E-Discovery Initiatives at the Antitrust Division[1] Tracy Greer

Like the rest of the bar, the Antitrust Division experienced exponential growth in the amount of documents and information responsive to its Second Requests and CIDs. Recently, the Antitrust Division has completed several investigations with productions just under or slightly above one million records. As a result, Antitrust Division spending on its electronic storage capacity has increased substantially. In 2003, the Antitrust Division had a 12 terabyte[2] electronic storage capacity. In FY2008, the Antitrust Division increased its electronic storage capacity to 70 terabytes.

In 2006, the Antitrust Division established an internal working group to address E-discovery issues in civil matters. Matthew Hammond of the Division's Telecommunications and Media Enforcement Group led this effort. Every civil section and the Division's Litigation Support staff has a representative on the working group. The working group met, compared electronic discovery practices across sections, discussed innovations as well as problems that had occurred in past investigations, and discussed possible best practices. The working group's goal was to create straightforward and practical guidance for staff attorneys responsible for negotiating Second Requests and CIDs. We also wanted to provide detailed guidance to law firms and their electronic production vendors about the optimal way to produce electronic data and documents to the Division. This guidance was designed to ensure that parties could avoid producing data multiple times and that the production is in a format that could be reviewed promptly.

The working group continues to revisit electronic production issues. By having members in every civil section, Division attorneys have ready access to a colleague who has devoted time and thought to these issues. In addition, Litigation Support Section now assigns support analysts by section. This has improved the Staff's ability to negotiate and process electronically produced documents and information. Much as attorneys develop commodity and company-specific expertise, Litigation Support analysts enjoy some of the same benefits. Companies in similar industries tend to use and store electronic data in similar ways. Thus, as attorneys and litigation support analysts gain expertise, many electronic production issues have been resolved and can be addressed in the same manner in subsequent investigations.

This e-discovery working group developed a variety of documents and training programs. One of the most important documents authored by the working group is a template for an electronic production letter (*See* Antitrust Division Electronic Production Letter, Attachment 1). The electronic production letter, which Staff attorneys can (but are not required to) use, is supposed to be sent to counsel to address a variety of electronic production issues. More than a drafting tool, the electronic production letter can also serve as a checklist of issues that should be addressed as soon as practicable in negotiations between staff and counsel. It is important to emphasize that the electronic production letter is just a baseline or starting point. In some instances, the target company may not have an extensive computer network, which would make the letter inappropriate. In other circumstances, a full-blown electronic production may not be required to resolve the key issue in an investigation. In appropriate circumstances, Staff will use the electronic production letter for CIDs directed to third parties, but often the limited nature of the CIDs make it unecessary. Staff is also encouraged to modify the electronic production letter with the assistance of Litigation Support staff to address any investigation-specific issues.

The electronic production letter cannot "solve" the e-discovery issues. This area is evolving, technology is changing, and the electronic discovery working group continues to receive feedback about the electronic production letter, making modifications or additional suggestions as needed. Although the electronic production letter has been used in multiple investigations across a number of sections, we encourage staff to try different approaches depending upon the specific issues involved in a particular investigation. One of the particular advantages of this initiative has been the sharing of experiences across sections relating different approaches to solving e-discovery issues.

Based on staff's two-year practice with this process, I want to highlight several important issues and themes.

A. Techniques to Reduce the Size of Electronic Productions

First, there has been a shift from emphasizing that parties produce all possibly responsive documents and information to producing the most relevant material as quickly as possible. We also want to avoid parties producing redundant and marginally relevant material. As the volume of responsive material has increased exponentially, the available tools for reviewing documents have remained much the same. Therefore, we often want to minimize the production of technically responsive material that will not facilitate the Division's analysis of the transaction. One simple example of this technique is filtering email by domain names. Staff can provides a list of domain names (*e.g.*, espn.com, cnn.com, *etc.*) that parties can automatically excluded from collected emails. By allowing parties to exclude all emails from a given list of domain names, parties can reduce the amount of email that is ultimately produced to the Division and staff need not review what is likely to be irrelevant material.

Another technique to reduce the size of electronic production is deduplication. For a number of years, the Staff has typically allowed parties to vertically (*i.e.*, remove duplicates within a custodian's files) deduplicate their productions. More recently, staff in some instances has permitted producing parties to deduplicate their electronic productions both vertically and horizontally (*i.e.*, removes duplicates from multiple custodians' data sets). This development reflects a growing comfort among staff with deduplication techniques, a recognition that it is ineffcient for staff to review multiple copies of the same document, and a technical solution to the issue of potentially losing information by allowing deduplication.[(3)] The Division estimates that deduplication can reduce production volumes by as much as 40%. Staff continues to consider additional techniques such as whether to allow parties to "near-deduplicate" (*i.e.*, eliminate close, but not exact duplicates) their electronic productions. To date, staff has not permitted near-deduplication (*i.e.*, eliminating all e-mails in a string except for the final one unless the string diverges into different threads or attachments are eliminated) but parties who are interested in using this technique should raise the issue. However, counsel (or their vendor) should be prepared to describe in detail how this process would work, its impact on the production, and, preferably, run a test on a sample data set.

Methods to reduce production volumes need not involve technology. Even "old" practices that worked for paper productions can be helpful for electronic productions. Techniques such as producing one set of recurring reports, industry publications, prioritizing custodians, or prioritizing production based on dispositive issues can assist both Division staff and counsel in

limiting the scope of responsive documents. (These approaches are often used in conjunction with the Division's Merger Process Review Initiative which encourages staff and the parties to focus on potentially dispositive issues as early as possible.)

B. Early, Detailed Negotiations Necessary

The Division's electronic discovery working group has also encouraged staff and counsel to engage in early and detailed negotiations regarding Second Requests and CIDs. To make these early negotiations fruitful, attorneys must have a thorough understanding of their client's personnel, operations, offices and, most importantly, the client's corporate information management systems. Second Requests include specifications that seek databases and information on corporate intranets and other shared corporate resources. However, many databases and other corporate information systems contain redundant data. If counsel or a suitable company representative has a clear, detailed picture of the content and format of client data, staff can negotiate significant modifications that can substantially limit the burden and cost to a producing party. To facilitate these discussions, the Division's Litigation Support staff developed a document titled "Questions about Electronic Systems and Back-up/Archiving Policies," which staff routinely send in order to highlight issues and topics that we would like to address during our negotiations (Attachment 2). This information typically resides not with the corporate IT departments, but with the employees who maintain and/or use the database or other corporate information systems. For large companies, multiple employees may need to be consulted and/or participate in negotiations with staff.

It is impossible to overemphasize the level of detail required for these negotiations to be fruitful. Counsel and/or corporate representatives should be prepared to discuss issues such as: every field of information captured, how the underlying data is collected and formatted, the frequency of collection, what software is used, the size of the data set, what reports are routinely generated from that database, etc. Staff recommends that counsel provide any written documentation and/or training materials to staff in advance of any discussion so that staff can educate ourselves about these corporate resources. Typically, staff negotiates the production of a subset of the data from a subset of the available databases. The more counsel can provide an accurate and complete information about corporate databases and other shared resource, the more staff may be willing to limit production. This approach will also limit the risk of misunderstandings and disagreements between staff and counsel.

C. Preference for Company Data Versus Interrogatory Responses

As staff has grown more comfortable negotiating electronic productions that comprise selections of corporate databases or other information, staff, particularly economic staff, has begun to prefer the underlying corporate data as opposed to interrogatory responses. Staff has more confidence in the underlying company data because it is in the form that furthers the company's business purpose and can be analyzed by economic staff.[4] Using the underlying data, the Division's economists can focus on understanding and using the data as opposed to spending time with counsel trying to understand how interrogatory answers were put together from that underlying data. This changing preference presents an opportunity for counsel to negotiate modifications to interrogatories or produce underlying data rather than having counsel or the

client compile the data. Staff hopes that this is an area where counsel can limit the burden and expense on their client.

D. Direct Access to Corporate Information Systems

In some instances, staff will forego production of corporate data or information if counsel can provide electronic access to the underlying data. For example, in the past staff has negotiated direct electronic access to due diligence material, corporate intranets, and internal databases. Obviously, this kind of access raises a number of issues that must be carefully negotiated *(e.g.*, providing access that will not alter the underlying content, how to produce subsets of the material, etc.), but it can provide an inexpensive and fast way to provide access to large amounts of data. Limited direct access can also provide a means to resolve a disagreement between Division staff and counsel about the value of certain corporate information. The Division's electronic discovery working group has a draft template for direct access (*See* Attachment 3). The draft template addresses issues including the availability of technical assistance, the ability to print or otherwise retrieve the data, the number of log-ins the company should provide, assurances that staff's activities will not be tracked, that underlying data will not be removed without agreement of the staff, and, most importantly, continued access throughout the entire course of the investigation and any subsequent litigation.[(5)] We are acutely aware that all these issues must be addressed at the outset, rather than"fixed" at a later date. In even more limited instances, when providing direct access to corporate resources has proved unworkable, staff has agreed to submit an agreed-upon set of queries to the corporation so that reports can be generated. Obviously, staff uses this option only as a last resort.

E. Detailed Guidance About Production Format and Required Metadata

The Division's Litigation Support group has developed extremely detailed guidance about the format in which electronic productions should be made. *See* Summation Submission Requirements and Summation Database Specifications (*See* Attachment 4). Litigation Support has also developed a list of metadata fields that we typically request from electronic productions. *See* Metadata and Family Records Instructions (*See* Attachment 4). These documents provide guidance to staff and to the parties. Although they are subject to further negotiation, staff has have found that these specifications and fields should be readily available and do not impose extraordinary burdens on the producing party. Like our other templates and guidance, the list of metadata and family record information is routinely updated based on our experiences. Staff encourages parties to submit electronic files that contain TIFF and TXT records that can be loaded directly into Summation, which is the staff's primary tool for reviewing electronic productions. Although staff generally requests TIFF and TXT files for most electronic records, we also request production of native files for certain types of documents (typically powerpoint presentations and formulas in excel spreadsheets) in which the native file preserves formatting necessary to understand the document.[(6)] Staff hopes that by providing this guidance it assists counsel and their clients as they collect responsive materials. Staff understands that counsel needs early guidance on these issues because once material is collected and processed, it becomes much more difficult and costly to change the production. While these documents may seem very detailed, staff believes that productions that deviate from these specifications are more

difficult to load and review, which ultimately hinders staff's ability to review the documents and, ultimately, bring the investigation to a conclusion.

F. Summary and Conclusion

One overall theme that applies to all of these issues and, indeed, to the topic of e-discovery as a whole, is the need for frequent ongoing dialogue about everything surrounding production of all electronically stored information. Electronic productions that have gone most smoothly have involved frequent communication with counsel, company representatives, and the company's electronic production vendor. Division staff is receptive to discussing issues as they arise. One typical example involves search terms. As you know, some Division staff negotiates search terms. In some instances, staff requests words or phrases that greatly inflate the universe of responsive documents. You should feel free to raise issues in these circumstances. Staff, including Litigation Support, is extremely comfortable discussing a broad range of issues with vendors and making changes based upon the results the vendor sees in processing corporate information. To that end, I would encourage you to run samples, test your production and discuss the results with Division staff.

Electronic production issues will continue to evolve. Staff will continue to evaluate its practices to ensure that an appropriate balance is struck between our need to complete our investigations and the cost and disruption to the corporations who receive our civil process.

FOOTNOTES

1. The views presented in this paper are my own, they do not reflect those of the Department of Justice or the Antitrust Division.

2. A terabyte equals one trillion bytes of information.

3. If files are deduplicated horizontally and vertically, it may not be possible to determine which custodians had retained the particular document in their files if their file copy is the one eliminated from the production. To address this issue, staff requires the producing party to provide custodial metadata from the eliminated duplicates so that staff can still determine which company employees had access to a particular document.

4. The underlying corporate data may also be more easily admitted into evidence in the event of litigation.

5. Alternatively, temporary direct access to corporate databases or other information can be a first step in negotiating the production of a portion of a database or a selection of corporate reports. For example, staff could obtain temporary access to due diligence material and select a portion of the reports or information that counsel will ultimately produce.

6. While staff routinely agrees to accept parts of an electronic production in native formats, it remains extremely difficult for the staff to accept an entire production in native format. There are a variety of issues that create problems for us, including the small universe of software that the

Division has licensed for its staff compared to the number of applications used in corporate environments and the lack of a document identifier. Even more difficult, many corporations have designed a variety of proprietary applications that cannot be easily produced. Staff Division generally cannot agree to accept productions in unfamiliar formats that may be effectively inaccessible.