TOOL SUPPORT FOR LEGAL ENQUIRY
Focus on spatial-planning law

Eriks Sneiders

Department of Computer and Systems Sciences, Stockholm University / Royal Institute of Technology
Forum 100, SE-16440, Kista, Sweden
Email: eriks@dsv.su.se

Abstract: An ideal "law machine" comprises a number of tools and technologies for data processing, storing, exchange, and retrieval. The paper identifies classes of tools and sample tools that support legal enquiry where spatial-planning and environmental law has been taken as the exemplar. Not only the traditional means of text processing and retrieval (e.g., search, taxonomy building, text mark-up, etc.) but also geographic information systems (GIS) and related standards are examined. Furthermore, a number of accompanying tools help to better utilize and make sense of the retrieved legal information. Except GIS, which is particularly convenient for spatial-planning legislation, the identified classes of tools are not specific to any particular area of law.

1 INTRODUCTION

Citizens and organizations experience increased complexity of the legal framework of their day to day operations. Legal professionals are recruited to query regional regulations, national law, EU directives and to interpret the results of the enquiry. As the situation becomes financially prohibitive, many citizens and small businesses seek initial help from paralegal intermediaries, such as citizen-advice organizations, trade bodies, etc. They provide a limited level of support on legal matters based mostly on professional experience rather than formal legal education or active legal-information support available to them.

Add-Wijzer (2004) is an EU eContent (2004) project that combines existing legal information with existing technologies and thus creates a gateway between the user (advanced citizen or paralegal intermediary) and the legal databases. Spatial-planning and environmental law has been taken as the exemplar. Two test beds – JustCite (2004) for text-based and Legal Atlas (2004) for map-based law access – are employed. JustCite extends the Justis portal (2004) for UK law and EU directives, whereas Legal Atlas is built from scratch to handle regional and municipal spatial-planning law in the Netherlands.

While it is accepted that legal practitioners access legal information, little is known of their research practices (Leith & Hoey 1998, pp. 110-112). Therefore an extensive requirements analysis has been done during the project. Some 40 interviews were carried out with lawyers, planners, policy makers, non-governmental organizations, members of farming community, etc. Leith and McCullagh (2003) report on the first part of the interviews, as well as present a theoretical discussion on the use of information technologies in the legal community.

This paper continues the work started in (Leith & McCullagh 2003) and a large amount of still unpublished material. The paper identifies classes of tools and sample tools motivated by the requirements analysis. Furthermore, the paper discusses map-based access to location-specific law, which is relevant for spatial-planning legislation.

2 OVERALL ACCESS TO LEGAL INFORMATION

Until the late 1990s, CD-ROM was a well-established means of electronic publishing of legal information. Nonetheless, CD-ROM has disadvantages, such as no common user interface standard and material out of date. AustLII (2004) states: "Materials distributed via CD-ROM are expensive to produce, cannot be updated as

frequently and create further technical problems. AustLII is a small organization with few resources and the World Wide Web provides the most suitable method of distribution for our materials." The Internet indeed offers a means of accessing information which is inexpensive and not bound to any particular location. Therefore our test beds JustCite and Legal Atlas are two *web portals* that allow:

– publishing proprietary material
– linking to material published elsewhere
– access to regularly updated information
– combining textual and graphical information, which is especially important for spatial-planning law
– integration of the information with means of communication such as e-mail or discussion groups, or even Internet telephone
– access to information "as and when you need it", given access to the Internet

The overall features of the desired access methods are simplicity and functionality: one-stop-shop, "no flashing graphics", focus on usable information. From the interviews: "I have no legal training. I would appreciate an idiot-proof guide to accessing and understanding legal information." "A lot of policy-makers are aware of their knowledge limitations regarding law. This causes stress in the department. […] I am the first in-house lawyer, and now colleagues are stressed that they can't get instant answers from me."

3 PUBLISHING AND RETRIEVAL OF LEGAL TEXT

Text documents are the most common carriers of legal information. The JustCite portal, operated by Context Ltd., is our demonstrator of text publishing and retrieval tools. JustCite, which is the only provider-neutral legal research service available, combines content from different publishers and includes links to full-text material on leading online services such as Justis, Butterworths, BAILII, Westlaw, Informa's Lloyd's Law Reports, Casetrack, and EUR–Lex.

3.1 Keyword-Based Search

The search engine behind the front-end on JustCite is Hummingbird SearchServer. The search engine delivers information indexing and searching for desktop and web-based applications. It supports

searching by different fields marked up in the data and supports a wide range of Boolean operators and wild cards.

HummingBird SearchBuilder is a toolkit for building UNIX and MS Windows-based *search applications*. SearchBuilder can be used for cross-platform development with applications developed in C, C++, MS Visual Basic, and Java.

3.2 Text Mark-Up

Data for the JustCite portal is stored in XML and displayed in HTML or XHTML. CCDF (Context Compound Document Format) is an XML compliant database tagging system owned by Context Ltd. CCDF represents the formal structure of a document. Structures that are common to documents from different sources are given common structure tags, thus facilitating cross product searching through a common interface. The development of CCDF was based on the following requirements:

– the CCDF language must be simple enough to ensure an acceptable level of accuracy
– any information tagged in CCDF format must be convertible to other standard tagging systems such as XHTML without manual intervention

3.3 Taxonomy Building and Document Tagging

Search and navigation capabilities are enhanced by assigning subject terms (tagging) and categorization of legal documents, which requires appropriate software. JustCite uses GammaWare from GammaSite. The process of using this technology consists of the following steps:

– *Development of a legal taxonomy.* JustCite uses a new legal taxonomy developed by Context Ltd. GammaSite provides tools that enable constructing and editing the taxonomy, fine-tune its tree structure, insert where required links between categories.
– *Training of the GammaWare system.* This stage involves selection of at least five training documents for each of the taxonomy nodes. These documents are then used to train the system and the GammaSite tools are used to analyse the results, view the factors that affected categorization results and fine-tune the process.
– *Production phase.* In the production phase documents are classified using the GammaWare software without the need of manual intervention.

3.4 Cross-referencing

Pieces of law need to be cross-referenced. The portal should have links to EU legislation, domestic legislation, case law, government reports, and explanatory information regarding the intended aim of the legislation.

Conventional hypertext systems require each hypertext link to be identified when a dataset is created, with the link to the referenced document pre-set. For large data sets this would be a massive and uneconomical task. Context Ltd. has developed a reference recognition system known as J-Link. J-Link has been further extended into a generic system known as Syntalex, which is able to analyse digital text on the fly and convert reference information into active hypertext links.

The generality and flexibility of the engine has been achieved through a rule-based design, and the behaviour of the engine is controlled and determined by the use of rule files. Each rule file contains one or more rules, each of which comprises two distinct components:

− The *matching* component describes and defines a specific type of text pattern, such as a document citation or other type of reference.
− The *action* component specifies a type of action that is to be carried out once the conditions specified by the matching component have been satisfied. Examples include creating a hypertext link, outputting information to a database, retrieving specific information from an online resource, and sending an email alert.

3.5 Consolidated Legislation

The interviews show the need for consolidation of law documents. "If we were to pay for a computer system – we would like information to be consolidated e.g. consolidated versions of legislation – otherwise it is not a financially worthwhile improvement on existing materials." "A computer system [should provide] ways of knowing which provisions are in force (e.g. strike through of amendments). Consolidation of legislation would be an added advantage, especially consolidation of statutory instruments."

For example, a law can be passed, then amended some years later and amended again after that. The amendments just state the words and paragraphs to be changed, so one needs to consolidate them with the original legislation to be able to make any sense of them. UK legislation databases are not consolidated. Consolidated legislation means that the publisher has included all amendments that have

been made to an act. To do consolidation of all planning law manually by using legal editors would be very expensive and take a long time. Currently there are no tools that can fully automate this work.

JustCite provides links to all other acts that make the amendments so that people can do the consolidation of a particular section of an act if they wish to do so.

3.6 Summaries and Comments on Legal Documents

An ideal "law machine" should explain the history and function of legal documents, give interpretative summaries with hyperlinks to full decisions. "This provides quick, convenient access to legal information – in a profession where speed is of the essence." "Access to relevant case law would be useful."

For most of the case law products, Context Ltd. gets a data feed from law publishers. Their editors manually summarize the reports. There are no good tools for automatic summarizing of legislation. All attempts to introduce such tools at Context Ltd. have failed.

Adding comments to legal documents is related to consolidation (Section 3.5). An alternative way of obtaining opinions and comments is discussion forums (Section 5.2).

3.7 Space for Improvements

Within the Add-Wijzer project, a number of tests have been carried out in order to determine the usability of existing legal text retrieval systems BAILII, Her Majesty's Stationary Office, EUR-Lex, Justis, LexisNexis, and Westlaw. Tests were carried out with seven non-lawyers who came from a planning or environmental background but had no experience in law, and five lawyers who had a background in law but were not necessarily fully qualified lawyers. Among the discovered technical problems were:

− information overload
− confusion caused by layout and labelling
− recognition of icons
− too many mandatory form fields
− too many irrelevant items in search result
− poor recognition of plural forms and synonyms (e.g., "vs." and "versus"),
− poor ability of the system to retrieve a document by its title
− navigation inside the retrieved document

One test person was very frustrated when he entered the search term "Environmental Protection Act 1990" in BAILII and was returned 747 results that seemed to included everything but what he was looking for. "Regulation of Care (Scotland) Act 2001" was returned first noted as having 100% relevance. Another test person became frustrated when she was looking for "Directive 94/62" in Justis, typed the search term "Directive 94/62" into the search box, and the correct document was returned as result number 26.

People who are not computer professional appreciate simple input. About the interface on LexisNexis: "Too fussy… way too much stuff. You get used to search engines and like Google for example, it's just one great huge bar that you put the information into."

If legal publishers wish to improve the quality of their service, a good source to learn from is commercial Internet search engines which have put a lot of effort into their relevance-ranking algorithms. Web documents and legal documents share two core features – structured text and interlinked documents.

Search Engine Optimisation is a new industry, 2-3 years old, that originates from web-design. It does not deal with optimisation of search engines, as the name suggests, but rather with optimisation of websites for top positioning on search engines. Search engine optimisers are well-aware of the features that influence relevance ranking most. Their knowledge can help in the design and implementation of legal text retrieval systems.

4 MAP-BASED ACCESS TO LOCATION-SPECIFIC LAW

Legal Atlas, operated by Framfab Nederland, provides access to regulations through a map-view on the legislation (Peters & van Engers 2004). In this view, law is retrievable knowledge objects rather than hierarchical structure of legal text. Dutch spatial-planning standards support such objectification of law and further linking of law objects to geo-spatial objects on the map. The approach allows challenging the traditional question "I want this here… Is it possible?" by targeting the reverse question, "I want this… Show me where it is possible!" A typical application area for Legal Atlas is urban planning where the user interrogates the system, and the system responds via a digital map as well as more traditional representations such as text.

Legal Atlas is still in the stage of the first prototype. Nonetheless, interviews with intended users show that the system has a good market potential. The prototype has caused some changes in the meta-data strategies for local law applied by the Dutch Ministry of Internal Affairs.

4.1 DURP and IMRO

The corner stone of spatial-planning law in the Netherlands is zoning plan ("bestemmingsplan") – a map with legal information attached to spatial-planning objects. Zoning plans are legally binding, they are a form of legislation.

DURP (2004) is the initiative on digital exchange of legal spatial planning data started by the Dutch Ministry of Spatial Affairs and Environmental Issues. DURP invites organizations, including the large number of municipalities, to standardize their digital zoning plans, which facilitates interoperability between different layers of government while dealing with spatial planning.

DURP introduces a standard for exchange of spatial planning data called IMRO (2004). About one thousand IMRO codes formally define the purpose and functionality of spatial planning objects. Together with textual "voorschrift" (prescription), the codes tell the user what is and is not allowed in various IMRO-coded areas on the map. Legal Atlas makes use of the coding and enables querying of the underlying legal system using a detailed map representing legal constraints and possibilities.

4.2 Geographic Information Systems (GIS)

GIS software helps to create and visualize graphic information (usually geo-spatial maps), and to interact with it. A distinguished feature of GIS is representation of visual information in layers as a means of information filtering. A user combines the layers according to his or her information needs.

Legal Atlas uses Autodesk MapGuide. MapGuide Viewer, a web-browser plug-in, offers display, query, and analytical features for a variety of platforms and browsers.

Open GIS Consortium (2004) is leading the development of geo-processing interoperability computing standards to create open and extensible APIs for geographic information systems. In other words, Open GIS standards make sure that GIS tools from different vendors can cooperate. MapGuide is Open GIS compliant.

Legal Atlas uses ISIS FlexiWeb (2004) as the main user interface – a shell around MapGuide that enables the combination of SQL queries of legal texts with GIS retrieval of related map objects.

5 INFRASTRUCTURE AROUND A LEGAL INFORMATION SYSTEM

The users of law retrieval systems need more than just an ability to query a database. The most important bonus features are interpretation of law and access to related information. The tools presented in this section do not lie in the focus of the Add-Wijzer project. It is, however, recommended that a well-established law publisher consider using them as additional services for the convenience of its customers.

5.1 Help-Lines and FAQ Lists, Legal Self-service

A computerized system does not interpret law for a particular case, it does not do practical application of law. Therefore "there should be an option of contacting a human as a last resort if you can't find or understand information" (citation from the interviews). The "last resort" apparently is a lawyer or paralegal intermediary.

People tend to ask the same questions over and over again, therefore websites often organize compilations of frequently asked questions (FAQ). People may prefer consulting an FAQ list prior to asking a question to a human expert: "A telephone hotline/online service would be ideal. It should be a two-tiered service (1) an on-line list of FAQ's and answers (2) send an email/contact a person with specific query."

An emerging market is self-service and question-answering systems. A prominent representative is Kanisa, formerly known as Jeeves Solutions, which advocates the concept of pre-packaged knowledge as "a blueprint of knowledge for your specific industry." Kanisa's Site Search Industry Solutions (2004) contain "not only the words, terms and concepts unique to each industry, but also a deeper understanding of which questions are most commonly asked and how these questions are asked."

Another example of on-line help service is Psychology Free Online Medical Advice on Web4Health.info. Both Kanisa and Web4Health utilise the approach of template-based question answering (Sneiders 2002).

There is little experience of introducing self-service legal-advice systems. In fact, Legal Atlas is a prototype of such a system which would counsel its users before they see their expensive lawyers.

5.2 Discussion Forums

Not all information is written down. People may contact someone they know to find out what might happen next year, or what is going to be tolerated despite the law. That sort of information requires human networking, perhaps supported by mailing lists or discussion forums.

Interviews with domain experts confirm the role of tacit knowledge. An experienced event organizer is well aware of the processes and procedures involved, a lot of the information that he would have to gather is already stored in his head. He does use the Internet to "look for details and things like that" but he concerns himself more with political decisions, which greatly affect his business. "I do heavily lean on political decisions. For that I *float on the rumour circuit*. I depend a lot on this circumstantial information."

Looking for opinions is another case where discussion forums may help. In order to understand the proposed or existing legislation, one interviewed politician who had access to a research facility also seemed to utilize the expertise of both government departments and non-governmental organizations. There was a sense that he was looking for a *balance of views*, and was prepared to seek these out, but not really through looking at the sources themselves, so much as getting *interpretations from others*.

What non-lawyers seem to be looking for is some way of discussing common problems with other people who may have solved similar problems. Discussion forums do not provide legal advice. On newsgroups such as uk.legal, much of the advice is of the kind, "when you go to see your lawyer, ask him whether X applies to your case". Law-related information needs not actually to contain laws at all. E.g., an environment agency may have tolerated particular pollution levels without prosecution in the past, as part of a policy to prosecute only the most serious offenders. Knowing what those levels are is useful, but is not the law. In some countries, there are standard rates for bribing civil servants to get away with what is strictly illegal.

5.3 Newsletters, Alerting Services

JustCite offers an alerting service which automatically notifies about relevant updates to the database. The service enhances the Saved Searches feature. It works by recognizing whenever a new document that matches the criteria of one of the saved searches is added to one of the JustCite databases. An e-mail alert is sent out then. Email alerts contain the reference, title and summary of

each document together with a hypertext link to the full text of the document on JustCite.

5.4 Printing and Text Editing

The interviews show that legal information systems have insufficient printing and text editing facilities. Some desired features are:

– an on-line copy must have the same layout as a paper copy
– the user must be able to specify particular sections or pages to print out
– copy and paste facility ("dislike how you can't copy and paste within Adobe documents")
– printing and text editing integrated in the system, without a need to buy sophisticated software

One interviewee noted the usefulness of the PDF format as it meant that information could be kept "in a file". This appeared to imply – as with other interviewees – that they wanted to be able to build up (via cut and paste) their own collection of information.

6 CONCLUSIONS

The paper identifies classes of tools and sample tools that support legal enquiry with the focus on spatial-planning law. The selection of tools has been motivated by an extensive requirements analysis.

The following classes of tools and technologies have been distinguished:

– *User access* through web portals.
– *Publishing and retrieval of legal text*: keyword-based search tool; XML-compliant text mark-up language; tools for taxonomy building and document tagging; tools for cross-referencing, consolidation, and summarizing of legal documents.
– *Geographic information systems (GIS) and related standards*: Autodesk MapGuide, Open GIS standard, a web-based shell around MapGuide, Dutch-specific spatial-planning standards.
– *Infrastructure around a legal information system*: help-lines and FAQ lists, discussion forums, newsletters and alerting services, improved printing and text-editing facilities.

Except GIS, which is particularly convenient for spatial-planning legislation, the above classes of tools are not specific to any particular area of law.

There is a need but no good commercially available tools for consolidation and summarizing of legal documents.

ACKNOWLEDGEMENTS

REFERENCES

Add-Wijzer, 2004, http://www.addwijzer.org/

AustLII FAQ, 2004, http://www.austlii.edu.au/austlii/faq/#q4.7

DURP – Digitaal uitwisselbare ruimtelijke plannen, 2004, http://www.digitale-plannen.nl/

eContent – European Digital Content on the Global Networks, 2004, http://www.cordis.lu/econtent/

FlexiWeb, 2004, http://www.isis.nl/isisint/flexiweb.htm

IMRO - Informatie Model Ruimtelijke Ordening, 2004, http://www.ravi.nl/standaardisatie/imro/

JustCite, 2004, http://www.justcite.com/

Justis, 2004, http://www.justis.com/

Kanisa Site Search Industry Solutions, 2004, obtained via http://www.kanisa.com/register_todownload.shtml

Legal Atlas, 2004, http://www.addwijzer.org/ and follow "Demonstrator"

Leith, P. & Hoey, A. 1998. *The Computerised Lawyer*, 2nd revised edn, Springer-Verlag.

Leith, P. & McCullagh, K. 2003. Developing European Legal Information Markets Based on Government Information: First Findings from the Add-Wijzer Project. *18th BILETA Conference: Controlling Information in the Online Environment*, April, http://www.bileta.ac.uk/03papers/leith.html

Open GIS Consortium, 2004, http://www.opengis.org/

Peters, R. & van Engers, T. 2004. Legal Atlas: Map-Based Navigation and Accessibility of Legal Knowledge Sources. Accepted at the *5th working conference on Knowledge Management in Electronic Government (KMGov2004)*, Krems, Austria, May 17-19.

Sneiders, E. 2002, *Automated Question Answering: Template-Based Approach.* PhD thesis, Royal Institute of Technology, Sweden.

All URL links valid February 1, 2004.