Computers and California Law - A Time for Decision

Gordon W. Cook

Follow this and additional works at: http://digitalcommons.law.scu.edu/lawreview

Part of the Law Commons

Recommended Citation
Available at: http://digitalcommons.law.scu.edu/lawreview/vol11/iss2/4
COMMENTS

COMPUTERS AND CALIFORNIA LAW—
A TIME FOR DECISION

Since the inception of the term "jurimetrics" in the legal profession some twenty years ago, theorists in both the computer science and legal disciplines have awaited with eager anticipation the day when automated legal retrieval would become a commonality. It is evident in contemporary literature on the subject of computerized legal retrieval that even the most anxious authors are loathe to speculate on when such a system will become a viable tool in the hands of the practicing attorney. The topic of an economically feasible system has been variously approached as "... may be available in four or five years" to "... is a long way off in the future."

This comment examines recent developments in computer technology and their potential effects on the legal profession in California. It will reveal a new economic outlook on the feasibility of an automated legal retrieval system and suggest some alternatives for its most rapid and effective realization.

Perspective

Today's attorney is faced with the task of attempting to operate effectively in a quagmire of legal writings which would have dumbfounded his ancient predecessors. The much touted "population explosion" has taken its toll in the literary field with the advent of the "information explosion." These developments are causing the world's libraries to increase in volume at the rate of 35,000 words per second, or one million volumes annually. This is reflected in the legal profession by the number of reported judicial decisions increasing beyond the 2,300,000 level at a rate of approximately 22,000 per month.

The development of the current system of indices and cross-references has served to stave-off the time when the attorney can no longer effectively utilize the mass of information with which he

---

1 Loevinger, Jurimetrics: The Methodology of Legal Inquiry, 28 Law & Contemp. Prob. 5 (1963) [hereinafter cited as Loevinger].
2 Von Briesen, Status of Legal Use of Computers, 1 Law & Computer Technology, April 1968, at 9, 12 [hereinafter cited as Von Briesen].
4 Id. at 6.
5 Id.
is expected to deal. However, in the evolution to its present state, it has become obvious that there are severe limitations inherent in this system.

The inflexibility of the indices themselves serves to severely distort the logical sequence into which the report case should fall. Some of the present index headings date back to the early 1900's when a simple term like "negligence" or "automobiles" was sufficient to cover the then-existing profile of cases. Today, however, such a topic may cover hundreds of pages and even with the use of sub-headings is at best an ineffective and time consuming effort.

Further, the reluctance on the part of the indexer to add to the existing index terms results in a forced stratification which further distorts the true picture of case law. If the indexer finds a case without an appropriate sub-heading, he is apt to either leave out that which does not fit or force it into an inapplicable, but existing, index topic.

Another limitation somewhat akin to an indexer's reluctance to add index terms is the inherent human error factor. If the indexer should choose a term which is inappropriate or ineffective, his choice might well result in the case being "lost" to the researcher.

Finally, the exclusion of pertinent facts from indexed headnotes greatly diminishes their comparative value. Because headnotes normally recite only the legal principles involved, the researcher must lose additional valuable time by scanning the full text of a case which may not prove relevant.

In summary, we should see that today's legal research systems become increasingly ineffective because of the impending collision between client's pressure for more rapid and accurate answers and our inherited mass of case reports and statutes. As the number of reported cases increases each year, the utility of the current index system reaches an apex and can no longer be forced into higher levels of effectiveness. This results in frustration of clients because

---

6 Wilson, Computer Retrieval of Case Law, 16 Sw. L. J. 409, 410 (1962) [hereinafter cited as Widson].
7 Id. at 411.
9 Id.
10 See Wilson, supra note 6, at 410.
11 Id.
12 Id. at 409.
13 Id. Effectiveness in an indexed system is measured by its freedom from bias and its ease of utility. In order to increase this effectiveness, while increasing the size of the information base, more categories must be created. This creation of new cate-
of undue delays in litigation and frustration of attorneys who must spend an unwarranted amount of time and money on research.

In order to enhance the clarity of the forthcoming proposals, the author will briefly examine computer fundamentals and the current state of the computer art.

**Computer Fundamentals**

A prerequisite to an understanding of the potential and current capabilities of an automated information retrieval system is a minimal knowledge of computer fundamentals.

The basic configuration of any computerized system consists of three elements: (1) *Inputting* the data to be processed; (2) *Processing* this data; and (3) *Outputting* the results.\footnote{14} The Input function, in the legal research context, has been the major obstacle in producing an economical system.\footnote{15} This is due to the methods by which data has been made "machine-readable."\footnote{16} Since a computer's decision-making ability is derived from its rapid utilization of electronic impulses\footnote{17} (decision-time is measured in billionths of a second,\footnote{18} or nanoseconds), the data to be processed must be transformed into "readable" impulses. This has, in the past, been done by the use of a typewriter-like device which punches holes in an 80-column data card.\footnote{19} This card is then processed by the computer with the presence or absence of the hole in a certain location giving rise to an appropriate electronic impulse.\footnote{20}

Once the data has been transformed, \textit{i.e.}, input, into machine-cognizable form, the Central Processing Unit (CPU) controls the activities performed on the data according to specific instructions called a "program."\footnote{21} The program is written prior to the inputting of data and is "stored" in the computer until requested.\footnote{22} The prohibitions is a direct degradation of the system's utility because the index system becomes more cumbersome.

\footnote{16} See Beard, * supra* note 8, at 370.
\footnote{17} Id.
\footnote{20} Id.
\footnote{21} See Johnson, * supra* note 14, at 11.
\footnote{22} Id.
gram itself is written according to various rules inherent in any given language, e.g., COBOL, FORTRAN, etc. These languages, while not difficult to learn, require a knowledge of both the intrinsic constraints of the language and a relatively sophisticated knowledge of the problem to be solved.

Once the data is input and the program written, they are "read" and "stored" in the computer. There are several types of storage capabilities, each normally defined by their "access-time," i.e., how many nanoseconds it takes the computer to retrieve a given article of information. Within the CPU itself is an area termed "working-storage"—known colloquially as "core." It is within this sector that all "thinking," i.e., arithmetic and transfer functions, is performed. Core storage is a high-priced mode of storage because of its extremely short access-time. It works on the principle of detecting the presence or absence of magnetic polarity in minute iron rings located on circuits within the core storage element of the CPU. "Disc-storage" is the next most expensive storage device and has only a slightly higher access-time than core storage. Disc systems operate on the principal of detecting the presence or absence of magnetic "spots" on a rapidly revolving disc or series of discs stacked one on top of the other with room between each for a "read-write" head. This head detects the spots and relays the impulses to the CPU. The slowest form of storage, and naturally the least expensive, is the magnetic tape. The slowness of the tape is due to the fact that it is not a "random-access" device as are core and disc storage, but must be read sequentially, which means going over much irrelevant data to gain access to the desired information. The capacity of one reel of tape is roughly equivalent to 1,000 case reports.

The last of the basic elements of the computing system is the "Output Function." This normally is characterized by a high-speed line-printer capable of printing 1,100 132-character lines of information per minute. The primary cost of the output function does not

24 R. Brightman, B. Lusk & T. Tilton, Data Processing for Decision-Making 208 (1968) [hereinafter cited as Brightman].
25 See Beard, supra note 8, at 370.
28 See Columbus, supra note 23, at 11.
29 Id.
30 Id.
31 Id.
32 See Wilson, supra note 6, at 413.
33 See Brightman, supra note 24, at 231.
lie with the printing device itself, but with the tremendously expensive CPU which controls its operation.\textsuperscript{34}

Two other concepts which an information retrieval system must utilize relate to the relatively new concept of time-sharing. This technique permits as many as 120 programs to be run on a single computer at the same time.\textsuperscript{35} Because of the computer’s extreme speed capability it is able to devote fractions of a second to each program on a recurrent basis, so as to make each user believe he has unique control of the machine. Time-sharing has been made even more valuable with the adaptation of existing telephone lines for communication between the computer and remote “terminals.”\textsuperscript{36} These terminals are basically a combination of an input keyboard, \textit{i.e.}, similar to a typewriter, and either of two types of output. The slower and least expensive means of output is a computer-controllable typewriter.\textsuperscript{37} The more dynamic means, and one which offers unlimited potential for future use, is the Cathode Ray Tube (CRT) system.\textsuperscript{38} This system utilizes the principle of television to place information on a small TV screen which may, upon request, be transformed to a hard-copy configuration.\textsuperscript{39}

With this brief consideration of computer fundamentals in mind, an analysis of the various applications of computer systems to legal research should be more easily understood and thereby more fruitful.

**LEGAL RESEARCH APPLICATIONS**

**Retrieval Systems**

Computerized legal retrieval systems developed to date may be broadly categorized into two groups: (1) index systems; and (2) full-text systems.

\textit{Indexing systems}. Of the various hybrid types of indexing systems the “Headnotes or Reporters Abstracts” and “Partial Full-text” methods are representative and, hence, worthwhile of review.

The “Headnotes or Reporters Abstracts” method of retrieval inputs the abstract or headnote resulting from an analysis by the

\textsuperscript{34} Id. at 233.
\textsuperscript{36} Dennis, \textit{Shall We Put The Law Into The Computer?}, 1 LAW & COMPUTER TECHNOLOGY, Jan. 1968, at 25 [hereinafter cited as Dennis].
\textsuperscript{37} Note 24, supra at 344-345.
\textsuperscript{38} Note 26, supra at 107.
\textsuperscript{39} See Dennis, supra note 36, at 28.
editor of the most important or useful information to be gleaned from a particular case. Some of the various techniques which have been attempted include the "point of law" approach, concept of decision" approach and the "descriptor" approach. A New York firm, Law Research Service, Inc., has provided their subscribers with a series of glossaries termed "Computer Thesauri." Each thesaurus covers a specific topic such as "Corporations," "Negligence," etc. The thesaurus is a compilation of words and phrases, i.e., descriptors, under which each decision has been classified. Every descriptor is assigned a unique ten-digit identity number. The attorney consults the appropriate thesaurus and, with the use of a remote terminal, directly interrogates the computer with the ten-digit identity numbers representing the appropriate descriptors. The computer then retrieves up to ten relevant citations for each descriptor and informs the attorney if additional cases are on file. After acquiring the requested cases, the attorney then proceeds as in a conventional case analysis by reading the cited cases from his existing library, or by requesting a complete copy from Law Research.

While much can be said for the economies of this system, the effective accuracy of such a system is no less suspect than our existing manually-used index system. The attorney is still forced to rely on the interpretation of a case reached by a human, albeit professional, indexer. Therefore, if the descriptors chosen by the attorney are not identical to those chosen by the indexer the case may be "lost." Since the system uses a similar hierarchical indexing system, as was criticised in our current system, the distortion arising from stratification and inflexibility are to a large extent retained.

The sole benefit of this system is economic. These economies arise from working with a much smaller database, i.e., mass of information, and perhaps more significantly, the comparatively inexpensive cost of original input.

The "Partial Full-Text" system as construed by Dennis, is a derivation of the Full-Text system to be discussed below. In this system the editor excludes all the "unnecessary" or "unimportant" words from the text of the case and then proceeds to correlate the

---

41 Id. at 284.
42 Id. at 285.
43 See Beard, supra note 8, at 371.
44 Id.
45 Id.
46 Eldridge & Dennis, The Computer as a Tool for Legal Research, 28 LAW & CONTEMP. PROB. 78, 90 (1963) [hereinafter cited as Eldridge].
remaining 30 percent of the words with those requested by the re-
researcher. The difficulty with this system, as with abstracted sys-
tems, is the intervention of the human judgment factor into the
selection of the database. The "partial full-text" system utilizes
a statistical analysis approach to the selection technique which
greatly diminishes the volume of data needing to be searched. The
fact remains that unique or little-used words, which may for any
given case be extremely significant, are excluded and thereby the
case is lost for research purposes.

Full-text systems. The full-text technique of legal retrieval
places the actual text of the decisions and statutes, rather than an
index or abstract, on computer discs or tapes. The statute or
decision is assigned a unique document number by which it may be
retrieved. Each word of the statute or decision is given a four-part
word locator number which denotes the number of the document,
the number of the sentence, and the position of the word within the
sentence.

When the total database has been entered on the tape or disc
in this manner, the computer performs a "concording" function. This
function consists of scanning the entire database, eliminating
all common words such as, "the, "it, "of, "and, etc. This results
in eliminating approximately 50 percent of the database. Assuming
that what is left are significant "key-words," this extensively abbreviated database is then alphabetized. While alphabetiza-
tion is being done, all duplicate words are combined and their
respective locations are added to the representative word in sequen-
tial order. This process results in a concordance tape which con-
tains all the significant words along with their respective locations in
the database.

The search of a full-text system begins by the researcher choos-
ing those words which he believes effectively characterize his prob-
lem. These would be the same words he would use in a conventional
search and would not be constrained by an indexing thesaurus. The
researcher then requests these descriptive words to be correlated
against the concordance tape. When each descriptive word is "found"
by the computer, it relates the cites to the researcher; however, as

47 See Kayton, supra note 3, at 19.
48 See Eldridge, supra note 46, at 91.
49 See Kayton, supra note 3, at 14.
50 See Beard, supra note 8, at 372.
51 Hudson, Some Reflections on Legal Information Retrieval, 6 OSGOODe HALL L.
52 Id.
53 Id. at 274.
there are likely to be as many irrelevant cases having the same
descriptor as relevant cases, the researcher merely requests the
computer to "sift" those words having identical document location
numbers until he acquires a manageable number of citations. More simply, any document with two or more of the descriptors
would normally be of greater importance than a document with fewer descriptors.

The advantages in efficiency of the full-text system should be
obvious when related to the foregoing discussion. There is no need
to rely on any human judgment other than the judicial representa-
tion of the case; this is due to the lack of prejudice inherent in
machine-analysis. There is likewise no distortion of logical place-
ment of each case by the forced stratification of a hierarchical index.
There exists a total cross-classification of legal theory not reduced
to the ineptitudes of rigid structure.

The disadvantages of the full-text system have been primarily
economic. The cost of keypunching the total text of a 450,000,000
character reporter series has been estimated at approaching the one-
half million dollar mark. The scanning or processing-time involved
by the CPU is also excessive due to the greatly expanded database.
However, the author will attempt a more definitive comparison of
economic feasibility in a later section of this comment.

A more sophisticated technique of the concorded full-text ap-
proach is a statistical profile of each case within the database.
This system stresses the temporal or frequency association rather
than the more simple occurrence analysis. It is based on the assump-
tion that a word appearing with more than expected or average fre-
quency in a document, in some measure characterizes the content
of the document. Statistical association makes the further assump-
tion that the appearance of two words with greater than usual
frequency in a given sentence relation, implies a relationship between
those words or the ideas or objects they represent. Thus, if "base"
and "glove" occur with greater than average frequency, the docu-
ment probably concerns baseball; if "base" occurs with "submarine"
the document most likely refers to naval installations. Since match-
ing and statistically analyzing the word-occurrences of a particular
case is an eminently "machineable" activity, and since the result of
that process is a "profile" of the case, it is relatively easy for the

54 Id.
55 Id.
57 See text accompanying footnote 72, infra.
59 Id.
computer to retrieve cases on a “profile” basis rather than on a non-statistical key-word or index basis.60

There are two additional modes of information analysis which might prove computer-compatible in the future. While neither of these would circumvent existing search techniques, the “Key-Word-in-Context” (KWIC) system61 is another means of information presentation, and the utilization of computerized Shepardization processes will insure the efficacy of retrieval systems. KWIC is a means of displaying a key-word within the context in which it is found in the requested document. The computer will locate the document and display the key-word imbedded in the 60 or 70 words in which it is contextually found. By this technique the researcher can further refine his citation analysis to the most applicable of the presented cases.62

To assure the continued thoroughness and efficiency of computerized retrieval, recent techniques have been developed which enable the computer to “recognize” citations within the body of the case.63 Through the use of this capability, networks of cases can be created much like those in the Shepard’s Citator.

The Synonym Problem

The richness of the English language in terms of synonymy has proven a major, although not insurmountable, obstacle in the effective searching of case law by word similarity systems. Not only the extensive use of perfect synonym words, but “concept synonymy” impedes the task of the researcher.64 Thus, “piercing the corporate veil” might alternatively be termed, disregarding the corporate entity.” If the researcher does not use the expression or key-word used by the court, he will not receive a potentially relevant citation.65 There are several factors which serve to mitigate the chances of a “miss” due to this problem. The first is the tendency of the courts to either restate a holding in alternative form or utilize prior holdings which used alternative verbiage.66 Other factors are the attorney’s familiarity with the topic and ability to continually refine his request.67

60 Id.
61 See Beard, supra note 8, at 373.
62 Id.
63 Borkowski, et al., Structure and Effectiveness of the Citation Identifier, 3 LAW & COMPUTER TECHNOLOGY, Feb. 1970, at 42.
64 See Wilson, supra note 6, at 429.
65 Id. at 430.
66 Id.
67 Id.
Possibly the most important factor in diminishing the extent of the problem is the use of a thesaurus by the researching attorney. Two approaches have been suggested for deriving a meaningful thesaurus; one being a human correlation between existing commercial thesauri and the database, and the second being a computer-derived thesaurus from the existing database. The latter, in the author's opinion, is the only realistically efficient alternative. With the magnitude of the databases which we are to consider, the human discretion necessary in the former would make it uneconomical.

An additional element of the thesaurus capability in the computer is to develop word-stemming applications. For instance, if the descriptor were the word "amputation" in a personal injury case, the researcher would miss all those cases in which the term was stated "amputated." The thesaurus, or interactive program, must therefore be capable of reviewing the descriptor in terms of synonymy and word variation.

State of the Art

The discussion thus far should illustrate both the need for an extensive revision of our present legal research system and the availability of computer science principles to come to its aid. It would at this point prove beneficial to re-evaluate the feasibility of an immediate utilization of these principles in light of the more recent developments in the computer art.

In our prior perusal of computer fundamentals we have observed that the prime bottleneck preventing the widespread utilization of the available computer technology is the prohibitive cost of inputting the large legal database. It was estimated as recently as April of 1970 that to convert the database of a single jurisdiction would involve transforming some 50 million words at one cent per word or one-half million dollars.

Perhaps there is no better indication of the speed of technological development than to note that by early 1971 this cost will likely be reduced to the $50,000 mark. The reason for this tremendous

---

68 See Kayton, supra note 3, at 26.
69 Id. at 28.
70 Id. at 36.
72 The computation is based on the Dec. 2, 1970 interview with Dr. Dan Forsythe of Information International Inc., Los Angeles, California. Dr. Forsythe stated that their Graphix I system, which became available on a commercial basis January 1971, will rent for $40,000 per month.
Dr. Forsythe stated that while their average reading speed is approximately 2400 characters per second, the frequent font changes necessary to read legal data would lower the average speed to approximately 1,000 characters per second. Since the
reduction in conversion or input costs is the recent technological developments in Optical Character Recognition (OCR) devices.\textsuperscript{73} These devices eliminate nearly all of the preparatory steps and the actual keypunching of the database character by character.\textsuperscript{74} Instead of being "punched" by a keypunch operator, the entire database is fed through the OCR system page by page and, through the use of electro-sensitive "scanning" techniques, the characters are "recognized" and transformed to conventional magnetic tape or disc.\textsuperscript{75} This process may be accomplished at a rate of up to 2,200 textual pages per hour, which is roughly equivalent to the output of three keypunch operators in a normal seven-hour shift.\textsuperscript{76}

Until recently the objection to inputting case law with OCR has been the inability of the OCR devices to recognize a wide variety of type-fonts.\textsuperscript{77} While OCR's have been on the market since 1967, they were mostly of a single-font reading capability or at most three to five fonts.\textsuperscript{78} Since the average case contains upwards of eleven font styles there was no possibility of utilizing an OCR device.\textsuperscript{79} Currently, Information International, Inc. of Los Angeles, California, has manufactured an "omni-font" OCR which is capable of reading both fixed and proportional spacing.\textsuperscript{80} The dynamic spacing capability permits the system to read text printed on a linotype machine; previously these systems could only read data converted to a common typewriter-like spacing. This system became commercially available in January 1971\textsuperscript{81} and facilitates a totally

average word contains seven characters, this would equate to 142,857 words per second or 514,285.2 words per hour. Since 50 million divided by 514,285.2 is equal to 97.22 hours, and since Information International's rental rate is $55.55 per hour (i.e., $40,000 \div 720/hr. in one month, based on 30-day mo.), the actual reading would cost approximately $5,400.

However, since Graphix I is itself a complex computer, it must be programmed to format the machine-readable copy, i.e., tape or disc, in a design compatible with existing searching techniques. Graphix I must also be staffed for purposes of error analysis, of which there should be virtually none if reading a case reporter series. It is difficult to imagine how these costs would amount to much more than 10 times the actual reading cost, so the $50,000 figure might well be considered a conservative estimate.

\textsuperscript{73} Advertisement, DATAMATION MAGAZINE, July 1970, at 141.
\textsuperscript{74} Graphix I Specification; promotional literature of Information International Inc. Copies on file in the Lawyer office.
\textsuperscript{75} Id.
\textsuperscript{76} Note 73, supra at 141.
\textsuperscript{77} See Hamilton, supra note 15, at 102.
\textsuperscript{78} Gebremedhin, Optical Character Recognition—Performance Up, Prices Down, 12 DATA PROCESSING MAGAZINE, June 1970, at 42, 46.
\textsuperscript{79} See Hamilton, supra note 15, at 102.
\textsuperscript{80} Note 74, supra at 5.
\textsuperscript{81} Interview with Dr. Forsythe, note 71 supra. Dr. Forsythe reaffirmed Information International's promotional literature by indicating that Graphix I would be readily applicable to "read," i.e., convert to machine-readable form, legal case reports as of its inauguration date in January of 1971.
new technique in font-identification. The data base is initially micro-filmed within the system and new fonts are "learned" by being identified character by character via a Cathode Ray Tube (CRT) terminal keyboard. This learning process is said to take less than one minute per character.

It must be kept in mind that the cost of input is a one-time outlay. Once the database is put in machine-readable form, the cost of duplicating the existing tapes is negligible. This factor becomes important when we discuss the use of numerous timesharing databanks which will use the same database. Also, the subsequent costs of updating will be minimal. Since state appellate reports are public domain and the state is responsible for their being published, they have control of the means of publication. These means of publication utilize new innovations in typesetting techniques which include the use of typesetting by computer. Since the computer must be able to read what it must typeset, it requires that judicial decisions be made machine-readable. Therefore, using these source-data tapes for updating the existing database will all but eliminate the updating input costs.

Other recent developments in computer "hardware" include new modes of core-storage which significantly reduce the storage cost per character and reduction in prices of remote-access output terminals. Precision Instrument Co. of Palo Alto, California, has developed an electro-optical mass storage system that compacts information 1,000 times more densely than present magnetic tape. The new system uses microscopic laser beams to burn the metal coating from a mylar based tape which is then machine-readable with the use of a reading head similar to that used in magnetic tape systems. While total storage capability may be expanded to approximately 18 billion words, a system for use in a legal databank could use a smaller version which is currently available for $360,000. This development will cut the current rapid-access storage cost by approximately seventy-five percent.

Terminal technology has drastically reduced the cost of remote terminals. CRT terminals which would be readily adaptable to a legal retrieval system can now be rented for $88 per month with

---

82 Note 74, supra at 16.
83 Id.
84 See Skelly, supra note 19, at 32.
85 See Hamilton, supra note 15, at 103.
86 Id. at 104.
87 Advertisement, DATAMATION MAGAZINE, July 1970, at 117.
88 Id.
89 Id.
an optional ability to produce hard copies immediately upon request.\textsuperscript{90} Also currently available is a 30 pound portable remote-access terminal which may be carried as an attache case and merely connected to the nearest telephone line.\textsuperscript{91}

One of the not so recent, but equally important developments, is the "software" (series of programs) developed by the Judge Advocate Staff of the Air Force. The LITE (Legal Inquiry Through Electronics) system has utilized the full-text retrieval method described earlier for some time, and has developed what appears to be an effective interface with the practicing attorney.\textsuperscript{92} For the time being the most significant point of interest is their mode of instruction to the researcher. Their system is so formatted that a person having no background in computer fundamentals can be taught the art of search-framing, \textit{i.e.}, phrasing the appropriate descriptors for retrieval, in a very few hours.\textsuperscript{93}

In highlighting the various developments within the computer sciences, the author hopes to show that the coming of automated legal retrieval systems is no longer in the misty realm of the "future." The savings brought by the concepts of timesharing and optically scanned input make the cost-per-search factor so reasonable that to use older, less efficient methods would be purely an economic waste of effort.\textsuperscript{94} It is on the assumption that legal data centers have now become economically feasible, that the author bases the following proposals for the consideration of the practicing California Bar.

PROPOSALS

\textit{The Computer Advocate}

The most important point to be gleaned from this brief discourse should be the realization that during the period of rapid technological development in the computer sciences field, there has not been a commensurate development of scientific orientation within the legal profession.\textsuperscript{95} The majority of individuals who have advanced commentaries and done experiments in processing techniques did not realize that the economic capability would so soon foist it-

\textsuperscript{90} Promotional data for Hazeltine 2000 CRT terminal; Hazeltine Corp., Little Neck, N.Y.; on file in the Lawyer office.
\textsuperscript{91} Promotional data for Execuport 300 portable terminal; Computer Transceiver Systems, Inc., Paramus, New Jersey 07652; on file in the Lawyer office.
\textsuperscript{92} See Johnson, \textit{supra} note 14, at 21.
\textsuperscript{93} \textit{Id.} at 27.
\textsuperscript{95} See Hamilton, \textit{supra} note 15, at 107.
Because the ideas which they sought to advance could not be brought to immediate fruition, the legal commentators have dealt primarily with semi-esoteric topics. This has resulted in the failure to establish a viable program which would ensure the development of a class of persons capable of actually bringing the benefits of computer technology to the law office and court room.

In order to bring together in workable form a totally new retrieval system, we must strive basically for two goals. We must first utilize existing technology as quickly as possible to develop a workable retrieval system. We must then develop a law school curriculum which will promote the development of an individual who will make extensive use of the new tool in his case preparation, and who will continually seek to improve the utilization of the computer by the practicing bar; i.e., a Computer Advocate.

Needed: A Working Model

The author proposes that the California Bar Association establish an organization responsible for reaching these goals in the most effective manner. This organization's first efforts should be directed toward the creation of a legal research timesharing model. Once such a working model has been created, its formation should yield sufficient background and legal-electronic expertise to form an advisory board to oversee the commercialization of future facilities and the realignment of law school curricula to include appropriate course work.

The primary creative efforts in the development of the initial model, while coordinated by the California Bar's proposed organization or committee, should come from the law schools. The law schools would have a more intrinsic interest in such a model than would a commercial firm. This is due to the profound effect its evolution would have on the way the schools' students will practice law.

The composition and status of the schools in the local society would also offer inherent benefits. Law schools admit individuals from a wide variety of backgrounds. Most schools, especially those with evening divisions, have within their ranks a substantial number of individuals who are currently active in California's growing computer science industry. Programmers, Systems Analysts, or engineers

96 Id. at 103-07.
97 Id. at 107.
98 See Elardo, supra note 18, at 565.
would be invaluable in establishing a workable model. Many schools are part of a university which has significant computer hardware either in use by its engineering departments or its administrative offices. Access to this hardware would give the students the ability to write, debug, and analyze the programs necessary for a legal model. Likewise, it would permit them to gain valuable expertise in the various aspects of a legal retrieval system.

Most importantly, the interaction of the law schools with the Bar's organization in creating a workable model would create a body of knowledge and base of interest sufficient to establish an effective computerized legal retrieval course within the law schools' curricula.

If these systems are inaugurated on a commercial basis: (1) there is the possibility that a lack of legal knowledge within such firms may result in not providing optimum services to the attorney, i.e., the commercial time-sharing firm will be more concerned in selling additional time than allocating funds for developing improved search systems or new applications; (2) commercial firms will have neither the facilities nor the desire to educate the practicing bar to the extent or depth that will be effective. A basic education of how to run an automated search is not adequate. The attorney must have sufficient knowledge of the systems functions to bolster his confidence in it and thereby more fully utilize his new tool.

Commercially Directed

The Ohio Bar Association has chosen the commercial route toward an automated legal retrieval model and currently has the most advanced retrieval system in operation.

The Ohio Bar has formed a non-profit organization known as OBAR (Ohio Bar Automated Retrieval), whose board of directors is comprised of Ohio Bar Association Officials. This organization has worked with Mead Corporation, Dayton, Ohio, to create a workable facility whose database consists of 600 volumes of legal reports or roughly 200 million words. The data consists of the Ohio Supreme Court Reports, Ohio Appellate Reports, Court of First Decision Reports as well as the Ohio Statutes.

---

101 *Id.*
102 Interview with Mr. Robert Malone of the Marketing Division of Mead Corporation, Dayton, Ohio. Mead Corporation provides all hardware, technical and technological assistance required by OBAR.
103 *Id.*
currency of the database, OBAR maintains a "hotfile." This is a separate file into which the newest cases are placed within two to three weeks after being reported.104

OBAR uses the full-text machine-concorded type of searching system and, through a time-sharing capability, offers to each subscribing attorney, judge or county library a time-sharing terminal at the client's choice of location.108

Two types of services are offered by OBAR. The first is a problem solving service to which the attorney sends his problem by mail and receives copies of the relevant cases by return mail. This service is offered for $35 to $50 per search depending on computer time used.106 The second type of service, and the type recommended by OBAR, is for the attorney, judge or student researcher to do their own search with the time-shared terminals. The cost of this service is estimated to average $20 per inquiry.107 This cost is comprised of a $75 to $200 per month service charge for the terminal and a charge for computer time of approximately $16 to $30 per search, depending upon the efficiency of the searcher in framing his case.108

This system was to continue in a "qualitative analysis" mode until April of 1971 in which OBAR limited the number of terminals to 30.109 The purpose of this mode of operation was to acquire statistical information on the use-patterns of the various researchers. For this reason the restricted number of terminals were placed in both large and small law offices as well as law libraries and service bureaus.110

In the service bureau installation, the researcher may either actively solicit his queries from the computer or may request the aid of a legal assistant to "parameter" (place in key-word form) his inquiry. At this time it is estimated that the attorney-researcher can learn to effectively use the terminal on his own in less than four hours.111

OBAR's projection is to have the system fully implemented on an unrestricted state-wide basis by December of 1971.112

104 Id.
105 See Troy, supra note 94, at 62.
106 Id. at 68.
107 Id.
108 Id.
109 Note 102 supra.
110 Id.
111 Id.
112 Id.
Academically Directed

The achievements of OBAR must be given their well-deserved recognition as pioneering efforts in the field of computerized legal retrieval. In this author's mind, however, California is in a position to rely on the expertise of these pioneering efforts and recent technological improvements to eliminate two significant drawbacks in the OBAR system.

By promoting a commercialized model, OBAR has neglected any efforts to cultivate a confident user population through educating either the practicing bar or law students. As discussed earlier, the importance of absolute accuracy to the practicing attorney, plus a lack of knowledge concerning electronic devices, will prove a major obstacle in widespread utilization of Ohio's new research tool.

A second factor is that of economy. Professor Kayton reported in 1966 that keypunch-conversion of a one million word database would cost $20,000. Since OBAR's database approaches 200 million words, their conversion costs alone were approximately $4,000,000. While avoiding any hard and potentially imperceptive estimate, California's use of OCR technology should enable it to convert a like amount of data for one-twentieth of Ohio's expenditure. Further economic advantages could be obtained through the use of a consortium of law schools. Much of the second-party profits inherent in time-shared computer hardware rental could be eliminated by using the slack-time normally existent in a university's data processing facility. Such a consortium might also eliminate much of the salaried-technical expense by granting research credit to assisting students. A natural outcome of this student-law school participation in the creation of a model would be an appropriate realigning of the elective curricula to accommodate computer instruction.

In a July 1970 survey of law schools there were only six law schools in the United States which offered computer-oriented courses in their curriculum. These include Boston University, Indiana University, Cornell University, University of Michigan, Rutgers University and University of Virginia. These offerings ranged from basic "Computers and the Law" to "Symbolic Logic in the Law." Of these schools only the University of Michigan offered more than an introductory survey course.

\[\text{References}\]

113 See Kayton, supra note 3, at 16.
115 Id.
116 Id.
The absence of a California law school from those reported is particularly evident.

**Conclusion**

Because of the time and expense involved in current methods of legal research, the California Bar can expect to become the target of an increasing number of commercial research organizations. If the Bar fails to take timely action in the evolution of its researching techniques, it may well lose the control or participation it rightfully should have in the development of such an important factor in the future practice of law. At this point in the technological evolution, commercialization, while very possible, would be a tenuous venture without the academic and economic assistance of the California Bar Association. In the near future, however, the technological advances discussed above will render a full-text retrieval system so economically feasible that commercial firms will no longer need the Bar’s support, since they could afford the costs of experiment and error. At that point the Bar could no longer effectively control the direction of legal retrieval’s progress in California. The Bar would then have to rely on the vagaries of the market place to provide it with an effective tool to meet its growing needs.

The efficiencies of OCR, time-sharing, mass storage and pioneering intellect have provided the California Bar with a doubled-bladed decision. The technology necessary for an economical computerized retrieval system is now available and awaiting use. However, if no decision is made, the practicing attorney will be forced, through economic coercion, to use a system which is not his creation and of which he has little or no knowledge.

_Gordon W. Cook_