Teaching and Researching Legal Information Systems: the Newcastle Electronic Law Library

Michael A. Heather

Newcastle Electronic Law Library

The Newcastle Electronic Law Library (NELLY) was first planned in 1979 and commenced operation early in 1980 to provide an electronic resource of legal materials for the purpose of research and teaching law using the computer. At that time there were no full text commercial systems available in England although two were introduced shortly afterwards. These are however too expensive for any extensive use by students and the aim was not so much to provide students with the opportunity to familiarise themselves with keyboards, query languages, etc., for they do not seem to find these very difficult but rather the opportunity to practise the quite difficult and different art of manipulating legal concepts in the electronic medium. The second aim was to make available a system that could be «got at and into» for trying out different and improved modes of operation. Not unreasonably the operators of the commercial systems cannot be expected to allow users to experiment with the innards of their systems.

Of course electronic information systems are quite different from manual systems, they are not just computerised versions of the latter. They are much more varied in their type and include quite a number of ways of storing and retrieving information such as the knowledge base, expert systems, etc. The Newcastle Library now includes the following: a full text library of English property law, SPILAW; the expert legal document retrieval system LRS developed by Hafner using classical artificial intelligence techniques to represent knowledge of American law on negotiable instruments; the ABF (American Bar Foundation) Processor of Sprowl which is a powerful package for drafting legal documents; the comprehensive

---

Michael A. Heather is Solicitor of the Supreme Court of England and Wales and Senior Lecturer in Law, Newcastle upon Tyne Polytechnic, UK.


analytical data-base of crime established by the International Organisation of Justice to investigate the decriminalisation of all those parts of English criminal law that are regulatory in character 4; there is also a quantity of materials for computer-assisted learning in law.

NELLY is based at Newcastle Polytechnic which shares a large computer installation NUMAC (the Northumbrian Universities' Multiple Access Computers) with the Universities of Newcastle-upon-Tyne and Durham. NUMAC is accessible from other educational institutions in the UK over academic networks. The original intention was to provide primary and secondary legal sources in electronic form for use at the three sites and at other institutions on a do-it-yourself basis using the various hardware and software facilities of NUMAC. The cost of computing at institutions of higher education is not separately accountable at the level of the department or school as a matter of policy of the UK Department of Education and Science. While this does introduce anomalies between computing and non-computing activities, it does mean that it is not necessary to budget for each and every piece of computing. The result is that there are no financial constraints or pressures placed upon the user. This may present the student with a somewhat unrealistic picture of legal information systems for in practice utilising the most efficient and cost effective method is always an important element in the retrieval process. However, it is to be hoped financial considerations will be less important in the future when the present high costs will be reduced by cheaper hardware and more users sharing the storage overheads.

It seems that students who are going forth to practice legal information retrieval for the next thirty or forty years are best trained in more permanent principles than present pragmatics. Students should be encouraged to understand thoroughly the procedures independent of cost. So the Newcastle system is in effect free for academic users. This has made it somewhat attractive, but it is perhaps noteworthy that while the computing charges are free this does not apply to the telecommunication costs and a number of institutions outside of Newcastle have found that they can afford only occasionally to connect to the system over the public telephone system. This is evidence that such institutions are unlikely to be able to afford to pay an economic charge for the search. It is hoped that telecommunications charges will be greatly reduced in future now that NUMAC is connected to the British Telecom packet switching network PSS.

SPIRES

Among the programming languages and software packages available on NUMAC is SPIRES (the Stanford Public Information Retrieval System).

SPIRES has proved so powerful that it has tended to dominate NELLY. The full text retrieval system SPIFAL^2 operates under SPIRES. It has been found easier to convert the Justice CRIME Project to operate under SPIRES rather than to maintain the original COBOL programs. The CAI-CAL in Law has moved progressively from the language BASIC to the ‘C’ programming language and is currently being converted to SPIRES. SPIRES does not appear to be too well known in Europe and therefore a brief sketch should perhaps be given of it.

SPIRES is a generalised data-base management system based on hierarchical ‘B’ tree structures defined on meta-data principles which encourage a logico-semantic approach to the storage of data. It was designed at Stanford University by workers in artificial intelligence in collaboration with library specialists for the project BALLOTS which has developed into the American University research library information network RLIN. It therefore combines recent theoretical advances with very practical considerations. It has more recently been under investigation for use in MEDLAR III the next version of the medical literature data-base retrieval system. Fifty seven systems were evaluated and SPIRES came third. It was reduced to this position because in the method used for assessment SPIRES scored poorly on portability. This is because SPIRES has been implemented with the alternative emphasis on efficiency and has been written in PL 360 the IBM Algol type structured machine language which is only appropriate for machines with IBM architectures. On the other hand, it is very portable between SPIRES sites: these amount to about 30 in the world of which there are two in Europe the rest being in North or South America.

SPIRES is a high level systems package which in addition to the facilities of the operating system has its own help, explain, mail, news, tutorial, logging, charging, full on-line documentation, synonym, thesaurus, arithmetical, statistical, sorting and reporting facilities as well as its own internal editor. There are six separate language processors for each of the following distinct functions of data-base technology:

(i) the file definition language is used to map a schema of the data structure into a meta-data file and includes the facility to build indexes, sub-indexes and combined indexes;
(ii) the search and display commands provide flexible and extended versions of the basic elements to be found in most query languages;
(iii) the updating commands are an important set of commands to enable the data-base manager to maintain the data even in a real time environment;

(iv) the format language controls the way that the data is displayed and can provide each user with his own physical view of the data;
(v) the protocol language enables either the data-base manager or the user to program custom designed interfaces between any two points in the organisation or use of the data-base; and
(vi) the partial processing language is an advanced feature for investigating within records.

The use to which these features can be put particularly for full text data such as is important in the law can be seen in the various data sets which make up the Newcastle electronic collection but chiefly in SPILAW.

***SPILAW***

SPIRES is organised into files and subfiles. This embodies the PUBLIC INFORMATION concept and illustrates the integrity of knowledge. English property law is held as a single subfile. It has been given the title SPILAW, pronounced as «SPY-LAW», following the nomenclature adopted at other SPIRES sites for SPIDOC, SPISORT, etc. SPILAW consists of the whole of the groups for property (including conveyancing, land registration, etc.) and landlord and tenant as published by Her Majesty's Stationery Office (H.M.S.O) with some complementary case reports selected from the same branch of the law for their educational value. In total, the subfile amounts to about 13 Mbytes. This appears to be about the minimum size8 needed to provide the open-ended requirement for student practice machines and also large enough to provide a realistic model database for research purposes.

The file definition language has been put to good use in capturing the computer structure of English statutes. The H.M.S.O. provides the UK Acts of Parliament in machine-readable form as a by-product of the computer-type setting process used for producing these in printed form. Groups of statutes can be obtained therefore on magnetic tape. These tapes are stripped of the machine printing control commands but retain the information that is contained in the typographical form9. In order not to lose any of this information it has been found necessary to utilise some forty SPIRES elements and to traverse the hierarchical trees to the fifth level. It is clear that all computer-assisted legal information systems which have attempted to incorporate English statutes have found the task quite formidable. The achievement of H.M.S.O. in retaining the statutory structure in their tapes is really quite remarkable and to map this on to a data-base has

tested SPIRES to its limits in several directions. Indeed, it appears that the complexity of the SPILAW file definition is not exceeded in any other application at any of the other SPIRES sites and it is probably true to say that it is only the later versions of SPIRES that could have reasonably coped with their complexity.

Of the other language processors, both the format language and the protocol language have been put to good use. The importance of the physical presentation of textual information to guide the reader/user and to assist him in understanding the content cannot be over-emphasised but to date there seems to be a neglect of both software and hardware to achieve this. It has been possible to use the format language to transform the text on output to suit the particular terminal device of the user. The user also has the facility to define his own preferred output format or formats which can be changed at will or set by default when running under his own identifier. The protocol language has been found particularly useful for its computer assisted learning features. On-line tutorials written in the protocol language enable the novice user to be taken through a search session driven by the system which also provides suggested user input together with a commentary. The facilities also exist in the protocol language to intercept the search and display commands and interpose a set of commands which mimic the query languages of the commercial system. No user, so far as is known, has troubled to do this and it may be that the vagaries of a particular form of a query language is quite incidental to the conceptual problems of retrieval techniques.

It should be noted that in keeping with more recent full text DBMS and in accordance with its logico-semantic principles SPIRES does not store the physical position of each word within the record. For inclusive data bases no information is lost and there is a great reduction in storage overheads and increase in efficiency for searching and updating. The partial processing language is then available if more detailed searching is required so as to look for phrases, perform statistical tests on words, etc. An initial result stack can be passed to main memory for this purpose to provide very fast partial processing.

Usage

Use of the Newcastle electronic law resources has so far been divided equally between teaching and research. Practical work by students in com-


puter-assisted legal research has to date been an optional extra and a voluntary part of the syllabus. Clearly this is unsatisfactory for it does not enable the student to get proper credit for his or her work. Also it belies the need for all law students to gain experience in evaluating electronic legal sources. In the UK Court of Appeal the Master of the Rolls Sir John Donaldson has already castigated legal practitioners for the indiscriminate way they have utilised the system LEXIS to cite authorities 12. It has been decided therefore to make certain computer elements in the syllabus compulsory for all full-time students who take the BA (Honours) Law Degree awarded by the Council for National Academic Awards (C.N.A.A.) at Newcastle Polytechnic. The degree is validated jointly by the C.N.A.A. and the Polytechnic and a new syllabus incorporating these compulsory elements 13 has recently been approved to commence from September 1983. It is expected therefore that the educational use of NELLY will in future be the dominant one and this may cause further problems for research use. Already there are some conflicts. Because of the bulk storage requirement it is only possible to keep one set of the legal materials on-line at any given time. Researchers want to alter and experiment with the system whereas teachers need a stable constant environment for their students.

With no financial accounting necessary for academic users no detailed statistics are maintained of use and it is possible only to give a rough estimate of the use of the system. SPII.LAW is maintained permanently online as a public subfile of SPIRES and is therefore accessible to all those with a NUMAC identifier. These amount to over two thousand. The use of the system specifically for purposes of law is known to average 10,000 minutes of connect time per year over the first three years. Clearly this figure could rise quite sharply when practical work in the computer laboratory becomes a compulsory part of the law syllabus. There is in addition a wide casual use for demonstrations and other educational purposes in librarianship, text processing, computing in the humanities, etc. Research use includes work on the construction of a natural language interface 14, the use of collocations 15 and stochastic techniques 16 for law retrieval as well as for experimentation in catalectic graphics 17 and to explore the hierarchic model 18 simulating other architectures for future law machines.

---

13. BA (Honours) Law - F.T. BA Law P.T. Joint Validation Documentation (March 1983), School of Law, Faculty of Professional Studies, Newcastle upon Tyne Polytechnic.
Conclusions

Establishing the Newcastle Electronic Law Library has proved quite a major effort. To provide students and researchers with these basic tools of electronic law consumes quite a substantial part of the NUMAC resources. However, if reference is made to the Council of Europe recommendation concerning teaching research and training in the field of computers and law it will be seen that the facilities at Newcastle provide materials for only a small part of the suggested syllabus. It is hoped therefore to continue to add to the collection to provide a truly comprehensive range. Perhaps one of the most interesting aspects is the extra dimension that the computer has brought to the law and the much richer concepts of what an electronic law library and legal information retrieval system should include.

Acknowledgements. – The assistance of a number of persons in the preparation of the materials in the Newcastle Electronic Law Library is gratefully acknowledged. The Polytechnic Computer Unit helped to produce the case reports for SPILAW. H.M.S.O. kindly made available copies of their tapes of Statutes in Force and Justice, their crime database. Further work on the latter has been a collaborative effort with Mr. A. T. H. Smith, Reader in Law at Durham University. James Sprowl of the American Bar Foundation has provided invaluable advice on the MTS implementation of his ADF processor. Because of the important role of SPIRES a heavy debt is owed to Dr. B. N. Rossiter the NUMAC UK SPIRES adviser for both general and detailed advice as well as for his continuous support.