Legal Information on the Web. The NIR Portal for the Citizen

Costantino Ciampi

ITTIG/CNR – Istituto di Teoria e Tecniche dell'Informazione Giuridica,
Florence (Italy)

The paper presents an e-government project set up in Italy to build the Normeinrete ("Law on the Net") portal. It offers an information service to citizens providing them with unified access to Italian and European Union legislation published on different institutional web sites. The system is based on co-operative technological architecture, resulting in a federation of legislative data bases developed on different platforms. Co-operation is achieved by means of suitable application gateways which provide "loose" integration by adopting two standards for identifying the resources and representing the document structure and metadata by XML mark up according to ad hoc DTDs. The adoption of these standards allows automated dynamic hyperlinking among laws and semi-automated building of legislation in force. The approach adopted enables a good level of integration to be reached among different systems while maintaining the autonomy of the institutions.

Keywords: *e-government, legislative databases, information and telecommunication technologies.*

Pagrindinės sąvokos: *e.valdžia, teisinės duomenų bazės, informacinės ir telekomunikacinės technologijos.*

Introduction

The *Normeinrete* Project, promoted by the Italian Authority for Information Technology in the Public Administration (AIPA) and the Ministry of Justice in collaboration with ITTIG/CNR, aims at fulfilling the citizen's right to acquire knowledge about legislation and supporting the Public Administration (PA) in managing the legislative documentation life cycle efficiently [1] [2]. These objectives are pursued through the following actions:

- implementation of a specialised portal for the retrieval of legislative documents;
- definition of standards to represent data content in the legal domain;
- software distribution to support legislative document management and publishing;
- training and knowledge sharing among Public Administrations.

The system is based on a federation of legislative data bases developed with different platforms and it is built upon co-operative technological architecture.

This paper is organised as follows:
- Section 2 contains a general description of the Project;
- Section 3 describes the standards;

- Section 4 describes the system's architecture;
- Section 5 deals with current developments and future initiatives.

The *Normeinrete* Project (NiR)

The *Normeinrete* Project (NiR) aims at improving accessibility to legislation by providing a unique point of access to Italian and European Union legal documents published on different web sites through a specialised portal (www.nir.it) [1] [2] [4].

The portal runs a search engine that operates uniformly on distributed data sources. Its full text search index is selectively built to detect only legislative documents [5]. The achievement of a higher level of co-operation relies on the adoption of two standards, defined within the Project by *ad hoc* Working Groups in which major PA and research institutions have taken part. The standards have been issued as AIPA technical standards and published as regulations in the Italian Official Journal. The definitions make use of IETF Uniform Resource Names (URNs) (RFC 2141) and eXtensible Mark up Language (XML W3C Recommendation) standards.

Another Project goal is to support the PA in

the tasks related to law consolidation. Its achievement is also based on the standard definitions provided, because they enable the rules in a distributed environment to be to identified and amendments to be tracked, thus allowing for semi-automated consolidation. An objective of the Project is also to create a virtual space for knowledge sharing within the PA community, also offering dedicated services, such as e-learning tools and open source software download. So far, more than 45 public institutions have taken part in the Project with more than 140,000 documents indexed. There are about 150,000 search sessions monthly on the site.

Unifying Standards

Uniform Resource Name (URN)

Each law contains several references to other laws. The whole legislative corpus can be seen as a net, each law being a node linking, and linked by, several other nodes through natural language expressions. Manual activity is required in order to build a hypertext through the usual web link mechanism based on the physical addresses. The disadvantages deriving from URL approach include the significant editorial work that must be carried out before publishing a document and the subsequent activities needed to prevent or to limit broken links.

The URN is a persistent, location independent, resource identification mechanism. URNS are defined as a combination of elements according to a specific grammar [7]. Their basic elements are:
- name of the promulgating authority;
- type of legislative instrument;
- date of the legislative instrument:
- number of the legislative instrument; and
- a set of more detailed specifications when needed.

The adoption of a scheme based on the URN enabled an automated distributed hypertext to be built according to a model similar to the DNS (Domain Name System) used to convert the self explaining web site names into numerical HTTP addresses. This relied on the following:
- the natural language expressions used in the citation of laws usually contain repetitive patterns, making references automatically detectable;
- the URN is built by combining data (almost) always included in the reference;
- the cross reference between each URN and the list of corresponding URLs, needed for resolution, can be built automatically.

Document Type Definition (DTD) of Italian Legislative Acts

XML representation of legislative instruments improves efficiency in managing, publishing and retrieving legislation by electronic means [3] [6]. *Normeinrete* has defined the Document Type Definition (DTD) for Italian legislation, considering the peculiarity of legislative documents and other significant useful information. Italian legislative and regulatory acts can be divided into three categories:
- documents with a well defined structure (such as, state laws, regional laws, etc.);
- partially structured documents (such as, regulations, decrees, etc.);
- generic documents (such as, any kind of non-structured instruments, schedules, attachments, etc.).

To avoid a proliferation of DTDs, it was considered better to have the definition of a single DTD containing many elements capable of representing all types of documents. Given the variations in the structure of Italian legislative instruments, the mark-up language is very complex and the resulting DTD has three different versions, containing the same set of elements to represent all kinds of documents with different constraints. Documents validated against strict rules are also valid against looser ones. The DTD elements that have been defined can be classified as follows:
- structural elements, identifying the parts into which the document is structured (heading, preamble, articles, etc.);
- special elements identifying meaningful parts of the text in the legal context, (for example, references to other laws) or associating a formatted representation to text-embedded relevant entities (institutions, dates, places);
- elements containing metadata (for example, subject matter classification, publication data, procedures for enacting a bill, etc.).

Mark-up must be carried out by using only elements relevant to the kind of document under consideration.

The Architecture of the NiR System

The main components of the system are:
- NiR nodes: components belonging to administration domains containing legal database systems and related application gateways. Documents can be stored in the file system or within database/full text management systems. They are all acces-

sible through the Internet.

- Central registries: components in the co-operative layer publishing information needed to allow effective co-operation. They include:

1. standards repository (XML DTD and URN grammar definitions and tools);

2. the registry of official authority names, needed to standardise URN adoption;

3. the registry of NiR nodes, containing information needed to allow interaction between NiR agents and domain application gateways;

4. the norm catalogue, containing, for each norm: title, basic classification, URN and the list of known physical addresses (URL) where it is published.

The co-operative system NiR which is the component in the co-operative layer that runs all the specific applications, including the URN resolution service (at the moment managed centrally).

Current Developments and Future Initiatives

New developments are being carried out in the current stage. Some of these concern new software tools to support Administrations in the adoption of NiR standards. In particular, a specialised editor is being developed, that will be distributed as open source software. An XML Schema definition is also being developed. A parsing service will be available, that will return submitted documents with references to other suitably marked-up legislative instruments and with the corresponding URN inserted. A Working Group is being constituted to investigate additional metadata representation and automated document classification.

Future initiatives will include the implementation of distributed URN resolution. In the next stage, the opportunity to define access services to legislation in terms of web services will be evaluated. This opportunity could become more attractive in the event of a more extended adoption of standard languages and models, such as WSDL and UDDI, within the Italian PA co-operative model. The other major theme to be addressed concerns the certification of the authenticity of acts through digital signature technology.

References

1. ITTIG-CNR (ed.), *Studio di fattibilità, per la realizzazione del progetto "Accesso alle norme in rete"* / "Informatica e diritto", 1 (2000), Edizioni Scientifiche Italiane, 200 p.
2. ITTIG-CNR (ed.), *Progetto Nir – Fase 2 "Accesso alle norme in rete"* / "Informatica e diritto", 1 (2001), Edizioni Scientifiche Italiane. P. 266.
3. Boer A., Hoekstra R., Winkels R., *MetaLex: Legislation in XML* / "Legal Knowledge and Information Systems JURIX 2002".
4. Ciampi C., *Gli sviluppi del progetto NIR "Accesso alle norme in rete"*, in "Informatica e diritto", 1 (2001), pp. VII-XI.
5. Lupo C., *Norme in rete: un progetto di cooperazione per l'informatica giuridica*, in "Iter Legis", luglio ottobre, 1999.
6. Marchetti A., Megale F., Seta E., Vitali F. *Using XML as a means to access legislative documents: Italian and foreign experiences*, in "ACM SIGAPP Applied Computing Review". New York: ACM Press, 2002. Vol. 10(1). P. 54–62.
7. Spinosa P. L. *Identification of Legal Documents through URNS (Uniform Resource Names)* / O. Signore and B. Hopgood (eds.), Proceedings of the Euroweb 2001 Conference "The Web in Public Administration" (18 20 dec. 2001). Pisa: Felici Editore, 2001.

Costantino Ciampi

Teisinė informacija internete. NIR portalas piliečiams

Santrauka

Naujosios technologijos skatina ir viešąjį administravimą pereiti į kitą darbo bei paslaugų teikimo lygį. Straipsnyje pristatomas e. valdžios plėtrai skirtas *Normeinrete* („Teisė tinkle") portalas, įgyvendinamas Italijoje. Šio specializuoto portalo dėka piliečiams suteikiama tiesioginė prieiga prie Italijos ir Europos Sąjungos teisinių dokumentų, esančių skirtingose interneto svetainėse. Straipsnyje analizuojamos naršymo skirtingose sistemose integravimo galimybės, palaikant institucijų autonomiškumą, bei kaip numatoma sukurtą virtualią teisinių žinių erdvę pritaikyti viešajame sektoriuje. Atskirose darbo dalyse aptariami bendrieji projekto niuansai, standartai, sistemos struktūra, dabartinis technologijų išsivystymo lygis bei ateities perspektyvos.

Costantino Ciampi – Teisės teorijos instituto informatikos bei telematikos tyrimų vadovas
Adresas Via Panciatichi 56/16, 50127 Firenze Florence (Italy)
Telefonas +39 055 43995
Elektroninis paštas c.ciampi@ittig.cnr.it