
Volume 67
Issue 4 *Dickinson Law Review* - Volume 67,
1962-1963

6-1-1963

Information Retrieval in Law: Problems and Progress with Legal Computers

Follow this and additional works at: <https://ideas.dickinsonlaw.psu.edu/dlra>

Recommended Citation

Information Retrieval in Law: Problems and Progress with Legal Computers, 67 DICK. L. REV. (1963).
Available at: <https://ideas.dickinsonlaw.psu.edu/dlra/vol67/iss4/2>

This Article is brought to you for free and open access by the Law Reviews at Dickinson Law IDEAS. It has been accepted for inclusion in Dickinson Law Review by an authorized editor of Dickinson Law IDEAS. For more information, please contact lja10@psu.edu.

INFORMATION RETRIEVAL IN LAW: PROBLEMS AND PROGRESS WITH LEGAL COMPUTERS

BY JOAN M. COVEY*

A discussion of the future of legal research and the search for improved techniques of indexing and retrieval dictates some reference to the opposing arguments concerning the nature of the law itself. Many in the legal profession insist that law is an art,¹ conceding it is also a system of reasoning which permits the communication of all pertinent thought and materials to be utilized in solving a specific problem, based upon the facts in controversy or the application of a rule of law, from a network within which all possible theories and arguments are operative. The second school is convinced that, as a result of the goals and dogma of the field, perhaps, law is a science² with a vocabulary interpreted by those trained in it, the effect of which is guided by both inherent and agreed upon meanings of common terms. When the required elements are present in a situation, the rules of law, or terms and theories implicit in them, are applied—rigidly—and essentially automatically. This almost mathematical application lends the entire collection of the written creation and interpretation of the law completely adaptable to the ends desired both by the users of it and those responsible for assuring concise entry into the mechanical means of recalling printed legal thought.

It is important to dwell upon the characteristics of law in order to compare it to the non-scientific fields; law is one of the best recorded segments of knowledge, with a language uniformly understood and interpreted, while the humanities are open to controversy as extensive as those interested in them. This very precision of the law in its attempt to maintain equality and justice among all litigants demands that every item of legal thought, in legislation, statutes, cases,³ or treatises, must be preserved in correlation—or the imperative goal of adherence to precedent is compromised if not ultimately lost.

* Law Librarian and Assistant Professor of Law, Dickinson School of Law; A.B., 1955, University of South Dakota; LL.B., 1959, University of North Dakota; Library Science, Drexel Institute 1960, 1963, Rutgers University, 1961.

1. "Is the nature of law such that it can be subjected to the precise formulations of mathematical logic without distortion?" Stover, *Technology and Law—A Look Ahead*, M.U.L.L. 1, 5 (March 1963). (MODERN USES OF LOGIC IN LAW is the newsletter of the American Bar Association Special Committee on Electronic Data Retrieval.)

2. "Lawyers are just perverted mathematicians . . ." LAW AND ELECTRONICS: THE CHALLENGE OF A NEW ERA 42 (Jones ed. 1962) (proceedings of the First National Law and Electronics Conference).

3. "In all 51 American jurisdictions there are about two and a quarter or two and a half million reported court decisions and the increase is about 25,000 per year." (This figure does not include quasi-judicial opinions.) LAW AND ELECTRONICS, *op. cit. supra* note 2, at 26.

The present legal classification system, though time consuming, is almost foolproof (if one has ample time and energy) primarily due to three factors:

1. Concentration of indexing among a few publishers.⁴
2. Uniformity of abstracting.
3. Uniformity of use of hierarchical⁵ indexes and subject-headings.⁶

An attempt is made to index all cases from all aspects of law involved, through abstracting and hierarchical indexing, so that each searcher approaching a problem, from any angle and from any set of books germane, will be led to the same constant—the cases concerning his problem. This consistent and almost guaranteed contact must not be lost in any future system devised, but the time demand and costly burden of it must be circumvented if the legal profession is to keep up with the increasing problems presented to the courts. Settlement of cases out of court, though growing, is not the answer for the litigant who feels he has been outrageously aggrieved, of whom there are enough that the courts shall always be congested.

The present research system is an elaborate indexing arrangement offering substantial assistance to the manual searcher. Although it is superior to others, "our present system is far from perfect, and we have adapted our procedures to compensate for the deficiencies."⁷ "The failure of the system lies in the enormous increase in the quantity of material which has overloaded the device and rendered it unwieldy."⁸ All of these items mentioned, inadequate indexing methods, insufficient storage, and inability to cover vast bodies of materials—both mentally and physically—are in turn the critical problems for which each experimental center seeks a solution. Primary emphasis is and should be on choice of indexing methods.

Even without the aid of costly computer machinery, improved indexing methods could be helpful in some areas. The use of a uniterm⁹ card system—with a subject heading classification scheme derived from the documents themselves—or an edge-notched card index system with numbers reserved

4. On the negative side, there is an abusive amount of both calculated and uncalculated publishing duplication, both in abstracting and reprint of court decisions. Perhaps this enhances competition, but the code of the bookman is ill-served.

5. "[E]ach entry takes part of its significance from the preceding entries and contributes some necessary element to those following it." Progression is from general to specific. "Examples . . . are the Dewey Decimal System and the West Key Number Index-Digest." Wilson, *Computer Retrieval of Case Law*, 16 S.W.L.J. 409, 421 (1962).

6. "Subject headings are seldom used in pure form . . . because of their length Normally the type of index employed is a mixture of subject heading with a modified form of hierarchical." *Id.* at 422.

7. Freed, *Machine-Assisted Legal Research*, 47 A.B.A.J. 764, 765 (1961).

8. Biunno, *Searching Legal Literature—An Appraisal of New Methods*, 46 LAW LIB. J. 110, 112-13 (1953).

9. The word is applied to any system in which descriptive words are selected to represent the actual text.

to identify headings or documents would serve to integrate data of a limited nature, or as an efficient means of cataloguing either a case law or book collection.

One information source in which the ultimate manual retrieval has reached fruition is that of *Shepard's Citations*, single volumes covering particular jurisdictions of the law within limited periods of time. Each volume, for example, *Shepard's Pennsylvania Citor*, is an index which provides "access to information . . . [by] tracing . . . the genealogy of an idea" based on internal case issues into subsequently reported cases, thereby permitting "the discovery of related subject matter not ascertainable by any other means."¹⁰ The *Citors* include law review articles and annotated reports. The material actually is ascertainable by other means, such as the typical manual search, though laborious and subject to oversight. The great contribution of the *Citor* is the mechanical organization of all case citations in a column offering instant notice of subsequent change or variance through the use of a simple minor key code.¹¹ The ideal of presenting a few key words to a machine or searching a file of cards to recover the initial citations does not manifest itself in this unique work; rather the *Citor* presumes knowledge of the pertinent citation and carries later information beyond it. The editorial work has been absorbed and the searcher does not have the benefit of the initial steps.

Very few imaginative indexing techniques have been forthcoming from commercial publishers in the past forty years,¹² and many of the advancements have been developed by technological sources like International Business Machines Corporation. It is difficult to evaluate the reasons for this failure by the profession to make heavier demands upon book publishers.

Perhaps as a warning note, if only to stress the responsibility undertaken by a few in the projects being perpetrated throughout the country, it is wise to consider the statement of Mr. Reed C. Lawlor, a member of the California Bar and former chairman of the American Bar Association's Committee on Electronic Data Retrieval.

The social scientists are already active in the application of modern scientific methods to the analysis of the judicial process and its impact on society. With their perseverance, they may make important

10. Herner, *Methods of Organizing Information for Storage and Searching*, 13 Am. Doc. 3, 13 (1962).

11. For example, 84 F.2d 641 (*Federal Reporter, Second Series*) is followed on the next line in a column with a300US 515 (official Supreme Court reports) indicating that the Supreme Court affirmed the case.

12. The first key-number system was initiated in 1879 with the *Northwestern Reporter* by the West Publishing Company. Commerce Clearing House began their loose-leaf services in 1913.

discoveries about the judicial process which are now just barely perceived in the legal profession. There is no way that the law can avoid the scrutiny of science. If the lawyers and judges do not participate in this work, it will all be done by others.¹³

The important thing is not to delay in revamping the retrieval system at the time when inadequate future retrieval is apparent and imminent. It is a well known fact that the courts and quasi-judicial bodies though theoretically bound to appraise and consider all relevant preceding decisions germane to any issue do not do so in practice; therefore some device must be inaugurated to assist the ideal review of precedents in the common law system. The superiority of the new retrieval methods, pointed out as follows by John F. Harty's analysis and statistical results, can only increase: "In a comparison of a computer search versus traditional manual statutory searching, the machine search was superior every time—the documents not retrieved by the computer represented only four per cent of those that were relevant; on the other hand, fifty-three per cent of relevant statutes were not found by the manual search."¹⁴ (Part of the fault might lie not with the searcher's limitations but with the disparity in state indexing systems, even in the form published for profit by a specialized editorial board.) The ultimate goal is to "be able eventually to set up a system that is simple enough to need no middle man between search requests and computer."¹⁵ This will require universal familiarity with any new uniterm system revising the present subject-heading terms or if the key-word-in-concept system¹⁶ prevails a better appreciation by the bar of the nuances of terms in their field. Ralph Shaw, a pioneer in technical reform, would argue that the future comes first but that only partial retrospective machine searching is a possibility.¹⁷ On these points he would appear correct. However, the past body of legal material is still adequately handled under the present scheme and could be absorbed later. Future systems should inevitably promote a more language-conscious profession, particularly oriented toward technical terms. The subsequent interpretation of authority would not be limited¹⁸ but would remain as important as it is with the lax pursuit of authority which pervades today. A brief review of the current and past research projects should serve to point up major areas of

13. Lawlor, *What Computers Can Do*, 49 A.B.A.J. 337, 339 (1963).

14. Gibbs & Adams, *A Report of the Second National Law and Electronics Conference*, M.U.L.L. 215, 216 (Dec. 1962).

15. *Ibid.*

16. See text accompanying note 25 *infra*.

17. Shaw, *Electronic Storage and Searching*, *The Times* (London), April 6, 1962 (Lit. Supp.), p. 235.

18. As Professor George W. Brown put it, "The ultimate responsibility for analysis rests with the user." Gibbs, *supra* note 14, at 218.

progress and stalemate; the following are representative of the twentieth century's efforts to adhere to the past.

The University of Pittsburgh Health Law Center, directed by Mr. Horthy, now has on tape the full natural language text of all Pennsylvania statutes on health; those of eleven other states, Arizona, California, Florida, Illinois, Maryland, New York, North Dakota, New Jersey, Ohio, South Carolina, and Washington, selected as representative of diverse areas of the country; the Pennsylvania Attorney General's opinions dealing with education; and the New Jersey court rules, evidence rules, and constitution.¹⁹ This uniterm retrieval system in which the terms themselves are the coordinate index is designed to store and retrieve large quantities of documents with no indexing prior to keypunching and transfer to magnetic tape. "Although costly at the present time, this approach is feasible, Director Horthy believes, because by the time a search system is operational it may be possible to get the full text, both past and present, into the computer in ways that are less expensive."²⁰

A full-text system dealing with concepts rather than with words—the "point of law" approach—is in progress at Oklahoma State University. This approach "involves . . . the analysis of each case for the particular pertinent issues or pivotal points on which the case rests."²¹ Imminent is the project of putting the language of the newly adopted Uniform Commercial Code on cards to compare it with the old Oklahoma law to determine whether any sections of the Code are mere rephrasings of the old law. (In the future this could be done before a new code is adopted by the legislature to graph its effect.)

The Jonker Business Machines firm handles Project Lawsearch. Based on optical searching devices, it has three objectives:²²

1. Provide an inexpensive office device.
2. Determine whether electronic research can rapidly and effectively search for documents using principles of coordinate indexing.
3. Explore the possibility of not having to change the atmosphere in which indexes have been published.

Arbitration cases for five states have been stored on magnetic tape, document numbers have been assigned by the IBM 1401, and a "root-index number"²³ will be assigned to each significant word uniterm in the text of the

19. *Id.* at 215.

20. *Ibid.*

21. *Id.* at 216.

22. *Id.* at 217.

23. Demonstration by Robert A. Wilson of the Southwestern Legal Foundation, Dallas 5, Texas, at the annual meeting of the American Bar Association in San Francisco held August 4-9, 1962.

cases. Under this project conducted by the Southwestern Legal Foundation, electronic correlation of the document numbers with the index numbers will permit retrieval of both citations to document numbers and a reprint of the desired case.

Under the Western Reserve University (center of documentation) indexing system, a "compromise between the extremes of specificity and abstraction"²⁴ (as would occur, on the one hand, by translating every word of a case into machine-readable form and, on the other, by following conventional indexing rules and translating only the index into machine form) is achieved. Only "significant words"—similar to the West Descriptive-Word Index—"which convey the important concepts in the passage are selected for indexing."²⁵ The advantages of this system are threefold:²⁶

1. The total volume stored is reduced.
2. The several generic concepts expressed by a word are encoded.
3. The search for specific words or concepts, or both, is simultaneous.

Unfortunately, the sort of research discussed in the preceding paragraphs has not met with universal approval in the law. As expected, resistance to both the embryo manual retrieval experimentation and the computer is gathering with able spokesmen²⁷ holding forth for random reasons: fear of diminution of professional importance, regret at mechanical intrusion upon the scholarly domain, ignorance of the facts and purposes of electronic investigation, dissimilar philosophical orientation and general cantankerousness.

The readiness with which many greet the notion of using computer systems in legal work and the evident ease of applying available electronic techniques to much of the work now done reveals the extent to which law in the United States, in thought and in practice, has already become an exercise in technical reason and an agent of the technological system. In significant measure, the judicial element is no longer charged with pursuing justice or creating law, but is concerned solely with the application and refinement of rules. Its aims are characteristically technical—the maintenance of formal order and social efficiency, in this case according to the dictates of the sovereign's will. The fact that so many are astounded by the notion that law might, in any practical sense, have other aims reveals how far the transformation has gone. Law is being absorbed by the technological system and made nonprofessional.²⁸

24. Melton & Bensing, *Searching Legal Literature Electronically: Results of a Test Program*, 45 MINN. L. REV. 229, 233-34 (1960).

25. *Id.* at 234.

26. *Ibid.*

27. *E.g.*, Wiener, *Decision Prediction by Computers: Nonsense Cubed—and Worse*, 48 A.B.A.J. 1023 (1962).

28. Stover, *supra* note 1, at 7.

A highlight of misinterpretation of what the information specialists are trying to do is achieved in a recent article by Mr. Frederick B. Wiener,²⁹ commenting at length upon the August 1962 Program of the Special Committee on Electronic Data Retrieval, the program entitled "Using Computers to Predict Supreme Court Decisions." The basis of this study consisted of the Supreme Court decisions on the question of the right-to-counsel of defendants as determined by forty-five determinative facts gathered into seven groups, to wit, the severity of the punishment, the characteristics of the accused, the time when the petitioner lacked counsel, circumstances surrounding the lack of counsel and waiver. Wiener assails these as the basis for any demonstration of predictability inasmuch as any one of these classifications could "lead to reversal quite apart from the right-to-counsel aspects . . ." ³⁰ (Granted that this is true, but the computer was programmed from the standpoint of the influence of certain facts upon the right-to-counsel issue and the decisions of the individual judges. The fact is that as predicted the earlier cases have been overruled³¹ on the basis of reinterpretation of the Constitution in application to this situation.)

Employing this interpretation of the deficiencies of the demonstration, Mr. Wiener proceeds to discuss the lack of need for a computer at all, by emphasizing that the West key word or American Digest system will do as well as any new uniterm or key-word-in-concept indexing system, since all prior law is recorded in each generation of cases,³² implying that a twenty-year span will recapitulate all of the theories and decisions which have gone before. (This is strange talk for a practicing attorney well aware of that old and only case right on point.) The documentalist's point is well made through Mr. Wiener's example of the law initiate who, after searching all of the state digests comes up with no case in point, and is queried by the judge as to whether or not he has "thumbed the reporters?"³³ (That is a highspot in free testimonials.)

Mr. Wiener's main objection appears to be that he distrusts the experts who will scan the material for input selection and programming. Mr. Lawlor has recently set forth numerous authoritative arguments refuting, though

29. Wiener, *supra* note 27. "In the opinion of the present writer . . . the notion that a computer can predict the course of judicial decision rests on assumptions that are demonstrably untenable, does violence to the very nature of law, and is moreover certain to blunt the professional techniques of any lawyer who relies on machines rather than on his own powers of reasoning and advocacy." *Id.* at 1023-24.

30. *Id.* at 1027.

31. *Gideon v. Wainwright*, 372 U.S. 335 (1963) overruled *Betts v. Brady*, 316 U.S. 455 (1942), distinguishing the facts.

32. Wiener, *supra* note 27, at 1026.

33. *Ibid.*

scarcely mentioning, Mr. Wiener other than to refer to his opinion that "law is not an exact science."³⁴

Mr. Lawlor is quite explicit in emphasizing that the aim of the Committee on Electronic Data Retrieval is the achievement of a balance of man with machine in the belief that "if logic has not been the 'life of the law,' the law would have little life without it."³⁵ "Computers may give the law a new life which logic alone failed to supply. Computers may help us understand the nature of the law more fully."³⁶

The onslaught of all of this contention and work should resolve itself into pushing the vanguard of reform into the purview of the legal profession. It is this very awareness which the legal profession should be seeking and encouraging.

Perhaps the most immediate and tangible problem in the lawyer's view is that of the legal ramifications of adoption of some computer system, individualized or centralized. Some of these have already arisen or are predictable, but some are as yet unforeseen. A brief review of a few might foreshadow the catharsis in store for the legal profession. Where magnetic ink is employed in imprinting a depositor's number on checks for purposes of saving time by the employment of an automatic scanning device (and about eighty per cent of the checks drawn on federal reserve banks are now of this type³⁷), the practical value and implication of the depositor's signature would diminish and the apparent reliance upon the bank's cashier duty of checking the signature proportionately.³⁸ The bank's liability on a check bearing a forged signature³⁹ might well be altered to require notification by the depositor of loss of the numbered check before liability could be imposed upon it.

The United States Patent Office with a file of three million patents, each containing a short statement of what is new, presently makes 100,000 searches per year through a hierarchical classification system which grows with the art.⁴⁰ One product of this mode of search should be to render patent seekers less vulnerable to litigation due to ignorance of lack of novelty.⁴¹

34. Lawlor, *supra* note 13, at 339, citing WIENER, BRIEFING AND ARGUING FEDERAL APPEALS 146 (1961).

35. "[I]t is not certainty we see in the law, it is less uncertainty." Lawlor, *supra* note 13, at 339.

36. *Ibid.*

37. Federal Reserve Bank of New York, Press Release No. 124, March 6, 1963.

38. Johnson, *Automation, Forged Checks and the N.I.L.*, 14 BUS. LAW. 1008 (1959).

39. *Price v. Neal*, 3 Burr 1354, 97 Eng. Rep. 871 (K.B. 1762).

40. *American Documentation Institute State of the Art Symposium*, 13 AM. DOC. 1, 13 (1962).

41. Freed, *A Lawyer's Guide Through the Computer Maze*, Prac. Law., Nov. 1960, p. 15, at 44: There will be "fewer barriers in the form of invalid patents, which, despite their lack of merit, have tremendous exclusionary power as a practical matter because of the great expense and long duration of patent litigation."

Employment of technological advances as evidentiary proof has been confined in courts⁴² while quasi-judicial use in performing primary functions has been pervasive. Once adoption of electronic equipment and computers becomes a business custom it may become prima facie evidence of negligence when the fact is established that such custom was not complied with.

In the field of federal tax administration, the ideal of centralized records has been accomplished through automatic processing of paperwork⁴³ and integration of the side functions of storing facts through the use of magnetic tape, one reel of which can store hundreds of records easily modified or altered. Efficiency in record-keeping is further improved through the use of microfilm as the transfer of such output records to microfilm offers easy reference to a given record in a matter of seconds through the use of microfilm reading equipment. "Electronic Data Processing has made it possible . . . to pull . . . out of a deepening hole and halt what had become a dangerous and costly inability to keep up with the vitally important job of auditing returns. From the standpoint of revenue loss alone, this . . . has been well worth the expense . . ." ⁴⁴ Automatic data processing of course has advantages over the electro-mechanical or manual systems in that it "eliminates the need for physical movement of documents or punched cards from desk to desk or from machine to machine . . ." ⁴⁵ The method is less expensive and delinquency can be quickly ascertained.

In a relatively simplified area such as today's title search, the laborious operation might be improved upon by recording on cards all grantor and grantee deeds within a master file index, arranged by name, geographic subdivision, and by five or ten-year intervals. Each grantee card (and grantor) would reveal all prior recorded events affecting this tract. A micro-facsimile of the pertinent deed of the latest grantee would be placed on the verso side of the grantee card containing on its face the preceding chain of conveyances and any other relevant transactions.

42. See *State Wholesale Grocers v. Great Atl. & Pac. Tea Co.*, 154 F. Supp. 471, 497 (N.D. Ill. 1957), wherein a ten-question survey producing 4,600 responses tabulated by an IBM machine was admitted by stipulation.

43. "One of the more important characteristics of the electronic computer . . . is . . . its capability of performing a combination of operations on facts stored in the so-called internal memory of the machine and of producing results that satisfy many different needs. . . ."

. . . . From a single initial recording of the data on the tax return and its conversion to magnetic tape . . . it is possible to program an electronic computer" Surrey, *Automatic Data Processing and Tax Administration: The Potentialities of EDP*, 17 TAX L. REV. 165, 166-67 (1961).

44. Murphy, *EDP and Tax Administration in New York*, 14 NAT'L TAX J. 223, 226 (1961).

45. Surrey, *supra* note 43, at 169.

As a practical matter a computer processing center, financed by appropriate county support could serve the entire state. A photo of the master or present grantee card and a facsimile of the microfilmed deed could be returned by mail to be studied with a relatively inexpensive microfilm reader. The time element might be improved; the search burden of the attorney would be greatly relieved; and deeds of distant property would become more readily accessible. One problem would remain: the original recorder's copy of all deeds should still be maintained as the best and primary evidence. Perhaps the original recorder's copy will be made on microfilm eventually; two copies can be made, one kept and one sent to the processing center. Of course, one should not overlook the possibility that a computer processing center for this purpose might be operated profitably by a private concern.

Computers are already in use to aid the performance of clerical functions, the computations connected with estate work, in the courts of St. Louis County, Missouri.⁴⁶ The minutes were analyzed to determine how many entries could be standardized and pre-punched in cards. Three hundred and sixty-five index items, all but two per cent of the required information, were selected. Since "the basic operation of the court is to handle information, not figures,"⁴⁷ the accounting benefits, with a higher degree of control over the accounts receivable, are an enormous relief. With the use of the cards and equipment already on hand, a recent detailed analysis was made of the effect of proposed state legislation which would change the court fee structure.⁴⁸ Although data-processing, as used by this probate court, is currently limited to practical problems, it is inevitable that computer centers of learning based on the same hypotheses will tackle problems of theory.⁴⁹

It will be noted that micro-media and storage have not been discussed in more than cursory fashion because the problem of indexing appears most crucial at this stage. To be consistent with the theme generated herein, it is submitted that the highest contribution that the bar and nonspecialist could make to benefit the documentation cause would be to remain concerned with and educated to its problems while exerting pressure against the obstacles of apathy and conservatism.

46. Hensley, *Punched Cards Produce Progress in Probate Court*, 48 A.B.A.J. 138 (1962).

47. *Id.* at 139.

48. *Id.* at 138.

49. Freed, *supra* note 41, at 33.