The DataLex Legal Workstation- integrating tools for lawyers

• *Graham Greenleaf, Andrew Mowbray, Alan Tyree*

Graham Greenleaf is a Senior Lecturer in Law, University of New South Wales. Andrew Mowbray is a Senior Lecturer in Law, University of Technology, Sydney, and Alan Tyree is the Landerer Professor of Information Technology and Law, University of Sydney. *This article is an extract from a paper presented to the Third International Conference on Artificial Intelligence and Law (ICAIL-91) held on June 25-28 1991 at Oxford University. Complete copies of the paper can be obtained from the editors.*

Integrated computerisation of law

Computerisation of law has developed from a number of originally unrelated technologies: the development of on-line free text retrieval systems from the 1960s; the revival of artificial intelligence research in the form of expert systems in the 1970s, the related development of automated document generators, and the 'rediscovery' of hypertext in the late 1980s.[1] Lawyers are interested in the computerisation of a number of different aspects of legal practice, including retrieval of documents relevant to decision-making, other forms of research, the decision-making itself, and the generation of legal documents.

Most commercial applications have concentrated on only one of these paradigms. This lack of integration is not peculiar to law, but has been observed to be a general feature of the computerisation of information[2]. Vandenberghe stressed the importance of integration[3], and others have done so since[4]. However, there has been relatively little development of an integrated theory for all aspects of the computerisation of legal materials.

We refer to the comprehensive integration of the modes of computerising law as the 'legal workstation'. We describe our approach to such integration, the DataLex Workstation software, and its use in an application to privacy law (the 'Privacy Workstation'). Arguments concerning the practical and theoretical importance of integration are also advanced.

DataLex Workstation software

The DataLex Workstation software[5] combines expert systems, hypertext, and free text retrieval into one general-purpose tool. It has been developed for use in commercial applications, and to teach legal applications development, rather than as a pure research vehicle.

Workstation design and components

The Workstation software incorporates three 'engines' which process legal knowledge and data in different ways: an inference engine, a hypertext engine and a free text retrieval engine. Each communicates with the user through a common user interface (which is based in part on the hypertext engine).

Two elements of integration are that first, the system must appear to the end-user as an integrated whole where all elements interact in a consistent way; and second, it should make maximal use of shared knowledge and data. The extent to which these aims have been achieved will be explained in relation to each of the Workstation components.

Information representations

Each 'engine' requires its own form of representation of the legal 'knowledge' and 'data' which it manipulates: a rule-base for the type of inference engine used in the Workstation; a concordance for text retrieval; and a hypernet (network of nodes and links) for hypertext. Each representation is conceptually distinct from the legal texts which are usually one of its principal sources.[6] The expression 'knowledge representation' is most often used in relation to expert systems, indicating that a knowledge-base involves more than mere data. Concordances and hypernets are more easily seen merely as 'data', but the creation of text retrieval and hypertext systems does involve some addition of legal knowledge to the 'raw' textual sources of the law, such as knowledge of the structure of different

types of legal texts. In our view, the distinction between 'knowledge' and 'data' is often a very fine one. A list of related concepts, for example, is as of much relevance to an expert system inference engine as it is to a hypertext engine (where it can be used as a cross-reference) or a free-text retrieval system (where it can be used to expand search terms). Perhaps it is best to say that the Workstation involves various 'information representations'.

Privacy Workstation application

The Australian Privacy Act 1988 (Cth) and the office of Privacy Commissioner commenced in 1989. Because privacy law is an area of expertise of one of the authors, it presents a good opportunity to build a comprehensive computerised representation of an area of law from its inception. Despite its recent origins, Australian privacy law is developing from a very heterogeneous and complex set of source materials: statutes; regulations; Commissioner's 'guidelines'; determinations on exemption applications and determinations of complaints; Court cases; Parliamentary decisions; the Digest of government personal information systems, and academic commentary. It covers a range of closely linked subject matter, including public sector personal records, data matching, spent convictions, tax file numbers and credit reporting. Material from all of these sources is progressively being incorporated into the Privacy Workstation.

All material in the Privacy Workstation is accessible through hypertext and full text retrieval. The conversion of all of the significant statutory sources into expert system components is not yet complete. At present, the system provides advice on the potential applicability of the 11 Information Privacy Principles (IPPs) in the Privacy Act, and on the applicability of the spent convictions legislation. The Privacy Workstation is in commercial use by Australia's largest credit bureaux, is used on-line by the Australian Privacy Foundation (a lobby group), and is being evaluated by the Privacy Commissioner's Office.

The following simple example gives the general flavour of the expert system dialogue, but does not show the Workstation interface, such as the availability of hypertext links from prompts.

What is the name of the information ? Smith's Medical Record

Is it personal information ? Yes

Was it collected by a collector ? Yes

Was it collected for inclusion in a record or in a generally available publication ? Yes

What is the name of the collector ? Medicare

Is the collector a natural person ? No

Was Smith's Medical Record collected for a purpose that is a lawful purpose directly related to a function or activity of Medicare ? Why

This will help determine whether or not the exception provided in IPP 1(1) (a) and (b) applies.

Was Smith's Medical Record collected for a purpose that is a lawful purpose directly related to a function or activity of Medicare ? Yes

Is the collection of the information necessary for or directly related to that lawful purpose ? Yes

During a consultation, the user may also instruct the system to 'forget' a previous user-supplied value, causing conclusions relying on that value to be re-inferred.

The significance of Integration

Having given an example of the operation of the Workstation it is appropriate to consider whether this type of integration is of practical or theoretical significance.

Open-ended expert systems

Should we aim to automate legal decision-making, or to provide support for legal decision-makers? One of the most difficult problem in the development of legal expert systems is that caused by the open texture of legal language. The problem arises in a number of ways[7], and may be resolved by one or more of three possible reasoning agents: the inferencing engine; the knowledge-base developer; or the end user. Differences over the most appropriate response to this problem leads to different models for the development of legal expert systems. One approach is to develop software and knowledge representations which will suggest solutions to open texture issues[8]. Such research presents challenging theoretical problems, and is as yet difficult to implement in commercial systems.

Another approach, concentrating on supporting and augmenting the decision-making abilities of the system user to resolve open texture problems, has been advanced by a number of researchers[9], one of whom sees 'less need for a cognitive legal machine than for a less sophisticated but more humble product to support intelligent human interaction' [10]. Taking this approach, it becomes crucial for legal expert systems to allow access to as rich a collection of support materials as possible, so as to support intelligent choices by the user when interpretation of an 'open textured' predicate is required (ie when the user has to make a choice which the system is incapable of

making). The user needs to be given open-ended access to the relevant supporting materials, rather than for the system to simply direct the user to a few definitions which may assist interpretation ('closed ended' assistance).

We take this second approach in the Workstation, by providing various methods by which a user may move from an expert systems consultation to the interpretive materials relevant to that point in the consultation (for example, a statutory definition of a term used in a prompt). Because these relevant materials are presented as hypertext. the interpretive process is open-ended, with the user able to pursue associations, or to conduct free text searches, until the interpretive resources of the system are exhausted. One advantage of this approach is that it may help overcome some aspects of the 'brittleness' of expert systems: a user can use the other resources of the system to help 'work around' factual variations not adequately dealt with by the expert system itself.

A useful way to view a legal expert system, from the perspective of the user, may be as an interaction between a semi-expert inferencing system and a semi-expert user/ interpreter, with control over the course of the problem's solution alternating between the two parties to the interaction. Each does what (s)he or it does best, then hands back control to the other. The program controls those steps in the solution process that are capable of being embodied in a computerised inferencing agent, given existing technology. The user controls those steps of the solution process which involve abilities which cannot (at least as yet) be so embodied, including the lawyer's various interpretive skills .

Susskind's rather pessimistic conclusion that research should concentrate 'on designing systems to solve clear and deductive cases[11] might be overcome by making the enhancement of the interpretive resources of users a central aim of legal expert systems research. A key practical question may be to find the boundary between those elements of open texture problems that legal expert systems can handle (given existing technology) and those elements that users must provide. Research into non-deductive methods of inferencing may, over time, push back this boundary.

Practical advantages

The main value of integrated tools in the building of 'real world' applications is that they can save application developers from attempting to use the techniques of one mode of computerisation for purposes for which it is not suited. Attempting to force square pegs into round holes is rarely satisfactory. It is sometimes difficult to anticipate at the outset of a project what combination of tools will be needed. Successful application development is aided by the availability of as wide as possible a choice of tools, and the ability to mix their use in ways which are easy to develop and transparent to the user.

Future development and research

Integration as a way ahead

Leading scholars of both legal text retrieval and expert systems have suggested that research and development have not advanced very far in the past decade.[12] One way forward, we suggest, is to give greater recognition to the importance of integration of the existing approaches to computerising legal information, at both the theoretical and practical levels.●

Footnotes

1 For brief histories: on text retrieval see G Greenleaf, A Mowbray & D Lewis Australasian Computerised Legal Information Handbook, Butterworths, 1988, Chapter 4, or more comprehensively, J Bing (Ed) Handbook of Legal Information Retrieval North Holland 1984 Part III; on expert systems see A Tyree Expert Systems in Law Prentice Hall 1989 Ch l; on document generators see S M Brooks Computerizing for Personal Productivity Butterworths, Toronto, 1989, Chapter S; on hypertext see J Smith and S Weiss 'Hypertext' (1988) 31(7) Communications of the ACM p8 1 6.

2 H P Frei 'The Future of Information Systems', Seminar paper, CIRCIT, Melbourne, 1990.

3 G Vandenberghe 'Software oracles' in H W K Kaspersen and A Oskamp (Eds) Among Friends in Computers and Law Kluwer, Deventer, 1990; see also A W Koers et al 'Delphi revisited: The mythology of the lawyer's electronic workbench' in the same volume.

4 A Oskamp and P van der Berg 'Legal expert systems and legal text retrieval systems: how about integration?' in Kaspersen and Oskamp, op cit

5 DataLex is the name used by the authors for their joint research since 1985. The research presented here has been assisted by a grant from the Australian Research Council for the development of the Privacy Workstation. The Workstation software was designed by Andrew Mowbray and Graham Greenleaf, and implemented by Andrew Mowbray

6 See J Bing 'Rules and representations' in P Blume (Ed) Nordic Studies in Information Technology and Law, Kluwer, Deventer, 1991.

7 See R Susskind Expert Systems in Law, Clarendon Press, Oxford, 1987 Ch5

8 For different approaches see G-J van Opdorp and R F Walker "A neural network approach to open texture" in Kaspersen and Oskamp op cit; E L Rissland and D B Skalak "interpeting statutory predicates" Proc. 2nd ICAIL, 1989 p46.

9 See Taylor Brown "Supporting local office adjudication" and Tyalor "The DHSS local office demonstrator" (both papers Alvey DHSS Large Demonstrator Project, Dept of Systems, University of Lancaster, 1988); A Berg et al "Developing a KBS support system for handling social assistance", SAFAD Stockholm, 1988; and P Johnson and D Mead op cit

10 R Stamper, book review, The Times 1988

11 Susskind op cit p192

12 See J Bing 'Performance of text retrieval systems: the curse of Boole' (1987) 79 Law Lib J. and L T McCarthy 'Artificial intelligence and the law: How to get there from here' Ratio Juris Vol 3 No 2 1990, p189.