cirsfid.unibo.it/
AKResolver: http://akresolver.cs.unibo.it/
Akoma Ntoso subschema extractor: http://akn.web.cs.unibo.it/aknssg/aknssg.html
Validator: http://legixinfo.wordpress.com/2013/08/22/free-akoma-ntoso-validator/

Other Akoma Ntoso tools
Xcential Web Editor: http://legisproweb.com/

OASIS LegalXML-TC
Akoma Ntoso and LegalDocML TC-OASIS
LegalRuleML TC –OASIS
LegalciteM -OASIS