



Kenneth.Conour@dbr.com

12/15/2004 02:10 PM

To peter\_mccabe@ao.uscourts.gov

cc

Subject January 12 hearing on the Proposed Amendments to the  
Federal Rules of Civil Procedure

RECEIVED  
12/15/04

Dear Mr. McCabe:

I would like an opportunity to testify at the Committee's January 12 hearing in San Francisco on the Proposed Amendments to the Federal Rules of Civil Procedure regarding electronic discovery. I am a partner with the law firm of Drinker Biddle & Reath. The firm does a variety of civil litigation, and my office has a particular emphasis on nationwide litigation involving mass torts, including representation of pharmaceutical clients in the Diet Drug, Propulsid, PPA and HRT litigations. Through these representations and others, I have had substantial involvement with electronic discovery issues; this experience has including negotiating electronic discovery protocols and orders, litigating protective orders regarding electronic discovery, and consulting with clients on the scope of electronic discovery obligations and the mechanics of preserving and producing e-mails, databases, enterprise applications, back up media and other electronic materials.

Please let me know if there is anything else I need to do in order to testify at this hearing, or if you have any questions or need additional information. Thank you for your consideration.

**Kenneth P. Conour**  
**DRINKER BIDDLE & REATH LLP**  
50 Fremont Street, 20th Floor  
San Francisco, California 94105-2235

[www.dbr.com](http://www.dbr.com)

Telephone: (415) 591-7500

Facsimile: (415) 591-7510

[kpconour@dbr.com](mailto:kpconour@dbr.com)

04-CV-051  
Request to Testify  
1/12 San Francisco

RECEIVED  
1/26/05

04-CV-051  
Comments after  
1/12 San Francisco  
Oral Testimony

DrinkerBiddle

MEMORANDUM

TO: Advisory Committee on Federal Rules of Civil Procedure  
c/o Peter G. McCabe, Secretary

FROM: Kenneth P. Conour

DATE: January 26, 2005

RE: Proposed E-Discovery Rules Amendments

---

**INTRODUCTION**

I write to clarify and expand upon my comments made before this Committee at the January 12, 2005, hearing in San Francisco. I thank you again for the opportunity to provide comments on the proposed rules. As stated at the hearing, a significant portion of my practice involves electronic discovery and I know from experience that the changes to the federal rules that are being considered provide an opportunity for much needed uniformity, guidance and fairness.

However, I remain concerned that the proposed amendments, and the guidance provided by the Committee Notes, do not adequately address the challenges associated with discovery of "active", or routinely accessed, electronically stored information. As discussed below, I suggest (1) changing the proposed rules to extend the presumptive limitations on discovery to include active data that is not reasonably available for production in discovery or, at the least, clarifying the Notes to emphasize that the other limitations to discovery set-out in Rule 26(b)(2) may preclude discovery of active data, (2) modifying proposed Rule 37(f) to extend application of the safe-harbor protections to certain active data that is routinely and necessarily modified as part of ongoing business operations and (3) adoption of cost-allocation provisions requiring the requesting party to pay for all or part of the costs of electronic discovery in appropriate circumstances.

**DISCUSSION**

**I. PROTECTING CERTAIN ACTIVE DATA FROM DISCOVERY**

**A. The Problem with the Proposed Rules and Notes**

The proposed rules changes and accompanying Notes appear to be premised on a perception that information routinely used for business purposes is necessarily a fair target for preservation and production without sufficient regard to the costs and complexities involved with discovery of such active data. In this Committee's Report to the Standing Committee, a principle focus in articulating the need for rules amendments specific to electronic discovery is the substantial volume of electronic data that is generated and retained. The report also appropriately refers to

# DrinkerBiddle

other characteristics of electronic data that justify explicit treatment of electronic discovery in the rules. These characteristics include

- the dynamic nature of electronic information – the constant and, often, unintentional, automatic, deletion or alteration of data through routine operations;
- the hidden nature of associated data; and
- the difficulties in translating electronic information into a usable form of production.

These characteristics are not limited to off-line data. However, the Committee's Report, the proposed rules changes and the guidance provided by the Committee's Notes do not adequately recognize these challenging characteristics of electronic data in the context of active data.

Rather, the direction seems to be that *all* active data should be preserved and produced, without regard to the form, content or portability of such data. In making the phrase "*reasonably accessible*" information practically synonymous with *discoverable information*, and then defining reasonably accessible information as information that is routinely accessed, there is not much room to assert that at least some active data should not be discoverable, at least in the first instance. For example, in explaining the phrase "reasonably accessible", the Committee Notes state:

Whether given information is "reasonably accessible" may depend on a variety of circumstances. One referent would be whether the party itself routinely accesses or uses the information. If the party routinely uses the information – sometimes called "active data" – the information would ordinarily be considered reasonably accessible . . . Note to Rule 26(b)(2) at p. 12.

There are no referents or illustrations to suggest that in some circumstances data that is accessible by a company might not be reasonably accessible for discovery. Although the Notes do not foreclose the possibility of withholding active data on the basis of burden, cost or otherwise, it is troubling that the only discussion of active data is in the context of it having to be produced. While there are various examples of off-line data not being discoverable, there are no illustrations or comments supporting the proposition that some active data may also properly be withheld from discovery.

The problem is compounded by other Notes comments that seemingly suggest that routinely accessed data should be discoverable without regard to the costs and burden of producing it. For example, the comment is made:

The goal is to inform the requesting party that some requested information has not been reviewed or provided on the ground that it is not reasonably accessible, the nature of this information, and the basis for the responding party's contention that it is not reasonably accessible. But if the responding party has actually accessed the requested information, it may not rely on this rule as

# DrinkerBiddle

an excuse from providing discovery, even if it incurred substantial expense in accessing the information. Note to Rule 26(b)(2) at p. 13.

Surely, the proposed rules changes and accompanying Notes were not intended to diminish the other protections of Rule 26(b)(2) when creating a presumptive limitation against discovery of certain electronically stored information. Indeed, it has been expressly stated that the proposed rules changes and comments are directed at the concept of “reasonably accessible” information and should not be construed as altering the other limitations of the rule. However, it is unavoidable that these rules changes will be perceived as the response to the problems with electronic discovery, not just the problems with off-line data. By expressly recognizing the problems inherent in the discovery of all electronic data – including those of costs and burdens – but then providing protections only for off-line data, it will be argued that it was intended that active data should be discoverable without regard to those problems. In other words, if it was expressly recognized that active data did not require the protection offered for off-line information despite the fact that discovery of active data also carries with it significant burdens and costs, how can a responding party attempt to limit discovery of active data based on such burdens and costs?

## **B. Why Active Data Also Requires Protection**

There are significant differences between discovery of *active* electronically stored information and paper discovery beyond just volume. After all, when dealing with paper discovery, how often do you have to spend tens of thousands of dollars on consultants just to help you understand

- where your documents are, or
- how to retrieve the documents for review, or
- how to copy documents for production to the other side, or
- even how to stop the documents from disappearing.

It is also true that discovery of some electronic information is easier than others, and that there are very substantial problems with preservation and discovery of certain active data. A good example is that of dynamic databases. The fact that a responding party can easily and routinely access active data in databases does not mean that such information is reasonably accessible – *or reasonably available* – for production in discovery (*or* that such data can easily be preserved).

In my practice, I routinely see discovery requests seeking production not only of data, but also the databases themselves. I am never quite sure what the propounding party means when it asks for production of a database. Typically, these databases are large relational databases maintained on Oracle, SQL or other platforms. Use of the databases is made through enterprise applications licensed to a company at costs of up to *several* hundred thousand dollars. The databases often contain dozens of tables, with each table containing multiple fields and tens of thousands, hundreds of thousands or even millions of rows of data. The data in any one table is very limited to a particular type of information and the entirety of information that might be relevant to any

# DrinkerBiddle

one subject or transaction will be spread out across many different tables. Much of this information will be in coded data, that can be understood only by reference to other tables that contain coding definitions.

Duplication of the database itself, while perhaps theoretically a possibility, is in reality not an option. For a party to receive the database, at the least, they

- must have a compatible server box;
- must have the same operating system identical to that of the producing party;
- must have the same database version (*e.g.*, cannot take Oracle 9i database and duplicate into Oracle 8i versions);
- must license the same enterprise application *and* make modifications to that application to match any modifications made by the producing party in its use of the application; and
- must have the same other associated applications that provide functionality to the database (*e.g.*, viewers, report writers, print applications).

Rarely will a requesting party want or be able to pay out the tens of thousands of dollars needed for hardware, and the several hundred thousand dollars required for software. And rarely will they have the database expertise to build or even maintain this database.

Moreover, because of the size and complexity of these databases, there is no simple copy or export function that can be run. Instead, much time and effort is required to build a duplicate system, perform a data transfer and complete validations to ensure that data is not lost or corrupted.

The process is further complicated in litigation, however, because usually not all of the data in a database will involve or otherwise be relevant to the litigation. In order to limit the data to that which might be relevant, significant efforts must be made to identify all linked records across the database. And when this reduction of data is done, the transfer of data to a duplicate system becomes even more complicated with an increased risk of corrupted data or data structures.

In my practice area, another confounding factor is that we are prohibited by federal regulation from producing identifying information for patients and others that is often contained in these databases. This information can be found in free-form narrative text, which can require the review and redaction of millions of separate database entries. This type of review and redaction has to take place outside of the native database environment because, for obvious reasons, the active data should not be altered (and because those systems are not designed to accommodate this type of functionality). *The same is true for review and redaction of privileged materials – if privileged information exists in a database, there will have to be alteration of the data before it can be produced.*

Accordingly, at least with respect to the area in which I practice, there is no way to accomplish production of databases in their native format – regardless of cost and burden. I expect that the same holds true in many other practice areas.

# DrinkerBiddle

Because production of a database is not realistic, the alternative is the export of data into some other environment. This usually involves export and conversion of data now existing in linked, relational database tables, into either pure text – usually something like comma delimited ASCII text files – or flat, independent tables, such as in an Access database.

This process presents many significant obstacles. You still have the problems of identifying and extracting only potentially relevant information. Typically this will involve many, many hours of evaluating data and developing queries needed to isolate subject data for export. These queries will be significantly different from those generated for routine business operations – they must pull together *all* data from *all* tables relevant to the subject records, not just some subset of data needed for a discrete business function. And while business queries are generally designed to output substantive information, a litigation extract must also include dictionary, look up and other system tables.

Production of data in flat files also necessarily means that the requesting party is not going to get the functionality developed for the original database. While the data itself is transferable, the export will not include the forms, reports or queries built into the original system.

Another problem is that many database applications now involve use of distributed data across various systems, maintained in various locations. For example, a database routinely accessed by company offices here in California may be drawing upon data actually resident on a computer system in India or elsewhere. Discovery of routinely accessed electronically stored information from this database, thus, may necessarily require discovery of associated data from another system in another country. Capturing that foreign data, and processing it for production, may require significant involvement of information specialists and consultants in that foreign country.

Use of encrypted data further complicates things. More frequently these days, data is routinely encrypted using applications associated with the database. Production of this encrypted data outside of the native database environment will be meaningless. Processing of the encrypted data may impair data functionality.

All too often, requesting parties do not understand – or seemingly do not care – that the production of database data in flat tables is of little or no value to them unless they spend substantial time and effort rebuilding links across dozens of tables containing thousands or even millions of records. Although the databases may contain some information undeniably relevant to the litigation, and that discrete information may be produced in electronic form with significantly less difficulty, requesting parties do not limit their requests to seeking *that* information, but instead demand production of *all* associated data from the database.

## **C. Addressing the Problem With Discovery of Active Data**

The proposed rules changes and Notes should be revised to reflect that the presumptive limitation on electronic discovery is premised on the problems associated with discovery of that data in litigation, not the ease of use of that data outside of litigation. In this regard, the phrase “reasonably accessible” should be replaced with “reasonably available.” By reasonably

# DrinkerBiddle

available, I mean information that can be made reasonably available for discovery. In defining the scope of discovery of electronic information, the touchstone ought not to be how easily the information can be *accessed* by the producing party *in the course of business* but, instead, how easily the information can be *made available* to the requesting party *in the course of discovery*.

Referents that can be used to determine whether electronically stored information is reasonably available for discovery include those already articulated for identifying reasonably accessible information, and others that would involve the ability of the producing party to produce that information in discovery using existing business routines or processes that do not require alteration of the existing data structure. A part of this would be the narrowing of the target of discovery to that data that can be made reasonably available for discovery even where the data is part of a broader data compilation that as a whole would be much more difficult to produce. For example, in looking at databases, it might be that only certain data from the database is relevant and material to the litigation. This data might be readily copied or exported for production using existing queries or reports, or simple cut and paste functions, although broader production of associated system data and other information within a database that is routinely accessed by the company would be much more difficult to produce.

Even if the proposed rule is not revised, the Committee Note for Rule 26(b)(2) ought to be clarified. Specifically, comments should be included that state that even if active data is not presumptively protected from discovery using whatever two-tiered approach is adopted, the problems associated with discovery of active data make it particularly appropriate that the undue burden provision and other protections set-out in the rule can be invoked to preclude or limit such discovery.

## II. EXTENDING THE SAFE-HARBOR PROVISIONS TO ACTIVE DATA

Databases also pose a problem in the area of preservation. The proposed rules and accompanying notes primarily speak to loss or alteration of data through automated and often undetected means. But databases are also continuously being purposefully altered to meet business needs.

Databases are constantly changing as new information is captured. This new information can involve either entirely new events, or can be in the nature of updated information for previously existing records – for example, new contact information or the identification of additional medications used by a patient. This new information can either be additional data, or it can be a modification to prior data. And some databases, say for example, those tracking sales for a defined period, are constantly changing by design as they retrieve current information and discard older data outside of the time window they were intended to track.

When your business requires that your databases change on a daily basis, what do you preserve – the data as it looks today, as it will look tomorrow or as it will look next week? And *how* do you preserve it? Replication of the database involves the concerns discussed above. And you cannot do that on a daily basis.

# DrinkerBiddle

The rules and notes need to speak to the fact that even purposeful and designed alteration of electronically stored information should not always be sanctionable. While certainly provisions need to be included for the automatic alteration of data, it is just as certain that such provisions do not go far enough to protect companies in the real world.

Accordingly, I suggest that the safe-harbor provisions of Rule 37(f) be revised to include something along the lines of:

A party that has preserved discoverable electronically stored information need not continue to preserve subsequent changes to that information when such changes are made as part of and with the purpose of the conduct of repeated and routine business functions.

### III. PROVIDING COST-ALLOCATION IN APPROPRIATE CIRCUMSTANCES

Cost-allocation is particularly appropriate for electronic discovery. It is, of course, recognized that in many actions individual plaintiffs will have little or no electronic information while corporate defendants maintain vast amounts of data. In those circumstances, there is nothing to deter the individual from seeking to impose broad electronic discovery obligations. Additionally, parties with electronic discovery may not have the cost-saving option of making information available for inspection. In traditional paper discovery, a producing party can tender documents for inspection without having to incur significant processing costs. In many instances, electronically stored information is not available for access in its native environment, for reasons including risk of loss or corruption of data, interference with ongoing business operations or inability to prevent disclosure of protected information.

There is also the problem, mentioned above, that electronically stored information might have little or no value to the requesting party when actually produced in discovery. Although the requesting party might be able to make out an argument that electronically stored information is relevant to the litigation, and perhaps does not exist in any other format, production of the data may have to be in a format that makes the information virtually inaccessible or otherwise useless to the requesting party. Often the sheer volume of data renders the discovery meaningless; it does not matter that information is produced if the requesting party will never have the time and resources to go through it.

Cost-allocation provisions similar to those adopted in Texas or California would provide a necessary deterrent to unreasonable and overbroad electronic discovery demands. Although the California statute seemingly would require cost-shifting in most instances where electronically stored information must be prepared for production, application of the “necessary” and “reasonableness” prongs emphasized by the court in *Lexar Media* may result in the practical application of something more like the Texas statute – requiring cost shifting where “extraordinary steps” are needed. Regardless, a cost-allocation provision will require the requesting party to carefully consider the value of the information sought, and perhaps, accept appropriate limitations on the discovery.