



University of New South Wales Law Research Series

The Datalex Project: History and Bibliography

**GRAHAM GREENLEAF, ANDREW MOWBRAY AND PHILIP
CHUNG**

[2018] *UNSWLRS* 4

UNSW Law
UNSW Sydney NSW 2052 Australia

E: unswlrs@unsw.edu.au

W: <http://www.law.unsw.edu.au/research/faculty-publications>

AustLII: <http://www.austlii.edu.au/au/journals/UNSWLRS/>

SSRN: <http://www.ssrn.com/link/UNSW-LEG.html>

The DataLex Project: History and Bibliography

*Graham Greenleaf, Andrew Mowbray and Philip Chung**

3 January 2018

This background paper outlines the participation by the authors and other colleagues in the previous wave of enthusiasm, particularly from the early 1980s, for AI-based solutions to legal problems, then generally referred to as 'legal expert systems'. As 'the DataLex Project, we were active participants, both academically (research and teaching) and commercially, from 1984 to 2001. The DataLex Project continues, after a hiatus, and current work on the Project's tools is noted briefly [1.4]. The purpose of this background paper is to provide a convenient reference to the original DataLex Project, its participants and its publications.

1 The DataLex Project

The 'DataLex Project' (1984-2001) research into computerisation of law involved many participants over nearly 20 years.¹ Like the Australasian Legal Information Institute (AustLII) which grew out of it, the DataLex project was a joint project of academics from UNSW Australia and from the University of Technology, Sydney. Numerous publications, software and knowledge-based applications related to the DataLex Project from these years are listed in the DataLex Bibliography in inverse date order, as 'DataLex 1987a' etc. Other publications cited appear in the General Bibliography in alphabetic order.

Some of the DataLex Project research was assisted by research grants,² other parts self-funded through commercial licences.

1.1 Pre-web, pre-AustLII, pre-Workstations (1984-90)

DataLex began as a project focusing solely on legal expert systems, with development of the LES shell for procedural (decision network) inferencing.³ Our initial ambition was to assist community legal centres,⁴ but events led the project down other research, legal education and commercial paths. A case-based inferencing mechanism based on nearest-neighbour discriminant analysis (PANNDNA), developed by Alan Tyree, was added⁵. Further software development then added a full text retrieval system, AIRS, which emulated the STATUS retrieval system,⁶ and a hypertext engine (HYPE) which was an early pre-web development of this genre of software.⁷

* Respectively, Professor of Law & Information Systems, UNSW Australia and Co-founder, AustLII; Professor of Law & DataLex researchers included: Andrew Mowbray (1984-); Graham Greenleaf (1984 -); Alan Tyree (1984-89); Phillip Griffith (1985-87); Michael Barr-David (1987-1988); Patrick Gunning (1989-91); Peter van Dijk (1992-95); Philip Chung (1995 -), Geoff King (1995-99), Simon Cant (1996-98) and Russell Allen (1999-2002). A number of research assistants also assisted with knowledge-base development, including Karen Budna-Litic.

² Pre 1990 research was assisted by grants from the Law Foundation of New South Wales in relation to the development of X-SH and in the use of AIRS to provide the 'AIRS legal information retrieval training system' to law schools. Alan Tyree also received an ARC (then ARGS) grant to assist in the development of PANNDNA. Development of the DataLex software in the web environment, from 1997-2001, was funded by an Australian Research Council (ARC) Discovery grant.

³ DataLex, 1995, 1985b; The STATUS search software was used on the CLIRS system, subsequently Info-One, from 1985.

⁴ Assisting community legal services is where the DataLex Project intended to start in 1985, with a funding application to the NSW Law Foundation to develop a 'database of legal expert systems' to assist such services, based on the approach that 'expertise is relative'.

⁵ DataLex, 1995a - PANNDNA (precedent analysis by nearest-neighbour discriminant analysis) inferencing system

⁶ DataLex, 1986b

⁷ DataLex, 1989a

Initial applications were expert systems on intestacy law (INTEST⁸) and copyright law (COPYRITA⁹). FINDER answered questions about ownership of found objects based on case law.¹⁰ A legal information retrieval training system used AIRS to simulate the performance of the CLIRS commercial system, with small databases¹¹ and was purchased by Australian law schools. A textbook supported it.¹² A commissioned hypertext demonstration using HYPE, plus AIRS text retrieval, ran over a remote dial-up (LAWS OF AUSTRALIA Defamation Demonstration¹³).

1.2 DataLex Workstations (1990-95)

In the years shortly before the development of the World-Wide-Web (1990-93), the DataLex approach was based on the integration of inferencing (primarily rule-based and to some extent case-based expert systems), hypertext and text retrieval, with some document generation capacity as well. This first required development of rule-based inferencing software (XSH), subsequently refined by the addition of quasi-natural-language knowledge representation (YSH), influenced by the SoftLaw systems.¹⁴ Backward and forward chaining rule-based reasoning was the core of the inferencing component. The key software that was integrated into one package (XSH and later YSH, HYPE and AIRS) was developed by Andrew Mowbray. That became the 'DataLex Workstations' approach.¹⁵ From 1990 a commercial avatar of the project, DataLex Pty Ltd, developed and licensed a number of 'workstations', primarily the 'Intellectual Property Workstation' and the 'Privacy Workstation',¹⁶ which used all three technologies. Updates to the content were distributed on stacks of floppy disks, in the absence of any effective online alternative. The Workstations had modest commercial success, with licences to government agencies, law firms, patent attorneys, credit bureaus and collecting societies, and enthusiastic users. Support stopped in 1995, when a commercial publisher terminated DataLex's licence to include case law content.

The DataLex Project software and the approaches and techniques we advocated concerning the development of computerised legal information resources were also the basis of undergraduate and postgraduate courses on 'coding' for law students (though it was not called that) from 1985-2002, at UNSW and UTS Law Faculties,¹⁷ as well as for developing computer-aided instruction.¹⁸ This was the first hands-on teaching of 'AI and Law' application development in Australia.

DataLex Pty Ltd also carried out consultancy work on the re-development of the Info-One commercial legal information system,¹⁹ and the SCALE system operated by the Commonwealth government,²⁰ but AustLII's arrival soon made these systems increasingly redundant.

1.3 The AustLII context – Web 1.0 (1996-2001)

The DataLex project, and the 'Workstations' developed under it, had a very substantial influence on the techniques and approach implemented in the development of AustLII from

⁸ DataLex, 1985b

⁹ DataLex, 1986a

¹⁰ DataLex, 1985c, 1986, 1987

¹¹ DataLex, 1986

¹² DataLex, 1988b

¹³ DataLex, 1989b

¹⁴ Johnson, P and Mead, D 1991 'Legislative knowledge base systems for public administration' *Proc. 3rd ICAIL* ACM Press.

¹⁵ DataLex, 1992a, 1992b, 1995

¹⁶ DataLex, 1991a, 1992

¹⁷ DataLex 1992b, 1999

¹⁸ DataLex, 1994a

¹⁹ DataLex, 1995a.

²⁰ DataLex, 1993b.

1995,²¹ particularly the development of HYPE into a tool for automated generation of large-scale automated hypertext mark-up of legal documents. From 1995, development of AustLII required concentration on text retrieval and hypertext, and further development of those aspects stemming from the DataLex Project became part of AustLII's ongoing development.²²

For the first six years of AustLII (1995-2001) the inferencing aspects of the DataLex Project were transferred to the new web environment, the integration of inferencing (knowledge-bases and dialogues) with hypertext and text retrieval was further developed, and methods of 'collaborative inferencing' (distributed, multi-author knowledge-bases) were pioneered, all in the AustLII context.²³ The WYSH (Web-YSH) software²⁴ and the SINO text retrieval engine²⁵ were developed as part of this. Developer Manuals and User Manuals, primarily for teaching purposes, were updated. As far as we know, this was the first significant attempt anywhere to develop legal inferencing systems on the web.²⁶ To demonstrate this, the copyright law parts of the Intellectual Property Workstation were moved on to the web platform. Attempts to move from a propositional calculus to a form of predicate calculus had only limited success,²⁷ as did work on semi-automation of the construction of knowledge-bases.²⁸

From 2001 AustLII did not focus on inferencing and knowledge-bases, and its 'AI-related' work instead concentrated on the use of heuristics to improve text retrieval and hypertext mark-up, enabling AustLII's 'autosearch' and 'note-up' features, and also the use of heuristics for automated construction of an international case and journal citator (LawCite).²⁹ AI techniques have therefore continued to be essential to AustLII's work.

1.4 Future development – DataLex 20.0 tools and AustLII's platform

To conclude, we should indicate the current state of the DataLex tools on which the development to 2001 were based. As part of AustLII, the hypertext mark-up software and the SINO text retrieval software³⁰ originating from the DataLex Project have been developed continually, and now are much more highly integrated. They have been augmented by LawCite, the only international law-specific citator, which currently indexes over 5.3 million cases, law reform documents and journal articles.³¹ AustLII's Point-in-Time legislation system³² could in future be a valuable but complex addition.

The DataLex inferencing software remains fully functional, based on the YSH inferencing engine and the WYSH CGI interface. Applications developed using them prior to 2001 still run (though the law in the knowledge-bases is rather out-of-date), and new applications are being developed, including a demonstration application on Australian electoral law. WYSH is being updated to take advantage of some of the many new interface developments discussed above. Further development of the PANNDATA and AIDE inferencing components is not a current priority, but case-based and predicate-calculus-like knowledge representations may be developed more in future. The DataLex software has been used in 2017 courses at UNSW Law Faculty to teach legal knowledge application development, and will also be used in other law schools in 2018.

²¹ DataLex, 1995, 1995a

²² DataLex, 1995a

²³ DataLex, 1997, 1997a, 1997b, 1999, 2000

²⁴ DataLex, 1997 – part 5; 1997c

²⁵ DataLex, 1996

²⁶ DataLex, 1997 – part 4

²⁷ DataLex, 2001

²⁸ DataLex, 1997 – part 3(6)

²⁹ Mowbray, A, Chung, P and Greenleaf, G 'A Free Access, Automated Law Citator with International Scope: The LawCite Project' (2016) Vol 7 No 3, *European Journal of Law and Technology* (EJLT)

³⁰ SINO <<http://www.austlii.edu.au/techlib/software/sino/>>

³¹ Mowbray, Chung and Greenleaf, 2016

³² AustLII's Point-in-Time Legislation project <<http://portsea.austlii.edu.au/pit/>>

The AustLII Communities platform is a closed wiki which is being used by content developers located outside AustLII to develop free access text resources, such as the *Northern Territory Law Handbook*³³ and a jointly-authored textbook on legal capacity.³⁴ The AustLII Communities platform automatically inserts hypertext links to cases, legislation and other content from the author's text, and allows updating whenever authors wish. We see this platform as a possible location for the development of knowledge-bases by free legal advisory services, using new versions of the DataLex and AustLII tools. A supportive editing environment for knowledge-bases is being developed within the Communities platform and various DataLex knowledge-bases and being developed and tested there.³⁵

2 DataLex Project bibliography, 1985-2001

For SSRN readers: A version of this Bibliography with functioning links is at <http://www2.austlii.edu.au/~graham/expert_systems.html>.

(DataLex, 2001a) Allen, R and Chung, P 'AIDE inferencing system' – software (Unix).

(DataLex, 2001) Allen, R, Chung, P, Mowbray, A, and Greenleaf, G (2001) '[AustLII's Aide – Natural Language Legislative Rulebases](#)' at *Proc. International Conference on Artificial Intelligence & Law (ICAL)* 2001 (SSRN); also '[AustLII's Aide – Implementing Rulebase Systems](#)' [2001] *CompLRes* 32, as presented at 5th Law via Internet Conference, 2001. This paper concerns the use of predicate logic in the Aide software.

(DataLex, 2000) Allen, R, Greenleaf, G, Austin, D, Chung, P, and Mowbray, A, (2000), '[With a Wysh and a Prayer: An Experiment in Cooperative Development of Legal Knowledge Bases](#)', *The Journal of Information, Law and Technology (JILT)*, vol 2 (SSRN); shorter version also published in (1999) *Proceedings of the 7th International Conference on Artificial Intelligence and Law (7 ICAIL)*, Oslo, Norway (ACM Press).

(DataLex, 1997d) Greenleaf, G, Mowbray, A, King, G, Cant S, and Chung, P, '[Rule-based inferencing over the world-wide-web: A tutorial using AustLII's ysh and wysh software](#)' Australasian Legal Information Institute, Sydney, 1997 (SSRN) – This is the most detailed step-by-step 'developer guide', to be used with the *Wysh Developer's Manual* (1997).

(DataLex, 1997c) King, G and Cant, S 'WYSH (web-yshell) CGI interface to YSH' – software (Unix).

(DataLex, 1997b) Greenleaf, G, Mowbray A, King, G, Cant, S and Chung, P, (1997) '[More than Wyshful Thinking: AustLII's Legal Inferencing Via the World Wide Web](#)' *Proceedings of the 6th International Conference on Artificial Intelligence and Law*, Melbourne, ACM Press, New York (SSRN). This is the main paper setting out how the DataLex Project was migrated to the web (and the AustLII context) and the advantages this provided.

(DataLex, 1997a) Mowbray, A, Greenleaf, G, King, G and Cant, S, [Wysh Developer's Manual](#) [1997] COL 1, Australasian Legal Information Institute, Sydney (SSRN).

(DataLex, 1997) Greenleaf, G, Mowbray, A and King, G '[New Legal Services via the Web - AustLII's Research on Legal Inferencing](#)' (Chapter 5 of 'New Directions in Law Via the Internet: The AustLII Papers', *Journal of Information, Law and Technology (JILT)*, No. 2, 1997). This was the first paper published about WYSH. (SSRN).

(DataLex, 1997) Greenleaf, G, Mowbray, A, King, G, Cant, S and Chung, P 'Copyright law demonstration application' – computer application, using WYSH, SINO and AustLII Mark-up tools (Unix).

(DataLex, 1996-) Mowbray, A '[SINO free text search engine](#)' – software (open source) (Unix) (AustLII).

(DataLex, 1995b) Greenleaf, G, Mowbray, A, King, G, van Dijk, P "[Public Access to Law Via Internet: The Australasian Legal Information Institute](#)" [1995] *JLawInfoSci* 5; (1995) 6(1) *Journal of Law, Information and*

³³ *Northern Territory Law Handbook* (2017–, AustLII Communities) <<http://austlii.community/foswiki/NTLawHbk/NTLawHandbook>>

³⁴ O'Neill and Peisah *Capacity and the Law* (2nd Ed, 2017, AustLII Communities) <<http://austlii.community/wiki/Books/CapacityAndTheLaw/>>

³⁵ AustLII Communities: DataLex Web <<http://austlii.community/wiki/DataLex/>>.

Science 49 (AustLII). This is the first paper written about AustLII, and shows the influence of DataLex on AustLII's development.

(DataLex, 1995a) Mowbray, A 'Info-One alternative system demonstration' (consultancy project) – AIRS and HYPE application.

(DataLex, 1995) Greenleaf, G, Mowbray, A and van Dijk, P (1995) '[Representing and Using Legal Knowledge in Integrated Decision Support Systems: Datalex Workstations](#)' *Artificial Intelligence and Law*, vol 3, nos 1-2, pp 97-124 (SSRN). This is the major paper summing up the DataLex Project work to 1995 (ie pre-web and pre-AustLII).

(DataLex, 1994c) van Dijk, P (1994) *The Australian Insurance Law Workstation LEXeCOM*, International Business Communications, Sydney, 1994 (1st Ed, plus updates) – AIRS and HYPE application.

(DataLex, 1994b) Greenleaf, G, Mowbray, A and van Dijk, P *DataLex WorkStations Developers' Manual* (2nd Ed) DataLex Pty Ltd, 105 pgs (not on SSRN).

(DataLex, 1994a) Greenleaf, G, Mowbray, A and van Dijk, P, (1994), "WorkStations for different learning styles in law", *Proceedings of the Computer Assisted Learning in Law Conference*, Law Book Co., Sydney.

(DataLex, 1994) Greenleaf, G, Mowbray, A and van Dijk, P, (1994), *DataLex WorkStations User Manual*, DataLex Pty Ltd, Sydney, 40 pgs. (not on SSRN) – Included all versions (Windows, Mac, DOS, UNIX); for use with YSH versions of both Privacy Workstation and IP Workstation.

(DataLex, 1993b) Greenleaf, G, Mowbray, A, and van Dijk, P *Re-development of the SCALE computerised legal information system ('SCALE Mk 2) – Report to the Attorney-Generals' Department*, DataLex Pty Ltd, November 1993.

(DataLex, 1993a) Greenleaf, G and Mowbray, A (1993), "[Controlling and augmenting legal inferencing: ysh, a case study](#)" *Proceedings of the 4th Int. Conf. on Artificial Intelligence and Law* (Amsterdam 1993), ACM Press, New York, pp162-166 (SSRN).

(DataLex, 1993) Greenleaf, G (Ed), Griffith, P, Mowbray, A and van Dijk, P, (1993-94), "The Intellectual Property Workstation", Sydney, DataLex Pty Ltd, (1993), 40 MB - computer application (Unix, DOS/Windows, Macintosh) – sold commercially.

(DataLex, 1992c) Tyree, A (1992) 3(1) '[The logic programming debate](#)' *Journal of Law and Information Science* (AustLII).

(DataLex, 1992b) Greenleaf, G and Mowbray, A, (1992-94), "DataLex Workstation Software - Application Developers Manual (Student DOS version)", DataLex Pty Ltd, Sydney.

(DataLex, 1992a) Greenleaf, G, Mowbray, A and Tyree A, '[The Datalex Legal Workstation – Integrating Tools for Lawyers](#)', (1992) 3(2) *Journal of Law and Information Science* 219 -240 (SSRN) (also in *Proc. Third Int. Conf. on Artificial Intelligence and Law*, ACM Press, 1991). This 1991 paper was the first to set out the DataLex approach to the theoretical advantages of integration of approaches.

(DataLex, 1992) Greenleaf, G, '[The Privacy Workstation](#)' (1992) 6 *International Yearbook of Law Computers and Technology*, pp177 – 196 (SSRN).

(DataLex, 1991c) Greenleaf, G and Mowbray, A [The Privacy Workstation - User Manual](#) Version 1,07, May 1991, DataLex Pty Ltd, Sydney (SSRN).

(DataLex, 1991b) Greenleaf, G (Ed), Mowbray A, Gunning P and van Dijk P, (1991-94), *The Privacy Workstation*, DataLex Pty Ltd, Sydney, 6 MB - computer application (Unix, DOS/Windows, Macintosh) – sold commercially.

(DataLex, 1991a) Mowbray, A 'DataLex Workstation system' – software (Unix, DOS/Windows, Macintosh) (integration of YSH, AIRS and HYPE into one package).

(DataLex, 1991) Mowbray, A 'YSH (y-shell) inferencing system' – software (Unix).

(DataLex, 1990a) Greenleaf, G, Mowbray, A and Tyree A '[Legal workstations — integrating tools for lawyers](#)' *Australasian Law Teachers Association (ALTA) Conference*, Canberra, September 1990 (SSRN).

(DataLex, 1990) Greenleaf, G 'Computerisation of public administration and the rule of law' *Australasian Law Teachers Association (ALTA) Conference*, Canberra, September 1990.

(DataLex, 1989d) Tyree, A *Expert Systems in Law* (Prentice Hall (Aust), 1989).

(DataLex, 1989c) Greenleaf, Graham, '[Legal Expert Systems - Robot Lawyers? \(An Introduction to Knowledge-Based Applications to Law\)](#)' *Proc. Australian Legal Convention*, Darling Harbour, Sydney, August 1989 (SSRN).

(DataLex, 1989b) Mowbray, A and Greenleaf, G 'LAWS OF AUSTRALIA Defamation demonstration' – HYPE and AIRS application (Unix).

(DataLex, 1989a) Mowbray, A 'HYPE hypertext system' – software (Unix).

(DataLex, 1989) Tyree, A, Greenleaf, G, and Mowbray, A '[Generating legal arguments](#)' *Knowledge-Based Systems* Volume 2, Issue 1, March 1989, Pages 46-51 (SSRN). This is one of the most detailed papers about the FINDER application of PANND A.

(DataLex, 1988c) Mowbray, A, Greenleaf, G, Tyree, A & Lewis, D (1988), 'Teaching Lawyers Online Retrieval: The AIRS Training System.', pp. 96-102.

(DataLex, 1988b) Greenleaf, G, Mowbray, A and Lewis, D *Australasian Legal Information Retrieval Handbook*, Butterworths, 1988; [Ch 2 Basic principles of legal information retrieval](#); [Ch 3 Improved retrieval and storage techniques](#); [Ch 4 Legal information retrieval in Australia \(AustLII\)](#).

(DataLex, 1988a) Tyree, A, Greenleaf, G and Mowbray, A (1988) '[Legal Reasoning: the Problem of Precedent](#)' in J Gero & R Stanton (Eds) *Artificial Intelligence Developments & Applications*, North-Holland, Amsterdam, 1988, pgs 231–247 (not on SSRN).

(DataLex, 1988) Greenleaf, G, Tyree, A and Mowbray, A (1988) '[Communications Aspects of Legal Expert Systems — Incorporating Them in Shells for Lawyers](#)' *Proceedings of the National Conference on Law, Computers & Artificial Intelligence* University of Exeter, United Kingdom, November 1988, 10pgs (SSRN).

(DataLex, 1987c) Greenleaf, G and Griffith, P 'COPYRITA – copyright law inferencing system' (LES application).

(DataLex, 1987b) Greenleaf, G, Mowbray, A and Tyree, A (1987) '[Legal Expert Systems: Words, Words Words...?](#)' *Yearbook of Law Computers & Technology* Vol.3, 1987, 119–136, Butterworths (UK) (SSRN). This is a much-expanded version of the Boston ICAIL paper, including details of the COPYRITA, INTEST and FINDER applications, and the LES, AIRS and PANND A software.

(DataLex, 1987a) Greenleaf, G, Mowbray, A and Tyree, A (1987) '[Expert Systems in Law: The DataLex Project](#)' (*Proc. First Int. Conf. on Artificial Intelligence and Law*, ACM Press, 1987) (SSRN) The first paper about the DataLex Project, and the first paper presented (with a demo of LES) at the first ICAIL Conference, Boston, 1987.

(DataLex, 1987) Tyree, A '[FINDER: an expert system in law](#)' (1987?) (unpublished) (AustLII).

(DataLex, 1986) Tyree, A '[Will Justice Fall to Bits? Expert Systems and Law](#)' (1986) 62 *Current Affairs Bulletin* 13-18 (AustLII).

(DataLex, 1985c) Tyree, A 'FINDER – finding cases inferencing system' – FINDER application.

(DataLex, 1985b) Mowbray, A 'INTEST – Intestacy legal inferencing system' – LES application.

(DataLex, 1985a) Tyree, A 'PANND A (precedent analysis by nearest-neighbour discriminant analysis) inferencing system' – software.

(DataLex, 1985) Mowbray, A 'AIRS retrieval system' – software (Unix, DOS/Windows).

(DataLex, 1985) Mowbray, A 'LES inferencing system' – software (Unix).