

3-1-1977

The Use of Data Processing in Litigation

James D. Prendergast

Recommended Citation

James D. Prendergast, *The Use of Data Processing in Litigation*, 10 Loy. L.A. L. Rev. 285 (1977).
Available at: <https://digitalcommons.lmu.edu/llr/vol10/iss2/1>

This Article is brought to you for free and open access by the Law Reviews at Digital Commons @ Loyola Marymount University and Loyola Law School. It has been accepted for inclusion in Loyola of Los Angeles Law Review by an authorized administrator of Digital Commons@Loyola Marymount University and Loyola Law School. For more information, please contact digitalcommons@lmu.edu.

THE USE OF DATA PROCESSING IN LITIGATION

by James D. Prendergast*

I. INTRODUCTION

"Byte my baud." This quote probably means absolutely nothing to the average reader of this article. And well it should not. The quote uses terms that are part of the language of a group of people who work in air-conditioned rooms with raised floors, occupied by spinning tapes, discs and blinking lights. A computer programmer or systems analyst would know that the word "byte" is a generic term used to indicate an identifiable grouping of consecutive binary digits ("bits"), *e.g.*, an eight-bit or six-bit byte. "Baud" is a term used to signify a unit of data transmission speed equal to the number of information bits transmitted per second.

This article purposely begins with a play on words. The world of automatic data processing, comprised of esoteric academic studies such as cybernetics and mundane functions such as programming a payroll, has a language all its own. The language not only represents an integrated and comprehensive body of thought, but also reflects mental processes which are foreign to most lawyers. The lawyer and the systems analyst live in different worlds, think differently, speak different languages and react differently to the same sets of stimuli.

What makes the lack of communication between the world of law and the world of data processing so disturbing is that the legal profession and the practice of law in twentieth century America desperately need the assistance of data processing to remain economically viable. The practice of law is one of the last cottage industries in the United States. By history and practice, almost every legal function must be custom tailored to fit the client's specific needs. Economies of scale and effective production management are alien to the legal profession, a

* B.A., 1965 (University of California, Berkeley); M.B.A., 1967 (University of California, Berkeley); J.D., 1974 (Hastings College of Law). The author served as Publications Editor of the Hastings Law Review. He is presently associated with the law firm of Meserve, Mumper & Hughes.

The author wishes to thank Robert Argyle Keller, Esq., who provided extremely important editorial assistance in the preparation of this article.

profession that has not advanced too far from the quill pen and legal briefs bound by red ribbon. Whether it is because the legal profession has virtually priced itself out of the reach of the middle class or because of the nearly impossible human task of getting one's arms around an exceptionally complex litigation case, the legal profession is waking up to the fact that data processing and the use of modern technology may make the difference between a profession that can cope with the twentieth century and a profession that will be run over by it.

One thing which will surely cause a lawyer to take note of data processing and its crucial importance to the effective practice of his art is to encounter another attorney assisted by data processing. This could occur at a deposition when opposing counsel drags one's client kicking and screaming through the facts of a given case aided by a computer printout, which chronologically lists all of the relevant documents in the lawsuit with a supporting narrative of the contents of the documents and their relevance to the case. While one rummages through boxes of manila envelopes with paper-clipped documents, opposing counsel is merely glancing at a concise printout and asking penetrating questions. Perhaps it takes such a sock in the jaw to convince a lawyer that the technology which enables men to walk on the moon might be of assistance to the practitioner of the legal arts.

II. BASIC CONCEPTS

The purpose of this article is not to berate the legal profession for its aversion to data processing. Rather, this article presumes that the reader wants to learn about data processing and how it can assist him in the conduct of litigation, and that he is willing to wade through a minimum of technical jargon to understand automatic data processing. This understanding is crucial. To apply data processing effectively to litigation, the lawyer, as decisionmaker in the lawsuit, must understand the capabilities and limitations of data processing. It is the lawyer who must understand what computers can do in order to envision how the technology that is available can assist him in proving his case. The programmer and systems analyst can provide assistance toward a specific goal, but the definition of goals must be left to the lawyer. It is that lawyer who must understand the technology available in order to formulate goals which must combine what is legally necessary with what is technically possible through data processing.

Once the decision is made to employ data processing, the lawyer must

face the computer salesman and choose from the available technology the best mix of hardware and software to do the job. With only a vague understanding of what one wants to accomplish with data processing, coupled perhaps with an irrational reverence for the new and modern, and with no understanding of what product mix might best fit one's needs, the attorney is an easy mark for the data processing salesman. With a fancy soft-shoe routine, aided by slide shows and tours of customer installations, the salesman can take advantage of ignorance, confusion, and desire. The attorney ends up with hardware and software capable of putting a man on the moon, with a corresponding price tag, when all he wanted to do was keep track of documents in a lawsuit. The converse might also happen. The attorney, because of ignorance and a tight purse, might buy technology that cannot do the required job, with resulting wasted effort, expense, and futile results. It will therefore be the further purpose of this article to convey enough technical competence so that the attorney can ask the right questions of computer vendors, understand the answers, and acquire hardware and software suitable for the attorney's corresponding needs.

Insofar as the majority of the functions the computer will be called upon to perform for the legal community, data processing merely takes a human function that can be defined in human terms and performs that function at phenomenal rates of speed. Operations within a computer are measured in terms of nanoseconds—one thousandth of a millionth of a second—within which to perform a function. However, one should not confuse speed with accuracy or omniscience. When the lawyer puts the computer to work, he is not putting a man on the moon, controlling mass transit, an oil refinery or an early warning missile system. Rather, in most cases, when the attorney resorts to data processing to serve a need, he generally uses the computer to perform a clerical function. The computer is merely doing in an incredibly short period of time what it would take hundreds of individuals to do in a much longer time. The only difference is that, as the magnitude of a lawsuit increases and the documentation expands to thousands or hundreds of thousands of documents, the computer, performing a clerical function such as document storage and retrieval, can handle the job within the time frame required for litigation.

Even though the computer functions at incredible speed, it is still merely a machine which cannot think. A computer might be able to play chess, but only because a human being has programmed every conceivable move into the system. The computer, receiving a given move from

an opponent, can scan its memory banks and select the optimum response. The computer, however, cannot invent moves, and it is certainly not intuitive. Therefore, even though the computer does things at incredible rates of speed, it can only do what it has been told to do, no less, and certainly no more. As a result, one cannot be vague with the computer, incomplete with instructions, or provide it with only part of the information. It requires precise and complete instructions and information. Accordingly, before bringing in a computer to assist in litigation, the lawyer must understand his case and what he hopes to prove so that clear instructions and sufficient information may be given to the computer system.

The burden on the lawyer is two-fold. First, he must have a basic understanding of the capabilities of data processing so he can correlate those capabilities with what he hopes to accomplish in his legal context. Second, he must know what he wants to accomplish in that legal context. Once he has adequately defined what he wants in the context of the law, he can then apply data processing to reach that goal. This is probably the most difficult obstacle to the effective use of data processing because most lawyers wait until the last minute to effectively grapple with the legal and factual contentions in a case.

This brings to mind the term *GIGO*. This term does not stand for "gospel in, gospel out," which some people enamored by data processing might consider it to mean. On the contrary, the term "GIGO" means "garbage in, garbage out." The computer will do exactly what it is told to do. Unless great care is taken at the outset in defining goals, in formulating the systems design, and in controlling the input to a data processing system, the output of such a system, provided at 2200 lines of print per minute, can be voluminous and absolutely useless. If this is the result, the attorney can only blame himself. Because of its speed, the computer can provide accurate information at incredible rates of speed and in great volume, or, conversely, it can provide garbage at incredible rates of speed and in great volume.

The *systems design* correlates what the attorney wants to do with the capabilities of data processing. The first step in systems design is to state in precise terms what the attorney hopes to accomplish through the use of data processing. Once the goals are defined, the attorney, with the assistance of data processing personnel, can work backwards from those goals and define, step by step and in minute detail, exactly how the computer is to use the information to reach those goals. This process results in the systems design.

Input refers to the information that the computer will process within the systems design. This information, or data, could consist of correspondence germane to a lawsuit which an attorney might want stored in the computer and then returned to him in a chronological, document number sequence, or grouped according to addressee. However, the computer will only return what has been provided. If the initial information is not reliable, or erroneous, the computer will merely reformat the information or put it in a desired sequence. Unfortunately, the computer cannot make unreliable information reliable or wrong information correct.

Regardless of the type of information that constitutes the "input" to a systems design, the attorney alone knows why each item of information is significant to the lawsuit and to the systems design. The attorney must make the initial decisions as to significance and relevance. The computer can correlate or order information according to predetermined parameters, but the computer cannot formulate its own guidelines as to importance. This designation of significance and relevance is the function of the *coding scheme* used to describe an item of input in the systems design.

Once the coding scheme has been developed, the next step is to transfer the information to the computer itself. The computer can ingest information only in machine readable form, *i.e.*, magnetic encoding on tapes or discs, holes in punched cards, holes in paper tape, and so on. One of the jobs incumbent upon the attorney is to translate the documents and other information to be put into the system into a machine readable form. Normally this is done by having clerical personnel code the documents according to the coding scheme. The information so coded is punched into cards or recorded via a typewriter onto magnetic tapes or other machine readable media to be fed into the computer.

Now, if we tie the above comments together, it becomes obvious that the linchpin of an effective computer system is an intelligent and thoughtful systems design which describes in detail the use to which the computer is to put the information it is given, the desired output from the computer and the purpose for which that output is to be used. Only by carefully designing the system to accomplish the end result, prior to taking the first step towards coding a document, may an attorney obtain from data processing the maximum benefits of this new technology. An incomplete, haphazard, or build-as-you-go system design will only create

headaches down the road and perhaps result in a system that, after much expense, is useless to the attorney.

III. USES FOR THE COMPUTER IN THE LEGAL ENVIRONMENT

Although this article directs itself primarily to the use of data processing in litigation, it is perhaps useful to mention other areas of legal practice where the computer can be of use. The use of data processing for more than one application in the legal environment can be crucial to the decision of whether to acquire in-house computer capability, to use a service bureau, or to forget data processing. Obviously, if more than one application can be found, the fixed costs of hardware and software can be spread.

Within the legal environment, the application of data processing can be broken down into six functional areas:

(1) *Accounting and office management functions*, which would include accounting for attorney time, generating accounts receivable monthly statements, and creating profit and loss statements and balance sheets for the law firm, and other functions generally associated with office management and accounting. The legal profession has started to use data processing in accounting and related functions, but data processing has not been fully applied in this area. The concepts of cost accounting are as applicable to a law firm as they are to a manufacturing firm. It might be useful, for example, to designate each client as a profit center with its separate profit and loss statement. The revenues received from each client would then be balanced against the amount of time charged to that client to determine whether or not the client, over a period of time, is generating a profit.

(2) *Document generation*, where the computer is used to store complete form documents; *e.g.*, joint venture agreements, wills, deeds of trust, and standard form interrogatories. The attorney can then take the form document and customize it to fit specific needs. The computer prints out a final, completed document through its ability to rapidly make changes, reform paragraphs, and correct typographical errors. This is an exceptionally useful function for the computer, because, as mentioned above, the legal profession must lower the cost of its product. A couple of points probably should be mentioned in this regard. The first is to consider the cost effectiveness of equipment. A docu-

ment has to be of a certain length, over five to seven pages, before computer document generation is economically justified. There are other types of equipment, such as mag-card typewriters, which are probably more cost efficient for the rapid production of small documents. The second point is the necessity for library management and control of documents stored in a data processing system. Documents must be constantly updated, the design of the form document must be related to the ease of the operators using the computer, an index of available documents must be established and attorneys made aware of how to properly annotate a form document for ease of processing. Once the procedural steps are established a data processing system can be of great benefit in the generation of documents.

(3) *Legal research*, where the computer is used for the storage and retrieval of statutory and case law. Although publishing companies and others have entered the field of legal research utilizing data processing, such application is still in its infancy. Current technology speaks in terms of trillion bit memories. With that amount of on-line storage, it is perfectly conceivable to store the entire statutory and case law of the United States, all services, and all secondary sources directly on-line to a computer system for access by the attorney. In the future, as the cost of hardware is reduced, an attorney could have a cathode ray tube (CRT) on his desk, with a hard output copying device, thereby conducting legal research without leaving his desk. The biggest stumbling blocks today are not hardware or technology, but designing the software for legal research and encoding the statutory and case law in machine readable form.

(4) *Document storage and retrieval in litigation*, which is the subject of this article.

(5) The most interesting and imaginative use of data processing in the legal environment is as an *expert witness*. Here a data processing system, assuming the evidentiary foundation for its programming, data and output formats have been established, provides the attorney at trial with statistical and other reports as an analytical expert. An example of the computer as an expert is discussed in the last section of this article.¹

(6) Finally, the computer can be used for the development of *client relations* and the *marketing of legal services*. Because the computer has the capability of storing and handling vast amounts of information in incredibly short periods of time, it is possible to use the computer as a client contact mechanism. For example, a law firm

1. See pp. 321-23 *infra*.

might store all of its wills in a data processing system. Programming could then be set up to divide the wills into functional areas of the law which change from time to time. As a legal change occurs, *e.g.*, in estate and gift taxation, the computer can be called upon to search all of the wills and determine which ones will be affected by the change in the law. The computer can not only provide a listing of all of the wills so affected, but can also type out letters, using law firm stationary, addressed to each client involved, advising of the change in the law and suggesting that it would be appropriate for them to contact their attorney to discuss this change. Thus, in the area of client relations, the computer can not only assist by lowering the cost of legal services, but can also aid the attorney in monitoring client needs and providing continuously updated legal services. Imagination alone limits the use of the computer in client relations, within the confines of the ethics of the profession.

IV. TECHNICAL TERMS

A modicum of technical awareness is useful to the attorney not only in relating his legal needs to available technology, but also in avoiding expensive mistakes in acquiring the technical support required. The computer salesman, like the lawyer, uses his own language to his advantage, capitalizing on the ignorance of the opposition to fill with his product line bigger needs than may exist. For these reasons, the following lexicon will help to assist the attorney.

A. Hardware and Software

The world of data processing is divided into two parts: hardware and software. Hardware is the physical equipment, including a computer and peripheral devices. Software is the totality of programs used to extend the capabilities of the hardware.

Conceptually, hardware can be divided into two units. The first is the central processing unit (CPU) of the computer which is the area where the action takes place. The second unit consists of peripheral input/output (I/O) devices, which, in simple terms, provide the link between the person and the computer.

The CPU contains the main storage of the computer, the arithmetic and logic unit (ALU) and other internal circuitry which gives the computer the capability of performing its function. The ALU performs arithmetic and decisional operations. Conceptually it is the hardwired brain of the computer. Main storage, or "core," is the blackboard upon

which the ALU, as instructed by a program, performs desired functions using variable data which has been fed into main storage.

Software operates within the confines of the hardware. For example, one can view the main storage of the computer as a blackboard, conceptually divided into horizontal sections. On the first section is written the operating system, an integrated collection of programs, which is the umpire and coach of a computer, enabling it to act, both as to hardware and software, as a computer.

The next horizontal gradation of main storage is committed to operational or "user" designed programs (sometimes called "application programs"). These programs are written to direct the computer to perform specific user-oriented functions, such as the preparation of client monthly statements.

The last horizontal gradation of main storage is allocated to variable input data that is used by an operational program to provide a desired result. As an example, consider an operational program designed to provide monthly statements to customers. The operational program would tell the computer the accounting steps to be followed in generating a monthly statement. Once the operational program is fed into the computer, the next procedure is to input the variable data, such as transaction items for a given customer showing the quantity of goods purchased and their price. This variable information is then processed by the operational program, and the computer prints out the monthly statements on the line printer.

Software can be divided into two conceptual groups: The operating system of the computer and user written operational programs. In general terms, an operating system is a collection of programs written by the computer vendor that provides for the preparation and execution of the user's operational programs (jobs). Most operating systems are stored on an input/output device such as a tape or disc drive and called into main storage as needed by a portion of the operating system which is always core resident, that is, internal to the computer.

The operating system consists primarily of a control program, which is the umpire of the computer, directing and supervising all functions of the computer, and processing programs which can be divided into three categories:

- (1) Language translators, called "compilers," that take a user's operational program, written in a specific program language, and translate this program into the language of bits and bytes, the internal language of the computer. If one were required to write a program in that

internal language, and to know the exact configuration of bits and bytes necessary to impulse the ALU to perform a specific function or to represent a specific alpha-numeric character, programming would be exceedingly tedious. To overcome this problem, programming languages were developed which are, to varying degree, somewhat similar to English and/or customary mathematics. For example, the Fortran compiler is a program which accepts a program written in the quasi-algebraic Fortran "source" language and produces a machine language "object" program in the internal language of the machine.

(2) Utility or service programs provided by the vendor to perform basic data processing tasks such as sorting and merging alphabetical lists of data.

(3) User-designed operational programs, that become part of the operating system library.

As an example of the relationship between the operating system and user written programs, consider writing a program to print out accounts receivable monthly statements. A programmer would sit down with a programming language that is business oriented, such as Cobol, and write a program that describes the functions that the computer must accomplish, in terms similar to English, which the programmer can understand. This user written operational program will define the format of the variable information, and client transactions, which will be processed by the program. It will describe in precise detail all of the functions that the computer must go through, such as adding up the item purchases for each transaction and multiplying by a standard percent to obtain sales tax, in order to arrive at an accounts receivable monthly statement. Once this program is written on coding sheets, the programmer will key-punch the program, normally into cards, and feed these cards into the computer. The computer will first have been told by the operating system, keyed by "job control" cards preceding the accounts receivable program, that there is a user written problem program, written in Cobol, coming in. The operating system will then call upon the Cobol language translator to read the user written problem program and translate it into the format that the computer must internally use. The translated user written problem program can either be put out on a random access device such as a disc for recall by name only when needed, or punched into cards by an output device, a card-punch, and stored external to the computer until use.

In summary, software can be divided into two sections. The first is the operating system which includes all of those items listed above, including, in sophisticated data processing systems, translated user written problem programs that have been stored internal to the system for

retrieval by name only and have become part of the operating system. The other section of software are those user written problem programs, as written in, for example, Cobol, or as translated, which are stored external to the system to be used when needed.

B. Internal Storage

Most computers store information in binary format. This means that information is stored internally in the computer in the form of groupings of electronic on-off switches or "bits," a given sequence of bits signifying a given alphabetic or numeric character. In most computers, data is stored and transmitted between main storage and the CPU for mathematical and decisional operations in multiples of eight bits. As noted at the beginning of this article, these eight-bit multiples are called "bytes." In this regard, when we talk of a "256 KB" computer storage capability, we mean that the computer has roughly 256,000 bytes (kilobytes) of main storage or storage for 256,000 letters or numbers.

Thus, one can look at the innards of a computer as an ordered series of eight-bit bytes, whose contents and locations are managed by the operating system. When a program is loaded into the computer, every instruction of the program is assigned a specific byte or string of bytes. Variable data coming into the computer is similarly placed in specified locations so that the computer knows where all items of data can be found. This allows the computer to have immediate, random access to all information, stored internally in the computer, whether operational programs or variable data.

C. Input/Output Devices

Any device which sends information to or receives information from a computer is an input/output device. In simple terms, an input device is a mechanical unit designed to put data into a computer for processing, for example, a card-reader, a tape-reader, or a keyboard. An output device translates the electrical impulses representing data processed by the machine into a more permanent form, (such as magnetic writing on tape), or into humanly readable form, (such as printed forms or words on a display screen). Certain devices, such as magnetic tapes and discs, can be both input and output devices. Other peripheral equipment, such as line printers, serve only as output devices. As the industry advanced, the sophistication and complexity of input/output devices have

advanced correspondingly. They now include key-board terminals that may be carried in a briefcase.

D. Batch Processing, Interactive Processing and Teleprocessing

The attorney considering employment of computer assistance should be aware of the distinctions between batch processing, interactive processing and teleprocessing. Batch processing is a technique whereby items to be processed and variable data to be used by an operational program, are coded and collected into groups prior to processing. In the example of generating monthly statements, batch processing would require that all customer transaction cards be collected, coded and fed into the computer at the same time. The computer would respond to all inquiries or all variable data at once. Batch processing is normally done at the physical location of the computer.

To be distinguished from batch processing is interactive processing where an operator sits at a console typewriter, or other form of terminal such as a display tube, and queries the computer question by question, receiving a response to each specific question, before the next inquiry or item of variable data is addressed.

Teleprocessing is basically using the computer as the hub of a network of terminals where, over normal telephone lines or private lines, information from the terminal is sent to the computer for processing.

E. Data Format

To understand the arrangement of variable data, which is used by an operational program to perform a specific function, terminology is important. Three terms are frequently used: *field*, *record* and *file*. A field is an assigned area in a record to be marked with information, such as a field to be used for the name of an employee in the employee's payroll record. A record is a group of related facts or fields of information treated as a unit, such as a given employee's entire payroll record including the fields of name, social security number, year-to-date pay, and the like. A file is an organized collection of information directed toward some purpose, such as all employee payroll records of a given business.

Records, and in turn fields, can be of fixed or variable length. Consider a standard eighty column punch card, very akin to your telephone bill. There are eighty columns on this card running horizontally. Into each given column a number or a letter can be encoded by punching. A number is termed numeric, and a letter is termed alpha. If

letters and numbers are combined, the encoding is termed alpha-numeric. In simple computers, in order to work upon the information punched in the card, the computer must be told exactly where to look for certain information. For example, a document number might be punched into columns one through eight of each eighty column card, with each card representing a distinct document. Within the first eight columns, any eight digit number can be used to signify a specific document. This is a fixed length field, eight columns for a document number, in a fixed length record, eighty columns for all information encoded concerning a specific record.

More advanced data processing equipment can handle variable length fields and records as distinguished from fixed length fields and records. With a variable length record, the computer is instructed where the specific field or record begins and ends but not its length. The computer must be keyed by a certain symbol to indicate that this is the beginning of a specific variable length field or variable length record and there must be, at some point in a given data stream, other keys to signify the end and the beginning of the next known variable length field within a given variable length record, or signifying the next variable length record. Once the computer in scanning data comes upon such a key, the computer knows that this is a variable length field or record and will use the information between the beginning key and the ending key in a desired way. Variable records have the unique advantage of not limiting the user to a predetermined number of columns.

A sophisticated computer can "select" information contained in a variable field or record. For example, a record, with fixed fields for date and document number, might have a variable length narrative section describing the contents of the document. This narrative will describe what the attorney believes to be significant about a given document. Within this variable length narrative, the attorney can put, in English, anything he desires. Once this information is stored in the computer, the attorney can ask it to read through the variable length narrative fields searching for records which contain, for example, the word "misrepresentation" and print out all such records.

Once the computer extracts from the data file all of the records containing the key word "misrepresentation" in the narrative variable field, the records can be rearranged or sorted according to fixed fields such as date or document number. However, the computer cannot sort these records in a sequence based upon information contained in the variable field. For example, if significant dates have been included in a

variable field, the computer can pull out any narrative that contains a date by using its select capability, but it cannot then sort these dates into a chronological sequence. To do this, the records would have to be rewritten creating a new fixed field for the date selected from the narrative variable field.

F. On-line and Off-line

All information that is to be used by a litigation data processing system, such as narrative summaries of depositions, must be encoded into machine readable form. This may be done either on-line or off-line. On-line describes the operation of a computer system in which the input/output devices are under the control of the CPU, and in which information reflecting current activity is introduced into the data processing system as soon as it occurs. Teleprocessing and the use of terminals connected directly to the CPU are examples of on-line processing. Off-line refers to peripheral equipment or devices, or the storage of data not in direct communication with the CPU.

The distinction between on-line and off-line includes not only the manner in which information is stored for later retrieval and the manner in which this information is processed, but also the manner in which information is encoded in machine readable form. On-line encoding of information is the translation of data from hard copy to machine readable form directly into the computer by use of a terminal without an intervening step such as punch cards. Instead of ending up with a box of cards, the data is transcribed onto a magnetic tape or a magnetic disc by the computer directly from the terminal. The typical off-line encoding of information involves key-punching information which has been summarized according to a desired coding format, into the computer cards. Once the information has been encoded, a box of cards will represent a program or the variable information to be used by the computer. When processing is to occur, the computer reads the holes punched in the cards and the data is fed into the system.

Once the information has been encoded, the next step is to store this information for later retrieval. In an off-line system the variable data will probably be retained in boxes of cards to be used when necessary. If it becomes necessary to store the information in a more accessible form for speedy access, the punched cards can be fed into a computer, and recorded on a magnetic tape or magnetic disc. The tape or disc can then be removed from the computer and stored for later use by the system. Whether in cards or in some other storage medium, if stored

data is not directly and immediately accessible by the CPU of the computer, the data is said to be stored off-line.

On-line storage, on the other hand, is the storage of data directly accessible by the CPU and immediately usable by the system. As discussed above, main storage is the blackboard of the computer on which the computer handles information it is currently processing. However, main storage is generally limited to two to three hundred thousand bytes and is very expensive. Therefore, entire data files are not generally stored in main storage; rather, records are fed into main storage from on-line storage devices, such as magnetic tape or disc drives, when needed.

If data files are to be used by the computer frequently or in the regular course of business, on-line storage of the files, coupled with on-line storage of operational programs, allows for immediate processing. The trade-off to this immediate access is the inherent cost in committing a peripheral device, such as a disc drive, to the on-line storage of the particular data file. If immediate access is not necessary, for example, if inquiries to the file are batched and processed against the file on a periodic basis, the file can be stored off-line until needed. This cuts the cost of data storage because the peripheral device is released for other jobs.

V. THE NEED FOR A GENERALIZED SYSTEMS DESIGN IN LITIGATION

It cannot be overstressed that the computer can only do what it is told to do, and it must be told what to do in extreme detail. Every step the attorney wants the computer to take, every use of information, every output format must be specifically described to the programmer and systems analyst so that they can design the operational required programs. The effective use of data processing is jeopardized when an attorney believes that a computer is intuitive, and assumes that, if he merely codes information, dumps it into the system, pushes a button and asks the appropriate question, the computer will provide the appropriate response. Nothing could be further from the truth. One has to know exactly what output one wants from the computer, exactly what the output is going to look like, exactly what information is available to the computer to generate the output, and, in extreme detail, exactly how the computer is to work with this data to generate that output.

This necessity for painstaking detail is probably the greatest stumbling block to the effective use of data processing in litigation. It is, of course, detrimental to the effective use of data processing if the attorney does not take sufficient time to understand data processing so that he can think in terms of the regimen of processing, and design accordingly. A much more fundamental concern, however, is posed by factors inherent in the process of litigation itself. Each case takes on a life of its own as it progresses. It is very difficult to prognosticate all constituent elements of a case. As the process of discovery evolves, the whole complexion of a case may change. However, if the attorney takes the time, at the outset, to understand the case and the application of data processing to litigation, the initial design of a litigation computer system can be sufficiently broad in scope as well as sufficiently detailed for use so that the changes which do occur in a given case can be accommodated. The goal is to broadly design a system, suitable for that first case in the office that is going to use data processing, but flexible enough to fit the needs of subsequent cases.

Therefore, the first step is for the attorney to sit down and learn something about data processing and how other litigation cases have been handled in the past using this technology. The next step is for the attorney to decide exactly what he wants to do with data processing in the context of a particular case. This must be in as much detail as is possible, considering the point in the life cycle of a case where the attorney finds himself. He must analyze his lawsuit carefully and decide upon the relevant information he wants from a data processing system and how he intends to use it. If he wants the system to organize his evidence for trial, he may have to decide, perhaps at the inception of the lawsuit, how he intends to try a case in which he has yet to depose a witness. The goal is to design a generalized systems approach to litigation and then customize it to fit the needs and requirements of the particular case as it develops.

A. Systems Design

The first step in the development of a systems design involves the identification of conceptually distinct categories of information that will be relevant to the lawsuit and their interrelationships. For example, if the attorney wants the lawsuit to be tried on the basis of factual contentions, and wants the computer to list all of the witnesses, documents, and depositions that support a specific factual contention,

he need not, at the systems design stage, itemize all of the factual contentions that he intends to use. This is probably impossible without some progress in the case. The attorney must, however, decide whether he intends to use the category "factual contention" and store and retrieve information on that basis. Once this is decided, a field within a record will be allocated to "factual contention." As discovery progresses, separate codes can then be assigned to each specific contention.

After defining the categories of information, one must define the interrelationships among categories of information. For example, the attorney may want to store documents according to date in order to develop a chronology of the documents in the lawsuit. This raises the question of what date is going to be used for the purpose of chronology, and whether the system must allow for multiple dates. A letter normally bears the date the letter was sent. However, a specific letter may speak internally of one or more dates that may be, in fact, more significant to the lawsuit chronology than the date the letter was sent. Therefore, in designing the system, the decision has to be made as to whether to allow only for the document date for the purpose of chronology, or to allow additional fields within a record for what might be called pertinent dates, which reflect specific factual points raised in the body of the letter.

Once the lawsuit is understood and the application of data processing to it is defined, the attorney will be in a position to sit down with computer experts and design the system. This step in the formulation of a systems design can be a time consuming and expensive undertaking, at least the first time through. However, it is the foundation for everything else. It must not be done in a haphazard fashion. The design will require not only discussion of what the lawsuit is about, what the factual and legal contentions may or may not be, but also how discovery will be conducted, how information will be collected, how this information is to be processed through the system, what will be done with this information, how the trial will be conducted, whether summaries will be utilized and how these summaries will be formatted, and the interrelationships among information.

At this stage hardware and software constraints may limit the freedom of systems design. The attorney may want to do X, but the programmer or systems analyst might tell him that it is not possible; or, if he decides to go that way it will be very expensive; or, perhaps an alternative systems design would provide the same in-

formation at less cost. Through this give and take process, the attorney and his computer assistants will complete the systems design. It is then up to the data processing personnel to program what the attorney wants to do and the manner in which it is to be done.

Out of this cooperative effort will evolve a systems design which will include:

- (1) General analysis of what the lawsuit is all about.
- (2) The initial framework for the factual and legal contentions involved.
- (3) The types of documents and other information that will be obtained.
- (4) The quantity of such information.
- (5) The manner in which the information will be obtained.
- (6) General categories of data that the attorney wants to be stored in his data processing system, such as factual contentions, areas of damage, and authors and recipients of correspondence.
- (7) The initial framework for the specific entries under each category of data, *e.g.*, names of witnesses under witness category.
- (8) What the attorney wishes to do with each category of data; for example, if document date has been selected as a category of data, is a document date chronology required?
- (9) Any interrelationships among the categories of information. If the attorney asks the computer to retrieve a specific category, what other categories are to be automatically retrieved in conjunction therewith? If an inquiry seeks to retrieve all correspondence written by a specific company, will the system automatically retrieve all correspondence written by employees of that company by internally correlating companies and employees?
- (10) The types and formats of output required from the system; for example, does the attorney only want a narrative description of each document stored, printed out by document number, or does he want the computer to generate a pre-trial conference order?
- (11) The decisional and computational steps necessary to generate each desired output, based upon the variable data available to the

computer. This step-by-step analysis of how the attorney would manually prepare his output will be the framework for programming the computer.

Once the systems design has been formulated, the next step is to add a systems flow. This will basically be the movement of paper, information, coded information and output reports through the systems design. All members of the litigation team, whether they be attorneys, legal assistants or computer personnel, must understand how the variable data is gathered, and then how this information is handled by clerical personnel or the attorneys themselves for evaluation, review and encoding into the computer. An overall design of the flow of information must be related to the systems design so that everyone who is a part of the litigation team will understand how the variable data is to be managed, how documents are to be numbered, stored, retrieved and encoded, and how the documents are to be maintained in numerical order for easy access based on the printouts from the system.

B. Program Design

After the systems design and the systems flow have been established, the next step is a program design. This is the function of data processing personnel. However, the assistance of legal personnel is necessary to explain specific elements of the systems design. Program design refers to the preparation of the individual programs that will process the information to reach the results specified by the systems design. A given systems design may require as few as two or as many as 500 programs.

One product of the program design stage will be the definition of the form of input to go into each program and the form of output generated by each program. As an aside, many programs may in fact have no specific input or output that the attorney will ever see because the input to one program may be the output from another program, completely internal to the data processing system itself, carrying out different steps necessary to generate the final output which the attorney desires. The programming support personnel will constantly refer back to the attorney to discuss their design of input forms, *e.g.*, coding sheets, to determine whether or not the information which they have included on the coding sheets is that which is necessary for the program to do what is required by the systems design. Further, report

layouts, showing what the output of a given program will be, must be reviewed by the attorney.

This whole procedure, from the big thoughts department of the systems design to the nitty gritty of designing how many spaces will be allowed for a document number, is an interactive process; first among the litigation team to determine what they want from the system, based upon their evaluation of exactly how they are going to prove their lawsuit, and then between the litigation team and data processing support personnel to complete an appropriate, workable design and programming of the system. The result of this moderately arduous process should be a systems design manual. This will include not only the systems design and the scope of the job to be performed by data processing, but also a description of each significant program used, the format for the input and output for each such program, and a scheme for the general flow of information through the system.

C. Change of Design and a Generalized System

One serious unknown in applying data processing to litigation is the inability of the attorney, at the beginning of a lawsuit, to project six months or a year down the road what the legal and factual contentions of his lawsuit will be, what facts will be at his disposal, how he intends to prove the lawsuit, and, ultimately, what output he wants from his data processing system. Further, because of the cost and time involved in designing a computer-based litigation support system, it is preferable not to reinvent the wheel every time a case comes into the office which might benefit from data processing. Therefore, the goal is to develop a generalized systems design which can be customized to fit the requirements of a particular case, and allow for maximum flexibility to absorb changes that will occur during the life of a lawsuit. This can be accomplished by the attorney, when he first applies data processing to a litigation case, by broadly designing his system so as to develop a generalized litigation system simultaneously with the development of the specific systems design for the immediate case.

In the application of data processing to litigation support, the similarities between lawsuits greatly outnumber the differences. The similarities include the storing and retrieving of documents, the providing of data summaries, the development of pre-trial conference orders, and the

preparation of trial books. Therefore, a systems design should be flexible enough to be adapted to almost any litigation. An attorney can then customize a generalized litigation systems design to handle the types of information peculiar to his lawsuit and add the specifics to predetermined categories of information. The end result will be a generalized systems design for the handling of litigation in general, to be used again and again, at minimal cost. The only variable will be the degree of customizing that is necessary to make the generalized systems design useful in a particular case. Finally, if a generalized systems design can be used, allowing for customizing, at the onset of a case, all of the data categories and their interrelationships will be established, so that as change occurs during the process of discovery and trial the system will remain effective and useful.

Although a generalized systems design can be developed as a working tool for the attorney, there may be situations where a particular lawsuit is, in and of itself, unique to the point where the generalized systems design cannot be of use. If so, and the case is sufficiently large to warrant the design of a radically different system, then the attorney should repeat all of the development steps discussed in this article to formulate a new systems design. This should not occur often because the similarities between lawsuits and the arrangement of information usually exceed the differences, so that a generalized systems approach to the use of data processing can be formulated and can be of use.

D. Cost of Change

Finally, the cost of data processing in litigation will, to a great extent, be determined by the frequency of substantive change. There is a basic cost floor for the use of data processing in litigation determined by the size of the data files and the frequency of inquiry. But once this floor is established, cost will be dependent upon the frequency of change. Change is expensive in data processing, because a change to program *A* may affect every other program in the system. If the facts change, *i.e.*, the interrelationships between information and the types of output required, then so might the programs. The costs rise proportionally. If a change occurs, then it must be brought to the attention of data processing personnel as soon as possible so that current programming can be halted, the change evaluated, and the design of the program altered to avoid wasted effort and minimize the cost of change.

VI. APPLICATION OF THE GENERALIZED SYSTEMS DESIGN

The economic justification for data processing is that once variable data is recorded in a machine readable form, it need never be rerecorded. Information is permanently available to the computer system to be retrieved in any of a number of desired output formats using the select and sort capabilities of the computer. The clerical effort required to encode and arrange information is limited to the initial step of encoding. As discussed above, data is formatted into fields (the date of a letter), fields of information are assembled into a record (all information concerning a particular letter), and all records involved in a particular application are collected into a file (all of the letters and other documents involved in a given lawsuit). Once the record format has been designated and the fields within a record defined, the mass of variable data to be utilized in a given application can be recorded once, and then retrieved and rearranged as often as desired without any additional clerical steps.

However, before coding documents in a particular lawsuit can begin, the lawsuit itself must be understood in great detail. This understanding of the lawsuit will determine the significant items of information to be put into a computer system. This determination in turn will define the fields comprising the format of the computer record. Without such forethought, an attorney might include only the obvious information in the computer system. But the real utility of data processing is its ability to assist the attorney in arranging information according to legal and factual contentions, areas of damage, and other legal issues relevant to a particular lawsuit.

Once the lawsuit is sufficiently understood so that the relevant categories of information can be defined, the attorney can develop a coding scheme applicable to all documents in the lawsuit and allocate the information they contain to fixed fields and to a variable narrative field. This allocation will be structured to take advantage of the select and sort capabilities of the computer.

Information contained in, or derived from, a document (the computer record) can be conceptually categorized as either objective or subjective information. Objective information is that data which is readily observable from the face of the document. Such objective information would include exhibit numbers, the date of the document, the person sending the document, the person receiving it, and the type of document.

On the other hand, subjective information is data pertaining to a specific document or record that is not readily observable from the face of the document. Rather, an analytical step is required between reading the document and drawing any conclusion as to its subjective significance. Subjective categories of information might include the various legal and factual contentions in the lawsuit, areas of damage, and an assessment of the importance of a document to the lawsuit.

Having divided information into subjective and objective categories, the attorney must then relate this division to the capabilities of the computer. The computer has the capability of sorting, that is, rearranging records in a desired sequence, chronologically by date or in document number sequence. In order for the computer to sort, it must key upon a specific fixed field of information contained in a given record within the file. A fixed field is an area of information of predetermined length whose location within the record is always known to the computer. For example, a fixed field allocated to date would be comprised of six digits, two each for day, month and year. Its fixed location in the record, for example, positions ten through fifteen of a 100 position record, would be known to the computer.

The computer also has the capability of selecting a specific item of information contained in the variable field, a field of indeterminate length in which the position of individual items of information is not known to the computer. For example, a computer record might contain a narrative description of the contents of a document that will not be limited in length to avoid constraining the narrative. Particular dates, derived from the body of a document might be included in this variable narrative. The computer can scan an entire narrative and retrieve the whole or a portion of it, if a word within the narrative matches the item of information requested. For example, if the attorney desires to review any record that contains the date January 1, 1977, the computer can, at machine speeds, read all of the narratives and pull out those narratives that contain the date January 1, 1977.

The computer generally cannot sort out an item of information contained within a variable field, unless, after records are selected, they are rewritten as a new data file and the criteria for the sort rearranged as a fixed field. For example, those records selected because the date January 1, 1977 appears in the variable narrative field, could not, on that basis, be incorporated with the other records in the file for a chronological sort. Document date is a fixed field upon which the computer could sort, but the date derived from the narrative would not

be in a fixed field. Therefore, in order to sort out the records retrieved by date within the narrative, the records would have to be rewritten and the date from the narrative formatted as a fixed field. This rewriting of files is expensive and requires specialized programming. It can be avoided by defining all fixed fields to be required in the lawsuit during the initial systems design stage.

In most litigation applications, the attorney will want to sort his records on objective data, such as document number or date. This will be the type of information assigned fixed fields in the computer record. Subjective information such as legal or factual contentions, will normally be assigned to a variable field, composed of a digest or narrative of the important points contained in a record. The attorney may want to select against specific items of information contained in the digest. This can be done through the select capability of the computer and the documents so retrieved can then be sorted by the objective information contained in the fixed fields. This becomes the generalized framework for the classification of information in the systems design.

One caveat must be mentioned as to the number of fixed fields. Because of the capabilities of standard programming packages provided by computer vendors, fixed fields must stand alone within the system in order to take advantage of the sort capabilities. Further, the standard computer printout is normally 132 characters in width and the more fixed fields that are used, the more compressed the area available on the printout for the digest or narrative field becomes. Figure 1² is a printout of computer records with three fixed fields, document number, date and a three position alpha code signifying the type of document, for example DEP denoting deposition. By limiting the fixed fields in the record to three, most of the printout area is available to the narrative making it visually appealing and easy to use. Should a specific lawsuit require an extensive number of fixed fields that so compress the printout area available to the narrative as to make it unusable, customized programming can accommodate those requirements. Most computer vendors have standard program packages for the storing of information and selecting and sorting upon the records contained in the data base, and, because these packages are standard, they are relatively economical to use. Once specific programming is necessary to customize data storage and retrieval, however, the data processing expense of the lawsuit increases dramatically.

2. See app. at pp. 325-26 *infra*.

Consider a lawsuit for breach of contract and fraud involving the installation of a computer system. A customer sues the computer vendor for promising certain results if the customer would install a computer system, the customer alleging that the computer vendor failed to live up to his promises. The computer vendor first installed a card system, then upgraded the plaintiff to a disc system, making additional contractual promises in the process. According to the customer, neither computer system functioned as promised, and the plaintiff incurred loss of business revenue and expenses during the period that the defendant was trying to get the computer system to work as agreed. The customer further alleges that, when the vendor realized that the card system would not perform as promised and warranted, he represented to the customer that a disc system was required because of work load, but that this was false and an attempt to cover up the problems with the card system.

As discussed above, the first step is to analyze the lawsuit and to determine the relevant and significant categories of information. Figure 2³ is the abstract input form used for encoding the source data in this particular lawsuit for input into the data processing system. This abstract input form is a codified representation of what the attorney determined to be the significant categories of information in the lawsuit. It is directly derived from the mental process the attorney must go through in analyzing his lawsuit and defining those categories of information to be put into the system, and the manner in which he wants the information retrieved.

The first caption on the abstract input form, "control information," designates three fixed fields to be used for the computer record. Entry number is synonymous with document number, *i.e.*, that number which is stamped on each document during discovery. The entry number has two major purposes: (1) the key to indicate to the computer that a distinct computer record follows, and (2) the link between the computer printouts and the original documents, which will probably be stored in file cabinets in document number sequence.

The next item on the abstract input form is the fixed field for a pertinent date, that date which is of significance in the document, but may or may not be the date of the document itself. The third fixed field designates the document type, a three position alpha code such as LET for letter or DEP for deposition. Space is provided on the form for

3. See app. at p. 327 *infra*.

alternatives should the document not fall within the six predetermined types.

The reason for using predetermined document types on the form, as well as other predetermined codes, is to provide for ease of encoding by merely checking a block. The key-punch operator will key-punch the three position alphanumeric code corresponding to the block checked.

On the abstract input form, (Figure 2⁴), after the control information is completed, all the remaining information required by the form is included in a variable length narrative or digest field. Information included in the narrative field consists of both subjective and objective information. Objective information includes the author of the particular document and the recipient. Other objective information derived from the body of a document would be included in the area labeled "digest information," where the encoder would write, in ordinary English, a general description of the contents of the document. Subjective information would also be included in the narrative analysis of a document, but certain key categories of subjective information have been given specific codes. This subjective information comprises the legal and factual contentions in the lawsuit as determined by the attorney. In this particular case that information was broken down into four different categories: liability subjects, fraud subjects, damage subjects and general subjects.

When the abstract input form was initially prepared at the beginning of the lawsuit, space was provided on the form for increasing the specific codes under each of the above categories as knowledge of the lawsuit and its relevant facts increased. For example, under the liability subject category, only three specific entries were initially thought relevant to the lawsuit and designated L/1 through L/3, referring to disc system, card system and bill-of-material processor, respectively. However, space was provided on the form for up to nine liability subject codes. Because the digest is a variable length field, additional entries could be added if necessary.

The next predetermined categories of information on the abstract input form consist of the author and recipient of a specific record of information. The "A/" and "R/" are codes whereby a given individual can be selected as an author only or recipient only or, by eliminating the "A" and the "R" during a select, by the name of the person only.

Following these predetermined categories of information, space is pro-

4. See app. at p. 327 *infra*.

vided on the form for a narrative digest of the pertinent information contained in a given record. Again, the computer record is merely a predetermined composite of information distinct from all of the records. It need not be a physically distinct piece of paper. A line or paragraph of a deposition can be a distinct computer record if it is recorded with a unique document number.

The information contained in the last line of the form, above "abstracted by," indicates whether the document is a plaintiff's exhibit, defendant's exhibit, or other cross-reference. The code Pex/A/Com would indicate that the record is not only a plaintiff's exhibit, but it has also been attached to the answer and to the complaint. Coding schemes like this can be designed to suit the needs of the case.

Figure 3⁵ is another example of a coding sheet, or abstract input form, that is somewhat more general in nature than is Figure 2.⁶ Conceptually, it is again derived from the same generalized systems design for storage and retrieval. Reading down the form, the control information includes document number, trial exhibit number, discovery exhibit number, and document date. In designing this form, it was determined that there would be no need, for sorting purposes, to assign a fixed field to document type, and that document date alone would be of significance rather than an alternative pertinent date. Trial exhibit and discovery exhibit numbers were added as fixed fields since the lawsuit involved numerous exhibits, and because, for trial purposes, sorting on the trial exhibit number, the discovery exhibit number, or both, after a select against a certain legal or factual contention, would be necessary, for example, in preparation of the pre-trial conference order. Further, in order to avoid reserving a separate record number for every page of a deposition, page and line were added to the control information as a continuation of document number in order to designate, if required, a given page or a given line of a deposition as a separate computer record.

The digest information follows the control information. The digest information is also a variable field containing all other relevant and pertinent information in a given record. The first code is an alpha code for document type. Following this code, space is provided for numerous legal and factual contentions. Because this coding sheet was designed for a very complex lawsuit, and because all of the legal and factual contentions were not determined as of the drafting of the coding

5. See app. at p. 328 *infra*.

6. See app. at p. 327 *infra*.

sheet, the code is merely listed as LC/— for legal contention. As the first legal contention was formulated, it was recorded in a systems manual and given the designation LC/1. This procedure was followed for all legal and factual contentions.

Following the legal and factual contentions on Figure 3,⁷ the coding sheet is divided into "I" for documents such as letters, and "II" for depositions and interrogatories. Under "I," the coding provides for author and recipient as to both person and company, and for those individuals who received courtesy copies of the document. "II" provides encoding for deponent by company and individual, and for listing individuals named by the deponent in his testimony. "I" and "II" could be combined using the document type code, (DEP for deposition), to differentiate a deposition from other types of documents. Following "II," space is provided which can be enlarged as necessary for the narrative digest of the document. This will contain the attorney's or paralegal's subjective evaluation of the document and its relevance to the lawsuit and interrelationships to other documents.

In the hypothetical lawsuit concerning fraud in the sale and installation of a computer system, consider a letter sent by J. Gray, the computer salesman, to B. Jones, the president of the customer. Gray wrote this letter to Jones to confirm their conversation wherein Jones requested that the computer vendor commit to a certain delivery schedule. In response to this request, Gray answered that there would be no particular conversion problem from the card system to the disc system so long as the preparation of the three major programs to be used by Jones was started prior to August 8, 1970. Let us assume that the date of the letter confirming the telephone conversation was July 31, 1969. Further, during discovery, this particular letter was stamped with the document identification number "18."

On the abstract input form, (Figure 2⁸), this document would be encoded as follows. For entry number, 000018 would be encoded, signifying document number 18. The date 07/31/69 would be encoded as the pertinent date. A check mark would be put next to the three position alpha code, "LET," designating the document as a letter. Because the letter concerns the disc system primarily, a check would be put next to the liability subject "L/1." Further, let us assume that additional fraud subject categories have been designated during the process of the lawsuit, and that "F/3" will be used to signify fraud

7. See app. at p. 328 *infra*.

8. See app. at p. 327 *infra*.

concerning the development of programming support. There may or may not be relevant damage subject and general subject areas derived from this letter, but for the purpose of this example let us assume there are none.

The next items of information would concern the author and recipient of this letter. For author, we would encode under last name, "Gray," and for first initial, "J." For recipient, we would encode "Jones" for the last name and "B." as the first initial. Under digest information, we might encode the letter as follows: "Confirming telephone conversation, requested commitments of delivery schedule. No particular conversion problems on three major programs provided they are begun by 08/07/70." Finally, we may have additional coding as to whether or not this letter has been used for a plaintiff's exhibit or a defendant's exhibit, and whether it is attached to the complaint or to an interrogatory.

All documents in the lawsuit, whether they are letters, memoranda or depositions, would be similarly encoded for input into the system. Once all of the documents have been fed into the system, the attorney would have a "data base" consisting of all the relevant information in the lawsuit, which he can now proceed to sort and select against to obtain computer printouts useful in the conduct of discovery and trial. As an example, let us assume the attorney wants to retrieve all of the documents which involve Jones as either author or recipient. Further, the attorney wants these documents printed out by the computer in document number order, so that he can find the original copies. This printout may be useful in the preparation of an additional deposition of Jones, or in preparing for direct examination at trial.

Figure 1⁹ shows the first page of a computer printout with all of the documents in the lawsuit which involve Jones, as either author or recipient. The documents are arranged in document number sequence. Reading from left to right across the printout, the first six positions of the computer record and the printout format show the document number. The next eight positions are the pertinent dates derived from the document. The next three positions are the alpha code for the type of document, such as DEP for deposition. Again, these three items of information alone comprise the fixed fields of the computer record. The remaining information contained in the computer record is the variable narrative field comprised of both objective and subjective information concerning a particular document/record. Again, because it is a varia-

9. See app. at pp. 325-26 *infra*.

ble field, one normally could not sort on the information contained therein; but, one can select information from the variable field and, in turn, sort the records so selected by a fixed field. In this case, the documents were selected against "Jones," without the A/ or R/ so that Jones would be retrieved whether author or recipient, and the documents so selected, sorted by document number.

By looking down the document number column, sorted in document number sequence, one can see that certain document numbers do not appear. This is because not all documents in the data base involve Jones. Rather, this particular inquiry into the computer was a search of all records and a select against only those records that contain Jones in the narrative, regardless of whether Jones is qualified as an A for author or an R for recipient. After the select has been completed, the command to the computer was to arrange all the documents so selected in document number sequence. Therefore, documents numbered 000009, 000010 and 000011 contain no reference to Jones. Finally, document number 000018 is the document that was given as an example of the use of the abstract input form.

This is just one example of how information encoded into the system can be sorted, selected and retrieved in a desired order for use by the attorney. This same information, recorded once, can be selected again on the basis of criteria other than Jones, and then sorted out in either document number or chronological order. For example, the attorney may want to know every document that was authored by Gray, or every document where Smith is mentioned in the narrative, as in document number 000002. The computer would go into the narrative and retrieve, utilizing its select capability, all such documents. Further, the attorney may want to retrieve all documents pertinent to liability subject area 1 or damage subject area 2. The same approach would be used.

Effective combinations of sort and select as the basis for retrieval require the imagination of the attorney. The only limitation is the prior planning that determined the categories of information that were going to be encoded. The usefulness of the system is directly related to the coding scheme, and the coding scheme is a direct derivative of the attorney's understanding of his lawsuit and his ability to formulate the significant categories of information.

VII. ALTERNATIVE HARDWARE AND SOFTWARE TO SUPPORT THE GENERALIZED SYSTEMS DESIGN

The generalized systems design discussed above can be adjusted to fit

the hardware and software available. Computer vendors have a full product line and service bureaus have a full range of services. For both, cost and capabilities of hardware and software go hand in hand. How one approaches hardware will, to a great extent, depend upon two factors. The first factor is the potential recovery from the lawsuit. This will determine how sophisticated a use of data processing can be applied to the lawsuit because, as the system becomes more sophisticated, expense increases. Second, if a standardized systems approach is developed so that an established procedure can be set up to fit all lawsuits, allowing for a moderate amount of custom tailoring, high-priced data processing can be used for the small recovery lawsuit. Further, an understanding of available hardware and software will enable an attorney to moderate the sales efforts of the computer vendor and fit available hardware and software to his real needs.

The examples of the generalized systems design discussed above utilized moderately sophisticated data processing equipment which require that the data base be on-line to the CPU for select and sort capabilities. This approach is expensive because it requires a computer to do the information handling and retrieval. However, the concept of data processing, in that a record is recorded once, and equipment rather than clerical personnel is used for all successive rearrangements of information, can be applied without even using a computer.

The least expensive data processing system utilizes the eighty or ninety-six column punch card for storing information, but relies on only a card sorter and not a computer for retrieving information. The systems design, with fixed fields for document number and document date for sort capability, applies equally to this punch card scheme. However, it is practically impossible to select against a narrative field in a punch card only application. The narrative can be used for the purpose of describing a document but the manner whereby records are retrieved from the data file can only be based on document number, document date, or other fixed fields. This being the case, more fixed fields would be necessary in a punch card only system in order to effectively retrieve information, using numeric codes to conserve space. For example, the first five columns of the punch card might be allocated to document number, the next eight columns to document date, and three columns committed to a code indicating the witness referenced in a particular record. All the witnesses in a given lawsuit would be alphabetically arranged and then each assigned a three digit code which would be punched into a card to signify that witness. The next

four columns would be used for page number, and two columns would be assigned to allow for coding another witness numeric category. With an eighty column card, such fixed field coding would take twenty-two columns. Columns 23 through 80 would be committed to digest information or commentary concerning the document. One or two columns must be set aside to provide for a sequence code, a 1 punch indicating the first card in a series of cards describing a given document, 2 for the second card, and so on. This sequence code permits a number of eighty column cards to be linked together in a numeric sequence so that the commentary or digest information can be of variable length extending from one to n cards providing sufficient space for a meaningful narrative. This sequence of cards, whether one or n describing a single document, constitutes a record.

In a card only system, the documents would be encoded and key-punched into cards, the first card of any given sequence containing the header or fixed field information of date, witness number, document number and so on, and including the beginning of the commentary. The second through n cards, containing the balance of the commentary, would simply repeat the document number and contain a sequence code. Once all variable data or documents are coded and key-punched, there is a data base in card form that can be rearranged in a desired sequence by sorting fixed fields. The file of records, contained in a box of cards, would be passed through a card sorter and rearranged in either chronological order or numerical order by witness number or document number. The cards would be kept in the sequence order within record so that the commentary would flow when printed out. This system does not allow for the retrieval of information from the commentary. The commentary is basically a fixed group of information that is associated with a date or document number without independent access. Once the cards have been rearranged by the card sorter in date or document number order, the cards can be taken either to a computer or to another piece of equipment that has the capability of reading cards and listing the information.

Because this simple card oriented system does not allow for selects against the narrative, and because of the limited number of fixed fields available, it might be necessary in a particular lawsuit to create more than one data base. If additional fixed fields, such as legal contentions or factual contentions were needed, a separate set of cards could be encoded and punched replacing fixed fields in the first set with the additional required fields, and duplicating into the second set the bal-

ance of the first, such as document number and the commentary. This process would create a separate data base.

A card oriented approach would provide the attorney with a printout similar to that of Figure 1.¹⁰ If a simple printer is used, one that only reads and prints one card at a time, the width of the printout will be limited to eighty characters. However, the printout in Figure 1¹¹ was prepared on a computer capable of reading more than one card at the same time. Therefore, the printout covers a width of 132 characters. The basic difference between the card only system and the computer system is the ability of the computer to access the narrative, seize upon a particular item of information, and retrieve the entire record or a part thereof in document number or date sequence. All the card only system can do is sort the records in number order, date order or based on another fixed field. The narrative is simply available to the attorney but not accessible by the system.

Finally, a card only system is limited not only by the physical space of the punch card and the inability to select, but also because sorting with a card sorter is a relatively slow process. The entire card file must be passed through a card sorter once for each column of numeric data to be arranged and twice for each column of alpha. If numerous columns are to be sequenced or the data file is large (over 500 records), sorting by means of a card sorter can be very time consuming.

The next level of sophistication above a card only system is a card oriented computer system. Basically, the data file is prepared as with the card only system. However, a computer is used, rather than a card sorter, for rearranging the information in the desired sort sequence. Depending on file size relative to available core, entire card file can be read into the system and stored internally, rearranged by the system at machine speeds, and printed out in the desired sequence. The major drawback of most card oriented computer systems is the lack of a capability to select against a variable narrative field because of insufficient software support or internal storage. The trade-off is between the cost of internal sorting and the time involved in sorting on the card sorter.

The next level of sophistication is to maintain the data file in a larger computer system with the hardware and software necessary for searching against data files. Larger systems allow for on-line storage of data

10. See app. at pp. 325-26 *infra*.

11. See app. at pp. 325-26 *infra*.

files which permits the system to access the entire data file simultaneously and select against the variable length narrative fields. This provides the full capability of the generalized systems design discussed above. On-line storage during operation and the software support required are relatively expensive, but the usefulness of on-line data storage and retrieval, in a moderate to large size lawsuit, clearly warrants the costs involved.

One method of reducing the cost of the on-line storage computer system is to remove the data file off-line when not in use. Most service bureaus operate in this manner. The data file is kept on a disc pack or a tape until a request is received for information search and retrieval. The file is then mounted onto the system for access. The degree of on-line use will depend upon the frequency of inquiry against the system. The more a data file is on-line, the more costly the use of the system becomes because service bureaus charges are partially based on average monthly on-line storage. To keep costs at a minimum, inquiries to the system can be batched and submitted in group to the service bureau. The data base is then mounted on the system and all inquiries processed at the same time. This may be inconvenient in that response to an inquiry is not immediate. The trade-off again is convenience and timeliness of response versus cost. At some point it becomes more economical to keep the data base on-line at all times to have the convenience of immediate response to an inquiry.

Once the data base is on-line continuously, teleprocessing terminals can be utilized. If the computer system is at a service bureau and frequency of inquiry against the system requires immediate turnaround, terminals can be installed in the law firm linked by telephone lines to the computer for immediate response upon inquiry. This is the most sophisticated approach that most law firms will use, and correspondingly the most expensive. It provides a typewriter terminal always available to the attorney for access to the data base. If the attorney is frequently requesting large printouts, a high-speed line printer can be used in place of the typewriter terminal to reduce the cost of connect time to the system.

The next variable in software selection is how the data file is to be fed into the computer. As discussed above, in a card oriented system the data file is key-punched into cards and the cards, in turn, are fed into the computer. However, by using an on-line terminal to input variable data, one is not constrained to an eighty column or a ninety-six column format and the requirement of a sequence code. This permits use of the

computer, and its text editing and data formatting capabilities, at the input stage.

When a data base is being created, the coding sheets, can be sent to the service bureau for encoding into the system by means of their terminal. Because the cost of connect time between the terminal and the CPU is normally a constant, if large data files are required, then it might be more economical for a law firm to acquire its own terminal, for access retrieval, and input. An employee at the law firm, trained in the use of data processing, can transcribe the coding forms directly into the computer utilizing a terminal. Again, there are trade-offs as to cost and efficiency that can only be analyzed in light of a specific case, size of file, availability of trained personnel, and the frequency with which a law firm uses data processing to assist in litigation.

VIII. USING THE COMPUTER SYSTEM

How one uses the information stored in the computer, applying the generalized systems design considered in this article, is a function of the attorney's imagination. Once narrative information concerning each document/record is stored in the computer, it can be retrieved in any desired order within the limitations of the system. One can select against the narrative, rearrange the information in date or document number sequence and retrieve all or a designated portion of the narrative.

The first litigation stage where this information might be highly useful is during depositions. As documents are collected, they are encoded into the system to provide a chronological framework for the conduct of depositions. As each deposition is taken and documents are gathered on requests to produce, the data base is enlarged, providing an expanded base of information for the conduct of subsequent depositions. Further, if discovery exhibit numbers are used as an item of fixed field information, the printout can be sequenced to list all exhibits to date in exhibit number order for reference at deposition. In the preparation of the deposition, the attorney can go down the narrative and select documents relevant to the deposition and, if necessary, retrieve the original copy. If a deposition is to be conducted on the basis of chronological facts, the attorney can retrieve every document that involves the individual being deposed in date sequence. If the deposition is to be conducted on the basis of factual contentions, the same information for a deponent can be retrieved in factual contention order. The printout

itself can be used as a chronological reference because the narrative description of each relevant document will be in date sequence. However, if warranted, a separate printout in document sequence can be called out for gathering the original documents to be taken to the deposition.

In addition to assistance at depositions, the computer data base can aid the attorney in the preparation of pre-trial conference orders. Depending upon the categories of information stored, the data base will have, at least in outline form, the information necessary for the preparation of a pretrial conference order. As discovery progresses, the data base will be continuously enlarged and updated and will presumably include admitted and uncontested facts, contentions of fact and law to be tried, exhibits, names of witnesses, and a narrative analysis of all documents considered to be significant and relevant arranged by any ordered sequence derivative from the data base. With a minimal amount of effort, the attorney can expand upon the computer printout as is required by the rules of court and, with a vastly reduced amount of effort, especially in a complex trial, prepare the pre-trial conference order.

The computer data base can also be applied to the preparation of a trial book. With this in mind, as the trial date approaches, the attorney can begin to refine his data base by designating those documents which are key to supporting his case and those that will be submitted in evidence. In this process, the attorney is basically creating a refined data base that will become the trial data base. Sorts and selects can be made against this refined data base for the preparation of the trial book itself. For example, a command to the computer can create lists of exhibits that will be submitted into evidence. These can be arranged in document number or chronological order or by factual and legal contention to be proved. Similarly, a list can be called out of all significant documents authored by a given individual. These documents can be arranged in both chronological and document number sequence, and used to examine a witness by taking him chronologically through the documents that he prepared.

The system can also be of use during the conduct of the trial. If an event occurs during the progress of the trial which requires additional information for effective response, such as the impeachment of a witness, the attorney can immediately access his on-line data base and retrieve all relevant data. This information in turn can be rearranged by witness within legal or factual contention, and in document or

chronological order, or any other arrangement that is required. This information then can be immediately available over terminals at the courthouse or even in the courtroom.

IX. THE COMPUTER *Qua* EXPERT

The discussion concerning the use of data processing in litigation up to this point has involved the computer merely as the repository for masses of information, using the computer as a library. Although this is a highly important use of data processing, considering the capabilities of the technology available, it is remarkably mundane and does not take advantage of the computer as an analytical tool. Once the attorney gains an understanding of data processing, the outstanding capabilities of this technology can be available to him, with imagination one's being the only real constraint.

One way in which the computer can be of assistance, beyond the mere storage of document records and the printout of these records in a desired sequence, is to prepare summaries of masses of information. When we think of computer summaries, we often think of accounting summaries, payroll or accounts receivables, and additions of columns of numerical information portrayed in statistical or other tabular form. However, the computer can collate and analyze any information that can be numerically delineated and summarize it in tabular or graphical form.

Consider a lawsuit involving the building of an ocean liner. You represent one of the subcontractors suing the general contractor for breach of contract during the construction stage. Your argument is that the general contractor was required to keep all of the construction trades on time according to the construction schedule. Your client, the contractor responsible for painting the interior rooms of the ship, was delayed in construction and ran up sizable overtime costs in order to complete on schedule. You now sue the general contractor for this overage in excess of your bid on the ground that the general contractor breached his responsibility to pace the trades and keep them all on schedule. The defense is that your client mismanaged its own personnel, was unable to coordinate its work crews, and, therefore, the overage and overtime was its own fault. One key to proving your point might be an evaluation of where, chronologically during the construction of the ship, all of the construction trades whose work had to be completed before your client could paint, were working on the project. In other words, if steel decking or electrical or bulk head were late in completing

their assigned tasks in a given area of the ship, it was impossible for your client to enter that area of the ship and complete its task. If you could show that the trades that paced your client did not complete on schedule and therefore the only way your client could complete its task on time was to use additional personnel or overtime, you may prevail in the lawsuit.

However, the problem of analyzing the construction of the ship, conceptually simple, becomes inordinately complex because of the mass of information available to the attorney. He may have five to ten thousand progress reports generated by all of the trades showing daily where they were in the ship during construction. There may be an additional ten to twenty thousand documents such as correspondence, inspection reports and minutes of construction meetings that date events during construction. The question then becomes how to collate and analyze all of this information to present to the court a clear visual expression of the delays which occurred during the construction of the ocean liner.

Here is where the computer, acting as an analytical tool, can assist the attorney. The dating of specific events can be encoded into the computer, broken down by area of the ship, by trade involved and by date. Once all of this information is assembled into a data base, the computer can be asked to portray, in graph form, where, within a given time frame, all, or one or a collection of trades were working in given areas of the ship. Graphs can be called out by week, by day, by all trades, by client against a particular other trade, within one or all or some of the areas of the ship, and compared with the construction schedule and payroll data.

Because it is possible to attach to the computer a cathode ray tube (CRT), a terminal device similar to a television screen, the attorney can sit down in front of the CRT and key in a request to the computer to display what construction events were occurring in a given area of the ship within a given time frame. This time frame will be displayed visually on the CRT. Another request to the computer can expand the time frame either backward or forward in time, and other pacing trades or factual events can be added or deleted from the graphical presentation. This information can be manipulated as it is displayed, and the attorney can begin to analyze the facts that are available to him, using the computer, in order to prove his lawsuit.

Further, payroll information can be fed into the system and displayed in conjunction with the graph of a given time frame of construction to

show whether, at a particular point in time, your client was incurring overtime. Also, the construction schedule itself can be put into the system and overlaid against what was going on during construction to show where other trades were off the construction schedule and where their delay in turn caused your client to be delayed in its task.

After the attorney has determined what time frames and what events are significant to proving his position, he can then request that the computer provide him with graphical and statistical summaries of the information stored within it. This information can be summarized by bar charts, continuous line graphs, overlays of graphs, or other forms. If the proper foundation were laid, these summaries, a composite of what occurred during the construction of the ship, could be submitted into evidence, to provide the court and jury with a manageable representation of what occurred in a complex factual situation.

This construction case is only one example of how the computer can be used to assist the attorney in analyzing the facts available and collating these facts in some organized fashion. What makes this example perhaps revolutionary is that the computer and the attorney interact in analyzing the data and evaluating its significance. The computer and attorney are basically in a conversational mode whereby the attorney bounces ideas off the computer and the computer, with its factual data base, tries to support those ideas. If the computer cannot, the attorney tries a new factual hypothesis. It is this give and take that moves this example beyond the mere storage and retrieval of information to the use of the computer to its real potential. The computer is no longer merely used in an accounting sense, storing and retrieving, but is now simulating events and evaluating hypotheses. The computer becomes a powerful analytical tool, not merely an organizational tool, providing the ability to grasp and analyze thousands of documents as though they were ten or fifteen. Without this capability, the attorney would have to guess as to what were significant documents, events and dates because of the impossibility of grappling with such a mass of information. Without the computer, there is a great probability that highly significant information would be unavailable merely because of the impossibility of understanding it.

X. CONCLUSION

This article has attempted to provide the reader with a threshold knowledge of data processing and a generalized conceptual framework for using the computer in litigation. It cannot be viewed as a

how-to-do-it manual, but rather an attempt to provide that minimal amount of information which the attorney needs in order to ask intelligent questions of the experts in the computer industry. The attorney should not try to use data processing in litigation, or in any other application, without consulting systems engineers or other data processing specialists whose job it is to understand the capabilities of the computer. The world of data processing is as complex and sophisticated as any discipline. The attorney should approach the use of data processing on that basis. With a moderate amount of information, he can formulate what he hopes to accomplish and to address intelligent questions to computer experts. Further, he will be able to understand the answers of the computer experts as they try to design a system to assist in the conduct of his art.

With an understanding of data processing, the attorney will be able to provide better service to his clients at reduced cost as well as handle complex cases involving mass information. Using the technology available, he can effectively and efficiently marshal the mass of data at his disposal and correlate this data to his legal knowledge of the case. This combination should prove to be highly effective.

APPENDIX

4									
5									
6									
8									
9									
10									
12									
13									
14									
17									
18									
20									
21									
23									
24									
25									
26									
27									
30									
32									
33									
34									
35									
36									
38									
39									
40									
41									
43									
44									
46									
47									
000001	05/10/70	DEP	D/1, A/Jones, B. Moved from 15 East 50th, Los Angeles to 17 Kenneth Blvd., Glendale. Increased square footage. P/6-7.						
000002	01/00/68	DEP	L/3, D/2, A/Jones, B. M. Smith was office manager when EDP was introduced, had no prior EDP experience. P/9-10, P/11/18-24.						
000004	07/19/69	DEP	L/1, F/1, A/Jones, B. No recollection of anything said by defendant about comparison of defendant's computer to others. P/19/7-10.						
000006	01/00/69	DEP	D/2, D/11, A/Jones, B. Business was increasing while EDP was installed. Has progress charts in office to show 10-15-20% increase a year. Increased dock and office staff. Extra office manager because M. Smith devoted time to EDP. Released office manager since terminating EDP. P/24 to P/25/17.						
000007	01/00/70	DEP	D/11, A/Jones, B. Business decreased since 1970 because they had to impose higher costs. There was also a delay caused by the computer delaying the production of paperwork. Also decreased after computer removed in 1971, but doesn't know why. P/25/18 to P/26/22.						
000008	05/13/71	DEP	S/1, D/6, A/Jones, B. Paper cost went out of sight. Disc problems. 20-25 percent down time. No recollection of whether manual forms could be used. No change for disc. Increase caused by disc having to begin all over every time it went down. No recall of defendant recommending this procedure. Failure almost every night. Records kept that would enable mathematical computation of forms used. Keeps invoices for supplies for seven years. P/27/31.						
000012	07/19/69	DEP	F/1, A/Jones, B. Defendant promised that EDP would cut cost, be faster, use less people, a lot of peripheral information at no cost. Reason they went for data processing was the promise of magic. P/34/4-18.						
000013	01/00/68	DEP	L/3, A/Jones, B. M. Jones has no knowledge of EDP. Neither did anyone else in the office. P/34/3.						
000014	02/12/71	DEP	D/2, D/6, A/Jones, B. Was not involved in preparation. (Defex/N/MRD). P/36/22, P/44/10.						

Figure 1

000017	07/19/69	LET	L/1, F/2, A/Gray, J., R/Jones, B. Demonstration of computer. Can deliver card system and add disc subsystem. Help you convert from present system. Pex/A/COM, N1. Defex/A/INI.	50 51 52
000018	07/31/69	LET	L/1, F/3, A/Gray, J., R/Jones, B. Confirming telephone conversation, requested commitments of daily schedule. No particular conversion problem on three major programs provided they are begun by 08/07/70. Pex/B/COM, Defex/B/INI.	54 55 56 57
000019	08/02/69	LET	L/1, F/2, S/4, A/Gray, J., R/Jones, B. Delivery schedule. First step is direct conversion. Defendant will convert two major programs, manifest and freight bill, and will provide guidance to programmers in converting other programs. Step two is a programming and operational system change from card to disc. Convert two programs, guide others. Step three is equipment substitution. Any program changes made with guidance. Description of proposed system and specification sheets attached.	58 59 60 61 62 63

Figure 1

ABSTRACT INPUT FORM

CONTROL INFORMATION:

Entry No.: 0 0 Pert. Date: / /

Doc. Type: LET DEP INT MEM FIN TEC

Liability Subjects:

Fraud Subjects:

Mod Frt

Disk Sys L/1 L/4 L/7 Fraud F/1
Card Sys L/2 L/5 L/8 F/2
Bill Prog L/3 L/6 L/9 F/3

Damage Subjects:

Space D/1 O/T-Dock D/5 Perif rntl D/9 D/13
Reg Time D/2 Cptr forms D/6 Constructn D/10 D/14
O/T-Adm D/3 Othr Suppl D/7 Cust Loss D/11 D/15
O/T-Cptr D/4 Hrdw rntl D/8 D/12 D/16

General Subjects:

Downtime S/1 S/3 S/5 S/7
S/2 S/4 S/6 S/8

Author: A/ Last F.I. Recipient: R/ Last F.I.

Digest Information:

Horizontal lines for digest information input.

Pex/ Defex Pla/ Def/ P/ Page/line

Abstracted By: Date:
Entered By: Date:
Proofed By: Date:

Figure 2

Control Information

Document No. _____ P/_____ L/_____

Trial Exhibit No: _____ Discovery Exhibit No: _____

Document Date: ____/____/____

Digest Information:

Document Type Cd: Let _____ Int _____ Prr _____
Dep _____ Mem _____ Drw _____

Legal Contention: LC/____, LC/____, LC/____, LC/____, LC/____

Factual Contention: FC/____, FC/____, FC/____, FC/____, FC/____

I.

[Author] AP/_____, _____, [Recipient] RP/_____, _____;
[Person] Last F.I. [Person] Last F.I.

[Author] AC/_____, [Recipient] RC/_____
[Company] [Company]

Courtesy copy: cc/_____, _____, cc/_____, _____
Last Last

II.

[Witness] WP/_____, _____, [Witness] WC/_____
[Person] Last F.I. [Company]

Parties Involved
in Testimony TW/_____, _____; TW/_____, _____;
Last Last
TW/_____, _____; TW/_____, _____;
Last Last

Narrative:

Figure 3