

“Clearing the Fog of War”– Recommended E-Discovery Meet and Confer Guidelines

Deborah H. Juhnke and Joan E. Feldman
Computer Forensics Inc.TM

The recent hearings before the Committee on Rules of Practice and Procedure of the Judicial Conference of the United States have yielded a variety of opinions regarding both the content and interpretation of the proposed amendments to the Federal Rules of Civil Procedure. Although there are some areas where attorneys may agree to disagree, one recommendation with nearly universal appeal is an early meet and confer conference regarding e-discovery issues. This early and focused discussion between the parties of issues surrounding the collection and production of electronically stored information will greatly facilitate the speed and efficiency of discovery.

Recommendations Regarding Rule 26(f) Meet & Confer Conference

Rule 26(f) sets forth a recommended protocol for a meet and confer dialogue, including an examination of issues relating to: (i) the disclosure or discovery of electronically stored information, including the form in which it should be produced; (ii) whether the parties can agree to production on terms that protect against privilege waiver; and (iii) the preservation of discoverable information (including electronically stored information). We believe two components should be added to the list of topics to enable a more comprehensive discussion. These topics are (iv) the nature and volume of material to be produced, including data sources, data types, date and time frames, and stipulations as to what constitutes duplicate or “near duplicate” data; and (v) the drafting of a mutually agreed upon glossary of terms to be used throughout the discovery process.

(i) Disclosure and Form of Production

It is important during the meet and confer process to have qualified technical personnel or experts not only advise, but also actively participate in the discussion. Even for counsel who possess advanced technological skills, the utilization of in-house technical personnel, information technology specialists, and/or outside experts in the meet and confer process will ensure that the information exchange is accurate, usable, and complete. Because technology changes rapidly, and because parties may have computing issues unique to their environments, active participation by technology experts can facilitate development of an accurate inventory of potentially responsive data, as well as allow experts to address issues surrounding preservation, form of production, and accessibility.

Collection, review, and production of inventoried data are central topics of the meet and confer discussion. As noted in the Manual for Complex Litigation, (4th) § 11.423, the Court should encourage counsel to “. . . produce data in formats and on media that reduce the transport and conversion costs, maximize the ability of all parties to organize and analyze the data during pretrial preparation, and ensure usability at trial.” Many factors bear on how to interpret this guidance, however, including how the data are maintained in the normal course of business, what is a practical and usable form of production, and each party’s idea of how they can best organize and analyze the data. Ideally, parties will agree upon the format(s) of production appropriate to the case at hand, rather than attempt to apply a common standard to all data. For example, while conversion of electronic data to a common format (most typically TIFF or PDF images and text) may sometimes be desirable from a review perspective, it may also be cost-prohibitive because of volume, or impractical for certain types of data, such as databases. Additionally, production of the native (i.e., original) file is sometimes necessary to ensure complete and accurate production. This is especially true when dealing with certain types of files, including spreadsheets, complex or compound files (such as Lotus Notes and Microsoft Exchange), and any other files containing embedded data.¹

(ii) Review and Privilege

Although the method of review has traditionally been the sole purview of the producing party, increasing efforts at cost sharing have opened the door for the requesting party to become actively involved in the methods, services, and tools used to review electronic data prior to production. Options such as native file review, keyword searching, concept-based review tools, and on-line repositories may be considered. Here again, technological advances will continue to shape the process, so parties have an opportunity to work together to minimize the cost and burden of review by discussing them early on. Good technical and/or expert advice at this stage will nearly always help to minimize these costs. Consideration should also be given to mechanisms that will allow speedy review of electronic documents without creating an undue burden to review all metadata, including the option of a stipulation of non-waiver.

(iii) Preservation

A comprehensive inventory is the base element for discussion of preservation issues, and is also a key component in full disclosure by the responding party. Failures to agree on a preservation protocol, or claims of spoliation, have in many instances been caused simply by an inability of one party to understand the scope and nature of the data collection. The

¹ Embedded data is information that is generally hidden but is an integral part of a computer file, such as “track changes” or “comments.” Metadata, on the other hand, is information *about* a file, including its creation date, size, etc. While some metadata is routinely extracted during a conversion process to TIFF and text, embedded data is not. Therefore, it is only available in the original, native file.

inventory should contain enough detail regarding retention policies and procedures to enable the parties to reach common ground on preserving what is necessary while not unduly disrupting the responding party's ability to do business. This is particularly true when the responding party is a larger business entity. In this case, knowledge of enterprise computing systems² is also a pre-requisite to any discussion of burden.

(iv) Nature and Volume of Data

A comprehensive inventory of potentially responsive data lays a critical foundation for subsequent dialogue regarding timing of discovery, collection methods, and costs. The inventory should be focused on the issues, parties, and time frames pertinent to the litigation, and designed to capture information regarding all likely sources and types of relevant data subject to discovery. Time is of the essence in compiling this inventory, as computer data is volatile and can be transitory. Components of an electronic discovery inventory might include identification of both custodial and non-custodial data³, email, shared servers, web servers, desktop and laptop computers, databases, voice messaging, instant messaging, backup tapes, and other media, software, and data sources upon which the responding party relies in its daily business, *and as it relates to the litigation*.

Identification of the nature and volume of data is critical in the formulation of a realistic discovery plan. For example, an early inventory should form the basis for cost and timing estimates for collection and review of the data. A vendor estimate that is based upon verifiable assumptions will be more reliable in both cost and timing than one based on speculation or ill-informed assumptions. The parties should consider, however, whether the vendor estimate—or, preferably, estimates—should be based on the universe of potentially responsive data or upon a subset of that data. As we have seen in recent cases, there may be occasions when a sampling of data is more desirable and appropriate. See, e.g., *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309 (S.D.N.Y. 2003). Though cost allocation issues have historically arisen post-production, the increasing volumes and complexity of electronic discovery demand earlier attention.

Duplicate Documents

The need to deal with duplicate documents in the collection and review of electronic data is far greater than it was when discovery was primarily paper-based. The volume of electronic documents presented for review can easily increase as much as ten- or twenty-fold if duplicates and near-duplicates are not eliminated early. Duplicate documents are

² An enterprise computing system is most often a mainframe- or mini-computer-based system designed to manage major business processes. For example: transaction processing for a national retailer, personnel systems for multi-office corporations, or banking systems.

³ A custodian is typically a person. Non-custodial data is that data which is either shared by a number of persons or is “owned” or managed by the corporate entity. Examples of non-custodial data include enterprise databases, management software, and workgroup libraries.

typically defined as exact copies. The source of the document may or may not be taken into account. Studies have shown that roughly 70 percent of documents in a corporate environment are duplicate or near-duplicates. Technologies exist to eliminate both exact and near-duplicates, such as those that result from the common practice of adding a superfluous 'thank you' comment to an original email and passing it along. Parties are, therefore, encouraged to examine ways to cull out these additional documents well in advance of further processing and review.

Accessible vs. Inaccessible Argument

The parties should attempt to agree on what constitutes accessible data. Whether data are reasonably considered inaccessible depends in part upon the timing and reason for a transfer to backup media, as well as the intended use and/or frequency with which it is retrieved in the normal course of business. As data storage technologies evolve, media that might have been considered inaccessible according to the *Zubulake* standard may become accessible, so care should be taken to define accessibility according to use and purpose, rather than by type of media. Though there are certainly instances where discovery of the inaccessible, unallocated space on a computer hard drive is not appropriate, there are others where these data are clearly relevant.

(v) Glossary of Terms

The benefit of a common glossary of terms is self-evident. Parties may find themselves unable to reach agreement on discovery issues merely because they do not fully understand technical terminology. By first ensuring this common understanding, parties may move more quickly through the remaining elements of the meet and confer protocol. Once a basic vocabulary has been established, new terms may be added during the course of discovery as the facts and nature of the case warrant.

If parties are able to meet in good faith and reach an understanding regarding most, if not all, of these topics, the efficiency of e-discovery will be greatly enhanced. To the extent they cannot agree, they will at least be in a better position to muster facts and figures for the Court's enlightenment.