DOES DISCOVERY OF ELECTRONIC INFORMATION REQUIRE AMENDMENTS TO THE FEDERAL RULES OF CIVIL PROCEDURE?

Committee on Federal Procedure

2/15/01

SUMMARY

The Advisory Committee on the Civil Rules of the Committee on Rules of Practice and Procedure of the Judicial Conference of the United States (the "Advisory Committee") has been studying whether any changes to the Federal Rules of Civil Procedure are required to accommodate discovery of electronic information.

The Section recommends that no change in the Federal Rules of Civil Procedure should be made. The issues involving electronic information can be dealt with under the present rules. At most, it may be appropriate to provide guidelines or commentary in The Manual for Complex Litigation or elsewhere to guide analysis by practitioners and the courts as novel questions arise.

WHAT IS ELECTRONIC INFORMATION?

Some definitions and descriptions of electronic information are necessary before analyzing discovery issues relating to it.

ACTIVE DATA consist of information readily available and accessible to computer users through file manager programs.

EMBEDDED DATA or METADATA consist of information contained within an electronic version of a document that may not be apparent in a print-out, such as the date the document was created, the identity of the author, the identity of subsequent editors, the distribution route for the document, or the history of editorial changes.

REPLICANT DATA are copies automatically made and saved to the user's hard drive.

RESIDUAL DATA are deleted files to which the reference has been removed from the directory listings and the file allocation table, but which have not been overwritten.

BACK-UP DATA consist of information copied to removable media in the event of a system failure, usually only of data on a centralized storage medium or network, and frequently in compressed form.

LEGACY DATA consist of information stored on media that can no longer be accepted or organized in a format that can be read using current software.

In addition, in connection with the internet, there are a few specific types of electronic information.

BOOKMARKS are one-click shortcuts created by the user and stored on the user's computer.

CACHE FILES are a record of internet addresses visited by the user and graphic

elements on those pages created and stored automatically by the user's computer.

COOKIES are information about the user placed in a file by a web-site operator.

From these definitions and descriptions several issues emerge. One is a concern with the impermanence of active data. It can be altered easily and a print-out of the document or its appearance on a screen may not indicate any change or what existed before the change. Litigants can be justifiably concerned that important information may be altered deliberately or in the ordinary course in a manner that might not be noticeable on its face.

One response to this concern might be to retrieve embedded or replicant data. However, that is not routinely available and may require some technical expertise for retrieval. It also therefore may require some extra expense.

The question of retrieving embedded or replicant data also raises a metaphysical issue: Is embedded data or other electronic information created automatically by a computer, such as a cache file, a "document" subject to discovery under the Federal Rules of Civil Procedure? Judge Shira A. Scheindlin has suggested that computer-generated information is not a discoverable document under Rule 34(a)'s definition including "data compilations from which information can be obtained, translated, if necessary, by the respondent through detection devices into reasonably usable form." See Scheindlin & Rabkin, Electronic Discovery in Federal Civil Litigation: Is Rule 34 Up to the Task? 41 B.C.L.Rev. 327, 346-47, 371-74 (2000). The argument is that information appended electronically and automatically is not a "data compilation." We believe that any judge faced with this metaphysical argument will hold that any electronic information, no matter how generated, is a "document" within Rule 34. See Anti-Monopoly, Inc. v. Hasbro, Inc., 94 Civ. 2120 (LMM) (AJP), 1995 WL 649934, at *2 (S.D.N.Y. Nov. 3, 1995) (Peck, M.J.) ("today it is black letter law that computerized data is discoverable if relevant").

Another issue that emerges from the definitions and descriptions above is what role, if any, should be accorded an intent to remove or destroy a copy. By striking or clicking a delete button, instruction or icon, a person presumably intends to toss away or destroy the information in the document, whether it is an e-mail, a draft of an agreement, a spreadsheet with incorrect data, or a piece of correspondence that is never sent. However, "deleting" on a computer does not remove the information until it is overwritten or scrambled, and even then it may still be able to be recreated, albeit at a significant increase in the cost of retrieval. Withers, Computer-Based Discovery In Federal Civil Litigation, 2000 Fed. Cts. L. Rev. 2, at II.D. (2000). Should the intent of the user be honored? If not, under what circumstances should the information be retrieved, and at whose cost?

Related to the role of intent is the question of retrieving relevant information from back-up data and legacy data, the difficulty of which retrieval is compounded by the relative ease of storing huge amounts of electronic information. Back-up tapes or cartridges are regularly created for the purpose of restoring entire computer systems if they should happen to crash; they are not created for the purpose of preserving information relevant to a particular litigation. As a side effect, relevant information may be preserved along with a massive amount of other information necessary to the operation of a particular computer system. See, e.g.,

Linnen v. A.H. Robins Co., 10 Mass. L. Rptr. 189 (Super. Ct. 1999) (back-up tapes recorded all use of the intra-office communication capabilities, electronic mail communications, and any word-processing files, spreadsheets, and models). Therefore, one must search a massive amount of backed-up electronic information to locate those relatively few pieces of arguably relevant electronic information. Who should bear the burden of such a search in terms of time, personnel and cost?

For legacy data, there is the further complication of recreating the environment in which the original electronic information was created and stored. Sometimes computer museums must be consulted to locate the appropriate hardware to accept the outmoded media, and special programs must be w ritten to convert the data into usable form. Withers, Is Digital Different? Electronic Disclosure and Discovery in Civil Litigation, http://www.kenwithers.com/bileta/
elecdisc.htm, at III.D.1. (Dec. 30, 1999). Who should bear this burden?

DO THE FEDERAL RULES NEED TO BE AMENDED?

In our view, any analysis of the applicability of the Federal Rules of Civil Procedure to electronic information starts with a fundamental principle: if the information is relevant to the claims or defenses in a case it should be discoverable, whether the information is in traditional or electronic form. Rule 26(b)(1). However, in allocating the burdens of producing electronic information, the costs of producing the information should be weighed against its benefit and importance. Rules 26(b)(1), 26(b)(2)(iii). These principles, already ensconced in the Federal Rules of Civil Procedure and applied by the courts to disputes involving electronic and other information, are sufficient to deal with the issues arising in regard to electronic information. Further amendment is not necessary.

Preservation and Spoliation

Preservation of information is not specifically covered by the Federal Rules of Civil Procedure. Case law has developed the rule that when it is reasonably foreseeable that a claim may be asserted, a party must preserve relevant information. Shamis v. Ambassador Factors Corp., 34 F. Supp. 2d 879, 888-889 (S.D.N.Y. 1999); Wm. T. Thompson Co. v. General Nutrition Corp., 593 F. Supp. 1443, 1455 (C.D. Cal. 1984); Carlucci v. Piper Aircraft Corp., 102 F.R.D. 472, 485-86 (S.D. Fla. 1984); Bowmar Instrument Corp. v. Texas Instruments, Inc., 25 Fed. R. Serv. 2d (Callaghan) 423, 427, 1977 U.S. Dist LEXIS 16078, at *11 (N.D. Ind. May 2, 1977); see Gorelick, Marzen and Solum, Destruction of Evidence § 3.12, at 104 (1989). There seems little reason to change this standard just for electronic information. Further, adding a rule for preservation of electronic information to the Federal Rules of Civil Procedure without a more general rule covering all information would be anomalous. A proposed amendment to Rule 34(b) to require information to be preserved only "whenever information . . . is requested" is far too late in many lawsuits, especially for active data.

Thus, attorneys for corporate litigants should routinely advise their clients to preserve their electronic information when they become aware of a potential claim. The electronic information to be preserved might include data files created by word-processing, spreadsheet or other applications software, electronic mail, electronic calendars and telephone logs and data bases (including structural information). See Feldman, The Essentials of

Computer Discovery, at III.A., American Bar Association (July 2000). It should be (i) copied from file servers, desk-top computers, back-up tapes or cartridges or wherever else it is stored and (ii) preserved on tapes, cartridges or CDs. Such a back-up should avoid questions of alteration of active data and should preserve embedded and residual data.

For more complex claims, such as antitrust or securities claims,1 judgment may have to be used to determine what steps to take to preserve relevant information, just as should be done now with regard to non-electronic information. A reasonable retention policy for electronic information should be adopted by companies, just as is done for information stored on paper. Nonetheless, as with discovery of non-electronic information, the steps taken to preserve relevant electronic information for any litigation may be second-guessed by a court. See Lewy v. Remington Arms Co., 836 F.2d 1104, 1111-12 (8th Cir. 1988) ("even if the court finds the policy to be reasonable given the nature of the documents subject to the policy, the court may find that under the particular circumstances certain documents should have been retained notwithstanding the policy").

If spoliation should occur, there appears to be ample authority under Rule 37 to impose appropriate sanctions. For example, in Illinois Tool Works, Inc. v. Metro Mark Prods., Ltd., 43 F. Supp. 2d 951, 957, 962-63 (N.D. Ill. 1999), the district court under Rules 37(a)(4)(A) and 37(b)(2) imposed attorneys' and a computer expert's fees and costs upon a party which deliberately disconnected hardware and corrupted data in violation of an order to preserve the integrity of computers and information stored upon them.

Initial Disclosure

Beginning December 1, 2000, initial disclosure of information, including a copy or a description by category and location of all data compilations, supporting a party's claims or defenses (whichever is applicable) became mandatory in all cases. Rule 26(a)(1). Since one of the theories behind initial disclosure, which the Section has opposed in the past, is that it merely requires responses to a standard set of initial interrogatories and document requests, it has been suggested that initial disclosure be expanded to include a description of any computer system or media maintained by the disclosing party from which discoverable information might be obtained, including the nature of such discoverable information, the software program used to store such information, the nature of any back-up medium, and the nature of any computer network or e-mail system maintained by the disclosing party.

The Section concurs with Southern District of New York Judge Lewis Kaplan's comment at an October 27, 2000 conference on computer-based discovery held at the Brooklyn Law School under the aegis of the Advisory Committee that, since initial disclosure is a bad idea, it should not be expanded. Moreover, the need for discovery of electronic information is not yet so ubiquitous that a description of the computer systems on which it is stored should automatically be a subject of disclosure and discovery in every case. Most personal injury plaintiffs should not be expected in every case to describe their home computer systems to determine if there is discoverable material on them. On the other hand, multinational corporations should not have to describe their numerous computer systems in many locales in every contract action. The proposal assumes a condition that it is not yet evident will ever exist:

that all information will be electronically stored and electronically retrievable. When such a condition is true, it may be time to amend the federal rules to require automatic disclosure of the nature and extent of electronic information storage and retrieval systems. Until then, the parties should investigate electronic storage and retrieval systems only when required by the circumstances of the case.

Discovery Conference and Pre-Trial Order

Similarly to the proposal for initial disclosure, it has been suggested that the storage and retrieval of electronic information should be a mandatory subject to be discussed at the parties' initial discovery conference under Rule 26(f) and at the subsequent pre-trial conference with the court. Again, it does not appear to the Section that issues concerning retrieval of electronic information are so ubiquitous that it should be a required regular topic of conversation among parties to a lawsuit. If production of electronic information will be an issue, we feel confident that the attorneys for the parties will raise it during discovery and pre-trial conferences and the issue will be presented, if appropriate.

The Burden of Discovery

The crux of the issues that have arisen with regard to electronic discovery is the allocation of the burden and cost of retrieval and production of electronic information. While some have argued that the current rules do not provide a sufficient guide, the Section believes they do.

Under Rule 34, a party need only produce a document to be inspected, and the requesting party generally bears the cost of reproduction (by photocopier or handwritten notes). However, the Advisory Committee Notes to the amendment to Rule 34 in 1970 provides that the producing party "may be required to use [its] devices to translate the [electronic] data into usable form," thereby apparently shifting the burden and cost of producing electronic information to the producing party. Nonetheless, the courts have long had the authority (and have exercised it) to adjust the burdens and costs of production of electronic information depending on the benefits to be gained under Rules 26(b)(2) and 26(c).1

The Section believes that for active data, replicant data, and embedded data, the burden and cost of production should almost always be borne by the producing party. For back-up data and residual data, the costs and benefits should be weighed and the costs and burdens shared appropriately. For legacy data, the costs will more often be shifted in whole or in part.

(i) Active, Replicant and Embedded Data

For active data, the producing party should almost always bear the initial burden and cost of production. The producing party is responsible for choosing its own application programs and should therefore be responsible for using those application programs to produce relevant information. See In re Brand Name Prescription Drugs Antitrust Litig., 94 C 897, MDL 997, 1995 WL 360526, at *2 (N.D. Ill. June 15, 1995) ("if a party chooses an electronic storage method, the necessity for a retrieval program or method is an ordinary and foreseeable risk"). Further, imposing the burden to produce active data on the producing party will encourage it to take steps to preserve that information once it becomes aware of a potential claim. In addition, the producing party should bear the burden of reviewing the electronic information to screen out irrelevant and privileged information, which should be done using its application programs.

The same analysis and reasoning applies to replicant and embedded data where the producing party's application program creates and links additional data to the document residing in active files which is to be produced. Further, the technical expertise to access the embedded data or replicant data should exist with the producing party that is using the application program.

To take a simple example: if a document is created in a WordPerfect word-processing program, but the requesting party only uses a Word word-processing program, it is reasonable and appropriate for the producing party to produce the document using its application program. Further, to the extent that embedded data for that document is to be produced, it is the producing party's WordPerfect word-processing program that must be used to retrieve that information. As a corollary, the producing party should not have to bear the cost of "translation" from WordPerfect to Word. But see Sattar v. Motorola, Inc., 138 F.3d 1164, 1171 (7th Cir. 1998) (Seventh Circuit approved as entirely reasonable a district court's requirement that a party, which had produced e-mails in the form of 4-inch tapes inaccessible to the requesting party, who lacked the equipment and software to read them, to (i) download the data from the tapes to conventional computer disks or a hard-drive, or (ii) loan the requesting party the necessary software, or (iii) offer on-site access to its own system).

(ii) Residual or Back-Up Data

For residual and back-up data, the technical requirements of retrieval become more complex and costly and the separation of relevant information becomes more burdensome than for active, embedded or replicant data. To search for residual data, it may be necessary to make a mirror image of the hard drive of a computer. See Gates Rubber Co. v. Bando Chemical Indus., Ltd., 167 F.R.D. 90, 112 (D. Colo. 1996) (defendants' expert sanctionably failed to copy all relevant information by doing a file-by-file backup instead of creating a mirror image, which would have collected every piece of information on the hard drive). To search back-up data may require re-creating the computer environment in which the data was originally recorded. Moreover, the electronic information on a hard drive or on back-up media may include a massive amount of irrelevant information related to the operation of the computer system or matters unrelated to the litigation. The relevant information must be separated from the irrelevant, the difficulties of which search are compounded by the lack of any organizing principle for storage of electronic information on electronic media. Withers, Computer-Based Discovery In Federal Civil Litigation, 2000 Fed. Cts. L. Rev. 2, at II.E.1. (2000).

For residual and back-up data, courts should rely on the balancing test articulated in Rule 26(b)(2)(iii) to weigh the costs and benefits. For example, in Bills v. Kennecott Corp., 108 F.R.D. 459, 464 (D. Utah 1985), the court considered the following non-exclusive factors in exercising its discretion to deny the producing party's motion to shift the costs of producing electronic information:

(1) [t]he amount of money involved [$5,411.25] is not excessive or inordinate; (2) [t]he relative expense and burden in obtaining the data would be substantially greater to the requesting party as compared with the responding party; (3) [t]he amount of money required to obtain the data as set forth by the defendant would be a substantial burden to plaintiffs; [and] (4) [t]he responding party is benefitted in its case to some degree by producing the data in question.

Even where the retrieval cost is a more substantial $50,000 to $70,000, one court refused to shift costs because "the costliness of the discovery procedure involved is a product of the defendant's record-keeping scheme over which the plaintiffs have no control." In re Brand Name Prescription Drugs Antitrust Litig., 94 C 897, MDL 997, 1995 WL 360526, at *2 (N.D. Ill. June 15, 1995) (quoting Delozier v. First Nat'l Bank of Gatlinburg, 109 F.R.D. 161, 164 (E.D. Tenn. 1986)). However, another court distinguished between data that could be produced with no more than de minimis programming, for which the costs would not be shifted, and

production of data requiring special programming, for which it required the requesting party to pay.1 Anti-Monopoly, Inc. v. Hasbro, Inc., 94 Civ. 2120 (LMM) (AJP), 1996 WL 22976, at *2 (S.D.N.Y. 1996). And a third court determined that the costs of restoring residual data (256 hours and $15,675 per employee's computer) did not justify the expected benefit from restoring all deleted files. Alexander v. Federal Bureau of Investigation, 188 F.R.D. 111, 117 (D.C.D.C. 1998).

Courts have also retained computer forensic experts as officers of the court, see Rule 26(c)(5), at the requesting party's sole expense to create a mirror image of a computer's hard drive, which is then provided to the producing party's counsel, who undertakes the burden (and cost) of reviewing the documents on the mirror image to produce relevant, non-privileged documents and information. Simon Property Group L.P. v. mySimon, Inc., 194 F.R.D. 639, 641-42 (S.D. Ind. 2000); Playboy Enterprises, Inc. v. Welles, 60 F. Supp. 2d 1050, 1054-55 (S.D. Cal. 1999). Further, as of September 2000, it appears that software, called EnCase, has been developed "to review the evidence and perform basic analysis of the mirror-imaged drives." Patzakis, "Electronic Evidence Discovery: From High-End Litigation Tactic to Standard Practice," Federal Discovery News, Vol. 6, No. 10, at 4 (Sept. 2000).

(iii) Legacy Data

Legacy data will generally present the strongest case for shifting the cost of retrieval to the requesting party. The argument that the producing party has chosen its computer system loses some force, because the producing party has abandoned the computer system on which the legacy data was originally created. Therefore, the argument that the party seeking the information, which might require extraordinary efforts to retrieve, should pay the cost of that retrieval carries more weight. However, if both parties would benefit from the retrieval of the information, such as transaction data in an antitrust price-fixing case, there is an argument for dividing the cost of recreating the computer system on which the legacy data was first recorded. Further, it may make sense to sample the universe of legacy data to determine statistically whether the benefit of finding relevant information will justify the cost of retrieval, before determining to proceed with a full retrieval and search of the information available. See Linnen v. A.H. Robins Co., 10 Mass. L. Rptr. 189, 1999 WL 462015, at *5 (Super. Ct. June 16, 1999).

* * *

The Section concludes that no federal rule can predict technological or software developments concerning the retrieval of residual, back-up or legacy data. The balancing of costs and benefits embodied in the current Rules is the best approach for determining whether discovery of such information should be undertaken, and, if so, at whose expense.

Form of Production

The issues relating to the form of production of electronic information are: (1) May a producing party, at its option, produce electronic information in hard copy or electronic form? (2) May the requesting party require that information already produced in hard copy, also be produced in electronic form, or vice versa? (3) If the requesting party can require the production of electronic information in both forms, should it have to pay for the second copy?

Rules 34 and 33 appear to allow the producing party to choose the form in which to produce electronic information. Rule 34(a) places on the producing party the burden of translating data compilations into reasonably usable form. Rule 34(b) requires a producing party to produce documents as they are kept in the usual course of business or organized to correspond to categories in the request. Thus, it appears that Rule 34 allows the producing party to choose whether to produce electronic information in electronic form (as it is "kept in the usual course of

business") or in hard copy ("organize[d] . . . to correspond with the categories in the request"). Rule 33(d) allows parties responding to interrogatories to answer by allowing a requesting party to make copies or compilations of records from which an answer may be derived or ascertained, if the burden of deriving or ascertaining the answer is substantially the same for both. It would appear that this rule permits a producing party to choose the form of the information to be reviewed by the requesting party.

The case law appears to be split on the question of whether the producing party must produce electronic information in a second form, once it has been produced in one form. In In re Air Crash Disaster at Detroit Metro. Airport, 130 F.R.D. 634, 636 (E.D. Mich. 1989), the district court directed McDonnell Douglas Corporation to produce a nine-track computer tape of a flight director simulation program and data, even though it had previously produced a hard copy. Applying a balancing approach, the court found that it would not be unduly burdensome for McDonnell Douglas to produce the data in computer readable form, which would reduce unnecessary costs and delays of the requesting party to manually load the program and accompanying data.

In National Union Elec. Corp. v. Matsushita Elec. Indus. Co., 494 F. Supp. 1257, 1258, 1260, 1261 (E.D. Pa. 1980), Judge Becker required plaintiff to create a computer readable tape containing certain data previously supplied in paper computer print-out reports in answer to interrogatories. However, because "the defendants ha[d] expressed their willingness to pay the costs of whatever operations [we]re necessary to manufacture a computer-readable tape . . . the problem of allocating the burden of discovery expense, which might be significant . . . [wa]s nonexistent [t]here." 494 F. Supp. at 1262.

In Anti-Monopoly, Inc. v. Hasbro, Inc., 94 Civ. 2120 (LMM) (AJP), 1995 WL 649934, at *1 (S.D.N.Y. Nov. 3, 1995), Magistrate Judge Peck stated that "[t]he law is clear that data in computerized form is discoverable even if paper `hard copies' of the information have been produced." However, the court required the requesting party to pay the cost ($6,680 to $7,680) of "special programming" to retrieve the data. Anti-Monopoly, Inc. v. Hasbro, Inc., 94 Civ. 2120 (LMM) (AJP), 1996 WL 22976, at *2 (S.D.N.Y. Jan. 23, 1996).

In Sattar v. Motorola, Inc., 138 F.3d 1164, 1171 (7th Cir. 1998), Motorola produced e-mails in the form of 4-inch tapes, which were inaccessible to Sattar because he lacked the equipment and software to read them. The court of appeals approved the district court's exercise of discretion in not directing Motorola to produce at its expense 210,000 pages of hard copy of e-mails, but instead to download at its expense the data from the tapes to a conventional computer disk, to loan Sattar a copy of the necessary software or to offer Sattar on-site access to its computer system.

On the other hand, in Williams v. Owens-Illinois, Inc., 665 F.2d 918, 932-33 (9th Cir. 1982), the court of appeals upheld the district court's exercise of discretion in allowing the defendant to provide to plaintiffs in an employment discrimination case only hard-copy wage cards and not computer tapes with that data, while also requiring the defendant to process whatever computer runs the plaintiffs requested.

The Section concludes that, under the existing rules, the producing party in the first instance may choose the reasonably usable form of electronic information to be produced, but that the balancing test of Rule 26(b)(2)(iii) should be applied to determine whether a second copy must be produced, and, if so, at whose expense. This encourages the parties to try to resolve the form of the production of the electronic information in light of the needs of the case.

CONCLUSION

Perhaps surprisingly, the Federal Rules of Civil Procedure governing discovery currently do sufficiently cover the discovery and disclosure of electronic information. If the information is within the scope of discovery under Rule 26(b)(1) it should be produced, although courts should exercise their discretion under Rules 26(b)(1) and 26(b)(2)(iii) to allocate the burdens and costs of producing any electronic information in light of the benefits and importance of that information to the case. While these standards provide no bright line applicable to all electronic information, they do provide the greatest likelihood that substantial justice will prevail in connection with the production of electronic information during discovery in litigation.

New York State Bar Association February 2001
Commercial and Federal Litigation Section
Committee on Federal Procedure

Gregory K. Arenson, Chair*
Robert Edward Bartkus
John P. Coll, Jr.
Thomas S. D'Antonio
Thomas F. Fleming*
Neil P. Forrest
Margaret J. Gillis
Robert J. Jossen
Martin E. Karlinsky
Thomas J. Kavaler
Mitchell A. Lowenthal
Thomas McGanney*
Charles E. Miller
James F. Parver
Allan M. Pepper
Thomas J. Quigley
Stephen T. Roberts
David H. Wilder
Scott H. Wyner

* Co-authors of this report.