7-1-2006

# Legal Information on the Web: the Case of Italy

Ginevra Peruginelli
*University College of London, Institute of Advanced Legal Studies*

Follow this and additional works at: http://scholarship.law.cornell.edu/ijli
The *International Journal of Legal Information* is produced by The International Association of Law Libraries.

# Legal Information on the Web: the Case of Italy

GINEVRA PERUGINELLI*

## 1. Introduction

Accessing legal information is a primary requirement for a variety of communities: ordinary citizens, scholars, and professionals. The dissemination of legal information contributes to the rule of law and to the overall ideals of democracy in a number of ways. Many are the benefits of accessing legal information, such as the awareness of the applicable rule of law, the creation of conditions necessary to the equality and fairness of a legal system, while improving the functioning of democratic institutions, the development and improvement of social and economic conditions.

Without access to legal information, there simply is no real access to justice. Access to legal information should be available for every citizen on the basis of two main principles: the principle of objectivity which is expressed by the latin brocardo *Ignorantia legis non excusat* (ignorance of the law is no excuse) and the principle of effectiveness, which implies that citizens should know the rules which regulate their life. Ignorance of the law excuses no one, and citizens have the right to know the laws governing their conduct. Everyone should be provided with tools and services to gain knowledge of the law, and governments all over the world have an obligation to put forth legal knowledge by enabling access to the law using all available and reasonable media.

The legal world is on the brink of a fundamental transformation that is being precipitated by existing and emerging information technologies. It is claimed that legal practice and the administration of justice will be profoundly changed largely in ways that will be of benefit to citizens. Information

---

* The author holds a degree in law from the University of Florence. Currently, she is a PhD student at University College of London, Institute of Advanced Legal Studies. Additionally, she is in charge of the development and maintenance of DoGi – Dottrina Giuridica database (a major Italian information source of legal literature). She conducts research in access services to legal literature, law and legal language documentation, law and public information, and computing policy. Her most recent research focuses on approaches in the field of cross-language retrieval in legal information resources.

technology has opened up new methods of publishing law and official information; in particular the wide and ever increasing availability of the Internet has reinforced demands that law should be made freely available on this medium.

## 2. A glance to legal information's access in Italy

Italian legal information has specific features due to its nature, its different utilisation purposes and the intrinsic need for integration of its components, represented by legislation, case law and doctrine.

In Italy, as in the rest of the world, the actual scenario of legal information's access shows new opportunities and challenges as well as problems due to the increased provision on the web of electronic resources.

In general Italian legal information is delivered on print formats, but the increasing production of legal texts has brought about a crisis in information access and the traditional tools are no more sufficient to its distribution. Efforts are underway to use new technologies to overcome this problem.  At the same time print media are still considered relevant sources of information by legal users and it should coexist with electronic resources as in the law domain the integration of traditional, electronic sources, both off line and on line, is a major requirement.

To ensure access, commercial publishing is insufficient. In fact, by the experiences of some countries, despite the fact that commercial publishing can serve the needs of legal professionals, such services only indirectly benefit the general public. New forms of dissemination, such as Internet distribution, have made it possible to reach large segments of the population, while it is recognised that information should be conveyed by intermediaries before reaching their final recipients.

Legal online information in Italy started to be produced in 1995 and was initially found only on academic and research institutes' web sites. Only in 1998 legal information started to be offered on institutional web sites of the major producers of legal information sources, such as the Italian Parliament, ministries, and publishers.

Retrieval systems of legal information are mainly provided by two categories of producers and providers of legal information. Public institutions such as governmental and academic bodies and private providers like legal publishers and law firms, which have started to create web sites for

professionals and the general public. At the moment information providers meet some difficulties in setting up joint, effective legal information services.

## 3. The Institute of Legal information Theory and Techniques (ITTIG-CNR)

In this scenario the Italian Institute of Legal Information Theory and Techniques (ITTIG-CNR) plays a very important role in promoting new methodologies for making legal information accessible.[1] The Institute is one of the 103 scientific institutes of the National Research Council, which is the greatest Italian scientific institution carrying out training and development as well as co-ordination activities in the most important cultural, scientific and technological sectors. ITTIG's mission is to carry out scientific and technological research on the impact of information and telecommunications technologies on law and law-related activities. The Institute came into being in 1968 as a result of the reorganisation of the Italian National Research Council - CNR and it is a product of the fusion of the Institute for Legal Documentation - IDG, based in Florence, and the Research Centre for the Study of Roman Law and Legal Systems - CSDRSG, which operated in Rome.

### 3.1 Objectives and activities

ITTIG's priorities are the following:

- study and development of technologies in the law and public administration's sector;
- consultancy to public administrations as support for e-government policy;
- production and distribution of law databases and of specialised computer software.

ITTIG researches into and applies information and communication technology to law and legal language, legislative technique, legal decision-making, and training of lawyers. It extends its studies and investigations into law related themes, public information and computing policy. It creates advanced applications in the field of legal computing and computing law. The target audience of the Institute's activities is legal practitioners in the public and private sector and its natural interlocutors are scholars from universities

---

[1] ITTIG web site is visible at: **http://www.ittig.cnr.it**

and other research establishments dealing with the disciplines that pertain to the Institute's areas of expertise.

In particular, specific themes the Institute is working at are:

- techniques and methods for legal documentation handling and for the creation of legal information systems;
- techniques and methods for electronic analysis, production and evaluation of legal acts;
- law and legal language documentation including its historical development;
- technology for the teaching and evaluation of legal learning.

The Institute manages the production and distribution of various databases, mostly bibliographies, of current and historical material. Whilst some databases are no longer updated, they are nevertheless maintained on-line due to their wide-spread use. ITTIG is also committed to prepare and maintain a series of guides to facilitate the retrieval of websites of legal interest.

### *On-line current databases*

| | |
|---|---|
| **DoGi** | Abstracts of law journal articles |
| **BIG** | Bibliography on pregnancy termination |
| **VIPD** | National and regional legislation, jurisprudence on independent life of people with disabilities |
| **LLI** | Italian legislative language |
| **Circulars CNR** | Circulars and other directives of the Italian National Research Council |
| **APAM** | Selection of material, acts and documents produced by public administrations. |

### *CD-Rom Archives*

| | |
|---|---|
| **ITLaw** | Information Technology and Law. An International Bibliography |

**DTI**                 Collection of European and Italian laws in the field of ICT

**BEGA**                Bibliography of old Italian legal editions (1470-1800)

**Gride**               Laws of the Health Courts of the State of Milan (1576-1706)

**On-line databases (no longer updated)**

**STOP**                Abstracts of important legal articles published in newspapers (1975-1993)

**BIGA**                Legal bibliography on environment (1975-1995)

**DAUE**                Transposition of EU environmental directives in the national law (up to 1997)

**EURO**                International bibliography on European Federalism

**Guide to Law on the Internet**

**Diritto Italia**      Database of Italian law internet sites

**GIGO**                Guide to online legal information

**Legal libraries**     Legal libraries in Italy: information and catalogues

**Legal literature**    Guide to legal literature online

**ELIOS**               Environmental legal information observatory system


In order to meet its objectives the Institute:

- entertains collaborative scholarly relationships with similar international bodies as well as with single countries, through exchange of experiences;
- provides consultancy in its areas of expertise through conventions with public bodies and research contracts;
- offers training to academic and technical staff, in particular public administration officers, and collaborates with universities in teaching legal information technology.

A brief description follows of the types of Italian legal information, available, such as statutes, cases, public administration regulations and legal literature.

## 4. Access to Italian legislation

As concerns legislative information, the full text of statutes is offered free of charge through the web sites of the Parliament, of the various ministries and the Istituto Poligrafico della Zecca dello Stato (IPZS), which operates under the Ministry of Economy and Finance. This organisation publishes and distributes the Official Collection of Regulatory Acts and the Gazzetta Ufficiale della Repubblica Italiana (Official Gazette of the Republic of Italy), which represents the well established, official collection of new laws, regulations, public work contracts and job offers. It publishes not only all the legislative acts, but also community and regional legal sources, as well as the decisions and orders of the Constitutional Court and government circulars. It gives access to the full text of statutes published only in the last sixty days.

Hoping that the IPZS extends its services, a free of charge legal information system offering legislation is provided by the Jesi Municipality.[2] Other initiatives are provided by local public administrations usually giving free of charge access to their local regulations. It is to be noted that in case of free access such services provide only headings of statutes.

Legislation outputs is also delivered by a number of commercial publishers, the main being  Giuffrè, with its *La legislazione italiana* on print format and *Juris data legislazione* on Cd-Rom, and UTET, with its *Lex*, available both on print and on Cd-Rom. These products show much care in information retrieval facilities such as friendly interfaces and value added functionalities based on user profile.

A particular mention is to be made of the national project NIR, Norme in Rete, an e-government project whose aim is to improve accessibility to legislation by providing a unique point of access to Italian legislation published on different web sites.[3] The NIR project, started with a feasibility study in 1999, is coordinated by the Italian Authority for Information Technology in the Public Administration in conjunction with the

---

[2] http://gazzette.comune.jesi.an.it
[3] Information on NIR Project are available on http://www.nir.it/

Ministry of Justice and the Institute of Legal Information Theory and Techniques. Its objective is the consistent retrieval of national legislation, jurisprudence and literature. It offers free of charge, about 1.500 national documents starting from 1904.[4]

The project is based on a federative approach through the cooperation among a number of producers of legislative information. Providers of statutes work together by adopting specific national standards, established within the NIR project, and using tools developed to facilitate the adoption of such standards. This allows the establishment of a stable cross reference system among legal documents guiding users towards sites of the public authorities participating in the project. The initiative's main feature is to encourage transparency and accessibility to legal information for ordinary citizens.[5] Benefits arising from the project also concern the support to public administration in managing their legislative documentation life cycle and law consolidation in order to realize simplification and interoperability.

At the operative level, the XML language which defines a standard for representing the various parts of a legal document from the formal and semantic point of view has been adopted, while a specific DTD (Document Type Definition) has been designed.

Two national working groups have been set up, aimed at producing two main official standards:

1. a standard for cross-referencing legal documents by using the uniform name (URN) technique: an unambiguous identifier that allows the identification of resources in an effective way, independently of their physical location;[6]

2. a standard for legal document description by defining XMLDTDs (NIR-DTDs) of increasing degree of depth in text hierarchy description.

## 5. Access to case law

---

[4] Ciampi C. et al. 2000. "Introduzione al progetto Norme in rete." *Informatica e diritto.* 1:1-200

[5] Cherubini M. 2002. "Verso l'armonizzazione dei portali pubblici: il Portale nazionale Italia.gov.it e Norme in Rete". *Informatica e diritto.* 1:47-110.

[6] Spinosa P. 2001. "Assegnazione dei nomi uniformi ai documenti giuridici". *Informatica e diritto.* 1: 233-264.

Case law on print version is accessible through:

a) legal journals which are mainly specialised in publishing decisions and other judicial measures such as *Il Foro italiano, La giurisprudenza italiana* and the *Giurisprudenza di merito.*

b) *Massimari,* which are collections of maxims (case abstract) of the decisions of the Italian Supreme Court: *Massimario della giustizia civile, Massimario della giurisprudenza italiana* and *Massimario del Foro italiano.*

c) *Repertori* (Digests). They annually publish the maxims of all the decisions and other judicial measures published in collections of maxims. The most popular are the *Repertorio del Foro Italiano,* the *Repertorio della giurisprudenza italiana* and the *Repertorio generale annuale di legislazione, bibliografia e giurisprudenza.*

Online retrieval of case law always refers to the Italgiure legal documentation system of the Supreme Court of Cassation (CED)[7], which is a major Italian documentation source for legal information. The service was already activated in the early 1970s, but only for a limited number of courts. From 1980 onwards, there was an incremental consolidation of the system which now offers not only access to case law, but to a number of legal literature and legislation resources. Mechanisms have been implemented to link these legal sources together so to allow navigation among several databases and the retrieval of relevant law material. The system includes 42 databases and 5 million documents and is available on charge; only employees of central administrations and judges have free access to this rich source to case law reports. From late 2003 a new search interface is available, called Italgiure Web.[8] Regarding access conditions, a distinction is made among administrative accounts and constitutional jurisprudence which are free of charge on the web sites of the various courts. Instead, ordinary jurisprudence maxims and full text, when available, are on charge.

## 6. Access to legal literature

Access to legal literature is a primary requirement and it responds to the demand for understanding and interpretation of statutes and cases, an objective that law scholars and professionals greatly contribute to.

---

[7] Information is available at http://www.giustizia.it/cassazione/infoced.htm
[8] **http://www.italgiure.giustizia.it.**

In Italy the difficulty of delimiting the scope of legal literature's sources makes its access problematic in comparison to other legal information sources. Retrieving legal doctrine necessitates a long and cumbersome research activity across multiple sources as there is no one single information provider to which legal researchers can effectively gain access.[9] Users have to identify information sources where legal literature pieces of work are found and verify the availability of legal sources through ad-hoc services.

In order that legal information can be accessed effectively, legal literature must have specific characteristics. Furthermore the performance of access services is strictly connected to political, organisational and technical factors. Legal literature's features are:

- coverage, that is completeness of information sources, although their selection is highly recommended on behalf of users;
- currency of information in relation to its production and updating;
- quality of indexing of documents and related retrieval systems provided to users.

It is well known how complex it is, if not impossible, to simultaneously reach these objectives. National bibliographic centres, in their effort to pursue the goal of bibliographic control over their national output through the provision of national multidisciplinary bibliographies, have long since faced difficulties in fulfilling their tasks. Coverage, currency and quality are their main goals and they have always tried to adapt their service policies and standards to the requests of ever increasing demanding users. The third requirement, quality of indexing, is particularly important in the law domain as to allow successful, precise and exhaustive search and retrieval, documents have to be analysed in a way that the indexing language is consistent with that used by legislators and judges, while adopting appropriate tools such as thesauri and authority lists. The production of Italian legal literature is mainly in the hands of commercial publishers, and bibliographic access tools are provided by libraries through their OPACs, by groups of academic libraries producing indexes for certain journal articles, and by a few specialised publishers.[10] To these, the national indexing service DoGi must be added. All

---

[9] Peruginelli G. 2002. **"**L'accesso alla dottrina giuridica: strumenti e linee di sviluppo in rete." *Informatica e diritto*. 1:111-176.
    [10] Catalogo delle pubblicazioni periodiche delle biblioteche aderenti a ESSPER. http://www.biblio.liuc.it:8080/biblio/essper/period.htm

these services have various interfaces and adopt different classification systems, causing troubles to legal users.

Abroad the production of legal literature is, in general, institutionally based, mainly published by universities. Mostly used access tools are OPACs that point to collections of monographs and periodicals (including remote e-journals), whereas indexing services are commercially based, accessible on subscription by users of law libraries and external users who can afford to pay high rates for such services.[11] As regards electronic legal literature's production, it is still rather scarce in Italy, whereas legal electronic resources are quite widespread in a number of countries.

## 6.1. Users of legal literature

In general, users of legal literature share the characteristics, attitudes and needs of other users in seeking information, but they also have some peculiarities due to the sophisticated nature of legal information.  In particular, they belong to different categories, professional and non-professional, with different skills in using bibliographic and indexing tools. They have different interests according to their profession and specialisation, which bring them to make various uses of legal literature materials. In particular Italian faculty members make quite an extensive use of legal doctrine documents for their teaching and research activities, followed by students preparing their thesis and by law professionals. Lawyers, judges, administrators and ordinary citizens usually start from statutes and case-law reports as their primary sources, and later search for specific legal literature items, showing a special interest in reference mechanisms allowing access to legislation and jurisprudence.

Results of a recent survey on use made of DoGi bibliographic database[12] (further illustrated in paragraph 7) by legal users has helped in understanding their behaviour in seeking information and their purpose for searching. The questionnaire aimed at identifying user profiles in order to plan services meeting their requirements. The questionnaire, which is also available in English, has been returned by about 500 users whose suggestions and indications have proved very interesting.  According to the findings these

---

[11] The most important and highly used services are Lexis-Nexis and Westlaw, which are world wide, on charge accessible databases containing TOCs, abstracts, and digital texts, whose size is growing progressively.

[12] DoGi is accessible on the web, partially free of charge, and it represents, despite some deficiencies in coverage and currency, the most important source of articles of Italian legal printed periodicals. http://nir.ittig.cnr.it/dogiswish/Index.htm

main user profiles have been identified: 1) scholars, 2) law professionals 3) citizens.

New specific requirements have emerged, such as the need for analytics of articles of pre-1970 periodicals and for the indexing of contributions included in miscellaneous monographs, the possibility for each user to have its own personal online desk to collect domain specific information and private records. Furthermore the demand for the full-text of articles indexed in the database has been proven to be a priority. This implies that the ITTIG must concentrate on favouring new forms of cooperation with commercial publishers and on improving contacts and exchange of initiatives among various indexing services in the domain of law.

## 6.2. Legal literature on the web

As regards networked legal literature, the requirements of coverage, currency, quality of indexing are still of paramount importance; to these, straightforward access to the full text has to be added. The integration between access to bibliographic information and availability of searched documents is in high demand by users and is made possible by establishing a link between secondary information records and electronic legal documents which are to be directly visible and usable at the same workstation used for searching.

Today, tools for quick, easy and enhanced communication are available, but there are some major issues of a strategic, organisational and technical nature to be considered in implementing information access services of this type.

Discontinuous and obsolescent networked information is, in fact, of low utility and runs the risk of becoming extremely expensive in economic terms, as well as making information providers and users' wasting precious time. Legal professionals are very concerned about the rapid evolution of statutes and cases' production which requires adequate services to access legal literature outputs; they recommend appropriate measures to avoid that such information become obsolete and hence useless. Therefore, a fundamental function of legal practitioners and experts in their activity of law interpretation and application is challenged, as what is needed is updated and timely information resulting from law scholars' and professionals' intellectual processing.

An additional problem is the variety of search interfaces for accessing legal databases, catalogues, web sites and other reference information sources. Differences in search and browse options and terminology used in retrieval systems may often disorient users. Quite often, even when a standard interface is used across information systems, the different functionalities offered by these systems are an obstacle to precise and consistent retrieval of legal information. Beyond the diversity of interfaces, a major issue is the difference between law classification systems, due to the system-bound nature of legal classification and terminology reflecting the various legal orders proper to each country.

Electronic legal publishing is only now starting to appear as legal professionals are generally considered a traditional user category; for this reason publishing outputs are mainly paper-based or at least on CDROM.

Despite the very active role of special libraries, mainly academic, in the provision of secondary information sources like catalogues, indexes, databases describing the material held or made accessible through them, Italian stakeholders are in general still rather prudent in venturing into the online arena; the full text of legal literature is somewhat a novelty, while quite a number of countries all over the world are able to offer a wealth of products and services based on full text accessibility. Quite often in these countries there is a strong commitment of law schools in documenting their activity and intellectual outputs, while professional associations have a major role in monitoring and certifying law periodicals' content. Furthermore, commercial publishers of legal databases, aggregators and intermediaries play an important role by offering online access to documents. Most databases are charge-based and high costs are being paid by the user community which, on the other hand, can rely on high quality legal information retrieval services. In various countries legal information providers pay particular care in providing the most complete thematic coverage as possible, in aggregating information content in specific fields of law and legal topics, and in adopting bibliographic standards.

## 7. The DoGi database: a bibliographic repository of legal literature

The DoGi database, produced and distributed by the Institute of Theory and Techniques of Legal Information is, in the Italian legal landscape, one of the most precious sources for legal literature research.[13] DoGi was created in 1970, offering abstracts of articles published in the most important

---

[13] http://nir.ittig.cnr.it/dogiswish/Index.htm.

legal periodicals (more than 250). Its main goal is to provide law scholars and professionals with exhaustive and updated information as found in Italian law reviews.

Two types of information units are accessible through the DoGi database:

> 1) a comprehensive record including: *a)* bibliographic references of each article; *b)* abstract prepared by professional indexers describing the topics treated and author's arguments; *c)* one or more classification codes identifying the subject assigned based on an analytical, hierarchical scheme; *d)* references to the significant statutes and cases' sources as cited in the article;
> 2) TOCs: bibliographic details of each article.

| | |
|---|---|
| Body | At the moment about 250 periodicals are indexed. The list of indexed periodicals is updated over time in relation to variations in publishing and to law evolution itself, aiming at maximum coverage*.* Period covered*:* 1970 to date |
| Analytics of | Articles of periodicals, summaries of cases, commentaries to statutes, conference papers, book reviews |
| Number of records | about 250.000 |
| Annual increase | 14.000 |
| Updating | bimonthly for abstracts; fortnightly for TOCs |
| Services | From selected records, by clicking on classification code and author, a search is automatically launched for those strings on the whole database
The full-text of Italian and European legislation  is retrieved through links to respectively the free-access NIR-Norme in Rete Portal and to CELEX database
Retrieval of sentences from the Constitutional Court, through a link to the Constitutional Court's information system
From selected records, by clicking on the title of a periodical, the following information is provided: a) bibliographic record of the periodical: b) holdings of Italian libraries through a link to the National Union |

|  | Catalogue ACNP; c) status and holdings of analytics since 1990. |
|--|--|
|  |  |

*This table shows the main features of the database:*

## 7.1. Search modalities

<u>Searching by field</u>

The DoGi search interface is made up of four search forms (templates). Through each of these, different channels to retrieve records are available. All requests made in fields and subfields are automatically combined through the AND Boolean operator. Searching is offered by selecting:

a) the comprehensive record;
b) bibliographic elements;
c) type of document (summary of cases, commentaries to statutes, etc.);
d) words in abstracts or summaries;
e) legal sources (citations of statutes and cases);
f) classification codes.

Figure 1: DoGi Template

The template permitting a search by source is widely used and appreciated. It allows users to make a request using specific parameters within the legal source field.

| Channel | Content |
|---|---|
| Scope *(Dropdown menu)* | Source category: national, international, European, foreign and historical  law |
| Type of legal source *(Dropdown menu)* | Legislation, cases |
| Reference Code | Type of provision (decree, statute, act, bill, sentence, convention, protocol) |
| Year, Month, Day, Region, Town, Article number | Identifying elements of provisions |
| Keyword | Other information contained in the legal source (example: title of the provision) |

### *Browse*

The following lists are available at the moment. Lists of authors and legal sources will be accessible in the next few months

- Indexed periodicals. By browsing the titles of periodicals it is possible to display the bibliographic records and the area of law covered, to examine the articles indexed in the DoGi database as well as to have information on the holdings in Italian libraries' catalogues.
- 
- Classification. By browsing classification information and clicking on codes a search is launched on the whole database and all related documents are retrieved.

*Example of a DoGi record*

| |
|---|
| D'Elia, Giuseppe |
| **La definizione dell'oggetto della delegazione legislativa nel processo democratico di formazione della legge.**<br>(***Nota a C. Cost. 16 aprile 2003, n. 125***)<br>in **Giurisprudenza costituzionale**, 2003, fasc. 3, pagg. 1763-1765<br>(Bibliografia: a pie' di pagina o nel corpo del testo) |
| **COST.4.0.4. RAPPORTI STATO-REGIONI; Funzione di indirizzo e di coordinamento**<br>**COST.2.1.5**. **DECRETI LEGISLATIVI; Legge di delegazione** |
| Con la sentenza in epigrafe, la Corte respinge un'articolata questione di legittimità costituzionale sollevata in riferimento al d.lg. n. 469 del 1997 (conferimento alle Regioni e agli enti locali di funzioni e compiti in materia di mercato di lavoro) e alla legge di delegazione n. 59 del 1997, relativamente agli artt. 70, 76 e 77 comma 1 Cost. Secondo il giudice a quo ci sarebbe un eccesso dell'oggetto della delega, nonché una troppo ampia e generica delega nella legge di delegazione. La Corte stabilisce che, ai fini della determinatezza dell'oggetto, non è necessario elencare in positivo gli oggetti della delega. In relazione alla legge di delegazione, l'A. sottolinea che, per vagliarne la legittimità costituzionale, occorre vedere se il conferimento della delega al Governo non nasconda una tecnica per estromettere minoranze e Parlamento in genere dalle scelte politiche. |
| *art. 70 Cost.*<br>*art. 76 Cost.*<br>*art. 77 comma 1 Cost.*<br>*art. 1 l. 15 marzo 1997, n. 59*<br>*d.lg. 23 dicembre 1997, n. 469*<br>*C. Cost. 16 aprile 2003, n. 125* (massima/e, testo) |

## 7.2 The DoGi database classification

The language used to index DoGi documents is a controlled language, based on the areas of law as structured in the Italian law faculty scheme. Such classification is a valid tool not only to retrieve legal literature items in the DoGi database, but also to understand the structure of Italian law. This system is of ordinary use in academic and professional contexts and many special libraries base their indexing on this classification system. There are 24 areas of law considered, each designated by a code.

The classification scheme is hierarchically structured (up to three levels) and is composed by alphanumeric codes expressing specific concepts. Codes are associated with descriptors (6.600 at the moment). An authority list of descriptors is maintained and updated on the basis of indexers' suggestions as well as of statistic analysis of searches made by users. The classification scheme has been conceived as a dynamic instrument which is periodically reviewed and new codes are established reflecting additional topics dealt with in the literature.

### 7.3. Access conditions

The DoGi database has been freely accessible from its origin, in the 1970s, to date. The ITTIG has always conceived that this repository should be available without charge as it supports a wide category of legal users irrespective of their affiliation and use made of the information contained. However since a few months, this is not possible anymore, due to the lack of support by the Government.

The ITTIG, on the basis of its long standing role as witness to Italian legal literature's evolution and of its reputation in the scientific community, has sought the support of new institutional partners to carry on a service whose production costs are very high.

To this purpose an agreement has been reached between the ITTIG and the Centro Interateneo per le Tecnologie dell'Informazione e della Comunicazione nella Ricerca e nella Didattica (CITICORD) of the University of Rome La Sapienza, within the wider context of the Legal Open Community Project. According to the agreement, from November 2004 the DoGi database is distributed on charge on the Infoleges web site, as a version integrated with (and allowing cross references to) the Infoleges files containing Italian and European Union's statutes and cases.[14]

The DoGi database is still accessible on the ITTIG web site, but only 2 documents per day can be retrieved and fully displayed (abstract, classification, sources), while a particular limitation is placed on TOCs, which although accessible for free, show only the title of the article and the name author.

The database loses an important feature: free accessibility, which has been a major feature since its origin, but is enriched in that it is integrated

---

[14] http://**www.infoleges.it.**

with other databases made available through the Infoleges service. Therefore DoGi continues to exist and progress, maintaining its features and functions.

## 7.4. Perspectives on the DoGi database

In a multilingual and multicultural environment such as the one in which we live today, functionality such as shared knowledge of legal information, quick and effective information access and exchange among different legal systems is of paramount importance. The consolidation of the European Union and its enlargement have made internalisation of many social aspects a reality, and nowadays the need for legal information communication between countries is greater then ever before. With the rapid increase of globalisation, discovery of foreign law material through adequate multilingual tools becomes essential for success in modern law practice and research. Furthermore the principle of multilingualism in the domain of law not only ensures democratic transparency and equality of citizens' rights, but also guarantees legal certainty. An adequate integration of legal concepts pertaining to different legal orders is a major requirement, but it is a very difficult goal to achieve, complicated by the diversity of languages and the various ways to express legal concepts.

In this respect the DoGi database is ready to open up its services to a wider international user community and aims to venture in the provision of multilingual access to information. For this purpose a project has been initiated to translate the DoGi classification by using English as the vehicular language allowing international users to understand the Italian law structure.

Strategic measures are also under consideration concerning cooperative ventures between European institutions which are active in the field of classification and documentation of law, in order to jointly define and verify linguistic and conceptual correspondence between legal concepts of the various systems.

This project is part of a wider initiative within the ITTIG, aimed at fostering and facilitating communication in the legal academic world, in the legal professional and business sectors and in public administration services delivered to citizens. In this context the feasibility of diverse approaches to legal and linguistic harmonisation, while preserving the legal identity of different legal orders, will be verified. The main problem of legal terminology at international level mainly regards the difference of legal concepts of the diverse national legal systems.

In this context a research study is underway concerning topics such as the relationship between language and law, identification of legal language peculiarities, significance of legal translation in comparative law, identification of methodologies and tools for shared knowledge and exchange of legal information. The translation of DoGi classification falls within such wider projects and intends to respond to the need of providing effective international access to Italian legal literature, while experimenting with methodologies for enhanced exchange of legal information.

## 8. The portal to Italian legal literature

The Institute of Theory and Techniques of Legal Information puts efforts in developing a system to ensure a unified point of access to legal literature: the Portal to legal literature.[15] The project is intended to offer a gateway to multiple legal doctrine resources, provided by European countries, to be achieved by exploiting rich metadata and by providing tools to discover, select and use relevant material.

The project intends to give access to different level of information of legal literature (bibliographic references, abstract, full text) and to retrieve the amount of available legal doctrine in electronic format (single articles, summaries of cases and statutes, conference papers, etc.) in order to guarantee searchability, usability and interoperability of digital content within the context of the existing legal framework. The creation of a unified access point, realizing integrated access to heterogeneous quality resources, must be based on academic and commercial joint ventures at European level, aimed at the collection and distribution of sources, as well as on well-established tools and services.

The Portal is indented to cover both data coming from structured data repositories (specialized legal databases, library and publisher's catalogues) and web documents (namely HTML semi-structured documents). The aim is to integrate these two different data sources in a unique view using a uniform metadata scheme such as Dublin Core metadata set. In order to harvest data from structured repositories, the OAI (Open Archive Initiative) approach is recommended.[16] This approach enables access to web-accessible material through interoperable repositories for metadata sharing, publishing and

---

[15] Francesconi E., Peruginelli G. 2003 "Integration between structured repositories and web documents." *Proccedings of the DC Conference 2003*. 99-107.
    [16] OAI – The Open Archives Initative Protocol for Metadata Harvesting, http://www.openarchives.org/OAI/openarchivesprotocol.htm

archiving to foster the proliferation of open European knowledge pools of digital objects, on behalf of education and research communities, as well as citizens. Data of structured repositories are essentially bibliographic metadata which are harvested, at service provider level, using the OAI-PMH approach and mapped from the native format into the Dublin Core metadata scheme. On such data the Portal retrieval system allows users to search by subject metadata following a semantic approach, since the semantics of these data is clearly defined.

On the other hand, web documents usually lack specific metadata as well as reliable or uniform HTML meta-tags, which can help the qualification of documents. Such kinds of data are essentially represented by natural language text, whose structure and semantics are not univocally defined.

At the moment work is underway on the enhancement of the existing solutions oriented to semantic searching, as well as to the implementation of specific facilities to support legal users in semantic querying the Portal, trying to guarantee both precision of retrieval and recall efficacy.[17] In particular, to ensure retrieval precision, the Portal aims at enriching documents with high quality metadata, so that retrieval is more focused and able to better match the semantics of the query.[18]

Moreover, to guarantee recall in retrieval, the Portal aims at matching the related information needs. To obtain this, the query has to be formulated in a way that expresses at its best the semantics of such needs. For this purpose users are to be offered facilities to construct a query, browsing a hierarchy of legal categories, as well as to expand it with broader or narrower terms. Through such expansion of an original query, users can retrieve relevant documents which did not match the query as originally formulated.

For example, let us consider a query searching for documents of the criminal law category where the term "pedophilia," correlated to the chosen legal category, is included. If relevant documents for the query do not contain the term "pedophilia," a no hits response is returned. The expansion of the query with broader terms such as "sexual offence against children" or "crime against children" and narrower terms such as "child pornography," as found

---

[17] Francesconi E., Peruginelli G. 2005. "Retrieval of Italian legal literature: a case of semantic search using legal vocabulary." *Proccedings of the DC Conference.*

[18] Jacobson T., Fusani D. 1992. "Computer, System, and Subject Knowledge in novice searching of a full-text, multifiles database." *Library & Information Science Research.* 4(1):97-106.

in the legal controlled vocabulary and related to the chosen legal category, can lead to retrieve relevant documents. This service improves the expressiveness and completeness of the query, so to match at its best user information needs, as well as to achieve recall.

In conclusion the designed project aims at providing a single point of access into disparate repositories where categories of law, as content of the field *dc:subject*, automatically generated for web resources, are the essential metadata to point to relevant legal literature documents improving precision in retrieval. As shown above, facilities in query formulation are offered to users through the exploitation of a legal controlled vocabulary, improving recall.

## 8.1. Management and organisational aspects

Management and organisational aspects pertain to both the different phases of the project and to ongoing activities.

Preparation activities include a detailed definition of project's characteristics and phases, its approval by ITTIG, and allocation of resources in terms of staff and funds. The need for special funds to be allocated must be verified, mainly for upgrading some hardware and telecommunication components and for software tools.

It is essential to set up and train a team in charge of the project, with different specialisation, such as technicians and information professionals expert in user services building, indexing and metadata matters. Their role is to:

- validate the proposed integration model;
- analyse metadata created by content producers and prepare metadata mappings, while checking the work done in this field at the international level and examining available tools,[19]
- thoroughly examine integration models, OAI specifications, ongoing international projects in this area;
- analyse the different indexing systems used by data providers and establish correspondence between classification metadata

---

[19] *See*, for example, Metadata: UKOLN software tools. http://www.ukoln.ac.uk/metadata/software-tools/

> to allow for cross subject searching (this is likely to be a recurring process during the project's life);
> ▪ investigate availability of software (open source and non).

ITTIG should commit itself to co-operate with data providers, persuading them to use standard protocols and common conventions for preparing and exposing their data.

Marketing and promotion of the new system is another important activity, to be done through the Institute's web site and through advertising in legal reviews, in official legal web sites, law libraries and law firms offices. Training and help modules are to be scheduled.

Running activities include the regular maintenance of hardware and software components of the system and periodical analysis of system performance, its availability and reliability, security of access and data. Monitoring and evaluating the system is an important activity to be undertaken once the service is implemented and tested from the strictly technical point of view. In order to test its usability a plan has to be prepared, specifying the user sample and the evaluation methodology. It is proposed to select different categories of users, reach them by electronic and ordinary mail, telephone, informing them on the novelty and importance of the service.

The project's main concept, focused on the provision of integrated access to legal literature bibliographic sources, is not revolutionary, but what seems important is the adoption, in the legal literature domain, of an open approach by different stakeholders for the benefit of legal users in their cross searching activities. A changing perspective by an institution such as ITTIG is also an important element, for its commitment in the setting up and management of a service which combines externally produced remote data sources.

## 9. Analysis of approaches for multilingual access to legal literature

Particular attention has to be paid in the Portal to multilingual access allowing cross-language retrieval of legal documents.[20] Emphasis is put on two distinct requirements, both addressing the need for international access through the legal Portal. These consist in: a) opening up the system to a wider user community, including foreign patrons, who must be given the possibility

---

[20] Francesconi E, Peruginelli G. 2004. "Opening the legal literature Portal to multilingual access." *Proccedings of the DC Conference 2004.* 100-107.

to access Italian legal material in their native language; b) providing multilingual access to foreign legal resources.

These objectives are based on the belief that the development of strategies and tools enabling access to information regardless of geographic or language barriers is a key factor for the truly global sharing of legal knowledge. With the rapid increase in globalization, transnational issues can be expected to arise today in virtually any legal context. This is the reason why the retrieval of foreign law material through adequate multilingual tools becomes an essential requirement for success in modern law practice and research. Furthermore, the principle of multilingualism in the domain of law not only ensures democratic transparency and equality of citizens' rights, but also guarantees legal certainty.

For the implementation of the Portal's multilingual access function, an analysis has been carried out of cross-language retrieval approaches, their application to law and the multilinguality issues of metadata. Consequently a practical approach to the retrieval of multilingual legal resources has been defined. Based on the features of the Portal's federation system, where documents coming from structured repositories and from the web are qualified using Dublin Core metadata set in its XML version, query translation and word disambiguation techniques have been experimented.

**9.1 Obstacles to cross language retrieval of legal documents**

There are two main obstacles in giving access to legal information in a multi-language and multicultural environment. The first one is represented by the complexity and richness of each legal language as well as by the intrinsic relationship between law and language. The second one is the difference between concepts inherent to the diverse national legal orders.

As regards the first obstacle, in each legal language there are terms which are ambiguous and which need to be contextualized. For example if we consider the Italian word "canone" and we need to translate it into the English legal system, we need to distinguish between "canone" as a rule of the Church and "canone" as the rate for lease of estates. This process is very important because it avoids lexical semantic ambiguity, imprecise and irrelevant results and makes users aware of the various contexts pertaining to the diverse legal systems.

As regards the second obstacle concerning the difference of legal concepts, there are different situations which can be envisaged when we compare two different legal orders:

- the same institution, governed in the same way. This case is extremely rare, if not non-existent;
- the same institution, governed differently;
- an institution that exists in one legal system but no longer exists in the other;
- an institution that exists in one legal system but does not exist in the other.

It is evident that it is very difficult to find effective equivalents because the attribution of an equivalence to a legal term, for which no comparable concept exists in another legal system, can be the cause of ambiguity, confusion and miscomprehension. What is necessary is a compromise in trying to favour the integration of diverse legal cultures, while respecting each national legal system: this means to find functional equivalents of legal concepts across legal systems.

## 9.2. Equivalence between different legal languages

In order that cross-language retrieval of legal material be effective, priority must be given to translating one legal language to another legal language and not to the ordinary terminology of the target language; in fact the whole process of interaction between legal languages can be identified as finding equivalents across legal systems.[21] If no acceptable equivalents can be found in the target language, subsidiary solutions must be sought, such as the use of source terms, paraphrasing, creating a neologism with explanatory notes.

The research for equivalence implies both a comparative study of the different legal systems and adequate knowledge of technical legal terminology. Some examples are provided, highlighting the complexity of comparing different legal languages.

Pure linguistic problems are likely to be encountered due to legal false friends. The terms "administrative tribunals" cannot be translated in

---

[21] J. Kerby. 1982. "La traduction juridique, un cas d'espèce, in Jean Claude Gémar (Ed.)". *Langage du droit et traduction, Essais de jurilinguistique*

French as "tribunaux administratifs." The English word for the French tribunal is Court and the administrative tribunals are administrative commissions which are comparable, *mutatis mutandis,* to the French "autorités administratives indépendantes."

The so-called "procacciatore d'affari," typical of the Italian legal system, has no equivalence in other systems. Translating it into English as "broker" or into French as "pourvoyeur d'affaires" is incorrect. Unlike what happens in Italian law, this latter expression refers to an employee of a company for which he works and from which he gets a salary and not a commission.

Despite the difficulties in establishing the equivalence of legal concepts belonging to different legal systems, a compromise has to be adopted in trying to favour the integration of diverse legal cultures, while respecting each national legal system. What is needed is the identification of a common ground, namely common legal concepts and facts which, although not perfectly coinciding with those belonging to other systems, are conceptually close. It is up to legal users, once the material has been examined, to perceive the differences and peculiarities which make these resources unique. It is to be underlined that this does not necessarily lead to noise or unsuccessful searches, but allows for a first-phase search in context, useful to give evidence of the existence or non-existence of a specific concept in the various legal systems.

### 9.3. Possible solutions

In order that multilingual access be guaranteed, that is information can be searched, retrieved and presented effectively, without constraints due to the different languages and scripts used in documents and in metadata, both the users' native language and the multiplicity and wealth of world-wide languages have to be accommodated. What is needed is functionality like the thorough and proper handling of characters (their presentation, arrangement, and transfer), putting queries into a preferred language and script, retrieving resources irrespective of the language used in searching and indexing in a word enabling world-wide communication no matter what the language. As a consequence multilingual facilities must include both multiple-language recognition and Cross-Language Information Retrieval (CLIR).[22] Among the

---

[22] Peters C., Picchi E. 1997 "Across languages, across cultures: issues in multilinguality and digital libraries." *D-Lib Magazine, May.* Available on http://www.dlib.org/dlib/may97/peters/05peters.html**.**

basic approaches to cross-language retrieval, based on controlled vocabulary and on free (or full-text) retrieval, no optimal solution exists: this is widely demonstrated by the wealth of applications, research and studies under way in the field of CLIR.[23]

The use of controlled vocabularies and of multilingual thesauri in a cross-language retrieval environment, where selected terms from each language are related to a common set of concept identifiers, involves a labour-intensive work for developing and managing such tools, especially in applications where diverse domains and material are to be managed.[24]

Free-text searching and, connected with it, methods based on dictionaries or corpora (that is analysing existing collections of texts from which to extract information for application specific-translation techniques) are an alternative approach.[25] Here either the query is translated or the document when it is indexed, but limitations exist and variations in performance occur according to the techniques used. Difficulties are represented by the lack of equivalence in translation, the ambiguity which can arise from translating terms between languages when no context is provided, and the availability of parallel and comparable corpora. This is particularly true in the domain of law.

As illustrated above, in designing the legal Portal's services the approaches to multilingual access are considered with reference to the peculiarities of legal language, which is a strictly technical language. Like other disciplines, law has its own lexicon, it uses ad-hoc terms and attributes specific meanings to terms taken from ordinary language. It is a sort of internal code allowing communication between legal experts, where terms and technical expressions respond to economic criteria, making concepts understandable by using a restricted vocabulary.[26]

---

[23] Mayfield J., McNamee P. 2002. "Three principles to guide CLIR research". Available on http://ucdata.berkeley.edu/sigir-2002/sigir2002CLIR-18-mayfield.pdf**.**

[24] Fluhr C. 1996. "Multilingual Information Retrieval." *Survey of the State of the Art in Human Language Technology.,* Available on http://cslu.cse.ogi.edu/HLTsurvey/ch8node7.html**.**

[25] Oard W. D. 1997. "Alternative approaches for Cross-Language Text Retrieval." Available on http://www.ee.umd.edu/medlab/filter/sss/papers/oard/paper.html.

[26] Van Laer C.J.P. 1998. "The applicability of comparative concepts." *Electronic Journal of Comparative Law*, 2(2).

All this applies at a national level. At international level the complexity and richness of diverse legal languages make understanding and exchange of concepts expressed in such languages a very difficult task. However, the main problem of legal terminology at an international level mainly regards the difference of legal concepts inherent to the diverse national legal systems. In fact merely translating words is likely to confuse users when a concept does not exist in the target legal system.[27]

A basic difference between legal systems lies in the fact that, speaking in very general terms and without taking into account mixed legal systems, the Western world has long had two dominant legal traditions: common law, with its beginnings in England, and civil law, rooted in continental Europe.[28] At this stage of the project the Portal mainly deals with material drawn from these two basic systems. Both systems have taken on a variety of cultural forms. Civil law systems have drawn their inspiration largely from their Roman law heritage and, by giving precedence to written law, have resolutely opted for a systematic codification of their general law.

Common law systems are instead based on English common law concepts and legal organizational methods which assign a pre-eminent position to case-law, as opposed to legislation, as the ordinary means of expression of general law.

Quite a number of concepts belonging to different legal systems can hardly be transposed or even compared with one another. Unlike technical and scientific disciplines, serious difficulties arise in translating law material due to the system-bound nature of legal terminology. In fact each legal system is situated within a complex social and political framework which responds to the history, uses and habits of a particular group. This complex framework is seldom identical from one country to another, even though the origins of the respective legal systems may have points in common.

Consequent to this, categories of law (such as trade law, constitutional law, or criminal law) are specific within a particular legal system and for this reason the categories of law of a specific legal system represent the way that retrieval can be satisfactorily achieved. At this stage of

---

[27] Sacco R., 1998. "Droit et langue." *Rapports italiens au XV Congrès international de droit comparé.*

[28] However no modern legal system is a "pure" representation of any type of system; all jurisdictions today represent mixed systems, to some extent.

the study we consider a one to one mapping between a specific legal system and a given language.

As often there is no conceptual nor content similarity between the categories of law of the different legal systems, a mapping between such law categories is necessary to reach proper contextualisation of the query in the diverse legal systems. An example illustrates the need for such mapping. The concepts related to property rights, such as the development of property law, land law, property questions on insolvency, intellectual property, etc. according to UK law belong to the field of property law, whereas in the Italian legal system these legal facts are regulated respectively by private law, agricultural law and industrial law.

## 10. Perspective in legal research in Italy

With the increasing complexity of the law and its sources, what is needed is the provision of adequate information retrieval systems able to support qualified and versatile legal users. As stated above no unique information system on law exists in Italy where scarce is the cooperation between public and private providers in making legal information accessible.

Surveys have shown that legal users need relevant, precise, timely information; they must be informed on available products, be trained how to use new resources and how to evaluate their quality. For example, the lack of availability of adequate and current legal information at the immediate disposal of lawyers can lead to endless adjournment and prolonged litigation.

The assessment of quality of web resources is of paramount importance. Methods are needed for their evaluation and for coping with the ever-increasing number of legal material available on the web. The presence of a librarian or information professional in law firms is extremely rare in Italy and quite often this results in a time consuming searching activity by lawyers themselves on legal and judicial traditional directories.
When used, legal databases are often only partially exploited and the richness and complexity of facilities offered are missed.

What is needed is first of all a liaison between public and private providers: the first ones have the content, the latter offer the platform to distribute such content.

Secondly, it is of paramount importance to plan legal research courses to permit legal users to become acquainted with basic and advanced legal

research (core legal skills in using modern hand-held technology, common computer applications and use of the web), helping them in the research problems within the concrete experience of legal questions.

Research findings and personal experience confirm that law is an information-intensive discipline and practice where accuracy and currency of information sources to support legal users are vital.[29]

Surely research skills are continuously improved by experiences but there has to be some foundation to start. It is a matter of fact that in Italy the need for legal research learning is not always clearly perceived by the various categories of legal users. Teaching of legal research is far from optimal and although legal research is a subject included in a number of Italian universities' law courses, at present it is not mandatory, and so far it has been confined within the broader field of philosophy of law.

Some training in legal research is provided by Italian academic libraries: these courses, although offered to law students on a general basis, do not receive much attention and are mostly seen as instrumental for thesis preparation. The academic sector seems not to favour the inclusion of this subject along with the ordinary fields of law, considering it not as a discrete subject in itself, but an application of information sciences to the object of law.[30]

The low consideration that legal research teaching has received in Italy so far and the consequent lack of adequate information skills of the majority of lawyers, as compared to their actual needs, are the reason for proposing ad-hoc courses for legal users. The idea is to make legal users aware of the current available tools, with special attention to electronic resources, of content and bibliographic legal databases and related services, providing them with skills in using them effectively and applying research strategy techniques on the basis of their legal problems; after all allowing them to make the most of legal information sources in a quick and effective way.

---

[29] Tunkel V. 1997 "Law finding for lawyers." *The Law Librarian,* 28(2):8-10; Leckie G., Pettigrew K. E., and Sylvain C. 1996. "Modelling the information seeking of professionals: a general model derived from research on engineers, healthcare professionals and lawyers." *Library quarterly.* 66:161-193; Oppenheim C. 2001. "Judged to be lacking." *Information world review.* 169:10.

[30] Taddei Elmi G. 2003. "Corso di informatica giuridica". Napoli, Edizioni giuridiche Simone.

There is a trend towards an increase in specialisation because the volume and complexity of law multiplies, and statutes, regulations and related judicial reports are created at local, national, and international level. This also results in increased competition from greater globalisation, and legal professionals have to face the market equipped with high standard professional expertise, where knowledge and skills in conducting professional research are essential prerequisites.