Turning Obscure Bits Of Data Into Hard Evidence: A Proposal For The Unorthodox Use Of A Document Request To Capture System Metadata

Nolan M. Goldberg and Scott M. Cohen

PROSKAUER ROSE LLP

Nolan M. Goldberg

Scott M. Cohen

A weakness in the discovery process is that it is difficult to determine when a party has not met its obligation to voluntarily produce harmful documents. When it is available, system metadata can solve this problem by providing a different view of electronic evidence that is not readily susceptible to manipulation or concealment.

The term "metadata" is typically used to describe information automatically included in substantive electronic files by application programs such as Microsoft Word or Outlook. Of course, this information is of no help if the file in which the metadata is embedded is not produced. It is often overlooked that applications, operating systems, and file systems also generate numerous "system" metadata artifacts, located in dedicated files separate from the substantive files to which the metadata may relate, which, despite any concealment efforts, can reveal the existence of a file that has not been produced, or even that file's contents. Significantly, system metadata is particularly valuable in discovery as many of these metafiles almost always exist and are not easily manipulated or erased without special tools. Even when these tools are used, they typically leave detectable marks flagging the destruction. *See* Brian Carrier, FILE SYSTEM FORENSIC ANALYSIS 198 (Pearson Education 2007). Should this be discovered, the target is potentially in trouble regardless of whether the actual metadata or underlying substantive files can ever be recovered.

System metafiles can also themselves have substantive evidentiary value, or provide information useful to the authentication of other documents. See *Lorraine v. Markel American Insurance, Co*., 241 F.R.D. 534 (D.Md. 2007). ("[B]ecause metadata shows the date, time and identity of the creator of an electronic record, as well as all changes made to it, metadata is a distinctive characteristic of all electronic evidence that can be used to authenticate it under Rule 901(b)(4).")

System metadata typically enters the discovery process as a result of a request to enter and inspect under FRCP 34(a)(2). However, such requests are infrequently granted in the absence of evidence of discovery misconduct, largely minimizing the historic use of system metadata in discovery. *See* FRCP 34(a), Advisory Committee Notes to the Dec. 1, 2006 Amendment (there is no "routine right of direct access to a party's electronic information system"). Accordingly, a paradox exists where in order to collect some of the most relevant evidence of discovery misconduct you need to already have substantial evidence in your possession. Additionally, while inspections of computers that are central to a dispute have been allowed in limited circumstances, the general need for system metadata as substantive evidence has not yet been held to alone be grounds for an inspection, effectively rendering useful evidence off-limits. Gaining easier access to this untapped evidence source within the context of existing rules of discovery will open new investigative possibilities.

Instead of routine abandonment, system metadata can instead be requested using a basic FRCP 34(b)(1)(C) document request. System metadata is stored as dedicated files, thus falling within the definition of a document under the Rules. *See* FRCP 34(a), Advisory Committee Notes to the Dec. 1, 2006 Amendment ("Rule 34(a)(1) is expansive and includes any type of information that is stored electronically"). Additionally, the standardized nature of the names and locations of these files makes them relatively simple to "describe with reasonable particularity." *See* FRCP 34(b)(1)(A). Accordingly, assuming the general requirements of FRCP 26 can be met, there seems to be no reason why such files cannot be the subject of a properly crafted document request.

The primary considerations of FRCP 26 are relevance and burden. Clearly, blanket requests calling for the production of all system metadata are, more likely than not, inappropriate. *See e.g., The Sedona Principles: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, Second Edition p. 4 (*The Sedona Conference® Working Group Series*, 2007) ("In most cases … metadata will have no material evidentiary value"). However, narrow requests can be directed towards the particular file or files that are most likely to contain discoverable data.

With regard to burden, it is true that the extraction of certain system metadata may be more complicated than copying normal files. For example, a given metafile may be routinely in use by the Windows operating system and therefore cannot be copied during normal operation. Certain files may also be hidden. One approach to address any additional burden is to include instructions on how to extract the specific data files requested in the document request. It may be difficult for the party opposing production to maintain a burden objection when there are only a few additional steps to carry out the collection of a narrow subset of documents. This is particularly true as the Rules already contemplate electronic discovery as a partnership between counsel and information systems professionals. *See, e.g., The Sedona Principles,* Second Edition p. 19 (2007) ("The team approach permits an organization to leverage available resources and expertise in ensuring that the organization addresses its preservation and production obligations thoroughly, efficiently and cost-effectively").

Finally, burden objections may arise out of the added complexity of reviewing system metadata for privilege, particularly as many of these metafiles are not decodable without technical know-how or expert assistance. However, because most of these files contain no substantive information, privilege concerns would ordinarily not be implicated. In those circumstances where actual substantive file data may be recovered, the situation may be more complicated, and the producing party may need to employ its own expert to assist in the privilege review. The burden in such situations would need to be evaluated on a case by case basis. *See* FRCP 34(a), Advisory Committee Notes to the Dec. 1, 2006 Amendment ("The requesting party has the burden of showing that its need for the discovery outweighs the burden and costs of locating, retrieving, and producing the information.")

Each operating system, file system or application may generate different system metadata that can be discovered. As an introduction to the possibilities raised by the analysis of this type of data, below are examples of specific system metadata files that may exist in computers using the Microsoft Windows operating system.

The Windows Registry

The Windows registry is a hierarchical database containing all of the options, settings and preferences for a computer running any of the 32-bit versions of the Microsoft Windows operating system including Windows NT, 95, 98, ME, 2000, XP, and Vista. The registry is frequently and automatically updated with information generated as a result of the use of the computer, including a record of logins, network shares accessed, searches performed, applications used, files most recently opened, or connection to a wireless network. For example, an entry is created in the registry when a USB removable storage drive is connected to a computer that includes the identity of the drive and the time of connection. Other entries track opened and saved files, and those files last accessed.

The registry can be copied using the Registry Editor program (REGEDIT.EXE) already installed on all Windows-based computers.

Windows Log Files

Log files, typically identified by the *.log extension, are created by the Windows operating system or other installed programs as a means of recording events which occur during system operation which can be useful in creating a timeline of a computer's use. Log files are most often simple text files though some, like the Windows Event Log, are actually databases. Two items that you can expect to find in most log files are the description of some sort of event, such as the burning of a CD-ROM and the date and/or time on which it occurred. Additionally, IM conversations in some IM environments (Yahoo is one) are frequently logged by default.

Collection of log files can be accomplished via standard collection tools such as Microsoft's Robocopy. One word of caution: many E-Discovery processing tools have been configured to exclude system files. If such a tool is used, its configuration must be modified to include log files.

Temporary Data/Cache Files

Temporary and cache files, created for the purpose of conserving a computer's memory during the editing or processing of data or for recovery purposes, are intended to exist for a finite period of time, but there are many circumstances where they will remain on the hard drive indefinitely.

For example, when a Microsoft Word 2003 document is opened, a hidden temporary file will be created whose name is of the form ~wrdxxxx.tmp, which contains a copy of the viewed document. (In this example, xxxx represents a unique number generated by Word.) If Word is closed in a manner that causes it to lose track of the temporary file it has created, for example, a system crash or program error, then the temporary file remains on the hard drive permanently unless it is manually deleted. Over time many such temporary files can be "orphaned" in this manner.

As another example, Microsoft Internet Explorer uses temporary files to store items downloaded during browsing sessions in order to improve performance during revisits to web pages. If collected during discovery, the names, dates and contents of these files can be used to recreate a user's browsing history.

Most applications use the file extension ".tmp" to indicate a temporary file, and many are stored in one of several standard temporary file locations, such as C:\TEMP and C:\WINDOWS\TEMP. A simple search for all instances of hidden files with a *.tmp suffix, coupled with a normal copy operation, is all that is required for collection. It is important to note that certain types of temporary files, such as those created by Microsoft Word, contain substantive data that needs to be reviewed for privilege.

The above examples are illustrative and will not be relevant or appropriate in every situation. The requesting party should consult a forensic expert to determine what system metadata might exist that could be relevant in a particular case and the methodology for the extraction of those files.

Conclusion

Collecting system metadata in the manner described above is not without its drawbacks, as it violates the generally accepted forensic practice of making a static bit-map copy of a hard drive before harvesting metadata, thereby preserving it for future verification and protecting it from further alteration. Extracting metadata files from a live computer may have the unintended consequence of altering or destroying other metadata. *See e.g*., *Krumwiede v. Brighton Assoc*., Case No. 05 C 3003, 2006 WL 130862at *4 (N.D. Ill. May 8, 2006) (noting that the creation of a forensically valid copy of a hard drive allowed for the examination of the contents of the laptop without "disturbing" it more than necessary). Additionally, unlike a full forensic inspection, deleted files cannot be recovered using this approach. The likelihood of recovery of deleted files decreases the longer a target computer remains active. *See* Linda Volonino, COMPUTER FORENSICS, PRINCIPLES AND PRACTICES 21, 93 (Pearson Education 2007). Finally, certain metadata cannot be recovered outside of a full forensic inspection. Accordingly, collecting metadata through a document request is not a preferred substitute for a proper forensic inspection when such relief is available, but is a good second choice when an inspection cannot be obtained.

The collection of system metadata using document requests is a natural evolution of E-Discovery, providing additional insight to the tech savvy practitioner when preferred means of access to these files are not available. Such an approach can, in the right circumstances, turn obscure bits of data that would have previously not been discovered into case determinative evidence.

***Nolan M. Goldberg** is a Senior Associate in the patent group of New York-based Proskauer Rose LLP and a member of the Litigation Department's E-Discovery Task Force. **Scott M. Cohen** is the Director of Practice Support at Proskauer Rose LLP and a member of the Litigation Department's E-Discovery Task Force.*

Please email the authors at ngoldberg@proskauer.com or scohen@proskauer.com with questions about this article.