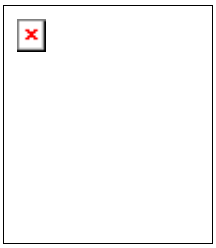

10th BILETA Conference Electronic Communications

March 30th & March 31st, 1995 Business School, University of Strathclyde, Glasgow

Legal Resources Via World Wide Web

Franco Zizzo, University of Oxford

Introduction

Data equivalent to 50 billion pages is transported every month through the various Internet networks. With other words, the data amounting to ca. 25 million books is transported through the Internet each month, equivalent to 5 times more books than the content of the Bodleian Library at Oxford. Through powerful software, the provision of information on the Internet is becoming possible for everyone, without great efforts. The development of the World Wide Web (WEB), a universal hypermedia-based method to access information, underlined a new departure in the provision of information on the Internet. Sounds, images, videoclips, clickable maps and much more can now be used by creating documents. This impressive development led to the growth of the Internet to a business worth billions, which is going to be used not only by the *sociatas academica* but by the whole society. In the USA the number of ".com" (commercial) addresses exceeded for the first time the number of ".edu" (educational) addresses in 1994. We are entering a new age of information technology. Naturally, lawyers have to play a leading part in the story. This paper analyses the power of the WEB in providing legal information having regard to the Internet problem of information chaos. It deals with how the efforts we are all making in providing this information can be co-ordinated, and what kind of co-operation is required to provide qualitative legal information which is structured and thus quickly searchable.

Section 1: Legal information on the World Wide Web, so far

1. What is the WEB?

For the purpose of this paper it may be useful to explain briefly what the WEB is. The WEB was developed in the early 1990s by researchers at the European Particle Laboratory in Geneva. It is based on the server/client principle(1). Two reasons may explain why the WEB is extremely successful and destined to become the most successful Internet facility (2). Firstly, it is designed around a system known as hypertext. A WEB document is defined through a collection of styles which are indicated by markup tags. Such documents are called HyperText Markup Language (HTML) documents. HTML documents are in ASCII format and can be created using HTML editors(3) or any other text editor. Hypertext documents use marked words or sentences which are linked to other documents via an Uniform Resource Locator (URL). The URL specifies the location of files on other servers and includes the type of resource being accessed, the address of the server and the location file. Thus a single HTML file on a local WEB server may include as many links to files on other Internet servers as are wished, required, or available(4). Secondly, the WEB is able to understand the numerous different information-retrieval protocols in use on the Internet today. For example it understands protocols like FTP, Gopher, Telnet, WAIS etc., but also data formats of protocols like ASCII, GIF, Postscript, DVI, TEXinfo etc. The WEB thus provides a single consistent user interface. The result is, on the one hand, that WEB clients (e.g. Mosaic, Netscape) can access almost every other Internet facility. On the other hand HTML documents can be created, including data such as graphics, pictures, sounds or videoclips.

On the whole the WEB and its software are in an early stage of development. It was developed by computer scientists whose primary interest was to fulfil their ideas and visions of an open information architecture able to connect millions of users. The openness of the Internet would not have been possible outside the world of research and education. Now several commercial companies have successfully started to develop and sell more powerful WEB clients/server software at a considerable price(5).

2. Structuring legal information provided on the WEB

The Internet is growing fast. In 1993 the amount of traffic data between the various networks on the Internet was 2.5 billion pages per month. By the end of 1994 this had doubled to 5 billion pages per month(6). This is ca. 17 times more information than the LEXIS-NEXIS legal database provides(7). The amount of HTML documents related to legal topics is already immense. All the legal documents previously provided by Gopher are accessible through the WEB. Recent developments and projects on the WEB underline the fact that a new departure in provision of legal information is seen in the WEB. One of the most ambitious WEB projects in the field of libraries comes from the Columbia University Law Library, with the aim to *"utilize the power of massively parallel super computers to provide text, images, sound and*

videos from remote and local workstations"(8) . WEB legal journals have so far also joined, for instance, the "Global Legal Studies Journal",(9) with a full-text search on all back issues. The "Roman Law in the WEB" project underlines a new departure in the use of non-copyright legal texts. Its aim is to provide the basic text of the Roman law, the Corpus Iuris Iustiniani, with the corresponding medieval comments. These comments are widespread and many variants exist so that normally their study takes scholars to several libraries often in different countries(10). In the USA law firms are presently on the WEB introducing their organizations and principal practice areas. Firm documents to cases of interest can also be accessed and searched with a WAIS searchable index(11). LawTalk at Indiana University is a service for playing files related to some aspects of law or legal studies. A few files are already available and authored by law faculty members(12).

In front of such developments the crucial question is how do we find the legal information we are looking for. Since the amount of data is immense, and increases every day, powerful search possibilities and an intelligent data structure are necessary, otherwise the information will, on the whole, be useless. An approach akin to commercial legal databases is required.

2.1. The technical approach

What computer scientists can do is devise services which administer the information traffic and search the various databanks. The WEB's predecessor, Gopher, is a versatile menu-driven information service. Within this service, Veronica searches automatically for indexed documents. WAIS is a service which allows users an "intelligent" search for information among databases distributed throughout the Internet. The WEB integrates all these possibilities, and is a universal hypermedia-based method of access to information. The search engine on the WEB is already powerful(13) and many Internet research efforts are made to improve the search possibilities(14). But it is unlikely that a service will be devised whose search covers the total number of documents stored on the Internet. There are simply too many databases and documents. We are, at the moment, far away from the commercial database system. Every document stored into a commercial database is added to the proper subdatabase, is indexed, and therefore, in the proper search context, is easy to find. We will naturally have to go a different way than commercial databases with a WEB database, at least because it implements a large variety of features. Search engines will be indispensable tools within a sophisticated structure providing a proper framework. The main problem is what a sophisticated structure should look like. We will consider below the efforts made so far to give some structure to the various legal information available on the WEB.

2.2. The Law Faculties approach

Though there is no "top", the WEB has a certain kind of roof, a "general overview" available at its home, CERN. This is an attempt to structure the whole available information by subject, list of servers and service type(15). Under the subject of 'Law', we find the Law Index held by Indiana Law School (USA)(16). The Law Index was the earliest attempt to bring structure into the provision of legal information. First, everything found with a relation to law was added to an alphabetical list, including documents stored on the Gopher system as well as FTP legal sites etc. Over time, perhaps almost all known legal information (free of charge) was added to this list. With the choice, for example of the link 'Belfast Law School', a small server I built up myself to search the potential of the WEB for legal purposes, more than 100 HTML original files can be found on local legal information. At this stage searchable engines were added on the top of the Law Index, as well as part of the information available alphabetically structured under the heading 'legal topics', 'type of source', 'State Governments servers' and 'law schools and law firms on the WEB'.

Further development in 1994 was to specify the information available on the Law Index. The Cornell law school project differed between administration law, commercial law, criminal law, articles, collection, organizations etc.(17). The P-Law Legal resource Locator, a recent project which joined the WEB (Jan 95), is going in the same direction without any obvious new idea, and with the aim "to serve as a starting point for research, be it by attorneys or non-legal professionals, into the myriad disciplines involved in today's increasingly 'legalistic' society"(18). JurWeb, also a recent interesting project at Bayreuth (D), has the aim to structure the information on the Law Index by continents, countries and legal topics(19).

Despite these attempts to give the available legal information a searchable structure the reality is that a normal WEB search involves a considerable waste of time. The fact that at the moment many projects are trying to differentiate the information available on the Law Index creates unnecessary confusion. In addition there is no certainty whether the available information has a sufficient standard of quality. Generally speaking, a co-ordination system is required which at least is able to give the provided information a quick searchable structure and which makes sure that the information is of sufficient standard. Before a new method able to meet these requirements in structuring legal information can be discussed in this paper, general principles behind the provision of legal information via WEB will be considered.

3. Principles behind WEB legal servers

Though more and more law faculties are joining the Internet with WEB servers(20), it will be sufficient to analyze the

principles of some older servers. The oldest European server is located at Saarbrucken(21). It is promoted and supervised by the only exclusive "Computers & Law" chair I know of in Europe and had as its precedent a Gopher server. The aim of this project is to provide a future new legal information source, including local dissertations, publications of the European Institute, as well as articles of members of staff, etc. Over time Saarbrucken wants to become a unique database in Germany. Tromsoe in Norway(22) has as its aim *"to investigate the potential of WEB as an information resource, with regard to legal research and education. This we plan to do taking a practical example, focusing on international trade law as a limited and vitally important area of law that is of international interest"*. At Belfast there was the clear aim to provide local legal information about staff, courses, graduate facilities, introductions, etc.

So far two different principles can be seen. Belfast and Tromsoe have defined their aim very narrowly; one on a limited legal field of international importance, the other on local information. This is not that ambiguous, but a realistic approach. Nevertheless a narrow approach outside a framework is likely to be overlooked and is thus of limited use. For example, nobody searching for information on human rights can know that a database on this topic that covers the whole world, and was developed over a number of years, is accessible through the link 'Belfast law school'. The other principle is to build up over time a unique database, something on the line of LEXIS. This approach is quite unrealistic since one chair will never be able to provide comprehensive and qualitative legal information like a commercial database. The Law Index at Indiana law school has an impressive amount of information, but this is the sum of the efforts of the whole Internet legal community. During a controversial discussion about the usefulness of legal information on the Internet in the e-mail discussion list EURO-LEX(23) Saarbrucken admitted that they are not making any considerable step toward their aims and that we would probably have to wait for the *"World Wide Web commercial wave"*. These examples show clearly that any principle in providing legal information must include a co-operation among people which are providing the information. The USA law faculties, many years ahead of Europe, suggested early on to establish *"among those of us maintaining legal oriented WEB sites"* (24) a communication and a co-ordination of the efforts which are made in putting information on the WEB. None of the many attempts like this have so far been successful. The striking fact is that the numerous USA law faculties have not reached a considerable co-ordination between each other . A reason may be the influential lobby of the printed legal press and commercial legal databases who may fear the WEB as a dangerous competitor. Antitrust investigations are taking place, e.g. in the case of WEST publishing, after a battle erupted over the Attorney General's proposal for a public domain citation system and databases of judicial opinions(25). On the other hand we have also to consider that computer scientists employed with a full job by the law faculties in the USA are doing most of the work. Computer scientists are naturally primarily interested (and trained) in exploiting and applying technology rather than establishing a workable framework on a non-local level.

A principle on providing useful legal information can only be successful if a framework between law faculties can be created. In the following, a concept will be presented which on the one hand will provide a law faculties framework through co-operation, and on the other hand a workable comprehensive information structure.

Section 2 : Towards a Legal Information Standard

1. The status quo

Before we can consider a new way of co-ordinating legal information which involves a considerable amount of efforts, we should once again sum up the advantages of legal information via WEB as well as possible alternatives.

The indisputable advantages of the WEB are an immense amount of legal information, covering almost every legal topic, accessible at all times and from any location, and the dimension of teaching through multimedia documents. Powerful Local Area Networks (LAN) are available at almost every university. The fixed costs of a LAN are large, but the marginal cost is effectively free. Whether 1 or 1000 persons are using the provided infrastructure is therefore not of great cost relevance for a LAN administration. Thus the connection as well as the provision of legal information for any faculty is available for a relatively low investment(26). Finally, the power of various world wide scientific electronic fora for the exchange of scientific ideas, and legal research must be mentioned. On the other hand, the vast amount of unstructured information and its unpredictable quality makes Internet legal resources almost useless at this stage. As we have seen before in this paper, a co-ordination system is required which is able to give the provided legal information a quick searchable, comprehensive structure. The administration of this system must also ensure that the information is of sufficient standard. Such a system can only be achieved if a framework can be created through co-operation between law faculties.

2. A co-ordination system

There are two institutions able to co-ordinate the legal information available on the Internet as an alternative to the status quo. Firstly, commercial companies and secondly the law faculties themselves. Both possibilities will be discussed in the following.

2.1. Commercial companies

Like the LEXIS legal database, a commercial company could start providing legal information on the WEB for a charge. There are many possible ways to achieve this(27). But such a development would be a critical result for many reasons. Commercial databases are in a position we may describe as similar to a monopoly; they can impose the terms they like on any contract. Take for instance, the German database JURIS. At the beginning the contract with the law faculty at Tuebingen imposed an inclusive monthly charge with unlimited access for faculty members. Now, a few years later, the Tuebingen law faculty has restricted access for few hours a day and an inclusive monthly charge several times that of a few years ago. The law faculty at Heidelberg did not renew the JURIS contract regarding student access in early 1994. The law faculty at Belfast allows few graduates to use LEXIS, and then only after comprehensive training. The Oxford law faculty is renewing the contract with LEXIS and part of the deal is the provision of an appointed LEXIS instructor. These representative examples show the kind of problems customers such as universities are facing when dealing with commercial companies providing legal information on-line. First favorable contracts are agreed, then when the product has established itself on the market, almost all of the privileges are abolished(28). In the case of the LEXIS or JURIS databases this behavior (which from an economical point of view may be justifiable) makes sense for "first class" customers, such as law firms, who are able and willing to pay large sums for the services and naturally are not willing to waste time by queuing for access(29). If we look at the information offered by such databases, the legitimacy may be doubtful. A big share of this legal information was produced at universities (of course, also by courts or by governments). Thus we have to pay real money for information we have produced previously.

There is another reason why a "commercial wave" can not be wished from universities. Commercial companies want primarily to make money and the information they want to put on-line reflects this aim. The university is concerned with all aspects of law, be it the history of the laws, comparative law or philosophical legal doctrine. Topics such as these will never receive sufficient attention from a charged database, since not enough money could be earned from its use.

2.2. Law Faculties

Today we understand under law at least a set of rules based on fundamental, common shared political principles. These rules work like a net thrown over every sector in the society and thus regulating them. Lawyers are trained to draft statutes, regulations, treaties etc., and to apply them, thus creating a net able to cover every kind of situation we may be faced with. Take for instance courts, no matter how unpredictable human behavior is, under any circumstance the set of legal rules has to be applied to a case directed to a Court. Neither in the civilian nor in the common law tradition can a judge say, "I don't know how to decide this case". A decision has to be given and the only alternative is to refer the case for decision by a higher court.

A digital system based on 1 and 0 sequences has created, together with powerful hardware, a virtual world unimaginable just a few years ago: the Internet. Lawyers are now starting to regulate this virtual world by applying and extending legal rules, be it in the field of intellectual property or in the field of contract. For the co-ordination of legal information on the Internet all we are required to do is, in my opinion, to define a set of legal rules. Every community requires a set of rules to regulate their relationships. The Internet community involved in the provision of legal information is no exception. The direction of any solution to make Internet legal information useful is to co-ordinate the information through a set of rules. Of course such an approach has to find a balance between individuality of WEB sites and subordination under a general concept. Such a concept requires an open, democratic architecture so that everyone interested in providing legal information can join it.

3. Law Faculties Association

A possible way to devise a workable concept for co-ordinating legal information is suggested in this paper on a sketch level and can be seen in the appendix. The main ideas are that every law faculty (but also other legal institutions) interested in providing legal information should join a Law Faculties Association (LFA). In providing legal information, members of the LFA observe rules stated in a Protocol which defines the architectural structure, standard forms, etc. The Protocol is part of a Statute which regulates the relation between faculties, a central country unit, and the European level unit. 3 electronic mailing lists, two main institutional on-line bodies and a set of rules would be sufficient to administer such an organization. A "top" LFA WEB Euro server would co-ordinate the "top" WEB Servers in countries which are members of the LFA. A country 'top' WEB server would co-ordinate legal information provided within this country on every Internet service. Promotion of joint projects, technology advice etc. could be easily co-ordinated. Everyone interested in providing legal information could thus join the framework provided by the LFA on a country level. All the effort within the reach of the LFA could thus be co-ordinated. Legal information would become transparent and useful. The fact that law faculties would regulate legal information on the Internet would be a step toward improved quality of information(30), allowing no links to useless documents.

4. Internet peculiarities

The famous case where a lawyer in the USA sent an unsolicited advertisement for his immigration legal service to many Usenet groups on the Internet and received over 30 000 replies, most of them outraged messages, shows that on the way

to co-ordinate information, the 'Internet culture' has to be observed. The LFA approach does not infringe the 'Internet culture'. What we are going to do is just to create a set of rules which enable us to co-ordinate information. An effort in this direction is urgently required. Other faculties are facing the same problem as lawyers. A workable concept to co-ordinate information would be a benefit to the whole Internet community and a great step forward. The LFA approach provides a prototype which could be adopted by faculties in any discipline. That the Internet is anarchic, uncontrolled and uncontrollable is a myth. At least every LAN has an administration which distributes accounts to access the Internet. Certainly the regulations differ from one LAN administration to another, such that there is no administration from the 'top', but there are, for instance, many thinkable reasons to withdraw someone's account on the Internet through a LAN administration e.g. as in the case of a repeated breach of the Internet "Netiquette"(31).

An effort to co-ordinate the information on the net by university faculties would be unlikely to go behind the border of secrecy or privacy. These are problems for every 'economic wave', which would probably lead to separate networks with a gateway to the Internet like Compuserve. The scope of an LFA is just to co-ordinate the efforts for those providing legal information on the Internet. The LFA approach does thus not infringe any rule of the 'Internet culture' and would be likely to gain a high acceptance within the Internet community.

Conclusion

The possibility exists to construct a database containing all the legal knowledge available (with the relative barrier copyright) from the various legal systems in Europe. It will be necessary to start an LFA from one country in order to convince other countries from the successful new departure. We have in Europe a difficult situation. France ignored for a long time the Internet, favouring its technically obsolete "Minitel" system, and it is now thinking about promoting a multi-lingual European network separate from the Internet. Germany has the most WEB legal servers in Europe and their legal community is becoming more and more interested on the Internet, but the policy of the leading WEB server at Saarbrucken, which is waiting for the "commercial wave", does not give much reason for hope towards a legal information standard. I am convinced that the only country able to make the first step in the right direction is the UK. The LTC at Warwick has a unique position in Europe since it reaches every law faculty in the UK and is an accepted authority.

Ten years ago it was very difficult to find a computer in a University. Today we have plenty of computers and in addition a powerful virtual world called the Internet. Where will we be in ten years time ? Nobody knows. I hope this paper has helped to contribute a few small stones to the mosaic of new information technology and to influence the path of change which lies ahead.

Notes

1 In principle a server connected to the Internet provides the information which is stored on files. A corresponding client is able to access this information from every point of the Internet.
2 The byte count shows that the WEB is due to supersede Gopher, see NSFnet Byte Counts statistics at FTP at: nic.merit.edu; See for a recent analysis from the USA, National Research Council, Realizing the Information future: The Internet and Beyond, 1994.
3 In the meantime many HTML editors have been written and offered on-line e.g. HTMLASST or HTMLEDIT running under MS Windows. Another example is a Microsoft Word for Windows version 6.0 macro template on-line called GT HTML which enables users to create HTML files from word documents. Location: use the nearest ARCHIE service ;-)
4 See for a comprehensive Beginner's guide to HTML at: http://info.cern.ch/hypertext/www/MarkUp/HTML.html
5 E.g. the WEB client Netscape. Netscape Communication Corporation was founded in April 1994 by two computer scientists and has in the meantime a team of more than 75 employees, see at: http://home.mcom.com/MCOM/mcom_docs/backgrounder_docs/background er_docs/founders.html
6 It is estimated that 4 million computers are connected to the Internet and perhaps 30 million users (Nov 94). Every month the Internet has an estimated 15 - 20% more subscribers. Thus in the near future a total number of 150 million users are supposed to be able to subscribe to the Internet; See National Research Council, op. cit.
7 The amount of stored information on LEXIS-NEXIS is 580 billion characters. For the comparison I assumed 2000 characters per page (typewriter). See for other statistics on LEXIS-NEXIS their WEB homepage at: http://www.meaddata.com /html/misc/background.html
8 Have a look at JANUS Digital Library at: http://www.janus.columbia.edu/ with a detailed account of the project.
9 To be accessed at: http://www.law.indiana.edu/glsj/glsj.html
10 See at: http://www.jura.uni-sb.de/Rechtsgeschichte/Ius.Romanum/english.html
11 See e.g. the Pepper & Corazzini law firm at: http://www.commlaw.com/pepper/search.html; or the UK law firm Clifford Chance at: http://www.cliffordchance.com/net_in.htm (since 29/Jan/95).
12 See at: http://www.law.indiana.edu/law/lawtalk.html, with an explanation what kind of hardware and software is required to be able to listen to it.
13 For the most useful search engines on the WEB in one place with searchable tools for servers, people, publications, etc. see at: http://www.cs.indiana.edu/inds/w3search.html
14 See e.g. the new version of WWWWais 2.5 documented at: http://www.eit.com/software/ wwwwais/wwwwais.html

or a new program which indexes WEB sites and searches for files using keywords, SWISH, at: http://www.eit.com/software/swish/swish.html
15 See at: http://info.cern.ch:80/default.html
16 Held by the Indiana University School of Law at Bloomington, USA, see at: http://www.law.indiana.edu:80/law/lawindex.html
17 See at: http://galaxy.einet.net/galaxy/Law.html
18 See at: http://www.dorsai.org/p-law/
19 See at: http://www.uni-bayreuth.de/students/elsa/elsa-home.html
20 In particular Germany, see the WEB server at Berlin for a complete list at: http://www.rewi.hu-berlin.de; But see also the law school at Strathclyde which joined the Web in Feb 1995, at: http://www.strath.ac.uk/Departments/Law/LawHome.html
21 At: http://www.jura.uni-sb.de/
22 At: http://ananse.irv.uit.no/law/nav/hp.html
23 Euro-Lex previous mail can be accessed at: gopher://gopher.law.cornell.edu:70/11/listservs, where the last 1000 mails can be found through a WAIS searchable index. Just type as key word 'zizzo' to be directed to the relevant mail.
24 Cited from an e-mail send to me by Will Sadler Indiana University early 1994, see his homepage at: http://www.law.indiana.edu/hyplan/will.html
25 See Cyber-Justice: 'West Fear of on-line changes groundless', Star Tribune, October 16, 1994; 'Computer aided legal research subject of probe, WSJ, October 3, 1994. On-line information distributed by James Love [jamie@tap.org] on e-mail request.
26 In the USA there are serious plans to make the Internet available to schools and use them as an universal tool for teaching. The main problem is an investment one. At the least a LAN with shared modem is required in order to be able to access services like the World Wide Web. For such plans a very large investment is required, see for a starting point in the discussion, US Congress Office of Technology Assessment, Global Standards: Building Blocks for the future, 1992.
27 See for a good example the Encyclopaedia Britanica Online at: http//www.eb.com
28 This is the case with the WEB client Netscape, which is available at the moment free for members of universities, see their homepage, cit.
29 LEXIS-NEXIS reported 1993 revenues of $551.5 million, an 11% increase from 1992. For 1994 similar figures are expected, 650,000 users have so far subscribed with estimated 200,000 searches a day (Jan 95) see their homepage, cit.
30 Of course there is no guarantee. Every commercial legal database excludes any kind of liability.
31 These are rules setting out how networkers should behave on the Internet, and are available on many locations throughout the Internet.

Abbreviations

The following abbreviations are used in the text:

WEB or WWW: World Wide Web.

FTP: File Transfer Protocol. A method of transferring files across networks.

ARCHIE: A network service devised to search FTP sites for files.

Telnet: A program which enables the connection to remote computers across networks.

MUDs: Educational/Professional Multi-User Dimension.

IRC: Inter Relay Chat, based on the client/server principle it is a multi-user chat system where groups of people meet to talk private or public.

URLs: Uniform Resource Locators. Standardized formatted entities which specifies location, type and address of a file.

Appendices