Chicago-Kent Law Review

Volume 52
Issue 3 *Law and Technology Symposium*

Article 3

January 1976

Evaluating the Credibility of Computer-Generated Evidence

James A. Sprowl

James A. Sprowl

EVALUATING THE CREDIBILITY OF COMPUTER-GENERATED EVIDENCE

JAMES A. SPROWL*

A neatly printed, computer-generated report can be a highly persuasive piece of evidence. Such a report has an air of authority comparable to that of a sealed document. Judges and jurors, knowing very little about computers, may be awed by such a report. They may feel ill-equipped to question its credibility, particularly when its credibility is supported by the testimony of one who purports to be an expert on computers. If a proper foundation has been laid, such a report will be admitted into evidence.[1] But an admissible report is not necessarily credible. Computer-generated reports can contain errors which can cause such reports to present a distorted picture of reality. Computers can also be programmed to slant such reports to favor one party or the other or to tell lies. Computer-stored evidence can be readily altered, and such alterations are very difficult to detect in a computer printout.[2] In appropriate cases, the trier of fact must be prepared to inquire into the credibility of such a report, and an attorney must be prepared to initiate or participate in such an inquiry.

In this article, I shall explain how one evaluates the credibility of such computer-generated evidence. In doing so, I shall draw heavily upon the experience and expertise of a small group of certified public accountants who specialize in the examination of computerized accounting systems.[3] The articles which follow, two of which are authored by these

* Research Attorney for the American Bar Foundation and Lecturer at IIT/Chicago-Kent and Northwestern University Law Schools; J.D., University of Michigan.

I wish to thank J. David DeHetre, Jerome J. Roberts, and L. Thorne McCarty for their helpful comments and suggestions.

[1] Several recent articles discuss how one goes about establishing the admissibility of such a report; this article will be confined to the narrower topic of credibility. Of course, a report that lacks credibility may be inadmissible so the two topics cannot be completely separated. A concise discussion of computer generated reports is presented in Younger, *Computer Printouts in Evidence: Ten Objections and How To Overcome Them*, 2 LITIGATION 28 (Fall 1975). *See also* Freed, Fenwick, McGonigal, *et al.*, *Mock Trial: Admissibility of Computerized Business Records*, 15 JURIMETRICS J. 206 (Spring 1975) and Roberts, *Practitioner's Primer On Computer-Generated Evidence*, 41 U. CHI. L. REV. 254 (1974). An excellent example of how one goes about laying a proper foundation for computer-generated evidence is presented in United States v. Russo, 480 F.2d 1228 (6th Cir. 1973).

[2] In this respect, computers are much like automobile odometers and voting machines as is discussed in the text *infra*. If the computer has not been kept secure from all possibility of tampering it cannot be definitely determined that the information stored in the computer has not been altered.

[3] I am particularly grateful for the assistance I have received from J. David DeHetre of Touche, Ross & Co.

certified public accountants, consider two aspects of this credibility evaluation problem in greater detail. The article by J. David DeHetre describes the types of safeguards that must be employed to insure the integrity of computer-stored information, and the articles by Carol Eastin and Martha Jenkins focus upon the use and reliability of computer models as accurate simulations of real life systems. We hope that this series of articles will supply most attorneys with sufficient insight into the nature of computer-generated evidence to argue intelligently for or against the credibility of a typical computer printout. But when the credibility of a printout is crucial to a cause of action, an attorney should consider retaining a certified public accountant with computer auditing experience to assist in discovery and trial preparation, just as he or she would retain other types of technical experts in appropriate cases.

I shall begin by discussing the flow of evidence to the trier of fact, emphasizing those features of computer-generated evidence that are unique. I shall then describe digital computers briefly, concentrating upon their role as conduits for certain types of evidence rather than their technical details. Next, I shall explain how one can judge the credibility of the evidence that flows from a computer by examining what happens to the evidence as it flows from its source to the trier of fact. Finally, I shall discuss a number of special problems that can arise when a computer is used to manipulate or simplify evidence prior to trial, as when it is used to prepare summaries.

THE FLOW OF EVIDENTIARY INFORMATION

Evidence is a form of information. Information, like energy, is an entity that may take on different forms at different times. Scientists and engineers have only recently come to recognize this fact, and they now frequently talk of "information systems" in which the flow of information is studied without regard to its form or mode of conveyance.[4] Many otherwise complicated problems become much simpler when considered from this point of view. To avoid having to explain the complex technical details of computers, I shall take this approach in the discussion which follows. I shall describe how evidentiary information flows through a digital computer to the trier of fact without describing the computer itself to any significant extent. It is the integrity of the flow, and not the technical details of the computer through which it passes, that determines the credibility of computer-generated evidence.

Information, broadly defined, is a pattern—typically one that conveys meaning to a human.[5] The printed characters of a book, the light and dark

[4] Claude E. Shannon originated the field of information science when he wrote his treatise on the mathematical theory of communication in the late 1940's. See C. SHANNON, THE MATHEMATICAL THEORY OF COMMUNICATION (1949).

[5] For those who are more mathematically inclined, the information content of a message that consists of a pattern selected from a set of N possible patterns is the

regions of a photograph, the fluctuations of a radio or telephone signal, and the arrangement of magnetic impulses on a roll of audio or computer tape are examples of such meaningful patterns. A pattern that does not convey meaning is also information, but I shall call such a pattern "distortion".[6] The hiss in the background of a telephone conversation, the snow in the background of a weak television picture, the grain in the background of a photographic enlargement, and the random marks and image deterioration in a second or third generation xerographic copy are all examples of such distortion. Distortion obscures and dilutes meaningful information and renders it more difficult to comprehend.

Evidentiary information is information that flows through time and space to the trier of fact, and the rules of evidence constitute a judicial control mechanism regulating this flow. With the possible exception of oral testimony, evidentiary information is always tangible. It may take the form of a pattern that appears on a piece of paper, in a photograph, or on a strip of magnetic tape, for example.

Frequently, a copy of evidentiary information is presented to the trier of fact in lieu of the information itself. Such a copy is never an exact duplicate of the original. For example, a photographic print is always more grainy than the negative from which it is made and frequently does not reproduce all the highlight or shadow details apparent in the negative. Similarly, a xerographic copy always contains far more contrast than does the original from which it is made, and a duplicate sound recording contains far more background hiss than does the original recording. When a sound recording is played back through a loudspeaker, the sound that flows from the loudspeaker is a "copy" of the recorded sound—typically a distorted copy, since most loudspeakers convert about five to ten percent of the sound conveyed into distortion. As a general rule, evidentiary information cannot be copied or conveyed without some loss of meaningful information and some increase of distortion. There is an important exception to this rule, as I shall explain shortly.

The "rule for documentary originals" (sometimes called the "best evidence rule") is a judicial formulation of this general scientific rule.[7]

logarithm of N. If the logarithm "base two" is taken, the information content of the message is measured in units called "bits" (an abbreviation for "binary digits"). For example, a message consisting of a pattern selected from a set of 64 possible patterns contains six bits of information and can be stored in a memory or storage device capable of assuming any one of 64 possible states. Such a memory device is said to have a six bit storage capacity. *Id.* at 3-4.

[6] An information scientist or an electronic engineer will speak of "noise" rather than distortion. I use the term distortion because attorneys who read early drafts of this article were troubled and confused by my use of the more technically proper term "noise."

[7] *See generally* WIGMORE, EVIDENCE §§ 1117-1282 (Chadbourn rev. 1972).

Under the new Federal Rules of Evidence, for example, the content of a writing, recording, or photograph must be proved by presentation of the original whenever a genuine question is raised as to its authenticity.[8] Of necessity, exceptions are made if the original is unintelligible to a human. Thus, because the trier of fact cannot examine the photographic negative directly, the new rules define both the photographic print and the negative from which the print is made to be originals.[9] And while there is no special rule on the subject, all courts permit loudspeakers to be used for the playback of sound recordings even though they distort the sound played back, since the trier of fact cannot comprehend the information contained in a sound recording without playing it back. Computerized evidentiary information, whether stored inside a digital computer's memory or outside the computer on machine-readable magnetic tapes, magnetic disks, or punched cards, must also be "played back" to the trier of fact—that is, printed out on paper or displayed on a suitable electronic display device—before it can be examined. The new Federal Rules of Evidence define such a printout or display to be an "original" if it accurately reflects the stored evidentiary information.[10]

As an exception to the above general rule, information that takes the form of strings of letters or numbers or both may be copied with any desired degree of accuracy, if one is willing to expend the necessary effort. For example, a blurry, fourth-generation xerox copy of a typewritten sheet may be retyped by a skilled typist and thereby restored to its original clarity with no loss of information whatsoever so long as the typist can distinguish the individual letters and numbers from each other and so long as the typist's work is carefully proofed.[11] Digital computers contain only information of this type, and they possess the ability to recopy that information hundreds and even thousands of times in the course of processing the information without altering it in any way. Because computer-stored information may be accurately copied, such information differs fundamentally from the nonnumeric, nonalphabetic information found in a photograph or sound recording

[8] FED. R. EVID. 1002, 1003.

[9] FED. R. EVID. 1001(3).

[10] *Id.*

[11] Typewritten or handwritten strings of letters and numbers may be copied accurately because the shape of the individual letters and numbers can be severely distorted without making the letters and numbers indistinguishable from each other. The shape of the individual letters and numbers may be thought of as containing a large amount of duplicate information much of which can be dispensed with without a loss of meaning. Within a computer, a small amount of duplicate information is added to strings of letters and numbers to enable the computer to determine when such a string has been altered and to take corrective action. For example, "parity bits" or "check bits" are small quantities of information that are routinely added to all computer-stored information so that copying errors may be detected and corrected. (A full explanation of "parity bits" goes beyond the scope of this article but may be found in any introductory text on digital computer design.)

that cannot be accurately copied. In appropriate cases, the trier of fact will want to examine a photographic negative to see if it contains highlight or shadow details not visible in a print made from the negative. On occasion, the trier of fact will want to examine a magnetic sound recording for visible evidence of splices or erasures in addition to having it played back. But nothing is ever gained by physically examining the data storage device of a computer, since all that is stored within that device is magnetic representations of numeric or textual information, and these may be printed out on paper with no loss of accuracy. Similarly, the information content of magnetic computer tapes, disks, and drums as well as that of keypunched IBM cards may also be printed out on paper with no loss of accuracy. Therefore, nothing is to be gained by examining such tapes, disks, drums, or cards directly. The provision of the new Federal Rules of Evidence defining a printout of computer-stored evidence to be an original thus has a scientific basis. The provision of those same rules defining a photographic print to be an original lacks any such basis and is simply a necessary expedient.

One must not assume that computers are the only entities able to generate accurate copies of information. The accuracy of a computer printout is attributable to the purely numeric or alphabetic form of the information printed out. Handwritten and typed copies of such information can also be highly accurate, as can oral testimony about such information when presented by a credible witness.

Digital computers are widely used for accounting purposes at the present time, and the most common form of computerized evidentiary information at present is accounting information. Because all accounting information is numeric in form, such information may be fed into a digital computer without any alteration whatsoever.[12] Computer printouts of accounting information are accurate, generally speaking, although some small and usually negligible errors can result from arithmetic computations that are rounded off. More and more frequently, digital computers are being used for text editing and storage, and thus computerized textual evidentiary information will become commonplace in the near future. Like accounting information, textual information may be fed into a digital computer without alteration, and computer printouts of textual information are extremely accurate.

Much evidentiary information is neither numeric nor textual in its original form, and such information must be converted into numeric form before it can be fed into a digital computer. Industrial process control computers, for example, measure such nonnumeric things as distances, time

[12] Of course, human errors can occur if the accounting information is manually typed or keypunched into the computer. I shall discuss human errors at a later point.

durations, pressures, weights, voltages, and currents, all of which may be relevant to tort litigation. For example, such computers are now used routinely to monitor the operation of electrical power generating equipment. If litigation were to result from a massive power failure such as the one that occurred in New England several years ago, almost all of the evidentiary information relating to the cause of the power failure would take the form of computerized records of what events happened at what times and in what locations. When used in this fashion to gather nonnumeric data, computers are performing two tasks that traditionally are performed by humans: making measurements and expressing the results of those measurements in numeric form.

Regardless of whether such measurements are carried out by a human or by a computer, the trier of fact should investigate two possible sources of error. The first is measurement error. Was the measurement equipment properly calibrated, and was it used correctly? Do repeated measurements give consistent results?[13] Do measurements carried out with different equipment give consistent results?[14] And if not, what was the possible range of error? The second is the error which results when any measurable magnitude that can take on virtually an infinite number of different values, such as temperature, is expressed as a number having only so many digits. For example, if only two digits are permitted, temperature must be expressed as either 66 degrees or 67 degrees even though the actual temperature may fall somewhere in between these two values, and the maximum possible rounding error that results from the use of only two digits is plus or minus ½ of a degree.[15] If three digits are permitted, then temperature can be expressed as 66.3 degrees, and the maximum possible rounding error is reduced to plus or minus .05 degree. But if the measurement error is ½ of a degree, the use of three digits can give an appearance of accuracy that does not, in fact, exist.[16] The trier of fact must consider both of these sources of

[13] The degree to which repeated measurements agree with each other is the "precision" of the measurements. A series of measurements that agree closely are said to be "precise". But, such measurements may lack "accuracy". *See* note 14 *infra.*

[14] The degree to which a measurement conforms to some recognized standard value is the "accuracy" of the measurement. Accuracy is to be distinguished from "precision", which is a measure of the repeatability of a measurement but not a measure of its accuracy on an absolute scale. *See* note 16 *infra* for an example that illustrates this distinction.

[15] The error which results when a continuous value such as temperature is expressed as a number having two, three, or any finite number of digits is called "quantization" or "rounding" error. Simply stated, two digits can define only 100 different states, and 100 states cannot accurately be used to represent the status of something that may take on an infinity of different states.

[16] It is not uncommon for a measurement to be "precise" (or repeatable) to within 0.05 units but "accurate" only to within 0.5 units due to improper calibration of the measuring equipment or other such factors. *See* notes 13 and 14 *supra* for definitions of "precision" and "accuracy."

error and their effect upon the credibility of the evidence submitted in every case where the validity of such evidence is questioned.

Whenever a human serves as part of the conduit over which evidentiary information flows, human errors will occur. A witness can mistakenly say fifty inches instead of sixty inches, for example, or a typist can strike the "5" key instead of the "6" key and thereby convert the value "60" into the value "50". Humans are frequently used to enter information into computers, and human errors are to be found in all computerized information files containing information entered by humans. The incidence of human errors in such files can be minimized, but not eliminated entirely, by careful proofreading, double entry of the same information by different humans, and like procedures. But the cost of eliminating additional human errors goes up rapidly as the number of undetected errors is reduced, so a small but measurable number of human errors is simply tolerated as unavoidable in most computerized information files. An error rate of .05 percent, for example, is considered acceptable in textual information files such as those used for automated legal research.[17]

Computer errors can also occur, but almost all computers contain elaborate error detection and correction circuits that reduce the computer error rate to a very low level.[18] Large computers, for example, rarely ever alter information without halting immediately or taking other appropriate remedial action. When evidentiary information is conveyed in part by humans and in part by computers, the errors attributable to the computers will almost always be negligible compared to those attributable to the humans.

To summarize briefly, evidentiary information may be thought of as flowing through time and space to the trier of fact. Whenever such information is copied, some meaningful information is lost and some distortion is gained. But information representable by strings of letters and numbers can be copied accurately if the copying is carefully done, and such information may also be fed into and through a digital computer without alteration. Information not so representable can be converted into numeric form, and then it, too, may be accurately copied and fed through a digital computer. But the conversion process always introduces some measurement and roundoff errors. And whenever information is copied by a human, copying errors invariably result. In most information flow systems, machine copying errors may be neglected, particularly when human errors are also present.

[17] This figure was supplied by a representative of Mead Data Central, Inc., promoters of the Lexis system of computer-assisted legal research.

[18] *See* note 11 *supra.* Transmission of information over a telephone or radio communications link, in the absence of special safeguards, can sometimes cause information to be distorted.

DIGITAL COMPUTERS

As the name implies, a "digital" computer is a machine that contains "digits" or numbers stored in an internal memory.[19] A digital computer accepts numbers from the outside world, moves them about, performs simple arithmetic operations upon them, and ultimately either prints them out on paper, displays them upon a viewing screen, or supplies them to an external device that is numerically controlled. When such a computer cannot contain within its internal memory all of the numbers it is called upon to process, it typically stores the numbers externally on reels of magnetic computer tape, spinning magnetic disks or drums, punched IBM cards, or some other "machine-readable" storage medium.

In a "programmable" digital computer, all operations are controlled by long sequences of numeric instructions called "programs" or "instruction sets". These are stored in the computer's internal and external memories along with the numbers being processed or stored. Pocket calculators and other so-called "special purpose computers" lack such stored instruction sets and thus lack the flexibility of a programmable computer. Since the instruction sets are man-made, they can and frequently do contain erroneous instructions.

Internally, a digital computer contains and processes only numbers. Devices external to the computer interpret certain numbers as code representations for letters, punctuation symbols, and the like. For example, when the "A" and "B" keys are struck on a keyboard attached to an IBM System 370 computer, the keyboard generates telegraphic codes for the numbers "42" and "43" and thereby causes these numbers to enter the computer's memory.[20] The computer may later retrieve these numbers from its memory and feed them to a printer, which responds by actuating the type bars carrying the letters "A" and "B." Hence, the number "42" represents the letter A within the system 370 computer, and the number "43" represents

[19] "Analog" computers do not contain digits or numbers—they are inherently nonnumeric. Analog computers represent pressures, distances, weights, and other such real-world variables by electrical "analogues"—voltages and currents. An analog computer is thus a simulation or model. It can be used to great advantage to simulate nonnumeric phenomena such as the vibrations of an airplane wing or automobile suspension, but is never used to store numeric or textual data and does not generate evidentiary information of the type discussed in this article. An analog computer could be used to simulate the stability of an airplane wing in tort litigation, so the comments on simulation and modeling presented at the end of this article are relevant to such computers.

[20] Programming experts will recognize that the numbers "42" and "43" are not ordinary decimal numbers but "hexadecimal" numbers. The decimal number system includes only the ten digits zero through nine. The hexadecimal number system is similar but includes sixteen digits instead of ten—typically, the ten digits zero through nine plus the six letters A through F. Hence, "FF" is a valid hexadecimal number that corresponds to "255" in the usual decimal number system. Hexadecimal numbers are used by computer programmers because they simplify the task of man-computer communication.

the letter B. It is possible for the computer to print out the numbers "42" and "43" instead of the letters A and B, and such a printout would be unintelligible except to a computer expert. As a dilatory tactic, a party might produce such an unintelligible printout in response to a discovery request. No judge should countenance such a response. When printed out, computer-stored information should always be intelligible, with all numbers in decimal form (not in octal, hexadecimal, binary, or some other strange form) and with all special computer codes for letters and other symbols eliminated and replaced by the letters and symbols themselves.[21] Such an intelligible printout can almost always be produced with little more effort than it takes to produce an unintelligible printout.

Sometimes a party will request a "machine-readable" copy of computer-stored information so he can process the information further on some other computer. Unlike a printed copy, a "machine-readable" copy is an exact copy of the numeric information stored within the computer—it is not normally intelligible. The requesting party should supply the necessary magnetic tapes, magnetic disks, or IBM cards to the party controlling the information, and the controlling party should have his computer copy the requested information onto the supplied storage medium. It costs very little to copy information in this manner, since the information does not have to be rearranged. The requesting party may then reorganize the information, at his own expense and using his own computer, if he desires to do so.

Computer programs or instruction sets are written out by computer programmers in special programming languages such as Fortran and Cobol. The original or "source" version of a program typically includes many helpful explanatory comments placed there by the programmer to make the program more understandable. This source program is typed directly into the computer's memory by the programmer or else it is punched into IBM cards which are then fed into the computer. In response to a discovery request, it is this original or "source" program that should be produced, regardless of whether the request is for a printed or machine-readable copy of the program. Once established within a computer, the source program may be readily printed out on paper or copied onto some machine-readable storage medium such as IBM cards, magnetic tape, or the like that can be delivered to the requesting party.

The source program is not, however, used to control computer operations. The computer translates the source program into a purely numeric form called the "object" program, and the object program is then used to

[21] *See* Greyhound Computer Corp. v. IBM, 3 C.L.S.R. 138 (D.C. Minn. Nov. 15, 1971), where one party's discovery efforts were hampered by an unintelligible printout. The "octal" number system includes only the eight digits zero through seven, unlike the commonly used decimal system that includes the ten digits zero through nine.

control the computer.[22] The final object program does not include any explanatory comments, and its purely numeric form makes it very difficult to follow. The object version of a program should never be produced in response to a discovery request unless it is specifically requested, and the production of an object version of a program unaccompanied by a source version should be considered nonresponsive.

Information other than programs is normally typed into a computer either directly or through the intermediary of punched cards or magnetic tape. More and more frequently, ways are being found to enter information into a computer without typing and therefore without the possibility of human errors. For example, almost all bank checks and many credit cards now bear magnetically encoded characters which can be scanned by magnetic sensors. Optically scannable bar codes are also coming into widespread use on such diverse things as freight cars and grocery products.

A computer used by industrial concerns to monitor temperatures, pressures, distances, time durations, weights, voltages, and the like and to control industrial machinery is called a "process control" computer. A process control computer is equipped with special measuring devices that generate numbers proportional to the temperatures, pressures, etc. being measured. The numbers so generated are periodically fed into the computer's memory. To control industrial machinery, such a computer generates numbers indicative of the desired control action. These numbers are fed into special numerically controlled devices that open and close valves, start and stop motors, and perform other necessary control operations. In most respects, a process control computer is similar to other forms of programmable digital computers, but evidentiary information taken from such a computer is subject to measurement and rounding off errors, as explained above, and these potential sources of error must be considered by the trier of fact in appropriate cases.

Computers rarely ever malfunction in such a way as to alter the information they are processing without warning the operator or taking some form of corrective action. Because computers are mechanically reliable, the trier of fact may generally assume that computerized evidentiary information has not been altered by a malfunction of the computer circuitry, and in most instances the mechanical integrity of the computer will not need to be

The "binary" number system includes only the two digits zero and one, and it is used internally by all digital computers. The hexadecimal number system is described in note 20 *supra.*

[22] Computer scientists do not use the term "translate" in this context. Instead, they talk of "compiling" a Fortran or Cobol program into "object code." Many computer scientists write their programs in machine-oriented languages called "assembly" languages, and assembly language programs are "assembled" into object code. The actual translation is carried out by the computer under the control of a special program called a "compiler" or "assembler," depending upon its function.

investigated either in discovery or in open court. But even when a computer is functioning perfectly, computer-stored information can be altered by the programs or instruction sets that control the computer. Computer-stored information can also be altered by a human who has direct access to the computer within which, or the tapes, disks, or cards upon which the information is stored. Whether or not done intentionally, such alterations typically leave no traces behind, and they are difficult to detect. Computers, voting machines, and odometers all share a common vulnerability to tampering because the tampering can be carried out quickly without leaving any telltale indications behind. In contrast, it is very difficult to tamper with a bank check without leaving smudges or other indications. Bank checks may therefore be freely circulated with minimal fear that they will be altered, while computers and computer-stored information must be kept secure from tampering at all times. Herein lies the principal vulnerability of all computer-stored evidence. If a computer is not secure from all possibilities of tampering, then it cannot be said for sure that the information stored within the computer has not been altered. Similarly, it cannot be said for sure that information stored on magnetic computer tapes or disks or punched into IBM cards has not been altered unless the tapes, disks, or cards have been kept in a secure place at all times.

Determining the Credibility of Evidence Taken from a Digital Computer

When evaluating the credibility of computer-generated evidence, one should examine the flow of the evidentiary information into, through, and out of the computer. At the point of entry into the computer, if the information was nonnumeric, it had to be converted into numeric form, and the conversion process may have introduced significant measurement and rounding-off errors. If the information was entered into the computer manually, as through keypunching or typing, the manual processing steps may have introduced human errors. Thus, one must carefully study the way in which the information entered the computer to determine how serious these errors may be.

Once within the computer, information is moved about and ultimately printed out under the control of computer programs. To insure the credibility of the information, it must be demonstrated that the programs are functioning properly and are free from errors. It must also be shown that no unauthorized programs were permitted to control the computer, since such programs could have altered the information. It must be demonstrated that unauthorized persons, particularly adverse parties, could not gain access to the computer or to the information itself while stored away from the computer in machine-readable form. Finally, it must be established that the information actually presented to the trier of fact is an accurate copy of the computer-stored information, and the information should not be accepted simply because it is presented in the form of an authoritative-looking computer printout.

The most likely point along the information flow path for errors to occur is at the point where information first enters the computer. Most frequently, computerized information originates in bills or invoices that are handwritten. Keypunch operators or typists extract certain information from these invoices, such as dollar amounts and account numbers, and manually feed this information into a computer. Other useful but nonnumeric information is not fed into the computer. For example, the shape of signatures is not computerized, and any evidence of forgery is not computerized. When such nonnumeric information is relevant to a courtroom proceeding, the original invoices should be presented to the trier of fact rather than a computer printout. Microfilm copies of the original invoices will usually suffice if the original invoices have not been saved, but such copies may not be adequate for signature identification or forgery detection. If numerous invoices are involved, the trier of fact can examine a randomly-selected sample of the invoices to conserve time.

If the information is manually typed or keypunched before it enters the computer, the human error rate will always be greater than zero. In most properly run computer installations, procedures are followed to detect and correct most but not all such human errors, as I have explained. Sometimes data processing personnel neglect to follow such procedures even when they are supposedly in effect. Numerous errors can result from such neglect, and the credibility of any information processed under such conditions is highly questionable. The trier of fact should be informed of the actual error rate, if it is known. The past error experience of the computer installation is good evidence of the actual error rate, and so is the theoretical number of errors that one would expect to result from use of a given manual information entry procedure. In some instances, actual error checks may have to be carried out to determine the human error rate.

When humans are not involved in typing or keypunching information into a computer, the likelihood of input errors is greatly reduced. A department store, for example, can print price and inventory information onto individual price tags using a machine-readable bar code, and they can issue magnetically encoded badges to their employees and magnetically encoded credit cards to their customers. A cash register of the conventional type is then not needed. At the checkout counter, the store employee inserts his or her badge and the customer's credit card into a magnetic scanner and then draws a special light sensing pen across each price tag. In this way the customer charges are recorded without error, and an accurate inventory turnover record is captured at the same time. Since transactions of this type generate computer stored information directly without any intervening written record, a computer printout of such information should always be accepted as an "original". But errors can result from use of the wrong card, errors or defects in the magnetically or optically scannable codes, and from failure of an employee to follow the proper procedures. Controls are also needed to prevent the occurrence of unauthorized transactions or the nonrecording of authorized transactions, and the presence or absence of such controls affects the credibility of this type of evidence.

Once information has been entered into a computer or is stored on a computer information storage medium such as magnetic tapes, magnetic disks, or punched cards, the information is moved about and copied under the control of computer programs. These will sometimes instruct the computer to alter the information. Large data processing installations contain hundreds and sometimes thousands of such programs, each containing anywhere from a few to a thousand or more individual instructions. Erroneous instructions are frequently found within these computer programs, particularly the more complex types of programs. Computer programs are written out by "computer programmers" in much the same way that a book is written out by its author, and computer programs are corrected in much the same way that a book is proofread. Except in the case of very simple programs, it usually takes longer to correct a program than it takes to write the program in the first place. Frequently, a computer program is placed into operation before all of its errors have been found, and sometimes errors are uncovered months or even years after a program is first put into service. Such undetected errors can cause a program to alter the information stored within a computer, and the alterations may be difficult to detect. The only way to insure that a computer program is operating properly is to test it thoroughly. If the programs in a computer have not been thoroughly tested, then there is no assurance that the information processed by the computer is accurate. Before trial, a party who must rely upon computer-generated evidence should verify that all the programs that processed the evidence have successfully processed test information without altering it and without malfunctioning. Long and satisfactory use of a program in an environment where errors will quickly be brought to the attention of management, as by customers complaining about improper billings, can create a presumption that a program is free of programming errors, and such a showing may be sufficient to establish the credibility of evidentiary information processed by that program.[23]

[23] Offering evidence of a lack of previous complaints as proof of a computer program's lack of errors is analogous to the offer of evidence in negligence and products liability cases of the absence of other accidents as proof of a safe situation. The trend of recent decisions seems to favor the admissibility of this type of evidence. The main objection to the admissibility of such evidence in accident cases has been that the persons passing safely did not expose the allegedly defective condition to the same use or test of its integrity as did the injured party who was suing. That objection has less

Computer-stored data may be altered intentionally. A skilled programmer who understands a given computer system and has direct access to the system can alter the data stored within the system, leaving no trace of the alteration. To protect against such alterations, a variety of safeguards—too numerous to be listed here—are required, particularly when large sums of money are involved. The accompanying article by Mr. DeHetre describes these safeguards. Through experience gained in working with a wide variety of computerized accounting systems, CPA computer experts like Mr. DeHetre have gained insight into the types of safeguards required to insure the integrity of a given type of computer installation. One standard rule that all computer-wise auditors insist upon, for example, is that skilled programmers should never be permitted near the computer. They should always be required to submit their programs to others whose sole function is to supervise the actual computational equipment. As an additional safeguard, almost all well-run data processing installations keep an accurate log of what programs are run and at whose request. These are just two of many safeguards that are possible. Some companies conduct audits of their electronic data processing systems at regular intervals to insure that the necessary safeguards are continuously in effect. Without such safeguards, one can never be certain that computerized information has not been altered. With such safeguards, there is still a possibility that the information may have been altered. The possibility, however, is greatly reduced, and the cost of the consequences that could result from the alterations still possible are less than the cost of adding further safeguards to the system. In other words, the safeguards employed will be those that are cost effective. In any installation, some desirable safeguards from the point of view of the trier of fact will not have been implemented simply because they did not prove themselves to be cost effective.

Another factor the trier of fact should consider when evaluating the credibility of evidence processed by a digital computer is who has custody of the computer. If the computer is under the control of one party and the computerized evidence is submitted by the opposing party to support the opposing party's position, there is little likelihood that the evidence will have been altered, since the party controlling the computer will see to it that the evidence is accurate. Similarly, data obtained from a computer controlled by an impartial third party is not likely to have been altered. But when a party introduces evidence taken from his own computer, the situation is much the same as when the party is testifying before the court himself. A party against whom such evidence is presented should be given wide latitude to investigate the credibility of the evidence, and any indication that the

weight in the computerized data processing environment. There, the computer programs make exactly the same set of data manipulations over and over and are not subject to gradual deterioration but remain constant. *See* McCormick, Evidence § 200 at 476 (2d ed. 1972).

proper safeguards are absent from the computer should be taken as an indication that the party proffering the evidence at least had the opportunity to alter the evidence while it was under his control. In the absence of adequate safeguards, it is just as easy to lie with a computer as it is to lie with oral testimony. Hence, self-serving computer statements should always be subject to close scrutiny.

The final factor which warrants consideration is the way in which computer-generated evidence is conveyed from the computer to the courtroom. It is tempting for the court to focus too much attention upon the form of the evidence at this point and too little attention upon the safeguards that have been taken to insure its accuracy. For example, in *North Carolina v. Springer*,[24] the Supreme Court of North Carolina rejected oral testimony to the effect that computerized records indicated $1,209.63 worth of purchases had been made with a stolen credit card because, the court held, the dollar amount should have been proved by means of the "best evidence"—in this case, a computer printout.[25] In so holding, the North Carolina court made two errors. First, the court held a printout of computer-stored credit card records to be admissible as the "best evidence" even though such records are incomplete insofar as they do not accurately depict the signatures and other nonnumeric features of the original transaction records. The court should have insisted upon in-court production of the original credit card invoices. Second, the court elevated an ordinary computer printout to the status of a signed contract by refusing to accept computer-generated evidence offered in any other form. While a computer printout might be preferable in a case where the information is voluminous, surely a witness can be relied upon to recite the dollar figure "$1,209.63" to the trier of fact without having to bring a computer or a printout into the courtroom. In some cases, it will be difficult or impossible for such a witness to obtain a printout, as when computer-stored information is displayed upon a display device having no printer. Insistence upon the production of an actual printout in every case may thus be unfair.

A computer printout lacks the uniqueness of a signed contract. Any computer may be readily programmed to print out any desired information in any desired format, and it is normally impossible to distinguish a genuine printout of evidentiary information from a fabrication printed at another time by the same computer and printer. One printout may be readily substituted for another prior to trial, and there is almost no way that such a substitution can be detected. The credibility of a printout must, therefore, be established by a witness who, for example, produces the printout and asserts that it has been kept in a secure place continuously from the time when it was printed until it was brought into the courtroom. An even better

[24] 283 N.C. 627, 197 S.E.2d 530, 5 C.L.S.R. 432 (1973).

[25] *Id.* at 633, 197 S.E.2d at 536.

procedure is to have the witness sign the printout immediately after it is printed and later identify the printout as genuine by reference to the signature. In the absence of such testimony, no printout should ever be considered credible except by stipulation. A printout should be regarded as nothing more than a set of notes that is brought into the courtroom by a witness to help him refresh his memory and to help him testify about computer-stored information.

The process of judging the credibility of computer-stored evidence thus involves four steps: studying the way in which the evidence was entered into the computer for possible human, measurement, and numeric conversion errors, and determining the frequency and magnitude of any such errors; evaluating the instruction sets or computer programs used to process the evidence to determine whether the programs may have altered the evidence in any way; investigating the safeguards employed to prevent unauthorized persons or adverse parties from having direct access to either the computer or to the tapes, disks, or cards that may have been used to store the evidence away from the computer in machine-readable form; and judging the credibility of the witness who conveys the evidence from the computer to the courtroom. No computer system is ever completely secure, and a thorough investigation will always raise some doubts about the credibility of any computer-generated evidence. These doubts should be taken under advisement by the trier of fact.

USING A COMPUTER TO MANIPULATE AND SIMPLIFY EVIDENCE PRIOR TO TRIAL

In the above discussion, the computer was treated as a conduit through which evidentiary information flows, hopefully without alteration. Computers can also be used to manipulate evidentiary information. If the evidentiary information is organized in such a way that it is difficult to understand, a computer can reorganize it for clarity. If the evidentiary information is voluminous, a computer can summarize it. If the trier of fact must render a decision that cannot be properly rendered without specific knowledge of how some highly complex system or entity functions, a computer may be used to simulate or model the complex system or entity. These uses of computers to manipulate evidentiary information should be encouraged, since they make it possible for the trier of fact to render more intelligent judgments than would otherwise be possible, and they also save time in the courtroom. Each of these uses, however, creates its own special problems.

When a computer is used to reorganize information, it is important to insure that no information is lost in the process. During trial, an impartial expert who supervised the reorganization process should demonstrate to the satisfaction of the judge that no information has been lost. Only one question needs to be asked of this expert: Can the reorganization be reversed and the information returned to its original form? If the reorganization is not fully reversible, then some information has been lost, and the information presented in court should be considered a summary and not a simple reorganization.

Frequently, one party may wish to enter into evidence information taken from the other party's computer but organized in a special way. In response to an appropriate discovery request, the other party might refuse to produce the information, stating that the information is not organized in the specified way and that it would be costly and "burdensome" to reorganize the information for production in the form specified. Such an objection may be valid, but it might also be dilatory. How should this situation be handled? Perhaps the best thing for the judge to do is order production of the information without reorganization in both printed and machine-readable form accompanied by a detailed explanation of how the information is organized. Such an order should not prove a burden upon the producing party, for the printed and machine-readable copies are readily generated, and the producing party should already have a detailed explanation available for his own use. The requesting party may then reorganize the information in any way he chooses, at his own expense and preferably under the supervision of an impartial data processing expert.[26]

Summaries of computerized information are readily prepared by a computer. Computers are widely used by social scientists for summarizing and simplifying information, and a number of special computer programs are available for their use.[27] Such programs are used routinely by the American Bar Foundation to process and summarize the results of the Foundation's many empirical studies of law. The Illinois State Bar Association recently used such programs to summarize the results of its recently commissioned economic survey of the Illinois Bar.[28] Attorneys in private practice should take advantage of such special programs when faced with the need to summarize computerized information for courtroom use.

A summary is a condensation. Of necessity, some information is lost when information is reduced to summary form. A summary of evidentiary information can favor either party, depending upon what information is lost and what is retained. Usually, the party who prepares a summary will attempt to make it favor his own position if he can do so without arousing

[26] *See* United States v. Davey, 404 F. Supp. 1283 (S.D.N.Y. 1975).

[27] For example, SPSS (Statistical Package For The Social Sciences) and DATA-TEXT are two such programs. N. NIE *et al.*, STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES (2d ed. 1975). D. ARMOR ET AL, DATA-TEXT PRIMER—AN INTRODUCTION TO COMPUTERIZED SOCIAL DATA ANALYSIS (1972).

[28] Illinois State Bar Association, *Economics of Legal Services In Illinois,* 64 ILL. B.J. 73 (1975).

the suspicion or hostility of the trier of fact. A case on point is *United States v. Russo*.[29]

The court in *Russo* permitted the jury to examine a computer-prepared summary of payments made by Michigan Blue Shield to Michigan doctors during 1967 for five specific medical procedures. The summary indicated, for example, that 17,747 payments had been made for a procedure called "subsequent aspiration of the bursa."[30] Other evidence indicated that 14,141 of these 17,747 payments had been made to the two defendants who operated a medical clinic in a Detroit working class neighborhood.

The summary evidence thus strongly indicated that the defendants administered the procedure called "subsequent aspiration of the bursa" far more times than did an average doctor in the state of Michigan—but the figure "17,747" was a summary figure, and it was highly prejudicial to the defendants. Many Michigan doctors undoubtedly do not treat arthritic conditions at all, and such doctors probably never use this particular procedure. Additionally, the defendants operated a clinic and had other doctors working in their employ. They undoubtedly treated far more patients and applied for far more payments of all types than would a doctor who practiced by himself. The defendants were also osteopaths and thus likely to use different procedures than those used by medical doctors. A properly prepared summary would have taken these and other similar factors into consideration and would have attempted to account for them. As things stood, the figure "17,747" probably should have been excluded from evidence, since it was highly prejudicial and not particularly meaningful when unaccompanied by additional information. As an alternative to requesting the court to exclude this evidence, the defendants could have made up their own summary of this same information, manipulating the data to slant it the other way as much as possible. In the *Russo* case, there was overwhelming evidence in addition to this summary that adequately supported the conviction;[31] thus the fact that the appellate court did not order a new trial was "harmless error." But in any similar case where the evidence for conviction is not so overwhelming, potentially prejudicial summary evidence of this type should be handled much more carefully by the court.

The defendant in *Russo* strongly argued that summarized evidence is inadmissible, and the appellate court was troubled by this argument. But after acknowledging that summary evidence must be treated with care, the *Russo* court noted that the summary in question had been prepared by Michigan Blue Cross for business and accounting purposes, not specifically for use in litigation. The court thus held the summary to be admissible under the shop book exception to the hearsay rule, stating that it was an

[29] 480 F.2d 1228, 5 C.L.S.R. 687 (6th Cir. 1973).

[30] 480 F.2d at 1236, 5 C.L.S.R. at 696.

[31] 480 F.2d at 1232-34, 5 C.L.S.R. at 690-91.

"original record and not a mere summary"—an absurd holding.[32] True, the summary was an original business record, but it was still a summary and as such excludable if it proved to be highly prejudicial to the defendants.[33] As a general principle, a summary exists whenever the volume of information has been reduced. To determine whether evidence has been summarized, the court should ask whether the original source of information can be recreated from the evidence. In *Russo*, the original computerized billing records of Michigan Blue Cross could not possibly have been recreated from the evidence placed before the jury, so the trial court should have investigated further to determine whether the evidence in summary form was prejudicial to either party.

Computer simulations or models, as described in the articles by Carol Eastin and Martha Jenkins, can be highly beneficial to the trier of fact, particularly if they help to illustrate how a complex real-world system functions. The use of a model or simulation as a teaching tool to help explain a complex phenomenon is its most useful role in the courtroom. But like summaries, simulations or models are almost always simplified representations, and they can prejudice either party by making the other party's position appear more favorable than it actually is. Modeling and simulation are particularly risky, since very small errors in the underlying assumptions that govern a computer model's operation can cause gross errors in the predictions which flow from the model. For example, an international organization recently used a computer model to predict that man was headed for unavoidable disaster during the next 50 to 100 years.[34] Others quickly pointed out that the prediction was unduly pessimistic because of some improvidently selected constants used in setting up the computer model.[35]

An excellent example of how a computer model may be used to advantage is presented in a book by Jay W. Forrester.[36] Mr. Forrester uses a computer to simulate the operation of a city, taking into account its housing, industrial capacity, population, and the way these factors interact with each other. He uses his computer model to demonstrate that many of the things politicians try to do to help a city often contribute to its deterioration over the long run. For example, Forrester demonstrates that the building of public housing can greatly speed the deterioration of a city,

[32] 480 F.2d at 1240-41, 5 C.L.S.R. at 698-99.

[33] "Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury . . . ." FED. R. EVID. 403.

[34] D. MEADOWS *et al.*, THE LIMITS TO GROWTH: A REPORT FOR THE CLUB OF ROME'S PROJECT ON THE PREDICAMENT OF MANKIND (1972), *reviewed in* McCarty, Book Review, 8 STAN. J. INT. STUDIES 154 (1973).

[35] MODELS OF DOOM: A CRITIQUE OF THE LIMITS TO GROWTH (Cole et al ed. 1973). Saunders, *Criticism and the Growth of Knowledge: An Examination of the Controversy over the Limits to Growth,* 9 STAN. J. INT. STUDIES 45 (1974).

[36] JAY W. FORRESTER, URBAN DYNAMICS (1969).

although one would expect the opposite result. He does not pretend that his computer model is an accurate representation of the dynamics of a city.[37] Instead, he uses the computer as a teaching tool that enables his reader to follow the workings of an extremely complex phenomenon—one that would be very hard to follow without computer assistance. One should never assume that the predictions of a computer model are any more accurate than the assumptions underlying its operation, and one should always remember that those assumptions are usually simplified in comparison to the real world system that is being simulated by the computer.

SUMMARY

Judges and attorneys do not have to become technical experts before they can handle computer-generated evidence. All they need to do is carefully trace the flow of the computerized evidentiary information from its source to the courtroom, concentrating upon the accuracy of the information rather than its form or the technical complexities of the computer through which it passes. Technical experts who testify in court about such evidence should be instructed to describe what happens to the information without explaining the technical complexities of the computers through which the information passes, just as I have done in this article, so that their testimony will be more understandable to all courtroom participants.

Attention should be focused upon the way the evidentiary information enters the computer, since most human errors occur at that point in its flow path. In addition, if the information originates as direct measurements of distance, time and the like, the magnitude of any significant measurement and roundoff errors should be considered. Attention should also be focused upon the reliability of the programs that move the information about within the computers, since program errors can adversely affect the credibility of the information processed. Also, the safeguards that prevent adverse parties from gaining access to a computer or to the evidentiary information when it is stored away from the computer should always be reviewed to determine the likelihood that such a party might have altered the information, if he had the inclination or the expertise to do so.

If the evidentiary information is reorganized by a computer prior to the trial, a check should be made to see if any information has been lost. If information has been lost, the evidentiary information has been summarized, and it needs to be checked carefully for any possible bias introduced by the summarization process. Finally, the assumptions underlying any computer model or simulation should be checked out with great care to determine the extent to which they accurately reflect the real-world entity or system that is being simulated or modeled.

[37] A highly critical appraisal of Forrester's Urban Model appears in Ackerman, *Regulating Slum Housing Markets on Behalf of the Poor: Of Housing Codes, Housing Subsidies and Income Distribution Policy,* 80 YALE L.J. 1093, 1141-43 & n.49 (1971).