

NOTE

THE REQUIREMENT FOR METADATA PRODUCTION UNDER *WILLIAMS V. SPRINT/UNITED MANAGEMENT CO.*: AN UNNECESSARY BURDEN FOR LITIGANTS ENGAGED IN ELECTRONIC DISCOVERY

Lucia Cucu[†]

INTRODUCTION	221
I. BACKGROUND	222
II. <i>WILLIAMS V. SPRINT/UNITED MANAGEMENT CO.</i>	227
III. ANALYSIS AND POLICY IMPLICATIONS OF THE <i>WILLIAMS</i> DECISION	229
A. The High Cost and Difficulty of Producing All Metadata	229
B. The Risk of Disclosing Privileged Information	232
C. The Unclear Definition of Metadata	235
D. The Relationship Between Metadata and the Form in Which Electric Documents Are Produced	237
E. The Pitfalls of Using New Technology in Electronic Discovery	238
F. A Better Approach: Metadata Should Be Produced Only when Relevant and Requested by the Adverse Party	239
CONCLUSION	242

INTRODUCTION

The amount of information stored electronically has increased dramatically in the past two decades.¹ Information that was not usually preserved in the paper age is now preserved in electronic format.²

[†] B.A., Amherst College, 2005; J.D. Candidate, Cornell Law School, 2008. I am grateful to Estella Chen for encouraging me to write this Note, and to Hiral Mehta, Ken Meyer, and Kyle Taylor for their excellent editing. I also wish to thank my partner, Jason Chang, and my parents and grandparents for their love and support.

¹ See THE SEDONA CONFERENCE, THE SEDONA PRINCIPLES: BEST PRACTICES, RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION 1 (July 2005), available at http://www.thesedonaconference.org/dltForm?did=7_05TSP.pdf [hereinafter SEDONA PRINCIPLES] (estimating that 90 percent of information is created in electronic format).

² See *Byers v. Ill. State Police*, No. 99 C 8105, 2002 WL 1264004, at *6 (N.D. Ill. June 3, 2002) (“Chief among these differences is the sheer volume of electronic information. E-mails have replaced other forms of communication besides just paper-based commu-

This increase raises a host of questions about which part of this information parties should be able to request through discovery. One issue is whether metadata—data stored in electronic files that is not apparent to the user and might not appear when a file is printed—should be discoverable and, if so, under what circumstances.³ Metadata may include information about the dates when a file was accessed or modified, the software that was used to modify it, or the name of the person who modified it.⁴ This Note analyzes *Williams v. Sprint/United Management Co.*,⁵ in which the district court of Kansas ruled that a defendant who produces electronic files during discovery must also produce their corresponding metadata.⁶

Part I of this Note discusses some of the issues that discovery of electronic documents presents and observes the scarcity of decisions involving metadata. Part II summarizes *Williams*. Part III analyzes the policy implications of the case and argues that the court's decision will make document preservation and discovery needlessly more costly and difficult. It also describes the problems in defining "metadata" and argues that the court's overly broad and unclear definition of the word will probably cause confusion. The Note concludes that the court in *Williams* correctly ordered the parties to produce metadata that was relevant to the case, but that the court should not have ordered them to produce irrelevant metadata.

I BACKGROUND

The discovery of electronic documents—commonly known as E-discovery—has created new legal issues with which courts must grapple.⁷ As more relevant information is stored electronically, courts

tion. Many informal messages that were previously relayed by telephone or at the water cooler are now sent via e-mail."); SEDONA PRINCIPLES, *supra* note 1, at 4 ("[T]he amount of information available for potential discovery has exponentially increased with the introduction of electronic data."); Richard A. Cirillo & Ann M. Cook, *A Bedeviling Little Subject Called Metadata*, N.Y.L.J., Apr. 17, 2006, at 1, available at <http://apps.kslaw.com/Library/publication/Metadata.pdf> ("One of the odd realities of the electronic age is that nearly everything is recorded somewhere.").

³ SEDONA PRINCIPLES, *supra* note 1, at 5.

⁴ See *id.* ("[E]lectronic documents, unlike paper, contain information that is known as 'metadata.' Metadata is information about the document or file that is recorded by the computer to assist the computer and often the user in storing and retrieving the document or file at a later date. The information may also be useful for system administration as it reflects data regarding the generation, handling, transfer, and storage of the data within the computer system.").

⁵ 230 F.R.D. 640 (D. Kan. 2005).

⁶ *Id.* at 656–57.

⁷ For example, given that it is easy to store information electronically, people often store a lot more information now than they did in the paper age. Often, corporations may have an enormous amount of data stored on backup tapes, and it would be very costly to produce all of the data and review it for privilege. Courts must then decide whether a

have tried to adapt the paper-age Federal Rules of Civil Procedure to address discovery issues unique to electronic information.⁸ The discovery of metadata is one such issue.

Metadata is information about a file that a computer automatically stores and is often not visible to the user.⁹ Examples of metadata include file creation and modification dates, authorship, past edits, hidden keywords used for finding a Web site in an Internet search, sender and recipient information in e-mails (including blind carbon copy recipients), and cookie data that can track usage information.¹⁰ A user can inadvertently delete metadata by moving a file or converting it to a different format.¹¹ Metadata can also be deleted intentionally, either manually or using specialized software.¹² Often, metadata can be inaccurate;¹³ where a user creates a file based on a

party should be forced to restore backup tapes, and if so, whether the parties should share the cost of such expensive discovery. See *Zubulake v. UBS Warburg LLC*, 216 F.R.D. 280, 287–90 (S.D.N.Y. 2003) (ordering the defendant to restore responsive e-mails from backup tapes and bear most of the estimated \$165,954.67 cost of doing so).

⁸ Theodore O. Rogers Jr., *Electronic Discovery: The Current Legal Landscape*, in LITIGATING EMPLOYMENT DISCRIMINATION & SEXUAL HARASSMENT CLAIMS 2006, at 171–72; see SEDONA PRINCIPLES, *supra* note 1, at 1 (“The same rules that govern paper discovery, such as Federal Rules of Civil Procedure 1, 26, and 34, govern electronic discovery.”).

⁹ See SEDONA PRINCIPLES, *supra* note 1, at 5 (“Metadata is information about the document or file that is recorded by the computer to assist the computer and often the user in storing and retrieving the document or file at a later date. The information may also be useful for system administration as it reflects data regarding the generation, handling, transfer, and storage of the data within the computer system. Much metadata is not normally accessible by the computer user.”); see also *Williams*, 230 F.R.D. at 646 (providing various ways of defining metadata); Marjorie A. Shields, Annotation, *Discoverability of Metadata*, 2006 A.L.R. 6th 6 (2006) (“Often, there is information that is hidden within a digital copy of [a] document, which is not rendered visible when the document is printed out into hardcopy. This information is generally referred to [as] ‘metadata.’ Metadata can be understood as ‘data about data.’ It refers to hidden data that usually can only be seen when a digital document is viewed in its native format using the program that originally produced the document. Often even the user of a program may not know it is there unless he or she knows how to find it. When a document is created by a particular program (such as [Microsoft] Word) there is hidden information (metadata) about that document that can only be viewed if the document is opened by that program.”).

¹⁰ SEDONA PRINCIPLES, *supra* note 1, at 5. For examples of metadata stored in Microsoft Excel spreadsheets, see Microsoft, How to Minimize Metadata in Microsoft Excel Workbooks, <http://support.microsoft.com/default.aspx?scid=kb;EN-US;223789> (last visited Sept. 22, 2007). According to Microsoft, Excel can store different types of metadata, including the user’s name, initials, company or organization name, computer name, network server or hard disk name where the user saved the workbook, other file properties and summary information, invisible portions of embedded OLE objects, document revisions, and hidden text or cells. *Id.*

¹¹ *Williams*, 230 F.R.D. at 646.

¹² See Cirillo & Cook, *supra* note 2, at 2; Microsoft, Find and Remove Metadata (hidden information) in your Legal Documents, <http://office.microsoft.com/en-us/help/HA010776461033.aspx> (last visited Sept. 22, 2007).

¹³ See *Williams*, 230 F.R.D. at 646; *In re Telxon Corp. Sec. Litig.*, No. 5:98CV2876, 2004 WL 3192729, at *17 n.17 (N.D. Ohio, July 16, 2004) (“[T]he appearance of an individual’s name in the metadata as having modified a document may be misleading. In some cases,

template created by another person, a software program can incorrectly record the file's author.¹⁴

In 2006, the Federal Rules of Civil Procedure were amended to address electronic discovery.¹⁵ Yet the amendments discuss electronic discovery only in general terms, leaving the courts to decide how to apply these rules to specific electronic discovery issues. For example, while the Federal Rules refer to "electronically stored information,"¹⁶ the Rules only suggest that the parties talk about discovery of this material in the initial conference.¹⁷ The minutes of the Civil Rules Advisory Committee reveal that the rule makers decided to remain silent on whether to require parties to produce metadata and preferred to leave the issue to the courts, presumably because electronic discovery was such a new and changing area of law that the Committee was not confident in setting down a firm and inflexible rule.¹⁸ Thus, the new Rules do not provide much guidance on whether a party must produce metadata during discovery if the parties cannot agree, leaving the courts to refine the Rules through case law.¹⁹

that individual may have prepared a document which served as a template for the document in question. . . . [I]n other cases, the appearance of an individual's name in the metadata as having 'modified' a document may indicate that the individual worked on the document in a previous year and the document was 'rolled forward' into the next audit year, carrying the individual's name in the metadata into the new audit." (citations omitted)); SEDONA PRINCIPLES, *supra* note 1, at 5-6.

¹⁴ See *Williams*, 230 F.R.D. at 646.

¹⁵ Dawn M. Bergin, *New Federal Rules on E-Discovery: Help or Hindrance?*, ARIZ. ATT'Y, Dec. 1, 2006, at 22, 24. These new amendments took effect on December 1, 2006. Federal Judiciary Rulemaking, <http://www.uscourts.gov/rules/newrules6.htm#proposed0805> (last visited Aug. 29, 2007).

¹⁶ FED. R. CIV. P. 26(a), (b), (f).

¹⁷ See FED. R. CIV. P. 26(f) advisory committee's note, available at http://www.uscourts.gov/rules/EDiscovery_w_Notes.pdf ("[P]roduction may be sought of information automatically included in electronic files but not apparent to the creator or to readers. . . . Whether this information should be produced may be among the topics discussed in the Rule 26(f) conference.").

¹⁸ See Civil Rules Advisory Committee Minutes, Apr. 14-15, 2005, at 18, <http://www.uscourts.gov/rules/Minutes/CRAC0405.pdf> ("A related question asked whether, if the motion should pass, the Committee Note would take a clear position on the question whether production in the form ordinarily maintained includes embedded data and metadata? It is important to be clear, lest the question be litigated continually and with conflicting results. Discussion of this question observed that however it may be for word-processing programs, there are real problems with requiring production of embedded data and metadata for other programs."); *id.* at 19 ("[U]nless the Committee is quite confident of what it should say, 'the less you say the better.'"); Thomas Y. Allman, *The Impact of the Proposed Federal E-Discovery Rules*, 12 RICH. J.L. & TECH. 13, 15 (2006) ("Neither default form is intended to mandate production of metadata or embedded data. The Advisory Committee discussed the competing concerns at some length but ultimately decided that the best course of action was to remain silent and leave the issue to individual case law development.").

¹⁹ See *Williams*, 230 F.R.D. at 649 (noting that the amended Federal Rules do not specify whether a party has to produce metadata).

There are relatively few cases that address the discovery of metadata.²⁰ A Westlaw search in the “all federal cases” database for the word “metadata” from January 2000 to September 2005 (when *Williams* was decided) yields only eighteen results, almost all of which are not related to electronic discovery. The *Williams* case is the first to discuss in depth the discoverability of metadata, but there are some previous cases that briefly address the issue.²¹

In *In re Priceline.com Inc. Securities Litigation*, the plaintiffs moved to compel production of electronic data in discovery.²² The defendant possessed computer files that were not in an easily readable and searchable format, partly because they had been archived for backup purposes.²³ The court ordered the defendant to convert the files into PDF or TIFF format, eliminate duplicate files, and produce a table containing metadata that would allow the plaintiff to search through and organize the files.²⁴ The court did not require the defendant to produce the files in their native format unless the files could not be read otherwise.²⁵

Although the court in *In re Priceline.com* discusses production of metadata, it seems to refer to a different kind of metadata than the *Williams* court. In *Priceline.com*, the defendant had kept a large number of disorganized files, so the court ordered the defendant to produce a table with “metadata” that would make the files organized and searchable.²⁶ Thus, the metadata at issue was presumably information about a file’s name, original location, and perhaps keywords describing the file’s subject matter.²⁷ By contrast, in *Williams*, the issue was not that the files were disorganized, but that they were missing certain data.²⁸ The *Williams* court did not want the defendant to produce a table that would help the plaintiff organize the files; rather, the court sought to provide the plaintiff with data stored in Microsoft Excel that related to the spreadsheet files defendant was ordered to produce.²⁹

Prior to *Priceline.com*, the district court of Louisiana ordered in *In re Vioxx Products Liability Litigation* that the parties should preserve all

²⁰ See *id.* at 650 (“[N]either the federal rules nor case law provides sufficient guidance on the production of metadata . . .”).

²¹ See Shields, *supra* note 9 (summarizing *Williams* and other cases in which courts ruled on the discoverability of metadata).

²² 233 F.R.D. 88, 88 (D. Conn. 2005).

²³ See *id.* at 89–90.

²⁴ *Id.* at 91; see also Shields, *supra* note 9 (summarizing the *Priceline.com* decision).

²⁵ *In re Priceline.com*, 233 F.R.D. at 91 (“TIFF or PDF format is the most secure format for the production of documents in this case. . . . Exceptions to this directive, however, may be applied for should production of a file in its native format be necessary to view or comprehend the information in the file.”).

²⁶ See *id.*

²⁷ See *id.*

²⁸ See *Williams*, 230 F.R.D. at 644.

²⁹ See *id.*

documents including metadata, but the court did not discuss the issue in depth or explain its reasoning.³⁰ In *In re Verisign, Inc. Securities Litigation*, the court upheld a prior order that compelled the defendant to produce documents in their native form along with their metadata instead of producing them as TIFF images.³¹ But the *Verisign* court did not decide whether the producing party must also produce metadata even without a court order. That issue would surface later in *Williams*.³²

Another case emphasized the importance of exchanging meaningful information before trial. The court in *Hopson v. Mayor and City Council of Baltimore* ordered the parties to talk in detail during their discovery conference about the characteristics of their electronic systems and agree on a form in which to produce electronic files.³³ The court did not decide whether metadata should be produced, but ordered the parties to discuss the issue and come to an agreement themselves.³⁴

When a party requests metadata, failure to produce it can lead to sanctions. The court in *In re Telxon Corp. Securities Litigation* sanctioned the defendant for repeatedly failing to produce certain requested electronic data, including metadata.³⁵ The defendants tried to explain the missing data, but the court remained skeptical.³⁶

Because few courts had dealt with electronic discovery issues, the district court of Kansas had very little precedent available when it decided *Williams*. The *Williams* court thus had to rely on its own analysis.³⁷ The court also looked to The Sedona Guidelines: Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age,³⁸ and The Sedona Principles: Best Practices, Rec-

³⁰ No. MDL 1657, 2005 WL 756742, at *3 (E.D. La. Feb. 18, 2005); see Shields, *supra* note 9.

³¹ No. C 02-02270 JW, 2004 WL 2445243, at *3 (N.D. Cal. Mar. 10, 2004).

³² See *Williams*, 230 F.R.D. at 650 (“While *Verisign* is helpful, it does not answer the question of whether metadata must be produced when the court’s order does not expressly reference metadata.”).

³³ 232 F.R.D. 228, 246 (D. Md. 2005).

³⁴ See *id.* at 245 (“At a minimum, they should discuss: the type of information technology systems in use and the persons most knowledgeable in their operation; preservation of electronically stored information that may be relevant to the litigation; . . . the format in which production will occur (will records be produced in ‘native’ or searchable format, or image only; is metadata sought) . . .”).

³⁵ No. 5:98CV2876, 2004 WL 3192729, at *34–35 (N.D. Ohio July 16, 2004).

³⁶ See *id.* at 34.

³⁷ See *Williams*, 230 F.R.D. at 649 (“In the few cases where discovery of metadata is mentioned, it is unclear whether metadata should ordinarily be produced as a matter of course in an electronic document production.”).

³⁸ THE SEDONA CONFERENCE, THE SEDONA GUIDELINES: BEST PRACTICE GUIDELINES & COMMENTARY FOR MANAGING INFORMATION & RECORDS IN THE ELECTRONIC AGE (Sept. 2005), available at <http://www.thosedonaconference.org/> [hereinafter SEDONA GUIDELINES].

ommendations & Principles for addressing Electronic Document Discovery³⁹—two documents published by a group of attorneys, judges, and electronic discovery experts who meet periodically at an event called the Sedona Conference to discuss electronic discovery issues.⁴⁰ The scarcity of authority on the topic makes the *Williams* decision important to commentators and attorneys alike.⁴¹

II

WILLIAMS V. SPRINT/UNITED MANAGEMENT CO.

In *Williams v. Sprint/United Management Co.*, the plaintiff sued on behalf of herself and others similarly situated and claimed that the defendant engaged in age discrimination when it terminated her employment.⁴² The plaintiff requested that the defendant produce Excel spreadsheets that contained data regarding the defendant's reduction-in-force decisions so that she could analyze whether age was a factor in terminating employees.⁴³ After the defendant delayed producing these documents for over two years, the plaintiff sought the court's intervention.⁴⁴ Initially, the defendant produced the spreadsheets in TIFF image format.⁴⁵ In a previous pleading, the plaintiff objected to receiving TIFF versions because the image format obscured some of the spreadsheet columns, hid the formulas, and did not allow the plaintiff to perform any calculations.⁴⁶ The court then ordered the defendant to produce the files in the form in which they were ordinarily maintained.⁴⁷

After the court's order, the defendant produced the Excel files but locked the cells to prevent the plaintiff from modifying the data.⁴⁸ The defendant also used a software program to delete certain metadata, such as the files' names, dates of modification, authors, history of revisions, printout dates, and other information.⁴⁹ The plaintiff argued that the defendant should not have erased the metadata or locked the cells.⁵⁰ The defendant stated that it did not produce the

³⁹ SEDONA PRINCIPLES, *supra* note 1.

⁴⁰ *Williams*, 230 F.R.D. at 646–47.

⁴¹ Only a year after the *Williams* decision, there are fifty-nine secondary sources and twenty-nine trial court documents citing the case, according to a Westlaw search. See Gary Blankenship, *Metadata and Other Electronic Realities Facing Lawyers Today*, FLA. BAR NEWS, Aug. 1, 2006, at 1 (discussing the importance of the *Williams* case and predicting that it will serve as a benchmark for future cases).

⁴² *Williams*, 230 F.R.D. at 641.

⁴³ See *id.* at 642 & n.1.

⁴⁴ *Id.* at 642 n.2.

⁴⁵ *Id.* at 642–43.

⁴⁶ *Id.* at 643.

⁴⁷ *Id.*

⁴⁸ *Id.* at 644–45.

⁴⁹ *Id.*

⁵⁰ *Id.*

metadata because it was irrelevant and may have been privileged, and because the plaintiff never requested it.⁵¹ The defendant argued that the law did not require it to produce metadata unless it was “both specifically requested and relevant.”⁵²

The court held that a party should produce documents with all metadata intact, even if some of the metadata is irrelevant to the other party’s claim, “unless [the producing] party timely objects to production of metadata, the parties agree that the metadata should not be produced, or the producing party requests a protective order.”⁵³ According to the court, the party charged with producing the metadata has the burden of timely objecting because that party is better able to determine whether the metadata contains privileged material.⁵⁴ If that party does not object to producing metadata, then it waives all such objections and risks having to produce all metadata, even if it is irrelevant⁵⁵ or privileged.⁵⁶

Because the defendant in *Williams* “should reasonably have known that Plaintiffs were expecting the electronic spreadsheets to contain their metadata intact”⁵⁷ and because the defendant did not object before producing the altered files,⁵⁸ the court ordered the defendant to produce the spreadsheets unlocked and with the metadata intact⁵⁹. But the court did not sanction the defendant for initially producing locked documents without metadata, partly because the defendant successfully argued that it did not act in bad faith.⁶⁰ As the court explained, sanctions were not appropriate given the “lack of clear law on the production of metadata.”⁶¹

⁵¹ *Id.*

⁵² *Id.* at 645–46.

⁵³ *Id.* at 652.

⁵⁴ *See id.*

⁵⁵ *Id.* at 653 (“[I]f Defendant believed the metadata to be irrelevant, it should have asserted a relevancy objection instead of making the unilateral decision to produce the spreadsheets with the metadata removed.”). In *Williams*, the court held that the metadata was in fact relevant. *See id.* at 652–53. Therefore, it is unclear whether the court ruled that the defendant had to produce the metadata because it should have known that it was relevant, because it did not object to producing it, or both. In either case, a defendant’s safest litigation strategy is to assume that the metadata may be relevant and object to its production, because no defendant will want to guess wrong on the question of relevance—as the defendant in this case did—and thereby waive the objection and subject itself to sanctions.

⁵⁶ *Id.* at 653–54 (holding that defendant waived any attorney-client privilege or work product protection regarding the metadata because defendant did not object to its production). The court, however, did not order the defendant to produce metadata “directly corresponding to the adverse impact analyses and social security number information.” *Id.* at 654.

⁵⁷ *Id.* at 653.

⁵⁸ *See id.*

⁵⁹ *Id.* at 656.

⁶⁰ *See id.*

⁶¹ *Id.*

III
ANALYSIS AND POLICY IMPLICATIONS OF THE
WILLIAMS DECISION

The *Williams* ruling is problematic for several reasons. First, it needlessly increases the cost and burden of preserving and producing evidence by requiring parties to produce potentially irrelevant metadata. Second, it increases the risk that producing parties will inadvertently turn over privileged information given the large amount of metadata that they must review, some of which might go undetected. This in turn creates an ethical dilemma for the receiving parties who must decide whether to look at the metadata if they suspect the opposite side inadvertently produced it. The *Williams* ruling also burdens the courts because they must decide whether a party waived privilege in such situations. Third, because metadata is not clearly defined, a rule that requires its production is vague. Fourth, it places the burden on the responding party to object to the production of metadata, instead of placing the burden on the requesting party to ask for what it needs. Because most metadata is often irrelevant, the *Williams* approach can unnecessarily burden the producing party to list all the irrelevant categories of metadata and explain why it is not producing them, instead of forcing the opposite party to ask for the few types of metadata it needs. Finally, the *Williams* ruling discourages attorneys from experimenting with new technology lest they end up like the *Williams* defendant who was almost sanctioned because it used software to scrub metadata from its files. A rule that does not require the production of all metadata, but instead states that the requesting party should specifically request the metadata it needs, would avoid or minimize these problems.

A. The High Cost and Difficulty of Producing All Metadata

Electronic discovery is often much more burdensome for the producing party than paper discovery⁶² because of the large volume of electronic information⁶³ and the difficulty of extracting and manipu-

⁶² See JERRY M. CUSTIS, LITIGATION MANAGEMENT HANDBOOK § 7:28 (2006).

⁶³ See Kenneth J. Withers, *Electronically Stored Information: The December 2006 Amendments to the Federal Rules of Civil Procedure*, 4 NW. J. TECH. & INTELL. PROP. 171, 173–76 (listing the causes that lead to “the tremendous volume of electronically stored information”). People store a much larger volume of electronic information than paper information for several reasons: electronic information is copied and stored in several locations, such as when an e-mail is stored both by the sender and recipient; computers automatically log certain types of communication, such as instant messages or Web-based meetings; electronic data might not be deleted when a user presses the “delete” key; electronic data is routinely backed up; and, finally, computers produce metadata used to organize and maintain the electronic information. See *id.*

lating electronic data stored in various formats.⁶⁴ The producing party normally bears the cost of discovery,⁶⁵ and the requesting party can gain leverage during litigation by forcing or threatening to force the producing party to incur high discovery costs.⁶⁶ Therefore, it is better to keep discovery costs low and not increase them unnecessarily.

Requiring the routine production of metadata increases the usually enormous costs of electronic discovery. Metadata can be difficult to preserve, both before exchanging documents and during the exchange itself.⁶⁷ A party has a duty to institute a "litigation hold" and preserve evidence as soon as it reasonably anticipates a lawsuit, which might happen even before a lawsuit is filed.⁶⁸ Preserving electronic data during the litigation hold is difficult because computers may routinely delete or modify such data and because computer users who are unaware of the litigation hold may modify or delete relevant data.⁶⁹ Preserving metadata can be even more difficult than preserving other types of electronic data because metadata can change when users perform routine tasks such as opening or moving files.⁷⁰ Furthermore, attorneys might not even be aware that certain types of metadata exist. Thus, an attorney who wants a client to preserve metadata must first consult technology experts to discover where the metadata is and how to preserve it.⁷¹ Then the attorney must instruct that client not only

⁶⁴ See CUSTIS, *supra* note 62, at § 7:28.

⁶⁵ See *id.* ("[T]he normal rule is that the cost of responding to discovery presumably is borne by the responding party.")

⁶⁶ See *id.* ("As with paper discovery, the litigating party that possesses the greater quantity of computer material is at a cost disadvantage because it can be put to the expense of searching for material pertinent to a case.")

⁶⁷ Steven C. Bennett, *Electronic Materials and Other Discovery Considerations*, in INSURANCE COVERAGE 2006: CLAIM TRENDS & LITIGATION 111, 126–28 (2006).

⁶⁸ See CUSTIS, *supra* note 62, at § 7:28 ("The duty [to preserve evidence] arises when a party 'has notice that the evidence is relevant to litigation or when a party should have known that the evidence may be relevant to future litigation.'" (quoting *Convolve, Inc. v. Compaq Computer Corp.*, 223 F.R.D. 162, 175 (S.D.N.Y. 2004))); David Chaumette & Michael Terry, *The World of E-Discovery or How I Learned to Stop Worrying and Love the New Rules*, HOUS. LAW., Nov.–Dec. 2006, at 30, 32.

⁶⁹ See FED. R. CIV. P. 26(f) advisory committee's note, *supra* note 17 ("The volume and dynamic nature of electronically stored information may complicate preservation obligations."); Bennett, *supra* note 67 (summarizing the difficulties of complying with the preservation obligation for electronic data and noting that while attorneys know which documents they have to preserve in the paper age, they may not be aware of their preservation obligations with regard to electronic data and, in particular, metadata); Withers, *supra* note 63, at 183–84.

⁷⁰ See Withers, *supra* note 63, at 183–84.

⁷¹ See Chaumette & Terry, *supra* note 68, at 32. ("Every time a division gets a new piece of software or an employee gets a new computer, someone at the company should be considering the preservation issues, including what needs to be preserved, how it should be preserved, and where it should be preserved."); David Hricik & Robert R. Jueneman, *The Transmission and Receipt of Invisible Confidential Information*, 15 PROF. LAW. 18, 20 (2004) (explaining that attorneys must understand technology to be able to locate metadata).

to preserve the text of important documents, but also not to open or move the files on the client's computer because doing so might change a file's path name or date of modification. Thus, the duty to preserve metadata can make the litigation hold even more burdensome.

Document production is also more difficult if metadata must be preserved. Here, metadata can be lost when a party copies files to a new storage medium to give to the adversary or converts files to a format that the adversary can more easily use.⁷² The *Williams* court states that a party would have to take affirmative steps to change or remove metadata,⁷³ but that is not always the case. The *Williams* court probably intended to prevent parties from deliberately scrubbing metadata. But by its ruling, the court also created difficulties for parties who might change metadata inadvertently or who might convert their files to a different format in good faith.

The Civil Rules Advisory Committee was aware of the difficulties of exchanging documents electronically across different computer systems.⁷⁴ Therefore, the Committee specifically provided that parties do not have to produce documents in their native form if they are in a form that is reasonably usable.⁷⁵ For example, if the producing party maintains files in a format that can only be read by using software that the requesting party does not own, then the producing party can convert its files to a format that the requesting party can read.⁷⁶ Neither party has the final say regarding form of production.⁷⁷ The parties must negotiate and agree on the form of production; if they cannot, the court will resolve their dispute.⁷⁸ The Advisory Committee was aware that metadata might be lost when converting files to a different format, but did not think this was a problem in all cases. In fact, the Minutes reveal that the Committee believed that the producing party could legitimately convert files to a different format even with the deliberate purpose of erasing privileged or irrelevant metadata.⁷⁹ Thus,

⁷² See Ronald J. Hedges, *Discovery of Digital Information*, in ELECTRONIC DISCOVERY AND RETENTION GUIDANCE FOR CORPORATE COUNSEL 2006, at 41, 97 (2006) (noting that if metadata must be produced, then the responsive party may not be able to produce documents in certain formats that do not preserve metadata); Withers, *supra* note 63, at 173.

⁷³ *Williams v. Sprint/United Mgmt. Co.*, 230 F.R.D. 640, 652 (D. Kan. 2005) ("Placing the burden on the producing party is further supported by the fact that metadata is an inherent part of an electronic document, and its removal ordinarily requires an affirmative act by the producing party that alters the electronic document.").

⁷⁴ See Civil Rules Advisory Committee Minutes, *supra* note 18, at 13.

⁷⁵ See *id.*

⁷⁶ See *id.*

⁷⁷ See *id.*

⁷⁸ *Id.*

⁷⁹ See *id.* at 17 ("The Committee was reminded that the comments expressed concern that a default calling for production in a form ordinarily maintained by the responding party might call for 'native format' production, including metadata and embedded data.

it appears that the Advisory Committee did not contemplate that the producing party would have to produce all metadata. Furthermore, removing this burden from the producing party would make it easier to convert documents to a more usable form if necessary, which appears to be what the Federal Rules and the Advisory Committee intended.⁸⁰

B. The Risk of Disclosing Privileged Information

The producing party must review metadata for privilege before providing electronic discovery to the requesting party.⁸¹ This is problematic because of the additional data to review and also because it can be difficult to extract the metadata.⁸² Privilege review can already be time-consuming and burdensome during paper discovery, and electronic discovery can make it even more costly.⁸³ The Civil Rules Advisory Committee observed that the burdens of reviewing metadata might be excessive because they have no counterpart in the ordinary world of paper discovery.⁸⁴ For example, if attorneys fail to review

But it was noted that at least in many circumstances the alternative default form would remain available—the responding party could strip out the metadata and embedded data and still produce the information in a form that is electronically searchable or that is reasonably usable by the requesting party.”).

⁸⁰ See *id.*

⁸¹ J. Brian Beckham, *Production, Preservation, and Disclosure of Metadata*, 7 COLUM. SCI. & TECH. L. REV. 1, 11 (2006) (quoting New York State Bar Association opinions that explain that an attorney must take care not to send to opposing counsel documents with metadata that could contain privileged attorney-client communications).

⁸² See Civil Rules Advisory Committee Minutes, *supra* note 18, at 15. For examples of how attorneys can inadvertently disclose privileged information through metadata, see Beckham, *supra* note 81, at 2–3; Andrew M. Perlman, *Untangling Ethics Theory from Attorney Conduct Rules: The Case of Inadvertent Disclosures*, 13 GEO. MASON L. REV. 767, 773–74 (2005) (“Imagine, for example, that you are negotiating a contract with opposing counsel through the exchange of an electronic document created in WordPerfect During the negotiations, your client instructs you to make an important concession in one of the contract’s provisions. You make the change in the electronic version of the document, but before emailing the proposed change to opposing counsel, your client decides not to offer the concession. You edit the document back to its original state and send it to the other party’s attorney. . . . Through the simple use of the ‘undo’ command, the adversary can view the earlier changes.”).

⁸³ See FED. R. CIV. P. 26(f) advisory committee’s note, *supra* note 17 (“The Committee has repeatedly been advised about the discovery difficulties that can result from efforts to guard against waiver of privilege and work-product protection. Frequently parties find it necessary to spend large amounts of time reviewing materials requested through discovery to avoid waiving privilege. . . . Efforts to avoid the risk of waiver can impose substantial costs on the party producing the material and the time required for the privilege review can substantially delay access for the party seeking discovery. These problems often become more acute when discovery of electronically stored information is sought. The volume of such data, and the informality that attends use of e-mail and some other types of electronically stored information, may make privilege determinations more difficult, and privilege review correspondingly more expensive and time consuming.”).

⁸⁴ See Civil Rules Advisory Committee Minutes, *supra* note 18, at 15 (“Reviewing [metadata] for relevance, responsiveness, and privilege and other grounds for protection

some of the metadata for privilege because they were unaware of the metadata, they might waive the privilege even if they unintentionally disclosed that information.⁸⁵ The volume of metadata and its hidden quality make it more likely that a party might fail to review it for privilege.⁸⁶ Accordingly, the Amendments to the Federal Rules provide a procedure for resolving waiver of privilege disputes when a party produces privileged electronic data by mistake.⁸⁷ But even under the new Rules, the court may still decide that the disclosing party has waived the privilege to the disclosed material.⁸⁸ It is true that parties could agree before they engage in electronic discovery that they will not waive privilege for mistakenly produced documents.⁸⁹ However, such agreements do not fully protect parties because courts might not uphold them in certain cases⁹⁰ or enforce these agreements against third parties.⁹¹ Thus, the producing party cannot adequately protect itself from the substantial risk of failing to find, extract, and review some of the metadata to avoid disclosing privileged information.

If the producing party does disclose some privileged metadata, this creates an ethical dilemma for the receiving party. Some commentators believe it would be unethical for the receiving party to look at the metadata if the producing party was unaware of some of the metadata and mistakenly failed to review it for privilege.⁹² But others

can add significantly to discovery costs. There is no close analogue to such problems with paper discovery, and the burdens may not be appropriate.”).

⁸⁵ See Beckham, *supra* note 81, at 9–11.

⁸⁶ See FED. R. CIV. P. 26(b)(5)(A) advisory committee’s note, *supra* note 17 (“When the review is of electronically stored information, the risk of waiver, and the time and effort required to avoid it, can increase substantially because of the volume of electronically stored information and the difficulty in ensuring that all information to be produced has in fact been reviewed.”).

⁸⁷ See FED. R. CIV. P. 26(b)(5)(B); Bergin, *supra* note 15, at 25 (describing the protections afforded by the new Rule 26(b)(5)(B)).

⁸⁸ See FED. R. CIV. P. 26(b)(5)(B) advisory committee’s note, *supra* note 17 (“Rule 26(b)(5)(B) does not address whether the privilege or protection that is asserted after production was waived by the production. The courts have developed principles to determine whether, and under what circumstances, waiver results from inadvertent production of privileged or protected information. Rule 26(b)(5)(B) provides a procedure for presenting and addressing these issues.”).

⁸⁹ See Bergin, *supra* note 15, at 25 (arguing that the defendant in *Williams* could have avoided waiving the privilege if it had entered into a “quick peek” agreement with the plaintiff during a preliminary conference).

⁹⁰ See Hedges, *supra* note 72, at 105 (“Such agreements may also lead to disqualification motions if, even after a privileged document is returned, the temporary possession of the document ‘creates a substantial taint on any future proceedings.’” (quoting *Maldonado v. New Jersey*, 225 F.R.D. 120, 141 (D.N.J. 2004))).

⁹¹ See *id.*

⁹² Beckham, *supra* note 81, at 9–11 (citing New York State Bar Association opinions that held that attorneys should not use technology to dig for hidden information that the opposing party did not intend to transmit); Brian D. Zall, *Metadata: Hidden Information in Microsoft Word Documents and Its Ethical Implications*, COLO. LAW., Oct. 2004, at 53, 56 (“Before using a metadata viewer or otherwise viewing metadata in opposing counsel’s

note that it might be unethical for the receiving party's attorneys to ignore the metadata and thereby disadvantage their clients.⁹³ After all, the sending party has a duty to be familiar with the technology it uses and to inspect the documents it produces in discovery.⁹⁴ Moreover, it might be inappropriate to forbid the receiving party from performing commonplace tasks such as checking a Word document's properties from the File menu, even if the action could reveal some of the document's metadata.⁹⁵

By allowing the producing party to turn over only a limited and specific part of the metadata, instead of all the metadata, courts could save time and costs and reduce the risk that parties might inadvertently release privileged information. However, under *Williams*, the producing party has to provide all the metadata, whether relevant or not. Thus, the producing party is burdened not only with reviewing even irrelevant metadata for privilege, but also with a higher risk of failing to review some metadata because of its sheer quantity. Moreover, the producing party might be forced to turn over electronic documents in their native format to avoid losing any metadata.⁹⁶ But files in their native format may contain different types of metadata, some of which are difficult to extract.⁹⁷ Thus, the producing party risks

electronic documents, attorneys should consider whether such actions violate applicable ethical rules.”).

⁹³ See Beckham, *supra* note 81, at 14 (“Whether failing to search for metadata is a violation of a duty to clients is unclear. However, as knowledge of the implications of metadata increases and tools to carry out such searches become more affordable and user-friendly, the duties of due diligence may increasingly require such searches.”); Hricik & Jueneman, *supra* note 71, at 20 (arguing that an attorney has a duty to use metadata to uncover fraud).

⁹⁴ See Hricik & Jueneman, *supra* note 71, at 20.

⁹⁵ See *id.* (noting that a lot of metadata, such as a document’s “title, subject, keywords, author, [and] company,” is not hidden and that opposing counsel could have easily removed it had counsel been familiar with the software); Jerold S. Solovy & Robert L. Byman, *Native Simplicity*, NAT’L L.J., Aug. 28, 2006, at 13 (“Even without special software, the Word and Excel programs allow you, with very little effort, to save a copy of a file that deletes prior history . . .”). However, Hricik and Jueneman also point out that not all metadata can be easily removed. See Hricik & Jueneman, *supra* note 71, at 18–19. So even if it is ethical for an attorney to look at the types of metadata that are easily accessed and removed, it might still not be ethical for the attorney to look at the more hidden metadata that opposing counsel could not have easily removed. See *id.* at 20 (“[W]e are not talking about opposing counsel using binary editors or other specialized forensic tools, any more than a consumer (as opposed to an art historian) would normally expect to use X-rays to reveal what mistakes an artist painted over. We are only expecting counsel to be reasonably familiar with the tools he or she uses every single day, if necessary by actually reading the manual or the Help files.”).

⁹⁶ See Hedges, *supra* note 72, at 97 (noting that if a party has to produce all metadata under the *Williams* decision, then it might not be able to convert its files to other formats such as PDF or TIFF).

⁹⁷ See Hricik & Jueneman, *supra* note 71, at 18 (discussing the difficulties of removing metadata from documents).

turning over metadata that it did not even realize existed.⁹⁸ By allowing the producing party to turn over only the relevant metadata, the party could then convert its electronic documents to a format that does not retain metadata (for example, from Word documents to image or PDF form, or from Excel to another spreadsheet application).⁹⁹ In this way, the producing party is less likely to turn over hidden metadata that it was not aware of. And given that a lot of metadata is often irrelevant,¹⁰⁰ courts could reduce the cost and burden of electronic discovery by requiring the producing party to review and produce only relevant metadata. Such an approach is consistent with the one recommended by the Federal Manual for Complex Litigation, which recommends that requesting parties frame their requests “as narrowly and precisely as possible” to reduce costs.¹⁰¹

C. The Unclear Definition of Metadata

The Sedona Conference defines metadata as “information about a particular data set which describes how, when and by whom it was collected, created, accessed, or modified and how it is formatted.”¹⁰² The *Williams* court also quotes other definitions for metadata, such as “data about data” or “information describing the history, tracking, or management of an electronic document.”¹⁰³ But none of these definitions draw a clear line between data and metadata. For example, are the titles of columns in a data table considered metadata because they describe the data in the tables and so are “data about data”? What if the titles can be seen on the screen and contain important words or sentences—are they then not metadata but part of the document because presumably the document’s author would have intended them to be read along with the table? Is underlined or italicized text in a document merely “formatting” and hence metadata, or is it actual data because it adds meaning to the text? Suppose a party converts a Word document into a WordPerfect document and some of the formatting is lost. Which changes are a loss of

⁹⁸ See *Williams v. Sprint/United Mgmt. Co.*, 230 F.R.D. 640, 647 (D. Kan. 2005) (noting that because metadata is hidden, it could cause attorneys to disclose confidential or privileged information inadvertently).

⁹⁹ See *Hedges*, *supra* note 72, at 105 (noting that some metadata is lost when files are converted from one format to another).

¹⁰⁰ See Alan F. Blakley, *Document Production in a Strange Native Land*, FED. LAW., July 2006, at 16, 17 (“[F]requently the metadata and embedded data may be irrelevant. Attorneys need to be careful not to request such information simply because it may exist unless they believe that it can be relevant.”).

¹⁰¹ MANUAL FOR COMPLEX LITIGATION (FOURTH) § 11.446 (2004), available at http://www.fjc.gov/public/home.nsf/autoframe?openform&url_l=/public/home.nsf/in-avgeneral?openpage&url_r=public/home.nsf/pages/470.

¹⁰² SEDONA GUIDELINES, *supra* note 38, at 94.

¹⁰³ *Williams*, 230 F.R.D. at 646 (quoting FED. R. CIV. P. 26(f) proposed advisory committee’s note).

metadata and which are a loss of the actual data in the document? Suppose the indentations in the text are a little larger—is that just metadata that has changed? What if the highlighted text is no longer highlighted—is that metadata or is it now actual data? What if some characters no longer display correctly?

The *Williams* court tried to give examples that distinguish data from metadata, but the court's distinctions seem arbitrary. For example, the court viewed spreadsheet formulas as metadata,¹⁰⁴ but it is not clear why such formulas are not a part of the document instead. In fact, the Sedona Guidelines seem to contradict the court by defining metadata as information about a file's properties that is not part of the actual informational content of the file.¹⁰⁵ By this standard, one could argue that formulas in spreadsheets are not metadata because they contain important data and hence are part of the "actual informational content of the files." But even this definition for metadata is confusing because not all data is equally important for all purposes; people might disagree on whether certain data is important, and thus part of the actual information in a file, or whether the data is just unimportant "formatting." For example, an accountant might consider the formulas in a spreadsheet the most important part of the file, while a person who is just using the file for some calculations might think that the formulas behind the table are "hidden," unimportant data.

Moreover, different courts have used the word "metadata" to mean different things. For example, the court in *Priceline.com* interpreted metadata to mean information that the defendant had to generate for discovery to organize and make searchable the files that the defendant was producing.¹⁰⁶ The court ruled that production in PDF and TIFF format was acceptable if the party receiving the files could view and comprehend them.¹⁰⁷ Thus, the court did not seem concerned with the hidden metadata stored in Excel files that was at issue in the *Williams* case.¹⁰⁸

Because there is no clear definition of the word "metadata," using it as a basis for law can lead to confusion. In fact, the word "metadata" is said to have been intentionally designed as a term with no meaning long before it was used in the electronic discovery context.¹⁰⁹ A good

¹⁰⁴ *Id.* at 647.

¹⁰⁵ See SEDONA GUIDELINES, *supra* note 38, at 82.

¹⁰⁶ See *supra* notes 24–26 and accompanying text.

¹⁰⁷ See *In re Priceline.com Inc. Sec. Litig.*, 233 F.R.D. 88, 91 (D. Conn. 2005).

¹⁰⁸ See *Williams*, 230 F.R.D. at 647.

¹⁰⁹ Solovy & Byman, *supra* note 95, at 13 ("'Metadata' is actually the registered trademark of Metadata Corp. Legend has it that company founder Jack E. Myers coined the term 'metadata' in 1969, intentionally designing it to be a term with no particular meaning. . . . Ken Withers of the Sedona Conference uses the term 'non-apparent information,'

rule of law would require the requesting party to specify exactly the information it needs instead of forcing the producing party to deliver “metadata” when the term has no clear meaning. Otherwise, disputes may arise. Even if a party has to produce all the files with their metadata, it may still leave out some information that it considers neither part of the file nor metadata, but which the opposing party considers to be metadata. For example, suppose that the producing party turns over copies of electronic documents on a CD. The opposing party might then claim that the producing party should have supplied metadata revealing the files’ location on the originating computer,¹¹⁰ while the producing party might argue that because metadata is defined as data inherently part of an electronic document,¹¹¹ the files’ locations were not metadata. This dispute could be avoided if the producing party did not have to produce all metadata by default, but the requesting party had listed the exact types of data it needed—such as “information describing a file’s location.”

D. The Relationship Between Metadata and the Form in Which Electronic Documents Are Produced

The *Williams* case does not distinguish between rules that govern which information may be discovered and rules that specify the form in which the information should be produced. The Federal Rules of Civil Procedure allow discovery of nonprivileged matters that are relevant to a dispute, including relevant metadata.¹¹² The Amendments to the Federal Rules also provide that the documents shall be produced as they are ordinarily maintained or in a reasonably usable form.¹¹³ The rules regarding the form of production are not intended to define the scope of discovery, but to ensure that the information is produced in a form that the requesting party can use.¹¹⁴

which might be more accurate . . . , but it’s a mouthful; and ‘metadata’ has a nice ring to it.”).

¹¹⁰ See *Williams*, 230 F.R.D. at 646 (stating that a file’s location on a computer can be considered metadata).

¹¹¹ See *id.* at 652 (“[M]etadata is an inherent part of an electronic document . . .”).

¹¹² FED. R. CIV. P. 34(a) & advisory committee’s note, *supra* note 17 (“The rule covers—either as documents or as electronically stored information—information ‘stored in any medium,’ to encompass future developments in computer technology. Rule 34(a)(1) is intended to be broad enough to cover all current types of computer-based information, and flexible enough to encompass future changes and developments.”).

¹¹³ FED. R. CIV. P. 34(b).

¹¹⁴ See Withers, *supra* note 63, at 194–95 (“One Advisory Committee member . . . expressed shock at the news that practitioners were asking for, and receiving, the metadata associated with discoverable electronic ‘documents.’ She took the position that only ‘documents,’ meaning the part of the files visible in printouts or screen images, were discoverable. This was symptomatic of the widespread confusion between the ‘scope’ of document discovery set out in Rule 34(a), which is supposed to track the scope of discovery in general expressed in Rule 26(b), and the ‘procedure’ for document production set out in Rule

In *Williams*, the parties disputed both the form and scope of production.¹¹⁵ Rather than analyzing these two complaints separately, the court considered locking the spreadsheet cells to prevent modification to be similar to erasing metadata.¹¹⁶ The unclear analysis in *Williams* makes it more likely that attorneys will continue to mix their complaints about form of production and usability with those about the scope of discovery, which only adds confusion to this area of law.

E. The Pitfalls of Using New Technology in Electronic Discovery

The *Williams* ruling may discourage attorneys from using new technology in electronic discovery. Before *Williams*, there was little case law on what to do with metadata.¹¹⁷ Many commentators advised attorneys to erase metadata, given the dangers of revealing privileged information.¹¹⁸ The law was not clear as to whether parties were obligated to preserve metadata during litigation if it was not specifically requested. The defendants in *Williams* might therefore have believed in good faith that scrubbing metadata was legal and that they had an ethical obligation to do so to preserve privileged information. Yet their experiment almost led to sanctions.¹¹⁹

The *Williams* court ultimately did not sanction the defendant, as it wisely recognized that “the production of metadata is a new and largely undeveloped area of the law.”¹²⁰ But the defendant barely avoided sanctions, as the court still appeared upset with the defendant for scrubbing metadata.¹²¹ Furthermore, the court’s ruling was harsh because it held that the defendant waived its objections as to relevancy and privilege because it did not assert them in time—even though the

34(b). It took a while for the Advisory Committee to come to the elegant conclusion that Rule 34(a) was about discovering information which happened to be recorded on tangible media, not tangible media that happened to contain information.”).

¹¹⁵ See *supra* notes 45–50 and accompanying text.

¹¹⁶ See *Williams*, 230 F.R.D. at 655–56.

¹¹⁷ See *supra* text accompanying notes 20–37.

¹¹⁸ See Beckham, *supra* note 81, at 13 (advising attorneys to scrub metadata routinely, but noting that there may be an obligation to preserve metadata during litigation); Hricik & Jueneman, *supra* note 71, at 18 (“To comply with their duty of confidentiality, lawyers should take steps to remove metadata from documents exchanged with opposing counsel or disclosed to the public.”); Sharon Nelson & John Simek, *Metadata: What You Can’t See Can Hurt You*, LAW PRAC., Mar. 2006, at 28 (advising attorneys to scrub metadata when they email documents and listing software designed for scrubbing metadata).

¹¹⁹ See *Williams*, 230 F.R.D. at 644 (“The Court then gave Defendant seven days to show cause why it had scrubbed metadata and locked data, ‘because my intent from the two previous Orders was to do as I said, produce it in the format it’s maintained, not modify it and produce it.’ The Court advised Defendant that if it could show justification for scrubbing the metadata and locking the cells, the Court would certainly consider it, but cautioned that ‘it’s going to take some clear showing or otherwise there are going to be appropriate sanctions’” (footnotes omitted)).

¹²⁰ *Id.* at 656.

¹²¹ See *id.* at 644.

defendant could not have known from the existing law that it had to object to the production regardless of whether or not the defendant believed the metadata was irrelevant.

Therefore, the *Williams* case sends a clear message to litigants: do not experiment with new technology regarding electronic discovery because the courts will be harsh if they do not agree with your reading of the law, even though this is a new area with no clear rules. But discouraging litigants from using new technology is unfortunate because technology can simplify and automate tasks, thereby reducing the burden of producing huge amounts of electronic data during electronic discovery.¹²² For example, if a litigant is allowed to use software to scrub irrelevant or privileged metadata, then such software can be useful because it automates a potentially difficult task and prevents litigants from inadvertently disclosing privileged metadata.¹²³

F. A Better Approach: Metadata Should Be Produced Only when Relevant and Requested by the Adverse Party

As the previous sections of this Note point out, a better approach to the discovery of metadata would be to require the producing party to preserve and produce metadata only if it is relevant and the adverse party requests it. The amended Federal Rules are consistent with this approach, and, in fact, the Rules Committee's notes show that the Committee did not anticipate that it would be necessary to produce metadata in all cases.¹²⁴ Rules 16 and 26(f) provide that the parties must meet and discuss any issues that electronic discovery poses, including whether metadata production is necessary.¹²⁵ The Rule 26(f) conference would give parties the opportunity to come to an agreement regarding which types of metadata should be produced and which are irrelevant.¹²⁶ If the parties cannot agree, the court will resolve the dispute. Such an approach is more flexible than the *Williams* rule because it does not force the producing party to provide metadata in cases when the metadata is irrelevant. Furthermore, both parties may be spared the expense of producing and reviewing unnecessary metadata.

Moreover, requiring parties to produce metadata is unusual when compared to paper discovery. During paper discovery, courts generally do not expect parties to produce the exact original of a paper file with all fingerprints intact. Nor do courts expect a description of the

¹²² CUSTIS, *supra* note 62 (describing the high cost of electronic discovery).

¹²³ See *supra* note 118 (describing the benefits of scrubbing metadata).

¹²⁴ See *supra* note 79.

¹²⁵ See FED. R. CIV. P. 16, 26(f) & advisory committee's note, *supra* note 17 (advising the parties to discuss metadata production during the Rule 26(f) conference).

¹²⁶ See FED. R. CIV. P. 26(f) advisory committee's note, *supra* note 17.

drawer in which the paper was located and a list of who had access to that drawer, unless that information is relevant to the case. So it is difficult to see why, during electronic discovery, a party should have to indicate who accessed an electronic document or provide the file path and file permissions intact, unless that information is relevant to the case.

Furthermore, the receiving party might not always want the metadata.¹²⁷ Preserving metadata usually requires files to be produced in their native format, which might cause difficulties if the receiving party does not have the software or the technical expertise necessary to use the files in that format.¹²⁸ Furthermore, files in native format are easily modified, which might make it difficult for the receiving party to prove at trial that the opponent's files have not been tampered with.¹²⁹ Admittedly, there are ways to check the integrity of data: the producing and receiving parties could agree to create digital fingerprints of the files when they exchange them, thus enabling either party to prove the files' integrity.¹³⁰ But this method requires technical expertise, adequate software, and an agreement between the parties. The receiving party might simply prefer to receive the files in a format that is more difficult to modify, such as PDF or TIFF, if metadata is not relevant and the files still contain all the relevant information.

The *Williams* court argues that it is appropriate to place the burden on the producing party to object to the production of metadata, rather than on the requesting party to request it, because the producing party is more familiar with its own files and the types of metadata

¹²⁷ See Solovy & Byman, *supra* note 95, at 13 (discussing discovery strategies for the receiving party and arguing that the receiving party might be better off without the metadata).

¹²⁸ See *id.* ("Do you really want native format? Unless you have all of that application and operating software, and unless you know how to use all of it, the files will not be usable.").

¹²⁹ See *id.* ("Do you really want the metadata? Do you really want the ability to manipulate the data you receive? The problem with getting files you can manipulate is—you can manipulate them. So when you go to use them at trial, how will you prove that the smoking gun electronic file you want to show the jury is the same file produced by the defendant? Moreover, the mere act of requesting production of native files can result in the inadvertent alteration of those files. When the responding party opens the file to review it for possible privilege, that innocent act can alter metadata or even substance.").

¹³⁰ See *Williams v. Sprint/United Mgmt. Co.*, 230 F.R.D. 640, 655 (D. Kan. 2005) ("Defendant could have run the data through a mathematical process to generate a shorter symbolic reference to the original file, called a 'hash mark' or 'hash value,' that is unique to that particular file. This 'digital fingerprint' akin to a tamper-evident seal on a software package would have shown if the electronic spreadsheets were altered. When an electronic file is sent with a hash mark, others can read it, but the file cannot be altered without a change also occurring in the hash mark. The producing party can be certain that the file was not altered by running the creator's hash mark algorithm to verify that the original hash mark is generated." (footnotes omitted)).

they may contain.¹³¹ However, this is not a strong argument because the parties must meet and confer about electronic discovery anyway, at which time they can talk about the types of files they have and the types of metadata they may contain. Furthermore, even if the requesting party does not know exactly what types of files or metadata the adverse party has, it can still phrase its request in general terms—for example, by asking for “any metadata related to the time a file is modified,” “e-mails with all header information,” or “any metadata that reveals a document’s authors.”

The Sedona Working Group has noted that in most (although not all) cases, metadata is irrelevant and does not need to be preserved.¹³² Thus, a default rule stating that a party must produce metadata seems counterproductive because in most cases, the responding party would have to then object to the production and prove that the data is irrelevant. A better default rule would not require a party to produce metadata unless the requesting party considers it relevant and asks for it. If in most cases the metadata is irrelevant, then the producing party could simply produce the documents without worrying about preserving metadata, and the requesting party would presumably be satisfied. Neither party would have to make any additional requests to the court, which would save time and expense.

The Sedona Working Group did believe that the responding party should produce metadata if it knows or reasonably should have known that the metadata is relevant to the dispute, even if the other party does not request the metadata.¹³³ The *Williams* court adopted this approach as an alternative ground for its ruling, but it also ruled that even if the defendant correctly believed that the metadata is irrelevant, it should have objected to its production instead of simply not producing it.¹³⁴ According to the court, the defendant should have

¹³¹ See *id.* at 652 (“The burden to object to the disclosure of metadata is appropriately placed on the party ordered to produce its electronic documents as they are ordinarily maintained because that party already has access to the metadata and is in the best position to determine whether producing it is objectionable.”).

¹³² SEDONA PRINCIPLES, *supra* note 1, at 4 (“On the one hand, it is easy to conceive of situations where metadata is necessary to authenticate a document, or establish facts material to a dispute, such as when a file was accessed in a suit involving theft of trade secrets. In most cases, however, the metadata will have no material evidentiary value; it does not matter when a document was printed, or who typed the revisions, or what edits were made before the document was circulated.”); *id.* at 46 (“Although there are exceptions to every rule, especially in an evolving area of the law, there should be a modest legal presumption in most cases that the producing party need not take special efforts to preserve or produce metadata.” (footnote omitted)).

¹³³ See *Williams*, 230 F.R.D. at 654 (“Of course, if the producing party knows or should reasonably know that particular metadata is relevant to the dispute, it should be produced.” (quoting SEDONA PRINCIPLES, *supra* note 1, cmt. 12.a)).

¹³⁴ See *id.* at 652–53.

known that some of the spreadsheets' metadata was relevant because the plaintiff claimed that the defendant revised the spreadsheets to make the numbers look more favorable to its case.¹³⁵

It is problematic for a court to require a party to produce metadata that it should know is relevant because a party might not know that some metadata is relevant—even if the court later decides it should have known. Thus, a prudent party will err on the side of producing more metadata than necessary to avoid being sanctioned for failing to produce material that it should have known was relevant. Moreover, this approach requires the producing party to try to guess the opponent's trial strategy to determine which metadata the opponent will consider helpful in proving its case. Again, the better approach would allow the requesting party to list the types of metadata that it expects will be relevant in establishing its proof. The requesting party is in a better position to judge which evidence it would find helpful because it understands its own case best.

CONCLUSION

By ruling that a producing party must automatically produce metadata during electronic discovery, the court in *Williams v. Sprint/United Management Co.* unnecessarily made the electronic discovery process more burdensome. Preserving and producing metadata is difficult because metadata is hidden, may require technical expertise to extract, can easily be modified by mistake during a litigation hold, and must be reviewed for privilege. Electronic discovery can be burdensome in many cases, even for a litigant who does not have to worry about metadata, because of the large volume of data involved and the technical problems that may arise. Requiring metadata production only adds to the already heavy burden of the producing party, which can give excessive leverage to the opposing party because the producing party may decide to settle to avoid discovery costs. And because metadata is irrelevant in most cases, the effort of producing it will usually be futile. Finally, the fact that there is no clear way to define metadata can lead to discovery disputes.

A better alternative to the *Williams* rule would have been to require the producing party to produce only the metadata that the opposing party asked for. This approach would reduce costs by eliminating the burden of producing metadata that is irrelevant to the case while still allowing the receiving party to receive any information it considers relevant.

¹³⁵ See *id.*