04-CV-001
Request to Testify
1/12 San Francisco

"Greg McCurdy (LCA)"
<gmccurdy@microsoft.com>
08/25/2004 04:33 PM

To <Peter_McCabe@ao.uscourts.gov>
cc
Subject Microsoft Comments on the e-Discovery Rules Change Proposals

Dear Mr. McCabe,

I just wanted to let you know that Microsoft Corporation would like to submit written comments and be heard at either the San Francisco or Washington, DC hearings that the Committee will be holding in January or February. I understand that you are maintaining the list of speakers for those hearings. I assume that there is still space available on the list of speakers. By when do you need a firm commitment from us as to which of the two events we will attend? How much time will be allocated per speaker?

Thanks for you advice on this.

Best,

Greg McCurdy
Senior Attorney, Litigation
Microsoft Corporation
425-705-2724 (office)
206-355-4464 (cell)
425-708-2155 (fax)

04-CV-001 Microsoft Corp. Testimony (Greg McCurdy) 1/12/05 San Francisco CA

"Greg McCurdy (LCA)" <gmccurdy@microsoft.com>

12/17/2004 12:30 AM

To <Peter_McCabe@ao.uscourts.gov>

cc <Lee_Rosenthal@txs.uscourts.gov>, "Tom Burt (LCA)" <tburt@microsoft.com>

Subject RE: Microsoft Comments on E-Discovery & Civil Rules

Dear Mr. McCabe:

Attached is a copy of our comments on company letter head. My apologies for the delay.

I am looking forward to seeing you at the hearings on January 12th. I have two logistical questions. Has it been determined yet how much time we will have for our presentation? Will we be able to project Power Point slides?

Thanks,

Greg McCurdy

-----Original Message-----
From: Greg McCurdy (LCA)
Sent: Wednesday, December 15, 2004 5:15 PM
To: 'Lee_Rosenthal@txs.uscourts.gov'; 'peter_mccabe@ao.uscourts.gov'
Cc: Tom Burt (LCA)
Subject: Microsoft Comments on E-Discovery & Civil Rules

Dear Judge Rosenthal and Mr. McCabe,

On behalf of Microsoft Corporation I would like to thank you again for the opportunity to participate in the public comment proceedings of the Advisory Committee on Civil Rules. Attached are Microsoft's comments on some of the proposals, signed by Tom Burt, Vice President and Deputy General Counsel in charge of litigation. Microsoft offers these both as a litigant and as a technology company some of whose products are intricately involved in electronic discovery issues.

Please do not hesitate to call on us if can assist with any technical or legal issues. We take a great interest in the rules revision process and believe that this is a unique area where law and technology intersect. Technology has created great benefits and efficiencies along with some very significant difficulties and costs for lawyers and litigants. We believe that a sound analysis of technology is an important foundation for the development of the law in this area and any solutions for the challenges of the information age.

Best regards,

Greg McCurdy
Senior Attorney, Litigation
Microsoft Corporation
425-705-2724 (office)
206-355-4464 (cell)
425-936-7329 (fax)

Microsoft FRPC Revision Comments on Letter Head.pdf

Microsoft Corporation
One Microsoft Way
Redmond, WA 98052-6399

Tel 425 882 8080
Fax 425 936 7329
http://www.microsoft.com/

04-CV-001
Microsoft Corp. Testimony
(Greg McCurdy)

Microsoft

December 16, 2004

Peter G. McCabe
Secretary
Committee on Rules of Practice and Procedure
Judicial Conference of the United States
Thurgood Marshall Federal Judiciary Building
Washington, DC 20544

RE: Proposed Amendments to the Federal Rules of Civil Procedure: Electronic Discovery

Dear Mr. McCabe:

Microsoft Corporation applauds the Committee's ongoing efforts to update the Federal Rules of Civil Procedure to address the serious problems posed by the discovery of electronically stored information. We believe that changes to the Civil Rules are necessary in order to provide needed guidance to both litigants and the courts to address the problems surrounding electronic discovery. Advances in computer software and hardware have greatly increased our ability to create, duplicate, disseminate, and store information. While these advances in technology have generally been a boon for productivity, they have also resulted in the exponential growth of electronically stored documents and information that may be relevant to litigation. This explosive growth has changed the nature of discovery in a way that needs to be addressed directly in the rules of procedure, most of which were drafted to address the much simpler world of paper documents.

The Impact of Technology on Discovery: The Federal Rules were designed to deal with a world where typewriters and carbon paper were the primary means to create and duplicate documents. The advent of the photocopy machine stretched the Rules and drove up the costs of discovery exponentially. The rise of computers with word processing, email, and other software programs has increased the costs and problems of discovery by further orders of magnitude. It is high time for the Federal Rules to catch up with this reality and adapt to the very different nature and quantity of electronically stored information that is the focus of so much expensive litigation and discovery. Here are two examples of the huge volumes of data and the associated costs incurred by any large IT network:

Email Volumes are Staggering.

The amount of email that Microsoft has received in 2004 is roughly double what it received in 2003. Most of that increase is due to "spam", *i.e.*, commercial junk email. Microsoft's IT network now receives 250-300 million messages a month from the outside. Internal messages sent and received average 60-90 million a month. Automatic systems are indispensable to deal with this flood. The

automatic filters in the "gateway/firewall" that protect our network now delete 85-90% of all incoming internet e-mail as spam. That still leaves 30-45 million legitimate emails per month that are delivered to employees. While email is by far the most commonly used data type, there are many others that substantially add to the volume of electronically stored information. The efforts of individual employees are not usually up to the challenge of managing this torrent of data. As a result, most large institutions must at least consider using automatic processes to dispose of data that is no longer needed for any business or legal purpose. The proposed amendments are a useful first step toward addressing this reality by balancing the need for the preservation and production of evidence relevant to litigations with the need of large organizations to manage their IT systems in a rational and efficient manner.

Backup Tape Operations are very Complex and Expensive.

Microsoft's IT network includes about 90,000 e-mailboxes for individual employees, vendors, and others. The email generated by all these users is stored and routed on over Exchange email 100 servers. Those servers are backed up daily by about 15 tape drives in our main data center plus others around the world. The main data center in Redmond generates between 150 and 200 backup tapes per day which hold about 15 terabytes of data. The daily tapes are recycled every 14 days. The cost of 14 days worth of tapes is about $65,000, not including storage and management costs. If we had to stop recycling backup tapes, the additional tape acquisition cost alone would be about $1.7 million per year.

Technical Background for our Comments

All of Microsoft's employees, and many of our customers, use our productivity software products such as Outlook, Word, Excel and Exchange to operate in a largely paperless world, making electronic discovery the central focus in our litigation. As both the creator of software that allows consumers and office workers to work and communicate in the electronic world, and a litigant in many major e-discovery cases, Microsoft would like to share its perspective regarding the challenges and problems that companies face in the discovery of electronically stored information. As background for our comments, it would be helpful to explain the evolution of computing in the last 60 years, how some of our products function and how they are deployed in the "client-server architecture" of most corporate and institutional networks, and, indeed, the Internet itself. To illustrate this discussion, we have attached a diagram of the major components of a typical corporate network. The diagram may seem complex but it vastly simplifies the complex network details of even a company with only a few hundred employees.

From Mainframe Computers to the PC Revolution. The first computers were developed in the World War II era. They were gigantic machines primarily used to process data. They were so big and expensive that they took up whole rooms, and only large institutions such as government agencies, big companies, and universities could afford them. This era of mainframe computers was dominated by companies like IBM and had a limited impact on discovery and the Federal Rules of Civil Procedure. The idea that a computer could be so affordable and small that it could fit on every desk and in every home was considered a joke. During the 1980s, the PC revolution turned that joke into a reality, greatly increasing productivity for information workers. Companies such as Apple, Compaq, Dell, IBM, Intel, and Microsoft developed hardware, micro processors, and software that led to both a vast increase in the computing power and data storage capacity of PCs, and a steady decrease in their cost and size. This stretch of incredibly rapid technical innovation made it easier and easier to create and store electronic documents, which in turn hugely increased the volumes of documents available for discovery in

litigation. Despite the rise of the PC very large and powerful mainframe computers dating back decades are still used to support legacy applications software and critical corporate data systems in many companies. The persistence of these older systems is often the source of electronically stored information that is "not reasonably accessible."

The Rise of Email, Client-Server Networks, and the Internet. In the 1980's distribution of electronic documents was still a bit slow and cumbersome. People still printed many documents on paper for distribution or exchanged the now largely obsolete "floppy disks." The 1990's saw the rise of corporate intranets, wide-area networks, and the Internet – which collectively served to connect many previously unconnected PCs. This connectivity revolution was made possible both by the development of communications protocols like TCP-IP, HTTP, or FTP that carried data across the wires, and by the rise of the client-server architecture as the most common design for computer networks.

As illustrated in the attached diagram, the general concept of a client-server network is simple, but its many implementations can be extremely complex. "Clients," in IT jargon, are most commonly PCs including desktops, laptops and tablet PCs, but can also include non-PC devices that individuals use – such as Blackberries, Pocket PCs, Palm Pilots, smart phones, and other PDAs (personal digital assistants). These clients can all be used to create and store data. Thanks to networks and the internet, they can also be used to send and receive data. "Servers" typically serve networks of many users and their client devices. As the name implies, they serve up data (*i.e.,* store, process, and transmit) to clients in the form of word processing documents, database information, emails, etc.

There are many different kinds of servers in terms of their hardware, and the operating systems and applications software they run. These servers are designed and deployed to perform many different functions and are sometimes used for specific tasks such as serving up web pages, emails, and other file and data types. Large organizations such as banks, airlines, government agencies, retailers, or software companies have very different needs and thus have very different networks running many different types of software and hardware on both their clients and servers. Mainframe computers also act as servers in networks. These networks have become extremely complex – often connecting thousands of servers and tens of thousands of clients. The composition of each of these networks is also constantly changing as the hardware and software are upgraded and adapted to evolving business needs. Over the course of a decade or less software and hardware products rapidly age until become out of date (a/k/a, legacy products) and eventually obsolete in the sense that their manufacturers reduce or stop providing support or updates. As these machines age and are retired, the data they support becomes increasingly difficult to access unless they are migrated to new and actively used systems.

Microsoft's Office Productivity Products. Microsoft is best known for its Windows operating system products that run on most PCs and a significant minority of servers. The popularity of Windows is a crucial factor in the rapid growth of PCs and electronic data over the last decade, but it is on some of Microsoft's other products that much of this data is created, stored, copied and distributed at lightning speed to users around the world. Two such Microsoft products are "Outlook" and "Exchange." Outlook is the email client software that is a part of Microsoft "Office" – a suite of productivity applications used by many PC users. In addition to Outlook, Office includes other popular programs such as Word, Excel, and PowerPoint. It is in Outlook on the PC that email is created. It is also where the most important email is often stored by users—because that is the email they are actively working on. "Exchange" is a server application that Microsoft and many of its customers install on dedicated email server computers

to distribute and store email generated by the Outlook client software. Outlook folders and their contents are "mirrored" onto the Exchange servers – meaning that the Exchange servers not only distribute the emails to the Outlook client mailboxes, but they also keep an additional "mirror image" copy of these live mailboxes as a precaution against loss of data on the PC. In addition to this system of email servers, many companies also use general purpose file servers for storing original or extra copies of many other kinds of documents that were originally created, and often still reside, on PCs.

This brief overview of the development of computers and how they are deployed provides essential background for understanding the implications of the proposed rules changes regarding early discussions of the architecture of IT systems, the identification of inaccessible data, just what makes data inaccessible, and why the automatic functioning of IT systems cannot be disrupted without very significant costs. It also explains the overriding reality of discovery in the electronic era: **there is no lack of documents and data. On the contrary, one of the biggest challenges faced by both requesting and producing parties is the ever growing mountain of data.** Search technology has helped us master some of this growth, but costs continue to spiral upwards because technology is expensive and, in the end, a human decision maker needs to look at anything that has significance and make the final call on whether it is relevant as evidence, or whether it is privileged in some way.

With that background, Microsoft submits the following comments regarding the proposed amendments to the Federal Rules of Civil Procedure relating to the discovery of electronically stored information. We have organized our comments to address, in turn, the five areas identified by the Subcommittee and its Advisory Committee on Civil Rules: (1) discovery of electronically stored information that is not reasonably accessible; (2) a limit on sanctions under Rule 37 for the loss of electronically stored information as a result of the routine operation of computer systems; (3) the application of Rules 33 and 34 to electronically stored information; (4) early attention to issues relating to electronic discovery; and (5) the assertion of privilege after production.

A. Discovery of electronic information that is not reasonably accessible

The proposed amendment to Rule 26(b)(2) is intended to clarify the obligations of a responding party to provide discovery of electronically stored information that is not reasonably accessible. The Committee has expressed particular interest in comment on whether further explanation of the term "reasonably accessible" in the Note would be helpful and what it should include. The Committee has also asked for comment on whether the proposed Rule 26(b)(2) and Note give sufficient guidance on determining the proper limits of electronic discovery and on appropriate terms and conditions, including allocating the costs of such discovery.

Rule 26. General Provisions Governing Discovery; Duty of Disclosure

* * * * *

(b) Discovery Scope and Limits. Unless otherwise limited by order of the court in accordance with these rules, the scope of discovery is as follows:

* * * * *

(2) Limitations. By order, the court may alter the limits in these rules on the number of depositions and interrogatories or the length of depositions under Rule 30. By order or

> local rule, the court may also limit the number of requests under Rule 36. The frequency or extent of use of the discovery methods otherwise permitted under these rules and by any local rule shall be limited by the court if it determines that: (i) the discovery sought is unreasonably cumulative or duplicative, or is obtainable from some other source that is more convenient, less burdensome, or less expensive; (ii) the party seeking discovery has had ample opportunity by discovery in the action to obtain the information sought; or (iii) the burden or expense of the proposed discovery outweighs its likely benefit, taking into account the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues. The court may act upon its own initiative after reasonable notice or pursuant to a motion under Rule 26(c). A party need not provide discovery of electronically stored information that the party identifies as not reasonably accessible. On motion by the requesting party, the responding party must show that the information is not reasonably accessible. If that showing is made, the court may order discovery of the information for good cause and may specify terms and conditions for such discovery.

Microsoft strongly supports the implementation of a "two-tier" approach to the discovery of electronically stored information, but we have a few reservations about the proposed language. The need for a distinction in the rules between accessible and inaccessible documents is clear. The evolving technical characteristics of electronic documents and the various stages in the e-document life cycle requires the presumptive exclusion of inaccessible electronically stored information from discovery absent a court order. First we will give some concrete examples of why this amendment makes sense and then will offer the a few critical comments regarding the wording of the proposed amendment to Rule 26(b)(2).

The Dumpster, the Landfill, and the Shredder: Paper Concepts in the Electronic World

Let us take deleted or fragmented data as an example, and consider how equivalent paper documents would be treated. In the old paper world people did not delete documents, they tossed them into the trash bin, the recycling basket, or into the shredder. There was often a time lag between the document custodian's decision to toss a document and when it ended up in the landfill, the recycling plant, the incinerator or in the shredding machine. Depending on how frequently the trash was removed, the users had a short grace period to change their minds and retrieve the document from the trash. After that, it may have been possible to retrieve documents from the dumpster or from the landfill, but that would have been extremely expensive and difficult. Consequently, in discovery discarded paper documents were generally considered outside the scope of discovery because they were relatively inaccessible. In fact, the notion that discovery obligations in the paper world required retrieval of documents from the dumpster or landfill was very unusual.

The proposed amendment on inaccessible data applies this familiar principle to the electronic equivalents of trash bins, dumpsters, landfills and shredders. There is a sliding scale of accessibility. Data becomes increasingly inaccessible with the passage of time and depending on the methods of their creation, storage, and disposal. For example, if a user of Microsoft's Outlook email program deletes an email, it goes into the Deleted Items "wastebasket" folder. While in this folder, the email remains easily and indefinitely accessible by the user until, by user decision or by automatic process, it is moved from the Deleted Items folder to a "Dumpster" type folder. The user can usually still retrieve the email from the Dumpster for a few days with the help of an IT professional before it is deleted from the Dumpster by the Microsoft Exchange email server. The deleted email remains in the Exchange Dumpster for three days. After that, the email can no longer be retrieved directly from the Exchange server by anyone. But that doesn't mean that the deleted email no longer exists anywhere.

Once deleted, a number of things happen to data depending on the hardware, the software, and how they are configured. The deleted email may be on a disaster recovery backup tape, at least for some period of time. Most companies would never restore an Exchange server backup tape just to locate a deleted email – the cost-benefit analysis would rarely, if ever, provide a business justification to do so. Additionally, fragments of deleted data remain on a hard drive until they are overwritten. But they can only be resurrected, if at all, by highly trained specialists using forensic techniques—typically at great cost. Both are the modern equivalent of sending detectives to the landfill or the shredder to recover and paste together whatever scraps of paper they can find. In the world of paper discovery, such efforts would be considered heroic or extraordinary and would not be required absent some very good reason. The proposed Rule brings that concept forward into the electronic age.

The Burden of Identifying Information that is Not Reasonably Accessible

First, Microsoft believes the proposed wording "A party need not provide discovery of electronically stored information that *the party identifies* as not reasonably accessible" should be revised to read: **"Electronically stored information *that is not* reasonably accessible need not be produced except on a showing of good cause."** We are concerned that the currently proposed wording may be interpreted to create a new and potentially very burdensome obligation on the part of the responding party. Requiring a party to proactively identify the information that is inaccessible is unnecessary and would either be very burdensome, or so general as to be meaningless depending on the specificity required. The Note does not clarify sufficiently the scope of the obligation to identify, and seems likely to lead to contentious motions practice. As currently drafted, the requirement that a party affirmatively identify the information which is not reasonably accessible would likely result in responding parties developing a boilerplate generic listing of categories of types of information they believe are not reasonably accessible such as deleted data, fragmented data, encrypted data, legacy data, or backup tapes. In order to preserve the protection provided under this Rule, a party is likely to be overbroad, listing every conceivable category of inaccessible information, without investigating whether or not the inaccessible sources are implicated in the instant case, and without distinguishing between those inaccessible sources it knows exist and are likely to have responsive information, and those sources it has included simply out of caution. The result is not useful to either party.

If the identification requirement is not to devolve into recitation of the type of boilerplate already so common in discovery papers, the responding party would need to undertake a significant investigation of its inaccessible data sources. This, however, would be very burdensome because

inaccessible data is, by definition, hard to access and is not used or accessed in the ordinary course of business. In the electronic world, as in the paper world, most users do not keep detailed records of what they toss out or delete. That means that to do more than telling the requestor that responsive documents might be found in the dumpster or in the landfill the responding party would have to search the dumpster or landfill, to see whether there are any inaccessible documents that might be responsive.

In addition to the inherent difficulties of "identifying" information that is not reasonably accessible, it is also important to keep in mind how complex and diverse even smaller or medium size corporate networks are. The attached diagram of a client-server network illustrates this even at a small scale. Then imagine that complexity for a large company like Microsoft. We have about 60,000 employees, plus vendors and others who use 90,000 email accounts. These people operate from 441 worldwide locations. Of those, 263 are outside of the U.S. (in 93 countries), and 178 locations are in the U.S. Many of our employees have multiple PC and non-PC clients that attach to the network and store electronic information. Any detailed investigation of inaccessible sources of data would be a gargantuan task.

It has been suggested that there must be some middle ground between the likely positions of most responding parties who will want to identify only very broad categories of inaccessible data and requesting parties who will move to compel a very detailed identification that will require significant investigation by the responding party. Theoretically, there will always be a mid-point between these opposing views. In practice, it seems very unlikely that litigants will reach agreement on such middle ground without significant argument and motion practice. Since one of the purposes of the rules revisions is to give clear guidance and decrease the need for motion practice, this would be counter productive. Therefore, we strongly urge the Committee to delete the identification requirement from the amendment to Rule 26(b)(2). This would be in keeping with the whole point of putting inaccessible data into a second tier that is presumptively not discoverable. If data are truly not reasonably accessible and not used for ordinary business purposes then there should be no requirement to "identify" them, certainly not on a level of detail that would require an investigation.

An Alternative Way to "Identify" Inaccessible Data

If the Committee chooses to retain the currently proposed wording, requiring a party to identify the information that it is neither reviewing nor producing on the grounds that it is not reasonably accessible, additional clarification should be added to the Note. In addition to the examples given in the Note at page 13, which are focused on certain types of inaccessible information of which the responding party is specifically aware, Microsoft believes there should also be an option that allows the responding party to identify electronic information that is not reasonably accessible in the negative. For example, instead of being required to list and/or describe each and every possible source of inaccessible information, many of which may be unknown to the party, the party should be allowed the option of affirmatively describing the various known reasonably accessible sources from which it intends to review and produce electronically stored information, and state that information that is not located in one of these reasonably accessible locations is considered not reasonably accessible for purposes of this Rule.

Suggestions for Further Explanation of the Term "Reasonably Accessible"

Microsoft also believes that further explanation of the term "reasonably accessible" in the Note would be helpful. We believe an appropriate description of reasonably accessible electronically stored information is information "in active use for the day-to-day operation" of the party's business. How the company actually uses the data in its normal course of business should be the primary factor in determining whether electronically stored information is accessible or inaccessible. To the extent that effort and expense are looked to in determining whether information should be considered inaccessible, we believe that there must be clear language in the Note clarifying that this applies to the effort and expense involved in the entire process of providing discovery of the information – including locating, retrieving, restoring to readable or searchable form, reviewing, and producing. The cost of retrieval alone should not be the determining factor since the other costs can often be very significant.

Disaster Recovery Backup Tapes. Microsoft supports specific mention of backup tapes as a type of data that should generally be considered to be inaccessible and available only by court order upon a showing of good cause. As mentioned in the proposed Note at p. 14, the good-cause analysis balances the requesting party's need for the information and the burden on the responding party. Towards this end, we think it is important to point out that the burden on the responding party who is asked to produce information from backup tapes will tend to be greater the older the backup tapes in question.

There are many ways in which the passage of time makes data sources, like backup tapes, increasingly inaccessible. For last week's backup tapes records may still exist, such as indices or catalogues of which documents were in what files and when and how those files were backed up, so that a rough guess of file location on a long backup tape may be made. As more time passes, the already unlikely scenario described above becomes less and less probable. The result is a considerable increase in the amount of time and effort needed to locate specific information. The party has no way to quickly zero in on specific documents or pieces of information – because backup tapes are not typically used to retrieve specific documents, they are not always created with an accompanying index, catalog, or other means to track the massive amounts of data stored on them. Even if such an index or catalogue is created at the time the tape is created, that index or catalogue typically resides in the active memory of a live server rather than on the tape itself. That can take up a lot of expensive live storage space and will therefore often be overwritten once there is no more business need to keep it. Additional problems arise for parties that have not always recycled backup tapes promptly when their usefulness for disaster recovery has passed (usually after a week or two), and/or who have stored these tapes off site. After a few years it is likely that the IT systems in use will have changed, and the software and hardware used to create the tapes may no longer exist. If an index or catalogue was created, it may have been recycled to free up memory space on the servers. Typically, the old hardware and software are eventually junked, or the IT staff that operated them change jobs or forget the complex details of their operations as they focus on new systems.

Deleted & Fragmented Data. Other forms of data similarly become increasingly inaccessible with the passage of time. Deleted data becomes harder to restore depending on how it was deleted, how many times it was overwritten, or how many times a disk was de-fragmented or wiped. A word of explanation: when computers store data files such as email on a hard disk, it is often done in "fragments" in order to squeeze as much data into the available space as possible. When a user "deletes" a file it is not actually removed from the hard drive. It is instead just marked "deleted" so that

it can be overwritten by new data if the storage space is needed. If the space taken up by the "deleted" fragments is not reused, then those fragments can persist for a considerable time until the disk is "de-fragmented" to improve memory retrieval speeds, or wiped clean to make room for new data. Until that happens—and sometimes even after such procedures—forensic specialists can, for a substantial price, retrieve such fragmented data and sometimes reassemble it into a useful form. Thus the passage of time and certain technical events determine how "accessible" deleted and fragmented data is.

Encrypted Data. The same is true of encrypted data which becomes more difficult to decrypt after the expiration, loss, or destruction of the relevant "key." Businesses are increasingly turning to encryption technology to preserve their confidential information. Microsoft and other software vendors provide products that make sophisticated and hard to crack encryption easy to apply to documents. The authors of such encrypted documents can determine who has access to them and for what time period. Encrypted files can persist on the recipient's PC long after their access keys have expired. At that point, they are inaccessible to the ordinary user – and with the passage of time they will become increasingly inaccessible even to the author of the document, or to technical decryption specialists.

The result is that within specific categories of inaccessible data (e.g., backup tapes, deleted data or encrypted data), there may be varying degrees of inaccessibility. So, while methods such as sampling can be helpful, they may not be entirely reliable and may still be quite costly. A court that looks only at the cost incurred in restoring and locating data from a recent backup tape may severely underestimate the burden and expense of trying to do the same for an older tape. Particular attention must be paid to the specifics of each inaccessible item being requested, and the devil is always in the details.

Treatment of Information a Party has "Actually Accessed"

In addition, Microsoft suggests either the removal or clarification of the language in the Note at page 13, which states that "if the party has actually accessed the requested information, it may not rely on this rule as an excuse from providing discovery, even if it incurred substantial expense in accessing the information." As written, this qualification lacks the flexibility to be applied fairly across all cases. Where information has been accessed and subsequently retained in an accessible form, this guideline may be reasonable. Similarly, a party that regularly accesses specific "inaccessible" information for their own purposes in the litigation should not be able to use the rule to avoid providing discovery of such information to the requesting party.

There are many situations, however, that would technically fall under the vaguely worded exception, with unfair result. For example, consider the situation where a party has made a general request for information from backup tapes. The fact that a backup tape has been accessed by the producing party at some point in its existence to restore data on a server that failed (e.g., for disaster recovery purposes) should not deny the party protection from having to produce relevant documents from backup tapes in general. Even being forced to provide discovery of the specific backup tape in question will often be unwarranted. For example, a specific backup tape that was accessed in the past may clearly be no more accessible than any other backup tape by the time of the request in litigation that arises weeks, months, or years later. There is a major difference between using a backup tape for disaster recovery purposes to restore an entire server, and looking for a specific document. Similarly, as written, this exception would apply to a situation where one court has determined that good cause exists to require a party to search certain backup tapes for a specific case. The result should not be that the party then forever loses the right to rely on this rule generally, or as to those backup tapes.

In situations where a backup tape has, at the time of the request, been accessed at some point in the past for matters unrelated to the litigation, but has not been "in active use for the day-to-day operation" of the party's business, the party should still be afforded the presumption that this information is not discoverable in the litigation depending on the specifics of how difficult and expensive it would be to access and search the data for litigation purposes. The reason behind this is that, particularly with certain types of electronically stored information, the extra steps of reviewing and producing information still present a substantial and often unwarranted additional burden upon the producing party. Every aspect of locating, retrieving, restoring to readable or searchable form, reviewing, and producing electronically stored information adds a burden of time and expense that needs to be considered in determining whether the information is reasonably accessible for purposes of this Rule.

Presumption of Cost Shifting when Production of Inaccessible Information is Ordered

Finally, Microsoft suggests the proposed amendment to Rule 26(b)(2) should impose a presumption of cost shifting for discovery of electronically stored information that is not reasonably accessible – a presumption that can be overcome by a clear and convincing demonstration of relevance and need. The presumption that the costs of retrieving, restoring, reviewing, and producing inaccessible information will be shifted to the requesting party would serve as an effective deterrent against overbroad, marginally relevant discovery, while the ability to overcome this presumption with a clear and convincing showing of relevance and need will ensure that the requesting party is able to obtain this information at the producing party's expense in those situations where it is truly warranted. Because of the substantial burden of reviewing and producing relevant information even from just the reasonably accessible electronically stored information, it is important that the entire cost of accessing, reviewing, and producing inaccessible information be presumed allocable to the requesting party. Otherwise, there is little incentive for the requesting party to be focused and/or reasonable in its requests, and the temptation will be great to impose significant discovery costs on the adverse party to increase financial pressure for a settlement.

B. Safe Harbor: Limit on Sanctions under Rule 37

The Committee has proposed amendment to Rule 37 which would provide a narrow "safe harbor" to a party that fails to preserve certain electronically stored information, under narrow circumstances. The Committee has expressed particular interest in receiving comments on whether the standard that makes a party ineligible for a safe harbor should be negligence, or a greater level of culpability such as willfulness or recklessness, in failing to prevent the loss of electronically stored information as a result of the routine operation of a computer system. The Committee has also expressed particular interest in comment on whether the proposed Rule and Note adequately and accurately describe the kind of automatic computer operations, such as recycling and overwriting, that should be covered by a safe harbor.

Rule 37. Failure to Make Disclosure or Cooperate in Discovery; Sanctions

* * * * *

(f) Electronically Stored Information. Unless a party violated an order in the action requiring it to preserve electronically stored information, a court may not impose sanctions under these rules on the party for failing to provide such information if:

(1) the party took reasonable steps to preserve the information after it knew or should have known the information was discoverable in the action; and

(2) the failure resulted from loss of the information because of the routine operation of the party's electronic information system.

The Need for A Safe Harbor

Ken Withers of the Federal Judicial Center has captured the realities of IT systems well in a recent paper.[1] He wrote that: "In the information age, information technology is synonymous with the production or business process. Any attempt to isolate, segregate, and preserve a certain set of potentially responsive information for pending or anticipated discovery is likely to throw a wrench into the entire information-processing assembly line. These systems are designed to be dynamic, ever-changing, self-updating, and responsive to immediate business demands, not to collect and preserve historical information. Among the clearest examples of this are enterprise-wide databases, which are constantly being overwritten and updated with current information; and backup systems, which are designed for disaster recovery purposes and routinely recycle outdated storage media."[2] This is very true for Microsoft and for many of our corporate and institutional customers. The safe harbor is a reasonable effort to recognize this reality and permit the automatic processes that are at the heart of large modern IT systems to keep operating without threat of sanction so long as the party takes reasonable steps to comply with its discovery obligations.

Defining the Routine Operation of an Electronic Information System

The Committee has stated that it intends that the phrase, "the routine operation of the party's electronic information system," identify circumstances in which automatic computer functions that are generally applied result in the loss of information. The Committee has expressed its concern that there be adequate guidance as to the aspects of an electronic information system that are within the proposed rule, without being limited to existing technology. Toward these ends, Microsoft favors a clear statement that this phrase should generally be understood to include a party's good faith operation of their disaster recovery systems – including the regular daily, weekly, or monthly rotations a party may follow to recycle the tapes that are commonly used to back up servers in corporate networks.

---

[1] Ken Withers, Two Tiers and a Safe Harbor: Federal Rulemakers Grapple with E-Discovery, August 23, 2004.
[2] *Id.* At 35.

It is also important to recognize that the routine operations of corporate IT systems[3] that warrant protection under this Rule extend far beyond the regular recycling of backup tapes. Today, many Fortune 500 companies, including Microsoft, struggle to find innovative ways to better manage the huge volumes of email generated every day. Three automatic systems that are essential to controlling the volume of email, and, in turn, enhancing employee productivity and preventing corporate IT systems from drowning in data, need to be mentioned: backup tape recycling, spam filters, and auto-deletion of unneeded[4] email. Enough has been said elsewhere about backup tapes so we will focus on spam and auto-delete systems.

Spam Filters. We would encourage the Committee to add spam filtering to the Note as another example of a routine operation protected under this Rule. As mentioned earlier, Microsoft receives an average of about 250-300 million external messages a month. This amount is nearly double the volume received by Microsoft from external sources last year, and that increase is mostly due to spam. "Spam" can be defined as electronic junk mail – an unsolicited, often commercial message transmitted through the Internet as a mass mailing to a large number of recipients. In addition to external messages, internal messages sent and received at Microsoft average 60-90 million a month. Faced with these huge volumes, it is crucial that Microsoft employees and its IT systems continually delete unneeded mail. Automatic systems are indispensable to deal with the flood caused by the spam epidemic. Of the 250-300 million external messages received by Microsoft in an average month, the automatic filters that protect our network delete 85-90% as spam. Only 10-15% of the external email received is identified by the system as "non-spam" and is actually delivered to our employees. The messages blocked as spam are never used by the company for any business purpose – they aren't used at all and in some sense were never "received" by the company.

It's unlikely that in the huge volumes of deleted spam there is much that is relevant to litigation, but with up to 270 million automatically deleted messages per month it is certainly possible, particularly given the complexity of filtering spam. Microsoft deploys third party commercial spam filters, but it is also in the business of developing software programs that provide spam filtering services. This is a very difficult task for a number of reasons. First, the fight against spam involves an arms race between the filtering software and the spammers who try to circumvent it, primarily by masquerading as legitimate email. This ever-changing landscape makes it hard to define what spam should be eliminated, and then to successfully target and filter out this material without also deleting wanted mail. Second, even if spam were not constantly morphing to evade detection, there is no one perfect and objective definition of spam. Most people and companies do not want to receive unsolicited advertisements for Viagra or pornography, or mail about fraudulent investment schemes in Nigeria. Others may want to receive ads for pharmaceuticals or news on investment opportunities in developing countries. The result is that spam filters are always going to be both under and over-inclusive in the hunt for spam – with the accompanying risk that, where over-inclusive, they could cause the loss of some material relevant to litigation.

Spam filters may also result in the deletion of potentially relevant information even if they are not over-inclusive. Consider the example of pornography. Most companies quite rightly try to block

---

[3] What applies to corporate IT systems applies equally to those of government agencies and any other institution or group large enough to have a network of PCs connected to server computers.

[4] That is, email not required to be retained due to business needs or legal requirements.

pornography from coming into their networks. However, it is possible to imagine a sexual harassment case alleging a hostile environment that involves an employee receiving pornographic images attached to Internet email and then distributing them internally. In such a case, if many of the emails with relevant images are captured and blocked by the company's spam filters, one could imagine that the auto-deleted pornography could be considered responsive to document requests in the litigation. Nonetheless, this automatic deletion of unwanted email for a valid business purpose should not give rise to liability for spoliation of evidence without a showing of intentional or reckless disregard of a preservation obligation.

Automatic Deletion of Unneeded Email. Microsoft and most companies that we are aware of do not provide unlimited email and e-document storage space on their corporate servers because to do so would be very expensive. One means of discouraging employees from indiscriminately hording all their email is to set size limits on their email inboxes. This forces employees to do two things when they hit their storage space limits to make space for new email: (1) file email needed for business or legal purposes in a longer term repository, and (2) delete unneeded email. We are aware that many of our customers, partners, and competitors have also implemented a technical solution to help their employees cope with the huge volumes of unneeded mail that pile up, and to free up storage space. The IT systems of these companies automatically delete unfiled email after some reasonable time period like 30 or 60 days—under the assumption that if the employee did not file the email away, it must not be important enough to retain. Once a party is on notice of a potential litigation it must, of course, take reasonable steps to preserve all potentially relevant materials. If there is no threat of litigation, or the party has taken reasonable steps to preserve what needs to be preserved, then there should be a safe harbor for the automatic deletion of email due to limits on the time or space for temporary email storage.

A practical illustration shows why this is necessary. A company learns of a potential lawsuit in a rather vague and brief demand letter. As a precaution it instructs 20 employees who deal with the subject matter of the anticipated lawsuit to preserve all relevant documents. Three months later the complaint is served with more specific allegations so the defendant's lawyers notify another 10 employees that they must retain documents because their work touches on the more detailed allegations. Six months later plaintiff files an amended complaint expanding and further detailing its allegation. At that point the defendant's lawyers add another 5 employees to the list that are instructed to preserve their relevant documents. In the nine months between the initial notice and the service of the amended complaint, the email that some 15 employees did not need for any business or legal purpose was automatically deleted from their inboxes because they did not designate it for longer term retention. In that case, the company took all reasonable steps based on information available at three different times, to preserve documents that it could have anticipated to be relevant to the litigation. It should enjoy the benefit of the safe harbor for the routine operation of IT systems.

What Does "Routine Operation" of IT Systems Mean?

This phrase is somewhat open-ended and that is generally a good approach. Nonetheless, it would be worth mentioning in the Notes that routine operation does not only mean the routine operations of computer software and hardware, but also the routine operations of the people who administer them. Some systems like spam filters run largely without human intervention once they are set up. Others like backup systems require technicians to load backup tapes into tape drives, move the full tapes to storage, and then load them back onto the drives when it is time to reuse them. It seems that

the Committee chose the word "routine" rather than "automatic" operation because it wanted to include human participation in these operations that might not be covered by purely "automatic" computer operations. We recommend clarifying that interpretation.

The Standard of Culpability

Microsoft believes that the current proposal is very helpful but it is too narrow in determining whether a party is ineligible for the safe harbor protection of Rule 37. The current wording, which applies a standard of negligence, is too limited in the protection it affords. We believe the appropriate standard is the one used in the alternate version listed in the footnote on page 13 – providing protection from sanctions so long as such data loss is not due to reckless or intentional behavior.

Rule 37. Failure to Make Disclosure or Cooperate in Discovery; Sanctions (alternate version)

(f) Electronically Stored Information. A court may not impose sanctions under these rules on a party for failing to provide electronically stored information deleted or lost as a result of the routine operation of the party's electronic information system unless:

(1) the party intentionally or recklessly failed to preserve the information; or
(2) the party violated an order issued in the action requiring the preservation of the information.

Negligence Compared to a Higher Standard

Litigants strive every day to take all reasonable steps to comply with their discovery obligations to preserve and produce relevant electronic information. Unfortunately, because litigation has become so expensive and contentious, parties seeking discovery know they can inflict considerable cost and institutional pain by making very broad discovery requests—and then challenging the adequacy of the response by bringing motions to compel and motions for sanctions. Even the threat of such challenges can have a substantial impact on the discovery value of a case – increasing the pressure to reach a settlement, even when the underlying case is meritless. The cases imposing sanctions for spoliation of evidence are only a small proportion of the spoliation motions that are brought or threatened. That kind of tactical extortion is not what Rule 1 was intended to achieve with its admonition that the rules "... shall be construed and administered to secure the just, speedy, and inexpensive determination of every action.[5]

This all means that the stakes are high for parties responding to discovery requests and the threat of sanctions for spoliation is a very effective deterrent for anyone foolish enough to even contemplate less than thorough responses. The Sarbanes-Oxley statute has further raised the stakes by providing that any knowing destruction of documents sought in litigation may be punished by lengthy prison sentences. Since the Andersen and Enron scandals everyone knows that a document retention policy that is implemented in bad faith can lead to the corporate death penalty – criminal indictment.

[5] FRCP 1.

These very effective deterrents have the undivided attention of corporate America and its counselors. A whole cottage industry of e-discovery conferences and consultants has developed because of the fact that so many corporate officers and lawyers are worried and want to make sure they are in full compliance with the law. That worry, combined with the strong need for automatic disposal of unneeded electronic documents to prevent information overload, has created widespread demand for a safe harbor from both government and industry parties that are continuously involved in litigation calling for massive productions of electronic documents.

We have no doubt that those calling for a safe harbor fully intend to act reasonably, but mere negligence should not automatically make them ineligible for safe harbor protection. What may seem like reasonable steps taken to preserve evidence by a company or government agency that received notice of a vague threat of litigation may in hindsight be easily painted as inadequate or unreasonable. Responding parties that operate their businesses and try to comply with their discovery obligations in good faith should not be sanctioned because a requesting party later thinks it was unreasonable that the responding party did not suspend the automatic operation of its IT systems in certain respects. Rather, litigants should be on notice that such systems will not be automatically suspended and the burden should be on the party seeking that discovery to demonstrate good cause why they should be suspended. Intentional or reckless disregard for discovery obligations should be punished but a safe harbor is not very safe and not much of a harbor if it does not protect a good faith actor from sanctions based on mere negligence.

For these reasons, Microsoft supports the adoption of the footnote version of the safe harbor proposal which would protect mere negligence while leaving intentional and reckless conduct sanctionable. We have one caveat, however, that should be addressed in the Committee Notes if this version is adopted. The "routine operation of [a] party's electronic information system" is usually designed to **intentionally** delete vast amounts of unneeded and inaccessible data such as spam, outdated or junk email, and outdated backup tapes. It must be made clear that a failure to suspend the routine systems that perform these intentional deletions is covered by the safe harbor unless the party decided intentionally or with reckless disregard of the facts that responsive data that they were ordered specifically to retain would be permitted to be destroyed by the routine operations. The point of the deeper safe harbor would be to protect the routine operation of IT systems unless a party intended to dispose of known potentially responsive documents that they were specifically ordered to maintain or recklessly disregarded this fact. It cannot have been the intent of the drafters of the footnote proposal that every deletion caused by the routine operations of IT systems are viewed as "intentional" so as to remove all such deletions from the protection of the safe harbor. With that clarification we believe that the footnote version of the safe harbor would best serve to protect the needs of businesses and other institutions to operate their IT systems while meeting their discovery obligations in good faith.

The Effect of a Court Preservation Order

Court orders must always be obeyed, and violations can always be punished. That goes without saying. Why then state the obvious in both safe harbor proposals? The notion that the protection afforded in both versions of the safe harbor is available only where the producing party has complied with all court orders is dangerous overkill since courts can always punish violations of their orders regardless of any safe harbor in the rules. If a preservation order is vague and merely directs a party to preserve all relevant documents that begs the question whether the order meant anything beyond

common law preservation obligations, yet might effectively neuter the protections of the safe harbor. If such an order has no independent meaning, it is mere surplusage that the requesting party obtained as an additional sword of Damocles to hang over the responding party. Beyond the usual sanctions available for spoliation of evidence the responding party can also be held in contempt of court for violating the preservation order.

If a requesting party has grounds to fear that a particular type of data may not be preserved then it should request an order addressing that particular circumstance. Faced with a specific order the responding party will know what it needs to do. If the order is vague then the responding party will risk contempt regardless of how careful or conscientious it may have been in its preservation efforts. This flies in the face of what we know about electronically stored information, and the routine operation of electronic information systems, which make it virtually impossible to ensure the preservation of all relevant electronically stored data.

Regardless of which standard of care is imposed upon a party, we believe the best approach to the issue of court preservation orders is to discuss in the Notes the effect a court preservation order may have on determining whether the necessary standard has been met, rather than denying protection altogether where an order is violated. For example, the proposed Notes should state that all relevant circumstances may be taken into account in determining whether the failure to preserve information should be considered intentional or reckless, clearly contemplating that a higher standard would be applied to a party's behavior when a court preservation order is in place.

If violation of a court order remains in the Rule as an exception to the protection afforded by the Rule, at the very least a qualifier should be added to the wording so that it is clear that this exception applies only in situations where there is a violation of an order in the action requiring the party "to preserve *specified* electronically stored information." This clarification is very important because of the tendency of some courts to enter somewhat vague preservation orders which, in their efforts to be all-inclusive, result in overly-broad wording that would render the newly proposed safe harbor language meaningless. Such vague orders do not give the responding party guidance on what needs to be preserved, but simply give the requesting party another club with which to beat the responding party when, despite its best efforts to preserve all relevant documents, some obscure source of data slips through the cracks.

Loss of Information Due to Reasons Not Protected from Sanction Under Rule 37

Finally, we observe that neither the proposed language nor the alternative option articulate a duty to preserve, but rather are aimed at limiting the risk of sanctions under certain circumstances that assume the duty to preserve that arises under common law. The efficacy of either of these provisions relies on the Committee making clear in its Notes that although there are many circumstances in which information might be lost or deleted that are not exempted from sanctions under the provisions, this is not meant to imply that sanctions should be routinely applied in those other circumstances.

C. Application of Rules 33 and 34 to electronically stored information– definition of document, form of production, and options for production of electronic information.

There are several proposed changes to Rule 33 and Rule 34 that would collectively serve to adapt them to discovery of electronically stored information. The Committee has expressed particular interest in comment on whether Rule 34 itself or the Note should specifically state that a responding party

should not avoid reviewing and producing electronically stored information because a production request did not separately seek it, and—if so—what language would be the most helpful and appropriate. The Committee has also expressed particular interest in comment on whether the proposed options for production of electronically stored information are suitably analogous to the existing options for production of hard copy materials.

Addition of Electronically Stored Information to Rule 33

Rule 33. Interrogatories to Parties

* * *

(d) Option to Produce Business Records. Where the answer to an interrogatory may be derived or ascertained from the business records, including electronically stored information, of the party upon whom the interrogatory has been served or from an examination, audit or inspection of such business records, including a compilation, abstract or summary thereof, and the burden of deriving or ascertaining the answer is substantially the same for the party serving the interrogatory as for the party served, it is a sufficient answer to such interrogatory to specify the records from which the answer may be derived or ascertained and to afford to the party serving the interrogatory reasonable opportunity to examine, audit or inspect such records and to make copies, compilations, abstracts or summaries. A specification shall be in sufficient detail to permit the interrogating party to locate and to identify, as readily as can the party served, the records from which the answer may be ascertained.

Microsoft's view is that Rule 33(d) already adequately covers electronically stored information, and no addition or change is required. However, if the Committee decides that the change to 33(d) is warranted, the only accompanying requirement in the Note should be that the electronically stored information be provided in the format in which it is maintained in the ordinary course or business, in a format mutually agreed upon, or in a "reasonably usable" format.

Addition of Electronically Stored Information to Rule 34

Rule 34. Production of Documents, Electronically Stored Information, and Things and Entry Upon Land for Inspection and Other Purposes

(a) Scope. Any party may serve on any other party a request (1) to produce and permit the party making the request, or someone acting on the requestor's behalf, to inspect, ~~and~~ copy, test, or sample any designated electronically stored information or any designated documents (including writings, drawings, graphs, charts, photographs, sound recordings, ~~images~~phonorecords, and other data or data compilations in any medium—from which information can be obtained, translated, if necessary, by the respondent through detection

devices into reasonably usable form), or to inspect, and copy, test, or sample any designated tangible things which constitute or contain matters within the scope of Rule 26(b) and which are in the possession, custody or control of the party upon whom the request is served; or (2) to permit entry upon designated land or other property in the possession or control of the party upon whom the request is served for the purpose of inspection and measuring, surveying, photographing, testing, or sampling the property or any designated object or operation thereon, within the scope of Rule 26(b).

(b) Procedure. The request shall set forth, either by individual item or by category, the items to be inspected, and describe each with reasonable particularity. The request shall specify a reasonable time, place, and manner of making the inspection and performing the related acts. The request may specify the form in which electronically stored information is to be produced. Without leave of court or written stipulation, a request may not be served before the time specified in Rule 26(d).

The party upon whom the request is served shall serve a written response within 30 days after the service of the request. A shorter or longer time may be directed by the court or, in the absence of such an order, agreed to in writing by the parties, subject to Rule 29. The response shall state, with respect to each item or category, that inspection and related activities will be permitted as requested, unless the request is objected to, including an objection to the requested form for producing electronically stored information, stating ~~in which event~~ the reasons for the objection ~~shall be stated~~. If objection is made to part of an item or category, the part shall be specified and inspection permitted of the remaining parts. The party submitting the request may move for an order under Rule 37(a) with respect to any objection to or other failure to respond to the request or any part thereof, or any failure to permit inspection as requested.

Unless the parties otherwise agree, or the court otherwise orders,

(i) a~~A~~ party who produces documents for inspection shall produce them as they are kept in the usual course of business or shall organize and label them to correspond with the categories in the request; and

(ii) if a request for electronically stored information does not specify the form of production, a responding party must produce the information in a form in which it is ordinarily maintained, or in an electronically searchable form. The party need only produce such information in one form.

* * * * *

Microsoft agrees with the Committee's decision to use the phrase "electronically stored information," and with the decision to add this phrase to Rule 34 in order to introduce the concept, rather than attempting to introduce it by adding a definition to Rule 26. The addition of this phrasing to the Rules is important because the definition of "document" under Rule 34 has long lagged reality when it comes to electronically stored data.

The Concept of Electronic **Information**

We support the idea of referring to electronic data using the more generalized concept of "information." Not only is this phrase a more accurate description for what any electronic data (whether text, digital images, sound recordings, metadata, embedded data, etc.) actually is—information—but it also provides both the guidance and the flexibility to deal with new technology that enters the market constantly. We believe that this shift in thinking will help alleviate the struggles faced by courts and parties in deciding what constitutes a document and how to address issues regarding "embedded data," "metadata" and "native formats." By recognizing that all of these items are simply information which, if relevant and non-privileged, may be subject to discovery, courts and parties can do what they are long accustomed to doing, balancing burden and utility, relevance and need for production of certain types of information on a case-by-case basis. We believe that the currently proposed wording in the Note at page 28 correctly and adequately clarifies that, despite the newly introduced concept of "electronically stored information," requests for production of documents should be understood to include electronically stored information, unless otherwise indicated.

The Importance of Focusing on Electronically "**Stored**" Information

It is important to note that the Committee wisely chose to use the word "stored" rather than electronically "created" or "transmitted." This recognizes the fact that a lot of electronic information is created and transmitted that is never stored. This has been true for a long time. The telephone, the radio and the television are classic examples of audio and video data that is created and transmitted over copper wires, fiber optic cable, or the airwaves. On the receiving end that flow of electronic information (digital or analog) is rarely recorded. While the technology to record all phone calls, radio and television transmissions has existed for a long time its use is the exception not the rule. Only exceptional industries and public services like stock brokers, or 911 dispatchers routinely record and store such electronic information. Those exceptions are determined by law or by business need. It is important that nothing in these rules be construed to change the fact that, in our society the recording and storage of electronic information is the exception rather than the rule. Technological advances are likely to make it possible in the near future to effectively have recording devices on all phones, and in all offices and meeting rooms.

PCs, PDAs and smart-phones are increasingly designed to be recording devices and they have the capacity to store vast amounts of electronic information. Since the digital multimedia boom of the 1990s, PCs and other devices increasingly, have not only built in speakers and video screens to display images and play sound, but they also have built in microphones and increasingly still or video cameras. The boom in camera phones is very recent and web cams are increasingly sprouting on PCs PDAs can be used as Dictaphones or can record conversations. Video conferencing once required expensive equipment but can now be done from any PC with a webcam and a broadband connection. All this electronic information can easily be recorded and stored, but that must remain the choice of the user. Only when recorded and *stored* should all this information be subject to discovery and the proper object

of these Rules. Email is generally subject to these Rules because by default the email software records and stores email until the user affirmatively chooses to delete it. Instant messaging is different because the leading software that creates it, including Windows Messenger and MSN Messenger, are designed by default not to record or store IM sessions. Just like with a normal phone call, when the IM session is over or the text window is closed the messages are not stored. They are automatically deleted, unless the user affirmatively chooses to save them.

Form of Production & Options for Production

Microsoft has concerns concerning the proposed revisions to Rule 34(b) which list the options for production of electronically stored information. It is our view that the two options for producing electronically stored information under 34(b)(ii) should be "a form in which it is ordinarily maintained, or in a reasonably useable form." We believe providing for production in "a reasonably useable form" is more appropriate than specifying "an electronically searchable form." The phrase "a reasonably useable form" better conveys the underlying goal—that parties receive information in a useful format—while refraining from making assumptions as to what that format will be.

The option of producing electronically stored information in a reasonably useable form allows greater flexibility to the courts and parties, and avoids making any unwarranted assumptions regarding the appropriate format for electronically stored information. It is not clear exactly what is intended by the phrase "electronically searchable," (i.e., Searchable for what – words, sounds, moving images, still images, numbers, notes? And how – searching within metadata, within content, etc.?), but much electronically stored information is not truly "electronically searchable" in a manner that would be of much use to parties in discovery. We see this already in today's world with file types used for data, sounds, and images such as .gif, .mpg, .jpg, .wav, etc., and have no way of knowing what new types of electronically stored information may exist in the future in which the core information is not electronically searchable in the sense most likely contemplated by the currently proposed language. The Note should contain clarification regarding this issue, including discussion of electronically searchable form as a potential "reasonably useable form" option.

A second, and somewhat overlapping concern, is that we believe it is important that the Rule should not favor or specify any particular format of production—and, in particular, that the Rules should neither require nor favor production of electronically stored information in its native format. It is our experience that the format of production is almost always the subject of legitimate discussion because the parties have different IT systems and litigation support software and databases that support different formats and different requirements for their systems.

We strongly urge against any wording in the Rules that could be read to favor production in native format (which we believe it currently does), because of the many complications that production in native format introduces. In many cases, production in native format can greatly add to the burden on the producing party, because additional review may be necessary to offset the increased risk of inadvertently producing privileged material. It can also add to the burden of the litigation on both sides because the dynamic nature of native format documents makes it difficult to ensure the integrity of the data and makes it impossible to apply page numbers (a.k.a. "bates numbers") and protective order designations on the page level for more effective data management. Anyone actively involved in complex litigation will attest to the importance of the unique number on each page of the document. Bates numbers help litigants track documents and ensure the integrity of their contents. *Data integrity*:

The native formats of documents created in programs such as Word, Outlook or Excel are designed to make it very easy to edit and alter a document. These documents can be converted into formats like PDF or TIF where the data are very difficult to alter or edit. From a data integrity point of view a preference for production of native format documents is like saying that in the paper world that a letter written in pencil should be produced in easy to alter pencil format rather than in difficult to alter ink format such as a photo copy of the original. *Document tracking*: Although it may be possible to place a number at the file level (*i.e.*, within the metadata)[6] of the electronically stored information, this is often not displayed when printed, and for anything longer than one page, is not a unique number. Document numbers are particularly important to Microsoft and many other litigants because they can also allow the user to identify the specific employee from whose files the electronically stored information came.

Protective order designations present a similar problem. We know of no way to automatically place the protective order designation on each page of a native format document, and manually altering each native file to place a footer with this information is not practical. It is also standard practice in large cases for protective orders to have more than one tier of confidentiality of production, in some cases restricting certain categories of documents to review by outside counsel only. The inability to systematically place unique protective order designations on each page of the native file is a critical reason *not* to require or favor native productions.

Finally, there is no good way to use the average native file in depositions, in motion practice, or at trial. Litigants want to see and use paper, or the electronic equivalent of paper – complete with pages, document page numbers, and protective order designations.

For all the above reasons, we urge the Committee to adopt a rule that does not favor a specific format of production, nor set a presumption in one party's favor to select the format of production. We believe that the Rule should require the parties to identify and discuss the issues at an early stage in the litigation. It is very important that the Committee makes it clear in its Notes that the Rule is not intended to dictate a particular format, but rather is intended to define acceptable alternatives for form of production, recognizing that native production may not be the most appropriate form of production in many cases.

D. Early attention to issues relating to Electronic Discovery.

The proposed amendments to Rule 16, Rule 26(f), and Form 35 are intended to present a framework for the parties and courts to give early attention to issues relating to the disclosure or discovery of electronically stored information.

Rule 16. Pretrial Conferences; Scheduling; Management

* * * * *

(b) Scheduling and Planning. Except in categories of actions exempted by district court rule as inappropriate, the district judge, or a magistrate judge when authorized by district court rule, shall,

[6] Processes are available to rename the native file to include a unique number. For example, a Word document named "Letter to John.doc" could be renamed, at the file level, "Letter to John_1000005.doc".

after receiving the report from the parties under Rule 26(f) or after consulting with the attorneys for the parties and any unrepresented parties by a scheduling conference, telephone, mail, or other suitable means, enter a scheduling order that limits the time

(1) to join other parties and to amend the pleadings;

(2) to file motions; and

(3) to complete discovery.

The scheduling order may also include

(4) modifications of the times for disclosures under Rules 26(a) and 26(e)(1) and of the extent of discovery to be permitted;

(5) provisions for disclosure or discovery of electronically stored information;

(6) adoption of the parties' agreement for protection against waiving privilege;

(75) the date or dates for conferences before trial, a final pretrial conference, and trial; and

(86) any other matters appropriate in the circumstances of the case.

The order shall issue as soon as practicable but in any event within 90 days after the appearance of a defendant and within 120 days after the complaint has been served on a defendant. A schedule shall not be modified except upon a showing of good cause and by leave of the district judge or, when authorized by local rule, by a magistrate judge.

* * * *

Microsoft supports the idea that issues relating to electronic discovery should be discussed early in the discovery process, and addressed in the discovery plan as appropriate. We believe the general wording used in the proposed amendment to Rule 16(b)(5), along with the accompanying Committee Notes, adequately alerts parties and the court to the possible need to address the handling of discovery of electronically stored information early in the litigation *if* such discovery is expected to occur. We believe this generalized approach is a more appropriate means of encouraging the parties to address electronic discovery issues early on than provisions that have been adopted in some jurisdictions that describe specific actions to be taken by the parties. Requiring a company "to investigate and disclose" specific information regarding its entire computer system will often be very burdensome and unnecessary to facilitate discovery regarding a narrow set of claims that relate only to a subset of a company's businesses, products or employees. As discussed in the introduction to our comments the client-server architecture of the IT systems of any large organization will be highly complex and constantly changing as technology and business needs evolve. Large organizations usually do not have any one person or department that is responsible for or has an overview of the organization's entire IT system. Those networks will be comprised of many thousands of PCs and other client devices connected to thousands of servers in order to meet the diverse operational needs of thousands of office workers in different occupations and lines of business.

While much of the litigation in the 21st century will involve discovery of electronically stored information this will not always be the case and the Rules should reflect that. Therefore, we see as crucial the Committee Notes' emphasis that if the parties do not anticipate discovery of electronically stored information, there is no need to discuss these issues.

Microsoft is concerned about the proposed new language for Rule 16(b)(6). As we will discuss again when considering the proposals for Rule 26(f)(4), below, we oppose any addition to the Rules that would influence parties to adopt agreements regarding privilege waiver – particularly insofar as such agreements may serve to pressure parties into the premature production of privileged material. In addition, we are concerned that the proposed language's subtle endorsement of agreements regarding waiver of privilege may have the unacceptable effect of influencing courts in their determinations of whether a party has waived privilege in regards to inadvertently produced privileged information in situations where the parties have not entered such an agreement. We are also concerned that courts will pressure a litigant who complains of cost to adopt an alternative review approach.

Rule 26. General Provisions Governing Discovery; Duty of Disclosure

* * * * *

(f) Conference of Parties; Planning for Discovery. Except in categories of proceedings exempted from initial disclosure under Rule 26(a)(1)(E) or when otherwise ordered, the parties must, as soon as practicable and in any event at least 21 days before a scheduling conference is held or a scheduling order is due under Rule 16(b), confer to consider the nature and basis of their claims and defenses and the possibilities for a prompt settlement or resolution of the case, to make or arrange for the disclosures required by Rule 26(a)(1), to discuss any issues relating to preserving discoverable information, and to develop a proposed discovery plan that indicates the parties' views and proposals concerning:

(1) what changes should be made in the timing, form, or requirement for disclosures under Rule 26(a), including a statement as to when disclosures under Rule 26(a)(1) were made or will be made;

(2) the subjects on which discovery may be needed, when discovery should be completed, and whether discovery should be conducted in phases or be limited to or focused upon particular issues;

(3) any issues relating to disclosure or discovery of electronically stored information, including the form in which it should be produced;

(4) whether, on agreement of the parties, the court should enter an order protecting the right to assert privilege after production of privileged information;

(53) what changes should be made in the limitations on discovery imposed under these rules or by local rule, and what other limitations should be imposed; and

(64) any other orders that should be entered by the court under Rule 26(c) or under Rule 16(b) and (c).

The attorneys of record and all unrepresented parties that have appeared in the case are jointly responsible for arranging the conference, for attempting in good faith to agree on the proposed discovery plan, and for submitting to the court within 14 days after the conference a written report outlining the plan. A court may order that the parties or attorneys attend the conference in

person. If necessary to comply with its expedited schedule for Rule 16(b) conferences, a court may by local rule (i) require that the conference between the parties occur fewer than 21 days before the scheduling conference is held or a scheduling order is due under Rule 16(b), and (ii) require that the written report outlining the discovery plan be filed fewer than 14 days after the conference between the parties, or excuse the parties from submitting a written report and permit them to report orally on their discovery plan at the Rule 16(b) conference.

* * * * *

Microsoft supports the changes to 26(f)(3) that add "issues relating to disclosure or discovery of electronically stored information, including the form in which it should be produced" to the list of items to be included in the parties' proposed discovery plan. We believe that this is the appropriate venue for the parties to discuss various issues surrounding electronically stored information, including the form of production, special requests for metadata or embedded data, etc.—and resolve these issues, if possible. We feel it is very important that the accompanying Note emphasize that the particular issues regarding electronically stored information that should be discussed will depend on the specifics of the given case. Microsoft believes the references in the Note to the possible need to discuss and become familiar with the "parties' information systems" or a "party's computer systems" should include a qualifier such as "relevant" – so that the Note is not read to imply that the entirety of a party's computer or information systems needs to be discussed. Often it is only a very small portion of those highly complex systems that is relevant to the litigation.

The Committee has expressed particular interest in receiving comment on whether the proposed amendment to Rule 26(f)(4) should be less restrictive, directing the parties to discuss and include in the discovery plan any issues relating to the protection of privileged information in discovery. Microsoft believes that either proposal may serve to increase the pressure for premature production of possibly privileged information. Accordingly, we would favor removing altogether this reference to the protection of privilege. However, if language regarding the protection of privileged information is included in Rule 26(f), we believe that the accompanying Note must make it clear that the provisions are meant to encourage discussion, but not intended in any way to influence parties to turn over material without first reviewing for privilege, to exert any other pressure for the premature production of possibly privileged information, or to imply that the absence of an agreed order regarding privileged material (or, in case of the Committee's alternative proposal, the absence of discussion in the discovery plan of any issue relating to the protection of privileged information), weakens a party's right to assert privilege. Finally, to the extent the revised Rule 26(b)(5)(B) establishes a procedure for asserting privilege when a party produces information without intending to waive its claim of privilege, the proposed paragraph 26(f)(4) appears to be redundant and unnecessary.

Microsoft has two concerns regarding the proposed change to 26(f), requiring parties "to discuss any issues relating to preserving discoverable information." First, we are concerned that this, the first time the rules allude to a "preservation obligation," comes close to statutory limits on the rulemaking power under the Rules Enabling Act because the rules address disclosure and discovery, not the preservation of information. Second, we are concerned that this language may be interpreted to encourage the entry of unnecessary preservation orders. If this language remains, it is crucial that the accompanying Note contain language that emphasizes that the measure is intended to encourage

consideration of preservation issues early in the discovery process, and should not be read to stimulate, encourage, or otherwise condone the entry of unnecessary or overly broad preservation orders.

E. Assertion of Privilege After Production.

The proposed amendment to Rule 26(b)(5) sets up a procedure to apply when a responding party asserts that it has produced privileged information without intending to waive the privilege. The Committee has stated a particular interest in receiving comment on whether a requirement that a party who receives notice that privileged material has been produced must certify that the material has been sequestered or destroyed if it is not returned.

Rule 26. General Provisions Governing Discovery; Duty of Disclosure

* * * * *

(b) Discovery Scope and Limits. Unless otherwise limited by order of the court in accordance with these rules, the scope of discovery is as follows:

* * * * *

(5) Claims of Privilege or Protection of Trial Preparation Materials.

(A) *Privileged information withheld.* When a party withholds information otherwise discoverable under these rules by claiming that it is privileged or subject to protection as trial preparation material, the party shall make the claim expressly and shall describe the nature of the documents, communications, or things not produced or disclosed in a manner that, without revealing information itself privileged or protected, will enable other parties to assess the applicability of the privilege or protection.

(B) *Privileged information produced.* When a party produces information without intending to waive a claim of privilege it may, within a reasonable time, notify any party that received the information of its claim of privilege. After being notified, a party must promptly return, sequester, or destroy the specified information and any copies. The producing party must comply with Rule 26(b)(5)(A) with regard to the information and preserve it pending a ruling by the court.

Microsoft supports the proposed amendment to Rule 26(b)(5). We believe that the rule should include a requirement that a party that receives notice that privileged material has been produced must certify that they have returned, sequestered, or destroyed all copies of the material. This additional

requirement is not overly burdensome, and is warranted in light of the ease with which a party could otherwise continue to use or circulate privileged material, with adverse consequences to the party attempting to assert privilege.

* * * * *

We thank you for the opportunity to provide comments on the proposed amendments to the Federal Rules.

Very truly yours,
Microsoft Corporation

Thomas W. Burt

By
Thomas W. Burt
Vice President &
Deputy General Counsel

Client Server Architecture Diagram