

Legal knowledge based systems
JURIX 92
Information Technology and Law

The Foundation for Legal Knowledge Systems

Editors:

C.A.F.M. Grütters

J.A.P.J. Breuker

H.J. Van den Herik

A.H.J. Schmidt

C.N.J. De Vey Mestdagh

R.M. di Giorgi and R. Nannucci, A Legal Hypertext System, in: C.A.F.M. Grütters, J.A.P.J. Breuker, H.J. Van den Herik, A.H.J. Schmidt, C.N.J. De Vey Mestdagh (eds.), Legal knowledge based systems JURIX 92: Information Technology and Law, The Foundation for Legal Knowledge Systems, Lelystad: Koninklijke Vermande, pp. 171-178, 1994 ISBN 90 5458 031 3.

More information about the JURIX foundation and its activities can be obtained by contacting the JURIX secretariat:



Mr. C.N.J. de Vey Mestdagh
University of Groningen, Faculty of Law
Oude Kijk in 't Jatstraat 26
P.O. Box 716
9700 AS Groningen
Tel: +31 50 3635790/5433
Fax: +31 50 3635603
Email: sesam@rechten.rug.nl

A LEGAL HYPERTEXT SYSTEM

R.M. DI GIORGI and R. NANNUCCI

Istituto per la Documentazione Giuridica, Consiglio Nazionale delle Ricerche, Firenze, Italy

Summary

At present legal operators have at their disposal many legal data banks documenting international and EC regulations, national and regional legislation, case law of the different courts and legal doctrine. This information is organized in non-standard documentary structures and has to be consulted in separate data banks utilizing different information retrieval systems. To face the problem of the consultation of these data banks, a joint project was launched to experiment the application of a hypertext system on a collection of legal documents including legislation, case law and legal doctrine, relating to the specific sector of environmental law. The hypertext approach is used to facilitate access to non homogeneous legal data not only for experts such as lawyers but also for non-expert users including public officials, members of environmental protection associations, economic operators and citizens.

1 . Introduction

Information in printed works is usually organized in a sequential way, in chapters or paragraphs, with a variety of indexes to allow more immediate retrieval. Besides this kind of approach, however, another type of retrieval can be identified which consists of cross references leading the reader from one section to another of the printed work, that is a logical structure between documents, a sort of metalevel network each reader can move along, on the base of his personal associations of ideas related to his own level of knowledge of a particular discipline.

One may argue whether this kind of retrieval can be reproduced in automated information systems and whether it may be possible to organize databases so to obtain a non sequential and semantically not too rigid retrieval, that is, to search data through conceptual associations.

The possibility of moving down from a reference to another with great liberty is of particular relevance in certain disciplines. When searching legal documentation , for instance, some problems arise which traditional information retrieval systems cannot afford: references between legal norms, between case law and norms, between legal doctrine, case law and legislation cannot be accessed.

Hypertext systems appear to be able to meet these demands for they allow databases to be structured on conceptual associations.

In the present paper the basic concepts of hypertext technology will be briefly described, focusing on its differences with respect to traditional information retrieval; the special and specific information needs of legal professionals will be analysed and a short presentation of Hyperlaw will be given, a hypertext system developed at the Istituto per la Documentazione Giuridica of the Italian National Research Council in Florence in collaboration with the University Department of Electronics and Informatics in Padua.

2 . Beyond traditional information retrieval

In automated legal documentation systems the most common approach to information retrieval is through boolean operators. Due to the peculiarities of legal documentation, many different issues arise in automated retrieval, such as the considerable volume of texts which is greater than in many other application domains. Most sources are stored in full texts and this implies long connection time to terminals. So, legal operators heed to specify the topic of interest exactly and to avoid approximations in query formulation as much as possible since it would mean otherwise wasting too much time and resources on-line [Davis, 1986].

It is, therefore, essential to be able to retrieve relevant data concerning a particular search through an efficient system. The legal user needs an information retrieval system based on efficient semantic structures capable of analysing his request, comparing it with the database documents and identifying those texts or subparts of them which are relevant to his search. Practically, the overall information power of data has so far been very restricted. Without a semantic approach inducing associative links, most pertinent information is often lost. Where there is no conceptual structuring of the application domain terms (thesauri, classification table, etc.) leading to specifically domain-oriented questions, the users might be confronted with a loss of information due to the non-handling of synonyms or with a certain degree of superabundance due to an extensive or ambiguous use of words - something which is very common in legal texts [Lloyd, 1986] [Bing, 1986].

Software and data banks producers are perfectly aware of these difficulties and have been attempting to limit them to a certain extent by building interfaces to facilitate man/machine communication. To date, however, such problems continue to exist and become particularly relevant when the computer is seldom used (it is quite commonplace to forget instructions and commands). They tend to multiply when it is necessary to consult various data banks, each with its own features and controlled by different systems.

3 . Hypertext technology

In the light of these considerations, new approaches can be tried which permit the user to venture more effectively into the area of automated legal documentation. We are referring to systems able to furnish the user with documents structured in a kind of semantic network, that is, a network of references between one document and another, within which each document fully acquires its individual meaning [Conklin, 1987][Swift, 1991]. The need to integrate even very different kinds of documents (legislation, case law, legal doctrine) is, therefore, a priority and a radical change of perspective in automated documentation is required where legal information is stored using the criterion of uniform documentary types.

This appears to be what can be accomplished by using hypertext technology.

Hypertext systems are capable of managing the full text of documents or parts of them as well as their relative network of relations. The document collection can be explored by utilizing very different associative criteria, thanks to the possibility of defining various kinds of connections (links) between the documents when building the system [Brown, 1988].

The information power provided by a computer system of this nature seems to be considerable: it is a software environment which tries to imitate the part of the human mind which stores and retrieves information through associative and/or intuitive relations [Ching-Lung, 1991][Urr, 1991].

3.1. Hypertext network

Current hypertext systems do not use the typical database record and file structures: their databases consist of physical work spaces (nodes), which can be filled with text, graphics, images or audio data. The links employed to connect the nodes operate according to the method currently adopted when there are notes, citations or phrases in brackets in printed texts: they are, therefore, the means for transporting the user from one node to another in the documentary base. The reader, through conceptual associations, is referred to the relative texts or to additional information [Kornbrot & MacLeod, 1990].

Inverted files of words, phrases or key words can be created and boolean or word searches can be carried out. Some programs allow hierarchic indices, such as tables of contents, to be created, whilst others permit indices based on semantic content, like thesauri, to be constructed.

When considering a thesaurus as a network of links of a hypertext system, this new technology seems to be very promising for the further development of information retrieval. Each node of the hypertext network could, in fact, be a term of the thesaurus which leads to the relative network of documents semantically linked according to typically thesaurus-type relations, namely, relations of preference, hierarchy and association [Hall & Papadopoulos, 1990] [Agosti et al., 1989a].

4. Hyperlaw, a hypertext system in environmental law

4.1. Aims of the project

The main objective of the Istituto per la documentazione giuridica of the Italian National Research Council in designing the hypertext prototype Hyperlaw was to offer the user of legal documentation systems modes of access directed towards his information needs and systems that enable him to retrieve the most complete information on specific domain in real time [Di Giorgi et al., 1989].

The originality of the hypertext system prototype developed at the Istituto per la documentazione giuridica is that it integrates a collection of unhomogeneous documents (legislation, case law and legal doctrine) connected to each other by the topic: Environmental Law in the specific field of Noise and Vibration.

The Institute designed the prototype in collaboration with the Department of Electronics and Information Technology of the University of Padua.

The Department was specifically responsible for the actual design and development of the system prototype, utilising a hypertext software (Hypercard by Apple Computer) that would meet the needs of this particular application domain, the law. It possesses a very powerful object-oriented language, HyperTalk, which allows data to be easily manipulated and also managed in a structured way [Agosti et al., 1989b].

The Institute provided the collection of pertinent documents and created an ad hoc classification table that would be useful for gaining access to the system, also indexing the material according to particular descriptors. It was also responsible for developing the links between the content of the documents, relevant for lawyers, that then enable them to use the semantic structure for retrieving the information relating to a specific argument.

The classification table was organized in topics into which the domain "environment" is usually subdivided. Each class, divided in subclasses, contains a set of descriptor terms.

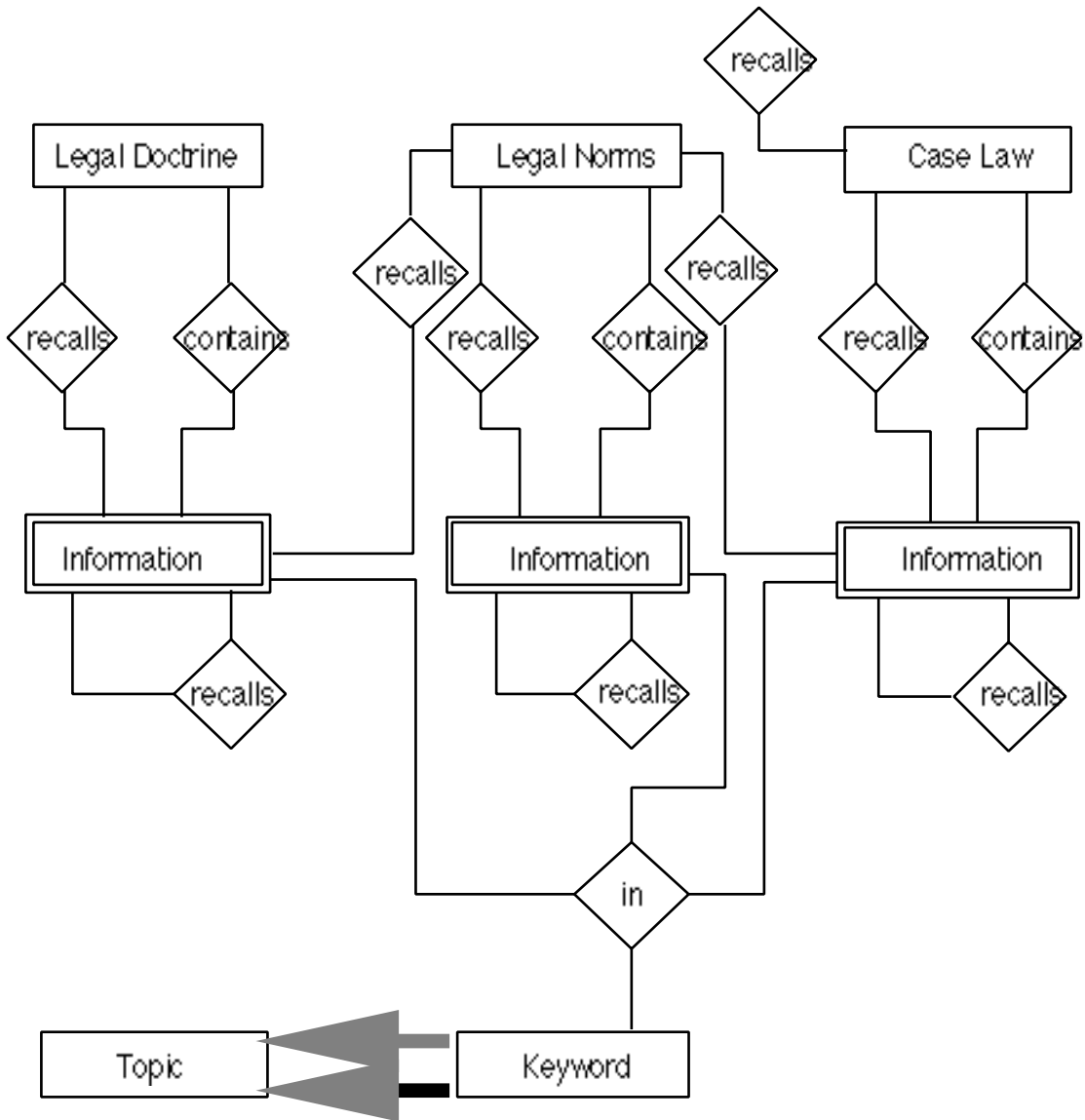


Figure 1: Conceptual Framework

4.2. User requirements

Hyperlaw is specifically directed to three types of legal users: it is meant to facing the needs of legal practitioners such as judges, barristers and solicitors, professionals in general, public servants and officials who use legal documentation for problem-solving or decision-making. It is also meant to reply to academics and theorists (such as university professors and students or researchers) who need to go deeper into general or specific aspects of their teaching/training function and/or to work out new theories or interpretations. It is also useful to non-expert users (individuals, companies and trade unions) as a support tool for identifying action strategies in specific areas.

4.3. System architecture

4.3.1. Hyperdocument and hyperconcept

The system architecture is based on a two-level abstraction. The first level is represented by the hyperdocument, that is, the hypertext which manages the documentary base. Each node of the hyperdocument represents a document in full text or a part of it. A network of structural and semantic links connect the information units.

The second level is represented by the hyperconcept, a parallel hypertext, which manages the semantic structure of concepts used for describing the documents' contents. This level, which is conceptually over the hyperdocument, operates as the thesauri or classification tables used in traditional information retrieval systems. Each node represents a concept or an aspect of it relevant to the information content of documents in the hyperdocument. Nodes are connected by links describing the various semantic relations between concepts.

Each level of this architecture represents a distinct network of nodes and links. Links specifically describe the relations between semantic terms in the hyperconcept and the relevant documents in the hyperdocument.

A controlled vocabulary allows "navigational" access to the hypertext documentary base. The hyperconcept represents the conceptual reference scheme the user utilizes to meet his own information needs. The user searching for semantically close documents moves along the hyperconcept structure up to the point when he succeeds in activating a link connecting the two parallel hypertext systems. He is, thus, brought to the hyperdocument, or, in other words, to the documents he needs.

This kind of architecture offers many advantages, such as the possibility of building a variety of hyperconceptual files on the same hyperdocument: the documents can be described in different semantic ways, that is, specifically oriented to different user groups.

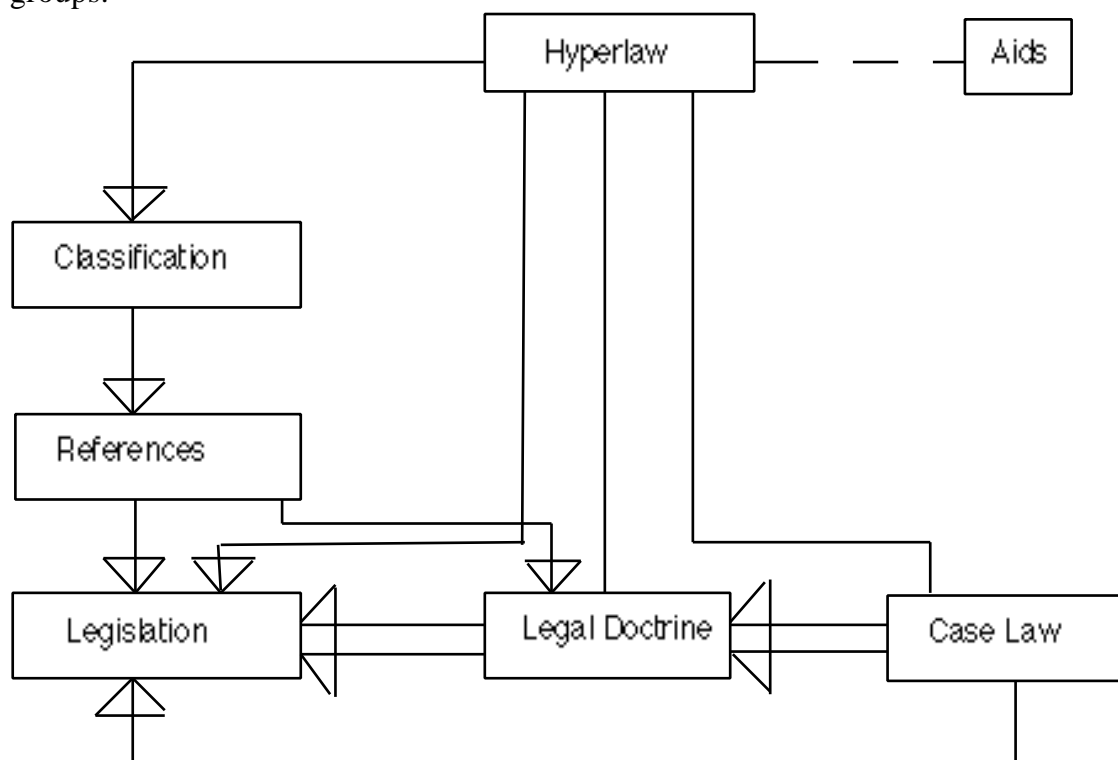


Figure 2: System Architecture

4.3.2. System database

The hyperdocuments of the prototype consist of three different types of documents: legislation (international agreements, directives or community legislation, national or regional law), case law (decisions of different Courts), and legal doctrine (monographs, proceedings and articles from journals).

The access to any specific collection of documents is achieved simply by "clicking" the relative icon present in the screen with the mouse button. For example, by clicking the legislation icon the user can gain access to the collection of laws, within which he can shift around and retrieve the various normative texts connected each with the other by appropriate links (direct access).

From the retrieved documents, furthermore, it is possible to move horizontally to the connected documents of the other two collections (case law and legal doctrine), by activating the related links. These can be distinguished into active or passive links. The former represent connections starting from a certain node (on the screen) and leading to its related nodes in the databases. Inversely, the latter indicate the connection starting from other nodes in the three collections of documents and leading to the considered node (reference access).

The most original function of Hyperlaw is that it provides access through a semantic route. The present version of the prototype contains a classification table, with several hierarchical level, whose lower terms, corresponding to the bottom branches of the structure, constitute the points of access to the documents. In this case, the user activates the icon relating to the classification table.

The list of headings representing the first hierarchical level, is, therefore, displayed. He points the mouse on the heading corresponding to his area of interest. According to the choice of he has made, the user can obtain the list of the term stored under that heading or else the display of a second hierarchical level.

By activating the term the user has access to an intermediate card, that enables him to decide on access to the documents represented semantically by that term. He can, in fact, opt for separate entry into the different files or he can choose to concentrate on only one documentary collection. The number of group of documents linked to the term appears at the bottom of the screen.

4.3.3. H-import procedure

The development of Hyperlaw takes advantage of a special external procedure, called H-Import, comprising several functions useful for increasing the application, namely: the importation of new documents, the insertion of new semantic terms in the classification table and the creation of new links.

The procedure permits new information units to be introduced into the hypertext network, by loading them from an external file. By means of H-Insert, a specially designed data description language, a new node and its relative links may automatically be created. This function is particularly valuable for managing information evolving dynamically overtime. The possibility to automatically link the new document solves the essential problem in managing a hypertext, namely the work load required for creating the links.

5. Future prospects

It is obvious that integrating the documentation relating to all branches of law seems, at present, an extremely ambitious project. One of the most common problems arising from

the use of hypertext technology is that of "getting lost" within the network of links in both the construction and search phases.

Unless such a difficulty is overcome by further technological development, the use of hypertext systems will probably be limited to specific sectors of a discipline, confining the subject matter and its related documentation to more easily manageable dimensions.

Even if this constitutes a considerable disadvantage to the use of hypertext systems, they, nevertheless, lend themselves to diversified information needs, thanks to the plurality and variety of links between documents they permit and to the greater flexibility in searching they offer as compared with traditional information retrieval systems. In fact, the management of the links allows search paths to be created specifically oriented to user groups with different qualifications and, therefore, different interests.

6. References

This article may only be cited as: Di Giorgi, R.M. and R.Nannucci, A Legal Hypertext System Prototype. In: Grütters, C.A.F.M., J.A.P.J. Breuker, H.J. van den Herik, A.H.J. Schmidt and C.N.J. de Vey Mestdagh (eds), *Legal Knowledge Based Systems: Information Technology & Law, JURIX'92*, Koninklijke Vermande, Lelystad, NL, 1992.

- [Agosti et al., 1989a] Agosti, M., G. Gradenigo, P. Mattiello, The Hypertext as an Effective Information Retrieval Tool for the Final User. In: *Pre-Proceeding of the Third International Conference on Logic, Informatics, Law*, Florence, IDG/CNR, 1989, vol. 1, pp. 1-15.
- [Agosti et al., 1989b] Agosti, M., M. Archi, R. Colotti, R.M. Di Giorgi, G. Gradenigo, B. Inghirami, P. Mattiello, R. Nannucci and R. Ragona, New Perspectives in Information Retrieval Techniques: A Hypertext Prototype in Environmental Law. In: *Proceedings of the Conference on "Online Information 89"*, Learned Information, Oxford, 1989, pp. 483-494.
- [Bing, 1986] Bing, J., Legal Text Retrieval Systems: The Unsatisfactory State of the Art. In: *Journal of Law and Information Science*, vol. 2, 1986, no. 1, pp. 1-17.
- [Brown, 1988] Brown, P., Hypertexts: The Way Forward. In: Vliet, H. van (ed.), *Document Manipulation and Typography*, Cambridge University Press, 1988.
- [Ching-Lung, 1991] Ching-Lung, W., Hypertext and Printed Materials: Some Similarities and Differences. In: *Educational Technology*, vol. 31, 1991, 3, pp. 51-53.
- [Colotti, 1991] Colotti, R., Hyperlaw: prototipo ipertestuale in ambito giuridico. In: *Informatica oggi*, vol. 11, 1991, no. 71, pp. 68-73 and vol. 11, 1991, no. 72, pp. 67-72.
- [Conklin, 1987] Conklin, J., Hypertext: An Introduction and Survey. In: *Computer*, vol. 2, 1987, no. 9, pp. 17-41.
- [Davis, 1986] Davis, D., Semantic Analysis in Legal Text Information Retrieval. In: Martino, A.A. e F. Socci (eds), *Automated Analysis of Legal Texts*, Amsterdam, North-Holland, 1986, pp. 473-481.
- [Di Giorgi et al., 1989] Di Giorgi, R.M., B. Inghirami, M. Ragona, R. Nannucci and A. Archi, Hypertext for Legal Documentation. In: *Pre-Proceedings of the Third International Conference on "Logic, Informatics, Law"*, Florence, IDG/CNR, 1989, vol. 2, pp. 283-291.
- [Hall & Papadopoulos, 1990] Hall, P.A.V. and S. Papadopoulos, HYpertext Systems and Applications. In: *Information and Software Technology*, vol. 32, 1990, no. 7, pp. 477-490.
- [Kornbrot & MacLeod, 1990] Kornbrot, D. and M. Macleod, Monitoring and Analysis of Hypermedia Navigation. In: *Proceedings of the XIII Conference on "Human-Computer Interaction"*, Amsterdam, North-Holland, 1990, pp. 401-406.
- [Lloyd, 1986] Lloyd, M., *Legal Databases in Europe. User Attitudes and Suppliers Strategies*, Amsterdam, North-Holland, 1986.
- [Swift, 1991] Swift, M.K., Hypertext: A Tool for Knowledge Transfer. In: *Journal of Systems Management*, vol. 42, 1991, no. 6, pp. 35-37.
- [Urr, 1991] Urr, C., Will the Real Hypertext Please Stand Up? In: *Computer Library*, vol. 11, 1991, no. 5, pp. 45-49.

