A legal drafting environment based on formal and semantic XML standards

C. Biagioli, E. Francesconi, P. Spinosa, M.Taddei
ITTIG - CNR
via Panciatichi 56/16, 50127, Florence - Italy
http://www.ittig.cnr.it
{biagioli,francesconi,spinosa,m.taddei}@ittig.cnr.it

The NIR project aims at making retrieval and navigation among normative documents in a distributed environment easier. To obtain this, XML and URN standards for normative documents have been established and the construction of tools working on these standards is promoted. Particularly, NIREditor, a specific drafting environment able to deal with new and legacy normative documents has been developed.

1. INTRODUCTION

Fragmentation of legal information and inconsistencies of document formats represent historical obstacles to a systematic organization of the normative corpus. In Italy the "Norme in Rete" (NIR) project has been proposed[1] to solve these problems and create a unique access point on the Web for normative documents with search and retrieval services. To this aim, NIR has established standards to represent normative documents and promoted the development of tools to make their adoption easier. The main one is *NIREditor* [3], a drafting environment for new and legacy normative documents. In Section 2 the NIR standards are introduced; in Section 3 the NIREditor main features are presented; in Section 4 the project of a module able to plan a new bill is discussed; in Section 5 some conclusions are reported.

2. THE NIR STANDARDS

The NIR project has established two official standards for normative documents:
– a URN standard for cross-references, allowing references to be expressed unambiguously, in a stable way and independently of target document physical locations [3];
– a standard to represent normative documents, defined by three DTDs (NIR-DTDs) of increasing degree of complexity [8].
Basically NIR-DTDs describe normative documents by two kinds of elements: Structural elements and Metadata.

[1]CNIPA (Italian National Center for Information Technology in the Public Administration) and Italian Ministry of Justice.

ICAIL '05, June 6-11, 2005, Bologna, Italy.

Structural elements are: *Generic document elements* (references to other laws, formatted text (tables, lists, etc.)) and *Specific normative text elements* (heading, sections, articles, paragraphs). Structural elements describe the form of a normative text (*formal profile*). Two kinds of metadata are provided: *General metadata* (subject classification, publication date; relationships among acts) and *Analytical metadata* [2] (*provisions* types (Amendments (Insertion, Abrogation, Substitution), and Rules (as Obligation, Definition, Penalties, etc.)) and by their *arguments* (for example the *addressee* of an Obligation)). General metadata provide information on the act, analytical metadata describe the semantics of the provisions (*functional profile*).
As the formal profile represents the traditional habit of organizing law texts in chapters, articles, paragraphs, etc., the functional profile is related to how the semantics of the text is organized.

3. THE NIREDITOR

The NIR-DTDs identify a wide and complex subset of documents. The production of new documents and the transformation of legacy contents according to the NIR standards, can be a hard problem to face without an editing system. Even though programs for XML drafting already exist, they have to be adapted to deal with NIR-DTDs. Other solutions (Microsoft Word, Open Office Writer) suffer from similar limitations: users need specific editing functions to deal with the standard as well; moreover these solutions describe a document using a proprietary format, therefore DTD constraints have to be mapped to this. For these reasons we have decided to develop a drafting environment (Fig. 1) handling NIR documents in their XML native format [3] according to legislative technique rules. Similar initiatives exist at European level, as MetaLex [4], a knowledge management system for legislative drafting providing users with both content management and decision support components. As compared to MetaLex, *NIREditor* is more focused on legal drafting to support users in adopting a legal standards. *NIREditor* is designed to process legacy normative texts, as well as to assist the drafting of new texts, using both manual and automatic facilities. In Section 3.1 facilities for legacy contents are presented; in Sections 3.2 functions dealing with the composition and the organization of new acts are described.

3.1 Legacy content handling

Particular attention has been addressed to automatisms for legacy content handling, as key-factors for promoting the

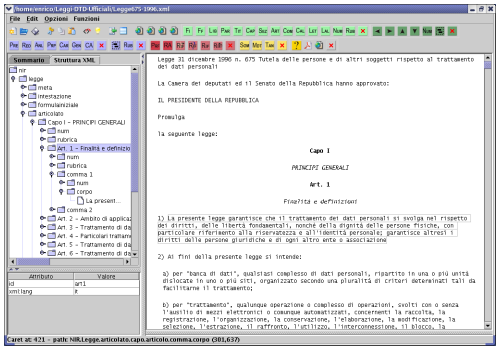

Figure 1: The NIREditor environment

adoption of the standards. Four modules have been implemented: a) a *Cross-Reference Parser*, able to detect cross-references and construct the URNs [3]; b) a *Structure Parser*, able to automate the NIR-XML conversion of legacy contents [3]; c) a *Provision Automatic Classifier*, which automatically classifies paragraphs into provisions according to the NIR provision model [3]; d) a *Provision Argument Extractor*, which automatically identifies the arguments of the provisions [1]. The first two modules detect the formal profile of a normative text, the last two detect the functional profile of it, producing its NIR-XML description.

3.2 Composition and organization of new texts

NIREditor is conceived as a visual editor, supporting users to produce valid documents according to the NIR-DTDs. No XML validation is necessary since it allows only valid operations. Specific facilities are: the insertion of partitions according to the insertion point context; the automatic numbering of the divisions; the updating of internal references in the event of text movements or variations; the external and internal cross-references construction by hand or using the *Cross-Reference Parser*; the analytical metadata insertions by hand or using the *Provision Automatic Classifier* and the *Provision Argument Extractor* as a support. Two possible text organization strategies can be followed: the *formal* and the *functional organization strategies* [2]. In the *formal organization strategy* text is considered as made up of divisions (formal profile): partitions of similar rank, to be grouped in a new partition of higher rank, are chosen exlicitly by the draftsman. In the *functional organization strategy* text is considered as composed by *provision* (functional profile): provisions to be grouped are chosen according to their content, affinities, etc., making queries on the analytical metadata; then it is decided where they should be placed, according to the preferences of the drafter and the customary procedure of presentation.

4. PLANNING A NEW ACT

Facilities to produce an organic and well-structured normative text are desirable. A well-structured normative text can be considered as the one where the sematics organization of text (functional profile) follows its formal organization (formal profile) [5]. For *NIREditor* a module has been designed able to guide the drafter at planning a new organic bill starting from a conceptual point of view, then constructing the best structural organization of the text able to effectively comunicate its semantics. The classical process of drafting (from structure to semantics) is inverted (from semantics to structure). The planning module is conceived as a visual editor of provisions: firstly the user is required to collect terms (manually or from an ontology (ex. JurWordNet [7])) representing entities of the domain to be regulated, then in a visual panel the drafter will insert objects representing the provision types of the new bill and collected terms will be used as values for the provision arguments. At this stage the functional profile of the bill is defined and users will be provided with visual facilities, as well as tools to express criteria (queries), to group sematically correlated provisions into formal partitions. So the formal profile is obtained and the XML skeleton of the new bill can be generated. Proposals of partitions wording can be generated on the basis of the defined functional profile [6].

5. CONCLUSIONS

A specific law drafting environment, *NIREditor*, working with the URN and XML standards established by the NIR project has been presented. It is able to deal with the formal structure and the semantics of a normative text. Manual and automatic facilities for legacy and new normative documents are provided. Finally, the project of a module able to help the drafter in planning a new bill has been presented and discussed.

6. REFERENCES

[1] R. Bartolini, A. Lenci, S. Montemagni, V. Pirrelli, and C. Soria. Automatic classification and analysis of provisions in italian legal texts: a case study. In *Proc. of the Second International Workshop on Regulatory Ontologies*, 2004.

[2] C. Biagioli. Towards a legal rules functional micro-ontology. In *Proc. of workshop LEGONT '97*, 1997.

[3] C. Biagioli, E. Francesconi, P. Spinosa, and M. Taddei. Xml documents within a legal domain: Standards and tools for the italian legislative environment. In *Proc. of IAPR Workshop on Document Analysis Systems*, pp. 413–424, 2004.

[4] A. Boer, R. Winkels, R. Hoekstra, and T. van Engers. Knowledge management for legislative drafting in an international setting. In *Proc. of JURIX 2003: Legal Knowledge and Information System*, pp. 91–100, 2003.

[5] L. K. Branting and J. Lester. Justification structures for document reuse. In *Proc. of the Third European Workshop on Case-Based Reasoning*, pp. 76–90, 1996.

[6] C. Burges. Towards a normalised language to clarify the structure of legal discourse. In *In Martino (ed.) Deontic Logic, Computational Linguistics and Legal Information Systems*. A.A. Martino eds., Amsterdam: North Holland, 1982.

[7] A. Gangemi, M. T. Sagri, and D. Tiscornia. Jur-wordnet, a source of metadata for content description in legal information. In *Proc. of the ICAIL Workshop on Legal Ontologies & Web based legal information management*, 2003.

[8] F. Megale and F. Vitali. I dtd dei documenti di norme in rete. *Informatica e Diritto*, 1:167–231, 201.