Saving Your Eyes and Sparing Your Memory:

Developments in Privilege Log Review and Implications for Log Preparation

BY BRENT MARSHALL AND ROBERT DRABA

IBM's *Jeopardy*-playing computer, Watson, has brought computer analysis of language back into popular attention. Since then, the use of computers to replace "armies of expensive lawyers" reviewing documents for discovery has made front page of *The New York Times*.[1] Everyone got to read about what e-discovery counsel have known for some time: Computers are playing an ever-increasing role. In addition to scanning for responsiveness, computers are also scanning for indications of privilege, even extracting information from metadata and putting it into privilege logs. But what about the other end of the process? Can computers help with privilege log review?

Of course, at the outset, one might question what difference it makes—whether reviewing privilege logs is worth doing in the first place. Several factors indicate that the answer is yes. First, documents on the privilege log are presumptively relevant and potentially significant to their subject matter. By placing a document on the log, a party implicitly represents that it is responsive to the request, and such communications are typically among persons with significant roles for the companies involved. Second, experience indicates that many privilege claims are not justified, that counsel often broadly interpret privileges in ways not supported by the case law. One sees this explicitly with inadvertently produced documents that counsel ask to have returned after they have been reviewed, as well as with taint review of seized documents in criminal investigations. Moreover, federal judges at e-discovery conferences have remarked that only about 10 percent of documents listed on privilege logs are actually privileged. Third, experience suggests that initial privilege review is often performed by contract personnel with comparatively little knowledge of the situations who consequently tend to mark as privileged anything mentioning an attorney, even as a cc or bcc recipient. This is illustrated by federal and state government personnel appearing on private-party privilege logs and by the apparent redaction of legal notices/disclaimers in signature blocks of email messages. Fourth, questions and challenges routinely result in opposing counsel, on the re-review of documents, withdrawing certain claims of privilege and producing additional documents.

For these reasons, the U.S. Department of Justice Antitrust Division reviews privilege logs, and—returning to the opening question—the division has employed computers in these reviews. This article provides a personal perspective on what has been learned from this experience and synthesizes this learning into a general approach. Reduced to its essence, this approach can be expressed in two principles, a mere eight words:

1. Focus on the persons.
2. Use a relational database.

To make that understandable and useful, however, we'll start with obstacles to efficient review and then propose a solution that addresses key challenges of large logs. Finally, building on this proposal, we'll outline an approach for reviewing large privilege logs. While this approach grows out of antitrust merger investigations, it applies to antitrust and other complex litigation generally.

Key Obstacles to Efficient Review

Why are privilege logs not reviewed efficiently? Reasons vary, of course. Maybe reviewing the privilege log is deemed less important—or at least less urgent—than other tasks, so as the team gets busier, privilege-log review does not make it high enough on the task list. When the task does get assigned, maybe the recipient is a junior attorney who lacks sufficient experience to review the log efficiently and is simultaneously balancing other responsibilities. Even if the task moves up the priority list and is assigned to someone with experience in reviewing privilege logs, another significant challenge remains: the size of the logs.

Over the past several years, the antitrust division has intensively reviewed the privilege logs in a number of major matters. These included many large logs, including logs with tens of thousands

of entries. It is worth noting that the entry count in a number of these logs would have been substantially higher if, as required by the document-request instructions consistent with applicable law, attachments had been listed separately.

At least three factors work against efficient review of such logs. First, the sheer size of the logs is an obstacle. Many privilege logs are not just large, but positively titanic. Some would fill multiple boxes of paper—how is that for a mind-numbing review session?

Not only are they large now, privilege logs are likely to stay large. Underneath is an electronic world in which documents are being created at increasing rates as computers, smartphones, and ubiquitous email turn most everything we do into electronic records. Not that much gets deleted. It is easier for documents to accumulate in this electronic world. Hard disks in the server room are out of sight and out of mind, and adding hard-disk space or other electronic storage is relatively simple and easy compared to finding space for more file cabinets, file rooms, or even file warehouses. Thus, we begin with a larger and still-growing base of documents. Add broad document requests characteristic of modern investigations and litigation and long list of hits from broad term searches, and the universe of potentially responsive documents grows larger and larger. The significant risks of disclosing privileged documents and the limited downside to making dubious privilege claims also contributes to longer logs.

Second, the structure of the privilege logs is an obstacle. The antitrust division typically receives privilege logs in two parts: the main document log and a separate name list. The document log includes document date, description, authors, recipients, and privilege type. The name list states the title and company of each person listed in the document log as an author or recipient. The instructions in our document requests typically ask for this division, replacing an older approach in which title and company information was included in the document log with every author and recipient name. Imagine how much larger and more cumbersome reviewing large paper logs would be if every name also had a title and company!

Yet separating the document and name listings has its own consequences, for one needs to tie that information together to analyze the privilege claims. A critical part of analyzing claims is to look at the authors and recipients and ask, "Who are these persons?" Answering that question requires linking author and recipient names in the document log with the corresponding title and company information in the name list. How is that most frequently done? Eyeball and memory—look at one list, look at the other, and try to remember enough information about enough persons to keep moving at a reasonable pace.

The weakness of this approach becomes readily apparent when one considers the number of persons and companies on the name lists. Figures 1 and 2 show the number of authors/recipients and the corresponding number of companies contained in a number of logs that the division has received and reviewed. Notice how often there are more than 2,000 names. This leads to another aspect of the problem.

Third, working memory is an obstacle. Humans use working memory to temporarily store and use information for learning, reasoning, or other such processing.[2] Here, the demands on working memory are tremendous because the amount of data is large—hundreds of companies and thousands of names. Yet the capacity of working memory is limited: Humans can maintain a reasonable focus on only so many details at a time. At least for those of us lacking photographic memories, eyeball and memory simply will not do. Another solution needs to be found.

Proposed Solution: Computer-Assisted Review

An efficient approach to privilege log review needs to address these challenges: It needs to handle large numbers of privilege claims, to link author/recipient names to associated titles and companies, and to do this in a manner that does not overload our working memories. One solution is computer-assisted review using relational databases. With this approach, the computer applies the same sort of decision rules[3] that are typically applied in the course of manual review, i.e., eyeball review.

How the Benefits Are Obtained

Computers have several key advantages over human reviewers. They are *fast*. Entry-by-entry manual review takes too long with large logs. A key aspect of the problem is that human reviewers can keep only a limited number of decision rules in mind at a time—working memory again—and thus need to move slowly, make multiple passes, or risk missing significant details. In contrast, computers can apply a decision rule to thousands of records in seconds. Further, computers are *consistent*. They do not miss details as human reviewers can when their eyes glaze over as they get tired or as they juggle multiple assignments at once. Computers apply decision rules

Figure 1: Entries in the Name Lists of 10 Recent Logs

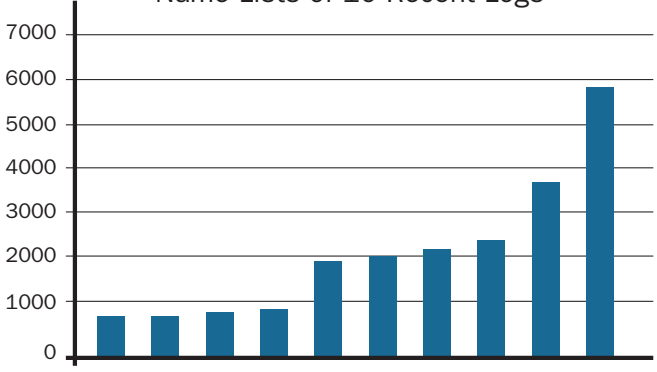
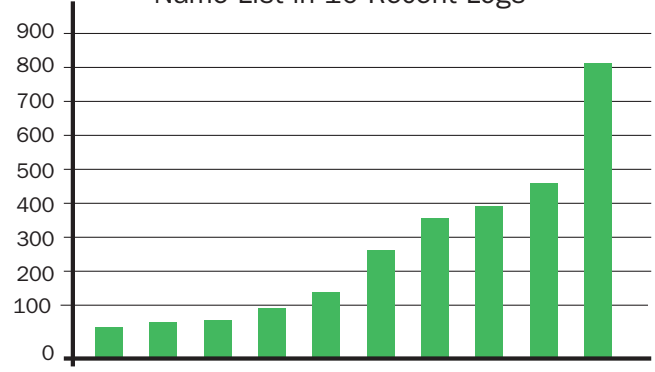


Figure 2: Count of Companies in the Name List in 10 Recent Logs



the same way every time. Finally, computers are *comprehensive*. They miss nothing and find everything specified in the decision rule. Of course, the rules must be formed well, just as our LEXIS and Westlaw searches must be formed well.

Relational databases build on these advantages. In brief, they "relate," or link, multiple tables of information (think different spreadsheets) based on data elements occurring in both tables. Putting the data into this form (sometimes called "normalizing" the database) will be handled by the IT department or other staff that set up the database. Once that has been done, the software dynamically links the information in the document log and the name list, joining it as work proceeds so that reviewers do not have to try to keep it all in working memory. This lets reviewers search the data together, view it together, and report it together and thus enables more robust and flexible decision rules. Ultimately, it provides greater control of the information in the privilege logs.

How the Work Is Divided

Some months ago, when discussing this approach with a colleague, he asked whether this approach could be made available as a turn-key, black-box system into which one would feed a privilege log and out of which would come lists of challenges and other ready-to-use output. As nice as that might be, that is not the case at this time—at least, not without a budget like IBM had for Watson. There is still important work for legal staff, which is why this is called computer-*assisted* review.

First, legal staff must interpret the information in the privilege logs. Computers can read and parse the language and contents, but they do not actually understand what is there. Legal staff must derive the meaning from the log entries. Second, based on what is found, legal staff must set appropriate decision rules. These are principles of the form *if <CRITERIA>, then <ACTION>*—when certain conditions or characteristics are recognized, then certain actions will be taken. Possible decision rules could be the following:

- *If* an author or recipient is listed as employed by a company viewed as a third party, *then* the privilege claim will be challenged as waived.
- *If* an author or recipient is not found in the name list, *then* flag the privilege claim and the recipient name for inquiry with the producing party.

Decision rules for privilege logs arise from substantive law and from experience, and they can vary from one log to another. Once one passes beyond the basics, creating the rules requires legal judgment and thus must be done by or under the direction of attorneys.

The role of the computer is to facilitate these tasks, and it does so in several ways. First, the computer provides a mechanism for running the decision rules, applying them to the log entries. This is what will come to mind for most persons when one says "computer-assisted review." Second, not only does the computer apply the decision rules, it can help us create them. The computer can produce data about the information contained in the logs—e.g., tables of word counts, lists of problem names, and counts of entire descriptions or parts of descriptions—that will help us identify possible issues and design appropriate decision rules. Third, the computer provides the means to store and use the results. So how might this be done?

Conceptual Overview: The Major Stages

Different persons can implement this approach in different ways. Conceptually, however, the process will involve three basic stages: Prepare and load the logs, analyze the entries, and code and report the results. Regardless of the exact implementation, however, the goals are the same: to identify missing data (e.g., a document's authors and recipients not listed), to flag questionable privilege claims (e.g., request for legal advice not sent to an attorney or apparent third-party recipients of the document), and to identify instances in which the description of the basis for the privilege is insufficient to allow evaluation of the claim (e.g., no litigation identified for a work-product claim).

Preparation

The first stage involves preparation. The key goals are to get the privilege logs into the database while checking them and identifying any major, systemic issues that might be present, e.g., missing and malformed data. As this work will typically be performed by the IT department or other support staff and the steps might vary based on the software used, the loading process need not be discussed further here.

One other aspect of this stage does warrant mention, however. During this effort, one can also gather initial data about the logs. Useful information includes counts of each type of privilege claim, counts of documents linked to each author and recipient, and counts of documents with common descriptions. In addition, one can count and list the words used in the document descriptions that explain the basis for withholding the document. Different logs have their own vocabularies and characteristic ways of describing claims. As a general matter, reviewing this data fosters familiarity with the log and helps build better decision rules. In particular, the word lists reveal the terms that need to be searched to find particular sets of entries, along with the spelling variations that need to be addressed lest the associated log entries be missed when term searches are run.

Analysis

The second stage is to analyze the log entries. This is the core of the process. A robust analysis has three key steps: Look at the privileges claimed, analyze the persons, and analyze the descriptions.

1. Look at the privileges claimed. Are privileges claimed? If so, what are they? One would think that a privilege claim would be a given. However, the antitrust division has received logs in which a number of log entries had no privilege claim. As a matter of law, privilege claims must be made explicitly, and the failure to do so is grounds for immediate objection.

Reviewing the privileges claimed gives a sense of the whole log. It helps one know what to look for in the logs. When there are joint defense claims, there should be grounds for an underlying attorney-client privilege claim. When there are work-product claims, there should be statements regarding the litigation that is anticipated. This orientation also lets us know what legal issues to consider and research, e.g., the law regarding joint defense in the relevant jurisdiction.

2. Analyze the persons. The second step is to analyze the persons: the authors and recipients of the persons identified in each log entry. The analysis of the persons is arguably the single most important part of the privilege log review process. The use of

relational databases—the other point of emphasis—is significant primarily because it facilitates this aspect of the analysis.

Why analyze the persons? The persons are critical to the existence of the privilege. Throughout the life of the privilege, the identities and roles of the persons involved in the communications are of central concern. First, the privilege comes to life only when certain eligible persons are involved in certain eligible activities. Second, a privilege ceases to exist when an ineligible person becomes a part of the communication. Consequently, there is no unimportant author or recipient on a privilege log. *Every person* that is an author or recipient of a purportedly privileged document is critical to the privilege claim. Thus, to know who *each person* is and to understand the relationship that establishes the right by which *each* is party to a privileged communication is a critical part of privilege log review.

As an aside, the critical role of the persons should give pause as to the proposed use of categorical claims in privilege logs and the use of sampling to test claims.[4] Categories are suited more toward describing the general nature of a set of documents and justifications related to the subject matter of those documents. However, a particular document is not privileged merely because it is of the same general nature and subject as privileged documents: The particular communication of which a particular document is a part must involve an attorney whose legal advice is being sought and must not involve third parties whose involvement would waive any privilege—that is black-letter law. Thus, for one party to substantiate a privilege claim and for the other party to test the claim, a listing of that particular document's authors and recipients seems essential.

Sampling presents a related issue. Given the nature of statistics, sampling is much more suited for demonstrating the existence of problems with a log than demonstrating the absence of problems. One intuitively recognizes that significant problems with a sample demonstrate a probability of systemic issues that warrants more intensive review and analysis. In contrast a limited number of problems in a sample provides less certain evidence that claims presented in a privilege log are generally justified. Similarly, while sampling can provide a reasonable level of confidence that tasks have been performed well, when tasks have not been performed well, sampling is weak at showing the exact nature and scope of the deficiencies. However, further discussion of these ideas must be left for another day.

The need for a careful analysis of the persons who are authors and recipients of documents listed on a privilege log is underscored by several systemic problems observed in some logs the antitrust division has received. Three key problems are unidentified persons, misidentified persons, and misused asterisks (*).

- **Unidentified persons**: An analysis of the persons requires more than each person's name. One needs to know who that person is, and that generally involves stating each person's title and company. When the person is not part of the entity to whom the privilege belongs, the relationship of that person to the owner of the privilege needs to be stated. When a person has multiple roles, either simultaneously or successively, during the relevant time period, these roles need to be distinguished. For example, since attorney-client privilege applies only to legal advice, not business advice, when a person simultaneously holds both business and legal positions, both roles need to be identified, and log entries need to state clearly the role in which that person is involved in each communication. Similarly, when a person moves from a legal position to a business position, the date of the change needs to known so that documents can be associated with the correct role. This information needs to be made available, and it needs to be made available in a form in which it can be readily associated with the correct persons.[5]

  This identifying information is often not provided, however. Experience indicates that high proportions of unidentified persons in privilege logs are not unusual. Who are these persons? Are these instances of third-party waiver? Without identifying information, how can one tell? How is one to judge?
- **Misidentified persons:** Similar problems occur when persons are not identified completely and consistently: The computer then cannot uniquely link a particular author/recipient name to the name list. One set of issues relates to the name list: The same name occurs more than once, or quite similar names appear and the circumstances suggest that the same person is intended. Reversed names, nicknames, middle initials, maiden names, and foreign names can pose issues. For example, consider this brief name list:

Brown, Shirley Williams
Doe, John
John, Doe
Jones, Alan
Jones, Alan E.
Smith, Bill
Smith, William
Williams, Shirley
Williams-Brown, Shirley

How many persons are represented here? When titles and companies seem to match, one begins to suspect that this is fewer than nine persons, possibly as few as four. In reviewing name lists, the antitrust division commonly finds multiple entries with the exact same name and scores of entries involving name variations. If reviewers are to associate documents with the right persons, these issues need to be identified and corrected.

Another set of issues relates to the manner in which author/recipient names are entered in the privilege log. In some logs only last name and first initial are used in the document log to identify authors and recipients. This poses particular trouble when multiple persons with that name and first initial appear in the name list. For example, does "Brown, J" refer to James Brown, Janet Brown, Janet M. Brown, Jennifer Brown, or Judy Brown? One cannot tell, and thus the author or recipient cannot be identified. A much more frequent problem is author/recipient names being entered in multiple ways. Experience indicates that logs often involve many instances of persons appearing with five, 10, even 15 or more variants. Even high-ranking corporate officers, presumably well-known to legal staff, have been spelled incorrectly, not just in the main log entries but even on the name list. Again, such issues inhibit identification of the persons involved.

Some may doubt the significance of these sorts of problems involving unidentified or misidentified persons. First, some may think that, even though there may be a problem for a computer, to the extent humans can recognize the variations, there is no problem. Yet when reviewing logs with tens of thousands of log

entries involving thousands of names at hundreds of companies, working memory is overwhelmed. Human eyeball and memory will not do. Second, some may think that these are simply the unavoidable outcome of data entry, typos that are just part of the burden that privilege log reviewers have to shoulder. Yet as discussed at the end of this article, the inability to identify these persons does not pose problems merely for reviewers. It also affects the persons preparing the logs because it implicates the basis for asserting a privilege claim in the first place. Further, a quick computer check can identify the vast majority of such instances, and it seems appropriate that they be dealt with by the producing party, which has much greater knowledge of the persons involved.[6]

- **Misused Asterisks:** In the logs the antitrust division receives, asterisks have been used to identify an author or recipient who is a licensed attorney and who is acting as an attorney with respect to that particular communication requesting or providing legal advice. The division has encountered several issues involving these asterisks.

  First, asterisks have been found appearing after the names of a great many business executives, finance officers, product managers, chief engineers, and industry consultants. Their titles indicate that these persons were not, in fact, licensed attorneys acting as attorneys with respect to the communications as to which they were authors or recipients. If not, then the asterisks were not properly placed after the names of those persons.

  Second, the division has encountered repeated issues involving lawyers holding nonlegal positions. One variant is the situation in which a lawyer moves from a legal position to a nonlegal position; for example, a lawyer in the general counsel's office changes position and becomes a vice president for business development. This is especially significant when the change occurs during the period of time for which documents are being produced. Asterisks have been found appearing after a person's name even after the person was no longer employed in a legal capacity. The other variant is when a person simultaneously holds both business and legal positions; for example, a person might be both general counsel and a vice president for business development. Whether these legal and nonlegal positions are consecutive or concurrent, the logs need to make these positions clear, along with the dates of changes of positions. When such a person's name appears as an author or recipient, persons preparing logs need to ensure that the entry is not marked with an asterisk simply as a matter of course; rather, they need to determine whether the person is, in fact, acting as an attorney with respect to that particular communication. Put differently, the decision to place an asterisk must be made on a document-by-document basis. This leads to the final issue.

  Third, the division has found clear evidence of the use of global search-and-replace functions to put asterisks in logs. Globally adding an asterisk after instances of a person's name in the privilege log seems hardly consistent with the fact that not every communication with a member of the bar is privileged and with the responsibility to judge document by document whether the lawyer is acting as an attorney with respect to the communication in each document.

Each of these problem categories involving unidentified persons, misidentified persons, and misused asterisks indicate practices that persons preparing privilege logs should avoid and that persons reviewing logs should detect. We reiterate, however, that the greater purpose of describing them was to emphasize the need for a careful, rigorous analysis of the persons listed as parties to the communications identified in the logs. These problems go to the heart of basic privilege log review, which involves a well-worn path. Persons new to privilege log review are quickly taught that third-party waiver and "no attorney" are the low-hanging fruit, two of the most useful lines of analysis and attack and thus two things that need to be checked carefully. Yet a reviewer cannot analyze third-party waiver without knowing who each person is. Similarly, checking for "no attorney" is typically done by searching for asterisks, which involves depending on their proper usage. Yet proper usage cannot be verified without knowing who each person is.

To do this in any sort of rigorous, comprehensive manner, the reviewer must have command of the persons. Put differently, this basic privilege log analysis requires that the reviewer, for each author and recipient of each document on the log, to be able to answer this key question: "Who is this person, and by what right do they communicate privileged information?" Yet because there are too many names to handle via eyeball and memory, one needs the computer. Specifically, one's methodology must give a command of the persons, and one's tools must support such a methodology. This is what impels the use of relational databases.

So what is involved? At a high level, the task is to *classify* the persons listed in the log by whether they are, for example, company personnel, merger counsel, other lawyers, or other third parties. One approach is to list the companies identified in the name list (one's software tools should do this readily), to classify the companies using one's existing knowledge supplemented by basic research (Internet searches are quite useful here), and finally to use these company classifications to begin to classify the persons associated with them. As part of this last step, these classifications can be used along with information provided in the name list, especially the titles, to provide a basis for judging which persons are attorneys.

This effort enables some significant lines of analysis. Of course, one can do basic checks for third-party waiver and documents with no attorney listed, but the power of a relational database is that this can be done far more robustly. For example, a single database query can find *every* document sent or received by an apparent third party … at once! For each of these documents, the query can, at the same time, identify each apparent third party upon which a claim of apparent waiver might rest. Similarly, one can validate the asterisks used in entries in the name list and in author/recipient entries of the document log and thus help determine whether all persons marked with asterisks are actually attorneys. A single query can then find all documents with an author or recipient who is not an attorney and yet who has been marked with an asterisk.

Yet the relational database enables additional categories of analysis of the persons. For example, one can identify all documents sent or received by unidentified persons, i.e., persons not included in the name list and thus as to which there is no title and company information. These documents warrant attention, for the unidentified persons might be third parties, which would imply a waiver of any attorney-client privilege. One can also count each document's recipients (attorney, nonattorney, and total) and thus identify documents with more than a certain number of recipients and documents with more nonattorney than attorney recipients. For example, too large

a number of recipients could indicate that the information was not confidential. Finally, one can identify documents sent to groups or lists of persons, which might waive any privilege or might imply that the information is not confidential, *especially in those instances in which the members of the group or list is not known*.

In sum, based on this approach, one can know who has been identified and who has not been identified, and one has some checks on the identification of attorneys. One can link persons and companies with documents working in either direction, i.e., either documents to persons/companies or persons/companies to documents. In other words, one knows and has a good measure of control over the persons named in the log.

3. Analyze the Descriptions. One is now ready to begin analyzing the descriptions contained in the log. While the analysis of persons tends to focus on the general eligibility of persons with respect to privilege, the analysis of the descriptions tends to focus on the specific basis of a privilege for this communication. In the language of the Federal Rules of Civil Procedure, the description should "describe the nature of the documents, communications, or tangible things not produced or disclosed—and do so in a manner that, without revealing information itself privileged or protected, will enable other parties to assess the claim."[7] The purpose is to "provide sufficient information to enable other parties to evaluate the applicability of the claimed privilege or protection."[8] Thus, a threshold issue is whether the description provides enough information to assess the claim. Suppose, for example, that the vast majority of the descriptions have one of six or seven entries along the lines of "email concerning legal advice regarding contract issues" or "email concerning legal advice regarding regulatory issues." One could reasonably conclude that this does not contain adequate information, and thus using the computer to identify this and similar issues would be helpful to the review process. For descriptions that do contain adequate information, more detailed analysis can then be performed.

A review of the logs submitted to the antitrust division reveals that descriptions tend to have a consistent structure with similar vocabulary. This is not surprising given the tools used to facilitate the process of performing privilege reviews and preparing privilege logs. Of course, there are notable differences and variations, which complicate the analysis, yet the same consistency that helps those preparing the logs also helps those analyzing them. While other divisions are possible,[9] the descriptions can typically be divided into two basic parts: the purpose and subject of the communication. These two, along with the privileges claimed, provide the three primary pieces of information upon which the analysis of the descriptions rests.

First, one needs to analyze the language used in the descriptions. This involves reading the descriptions. Because descriptions often recur, having the computer prepare a sorted list of unique descriptions, with counts, simplifies this step: One can read the list much faster than the entire log, especially when the log fills multiple boxes. The computer can help in two additional ways. First, the computer can easily prepare basic *word lists* of what words are used in the descriptions and how many times each occurs. This helps one better understand the terminology and other language used in this particular log. Second, the computer can frequently parse and divide the entries into its key parts, e.g., dividing the purpose and subject. This enables more focused analysis.

Second, one needs to design and validate appropriate decision rules of the form described previously. These rules are built on searches, such as those familiar from the legal research context. As with LEXIS and Westlaw searches, one's search terms must account for alternative terminology, e.g., "advice of counsel" versus "advice of an attorney." One must also consider the effect of misspelled words. A human will, upon seeing "provice" in context, readily recognize that "provide" was intended, while a computer will not do that so readily. Word lists are useful in identifying instances in which misspellings need to be added to the search terms in decision rules to help them be sufficiently comprehensive. Once initial rules have been created, they need to be run and tested for under-inclusiveness and over-inclusiveness: One wants to determine the degree to which the rules are catching records that are unwanted and missing records that are important. This requires some level of eyeball review and analysis of the false positives and missing records, and based on that work, one can refine the decision rules as needed.

Such decision rules enable much useful analysis. Looking just at the descriptions, one can identify work-product claims that fail to identify the anticipated litigation, documents that apparently should have been produced in redacted form rather than being withheld, and other vague or otherwise insufficient descriptions. For example, the division has received logs in which many thousands of descriptions contain vague identifiers like "Contract Issues" or "Regulatory Issues" as the only indication of the subject matter of the communication. Other logs have contained many descriptions with insufficient statements like "pending review by counsel" or "subject to legal review" as the basis for the attorney-client privilege claim. The computer can readily identify all claims whose descriptions have such characteristics.

Of course, such decision rules need not consider the description alone. Quite powerful analysis can be performed with rules that consider both the description and the authors and recipients of the communication. For example, one can identify requests for legal advice that were not sent *to* an attorney or that were sent to more nonattorneys than attorneys. Similarly, one can identify legal advice not sent *by* an attorney.

Code and Report

The final step is to code and report appropriate privilege claims. The codes applied to specific privilege claims identify claims that have particular characteristics. Codes can have a range of uses. Some, such as whether there is an attorney author, are not determinative by themselves but are used for further analysis. Others, such as whether the description refers to a particular topic of interest, might be used to flag a claim for further eyeball review. Still other codes, such as the presence of a third-party recipient, can be used to indicate grounds for an immediate challenge.

Based on these codes, one can readily prepare formatted reports listing all documents with a chosen set of codes (characteristics). For example, one report might be a formatted listing of every document sent or received by an apparent third party and, separately for each document, listing each individual upon which a claim of apparent waiver might rest. Another report might be a formatted list of each document with an author or recipient who appears not to be an attorney and yet who has been marked with an asterisk.

These reports may be used internally, such as to assign someone to review a particular subset of claims more carefully. They can also be used with opposing counsel, either to request further information

about certain claims or to challenge them. The reports range from simple reports that list only the Bates numbers or other identifiers of the selected documents to complex reports that contain additional information extracted from the logs. These reports facilitate further analysis and discussion and thus faster resolution of disputes as to the privilege claims.

Recurring Issues: Progress To Be Made

Notwithstanding the major improvement that this approach affords, issues remain. They are most aptly summarized by the well-known aphorism from the computer world: *garbage in, garbage out*. Problems with the information provided (or not provided) in the privilege logs will have a continuing effect on the analysis of privilege claims. Three categories of problems warrant brief mention.

First, privilege logs often have *missing information*. For example, the division has received a number of logs in which the privilege claims for redacted documents had some document identifier but not the Bates number of the produced documents. Without that information, one cannot link the privilege claims with the respective documents. Some missing information points back to the document requests. Counsel need to recognize that work on privilege issues begins with drafting the document request and the associated instructions. Instead of relying simply on the relevant rules and cases, counsel should specify the information that they want to see provided in the logs.

Second, privilege logs often pose *loading issues*, especially due to inconsistent formatting. A number of years ago, as increasing numbers of documents began to be produced in electronic, rather than paper, formats, it became necessary—and common—to negotiate electronic production formats for responsive documents. The time has come for similar discussions as to the electronic production formats for privilege logs. Even within Excel format (the most common option), a host of issues can arise, including the division of information into fields, the use of delimiters within fields, and how numbers and dates are entered. Uniform formats would be useful, but in the interim, early discussions are needed.

Third, privilege logs often display *poor quality*. Common problems include misspelled names, other failures to identify persons consistently, and omitting authors and recipients from the name list. Many approach this as a simple matter of typos, an inevitable burden to be borne by the receiving party. However, these issues impede the analysis of claims and thus implicate more substantive issues.

Moreover, these issues do not affect the receiving party alone. The analysis on both sides is impeded. Specifically, if the information produced in a privilege log is not adequate for the receiving party to analyze, then it seems doubtful that the producing party can analyze it either. Because withholding for privilege is an exception to the general rule of production, to withhold an entire document or a portion thereof through a redaction, a producing party must determine that a privilege exists and has not been waived. Yet the producing party's determination depends on the same information that the receiving party needs to examine the asserted privilege claims: *These respective efforts are two sides of the same coin*. If the information is insufficient for analysis, then to what extent does the producing party have good-faith grounds for asserting a claim of privilege?

All sides need to ponder these issues. Hopefully, just as electronic production of responsive documents is improving with time and experience, the same will prove true of privilege logs.

Conclusion

While large privilege logs may not be pleasant, they are manageable with computer-assisted review. The application of three principles will optimize the results of such a review. First, the persons are critical, and thus the use of relational databases is not merely desirable, but essential. Second, work on privilege issues begins, not when the privilege log arrives, but when drafting the document request, and it continues while negotiating the request and discussing the production. Third, large numbers of seemingly minor errors in the privilege logs have substantive implications for both sides. ⊙

Brent Marshall and Robert Draba are trial attorneys in the antitrust division of the U.S. Department of Justice. The views expressed herein are those of the authors and do not necessarily reflect those of the antitrust division or the Department of Justice.

Endnotes

[1]John Markoff, *Armies of Expensive Lawyers, Replaced by Cheaper Software*, N.Y. Times, Mar. 5, 2011, at A1.

[2]For a useful overview of working memory, *see* Alan Baddeley, "Working memory," *Scholarpedia*, 5(2):3015 (2010) (*available at* www.scholarpedia.org/article/Working_memory).

[3]Decision rules are principles of the form *if <CRITERIA>, then <ACTION>*—when certain conditions or characteristics are recognized, then certain actions will be taken. They are discussed further below.

[4]*E.g.*, Hon. John M. Facciola and Jonathan M. Redgrave, *Asserting and Challenging Privilege Claims in Modern Litigation: The Facciola-Redgrave Framework*, 4 Fed. Cts. L. Rev. 19 (2009).

[5]A standard form for documenting such multiple roles would be a most useful development.

[6]Presumably, the producing party should know the identity of these persons. The distribution of information to unknown persons is rather inconsistent with the notion of confidentiality.

[7]Fed. R. Civ. P. 26(b)(5)(A)(ii).

[8]Fed. R. Civ. P. 26(b)(5), 1993 Advisory Committee Notes. *See generally Victor Stanley Inc. v. Creative Pipe Inc.*, 250 F.R.D. 251, 264-65 (D. Md. 2008).

[9]A more granular way of dividing the entries is what type of communication is involved (e.g., email, memorandum, or presentation), whether the document is a draft, why the communication occurred (e.g., requesting legal advice or providing legal advice), what the nature of the issue is (e.g., contractual obligations, corporate organization, or regulatory issues), and what the subject matter is. A standard for providing this information in a parsed form would be another useful development.