

DJW/byk

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF KANSAS**

SHIRLEY WILLIAMS, et al.,

Plaintiffs,

CIVIL ACTION

v.

No. 03-2200-JWL-DJW

**SPRINT/UNITED MANAGEMENT
COMPANY,**

Defendant.

MEMORANDUM AND ORDER

Plaintiff Shirley Williams filed this suit on behalf of herself and others similarly situated, asserting that her age was a determining factor in Defendant's decision to terminate her employment during a reduction-in-force (RIF). Currently, 1727 plaintiffs remain in the case out of the 2354 plaintiffs who opted into this provisionally certified collective action pursuant to 29 U.S.C. § 216(b). The parties are presently engaged in discovery concerning the merits of Plaintiffs' pattern and practice allegations. This matter is presently before the Court on Defendant's Response to the Court's July 12, 2005 Order (doc. 3037), which ordered Defendant to show cause why it should not produce electronic Microsoft Excel spreadsheets in the manner in which they were maintained and why it should not be sanctioned for "scrubbing" the metadata and locking certain data on the electronic spreadsheets prior to producing them to Plaintiffs without either the agreement of the parties or the approval of the Court.

I. Background Information

Plaintiff Williams commenced this action in April 2003, and, to date, the docket reflects that over 3300 pleadings and orders have been filed. The case is assigned to Chief Judge John W. Lungstrum but is referred to the undersigned Magistrate Judge for pretrial proceedings, including discovery. Due to the highly contentious nature of this litigation, the Magistrate Judge has conducted discovery conferences twice a month since March 2005 to resolve discovery issues identified by the parties. One of the ongoing discovery disputes has been Defendant's production of spreadsheets that relate to the RIFs at issue in this case.

Plaintiffs raised the issue of the RIF-related spreadsheets at the May 5, 2005 discovery conference. Item 1 on Plaintiffs' List of Issues for the May 5, 2005 discovery conference was "Defendant's Failure to Produce Candidate Selection Spreadsheets and Other Basic RIF Documents."¹ Item 2 was listed as

¹The following explanation accompanied Item 1 on Plaintiffs' List of Issues for the May 5, 2005 discovery conference (doc. 3301):

These documents constitute the essential materials regarding the termination and other RIF decisions that are at issue in this collective action. Plaintiffs requested these documents in November of 2003 and again in December of 2004, and yet Defendant has produced only a few improperly redacted versions relating to some Plaintiffs and initial Opt-in Plaintiffs. At an April 12th meet and confer with [Plaintiffs' counsel], Defendant's counsel promised these documents would begin being produced in batches. At the meet and confer prior to the April 21st hearing, Plaintiffs noted that Defendant still had not begun producing any of these documents. Defendant's counsel acknowledged at that time that he had nine (9) boxes of these documents in his office. At the April 21st hearing, Defendant stated it would begin getting these documents produced. To date, Defendant has still not begun production of these documents which are of core relevance to this collective action. Defendant's counsel has been indefinite and imprecise as to when Defendant will be able to certify that it has produced all of these basic RIF decision documents and spreadsheets. At the May 5th conference, Plaintiffs seek the Court's Order that Defendant produce all such documents no

(continued...)

“Defendant’s Failure to Produce Basic RIF Documents Such as HR Notes and Other Documents Re: RIF Decisions.”² At the May 5, 2005 discovery conference, Plaintiffs requested that the Court enter an order requiring Defendant to produce candidate selection spreadsheets and other basic RIF documents by May 16, 2005.³ Defendant objected to a May 16, 2005 deadline and suggested a June 1, 2005 deadline instead.⁴ Plaintiffs advised the Court they would agree to a June 1, 2005 deadline.⁵ The Court accepted this compromise position and ordered Defendant to produce Items 1 and 2 from Plaintiffs’ May 5, 2005 List of Issues by June 1, 2005 or show cause why it could not produce them by that date.⁶

Two weeks later, at the May 19, 2005 discovery conference, Plaintiffs requested that Defendant be required to produce the actual electronic “active file” version of all the Excel RIF spreadsheets.

¹(...continued)
later than May 16, 2005.

²The following explanation accompanied Item 2 on Plaintiffs’ List of Issues for the May 5, 2005 discovery conference (doc. 3301):

Plaintiffs requested these documents in November of 2003, and yet Defendant continues to trickle such documents in and to indicate that it has not produced all such documents. Indeed, Defendant’s counsel recently advised Plaintiffs’ counsel that it had not produced various Sprint Human Resources officials’ notes and other documents regarding the RIF decisions at issue in this case and that Defendant would not consider producing such documents until a new, formal request for production was made. At the May 5th conference, Plaintiffs seek the Court’s Order that Defendant produce all such documents no later than May 16, 2005. (emphasis in original)

³May 5, 2005 Discovery Conf. Transcript (Tr.) (doc. 2905) p. 15, l. 11-21.

⁴*Id.* at p. 25, l. 18-21.

⁵*Id.* at p. 26, l. 10-12.

⁶*Id.* at p. 27, l. 3-7.

Plaintiffs' stated reason for requesting that the spreadsheets be produced in their electronic form was so Plaintiffs could perform "statistical or manipulative things without taking the spreadsheets and going through the laborious process of keying in all that data again."⁷ When the Court asked Defendant whether it was now producing the active file of the spreadsheets requested by Plaintiffs, Defendant reported that it was continuing to produce the TIFF⁸ images of the spreadsheets as previously agreed by the parties. The Court then stated that:

[G]enerally, when things are maintained in the regular course of business in electronic form, they should be produced in that form, unless there's an agreement otherwise. And it sounds like what you're telling me, there was an agreement otherwise, until May 11, and then it was pointed out you want them in the form they were maintained.⁹

The Court then asked Defendant why it could not produce the spreadsheets in their electronic form. Defendant responded by suggesting that it should be allowed to finish its review of the documents to be produced in TIFF image and then go back at a later time and review what it holds in electronic format. The Court commented on Defendant's reference to reviewing documents by stating that "on the information

⁷May 19, 2005 Tr. (doc. 2915) p. 11-12, l. 24-25; 1-2.

⁸TIFF (Tagged Image File Format) is one of the most widely used and supported graphic file formats for storing bit-mapped images, with many different compression formats and resolutions. A TIFF file is characterized by its ".tif" file name extension. *The Sedona Conference Glossary for E-Discovery and Digital Information Management* (The Sedona Conference Working Group Series, May 2005 Version), available at <http://www.thesedonaconference.org>. The Sedona Conference is a nonprofit legal policy research and educational organization which sponsors Working Groups on cutting-edge issues of law. The Working Group on Electronic Document Production is comprised of judges, attorneys, and technologists experienced in electronic discovery and document management matters.

⁹May 19, 2005 Tr. p. 12, l. 17-24.

that's in electronic form, the only review [Defendant has] to do is the privilege review."¹⁰ Following this exchange, the Court further clarified its position on electronic document production of the spreadsheets:

What I'm talking about is if you're talking about documents maintained on Excel, you've got that in some form, whether it's on disk or paper, whatever it's on. It's an electronic form of Excel containing the data. The only thing you would have to do is review it for privilege and then give it to them.¹¹

At the June 2, 2005 discovery conference, Plaintiffs again raised the issue of Defendant producing the RIF spreadsheets in their electronic Excel form. Plaintiffs explained why they needed the electronic version of the spreadsheets in the following exchange:

PLAINTIFFS' COUNSEL: There are, on each of these documents, things that show you why you have to have an electronic copy. These are spreadsheets. There are columns not there, that should be there. There are columns where the entry will have a sentence, and it cuts off in midsentence because the box ends. And we all know, on the computer, we click on that and we get everything.¹²

* * *

The electronic form of this document would reveal whether or not it had any actual other columns or types of information available on a spreadsheet.¹³

THE COURT: Okay. Before we get much further here, I thought it was clear from the last time we discussed this electronic issue, that you [Defendant] were looking for them and you were going to produce them. It's not an issue that you're not going to do it. It's a question of when.¹⁴

DEFENDANT'S COUNSEL: Absolutely. And the only caveat is the one that I mentioned to you, is that there may be the issue of Social Security numbers. And if there are privileged

¹⁰*Id.* at p. 16, l. 11-13.

¹¹*Id.* at p. 16, l. 19-25.

¹²June 2, 2005 Tr. (doc. 2940) p. 55, l. 11-18.

¹³*Id.* at p. 57, l. 15-18.

¹⁴*Id.* at p. 57, l. 19-24.

communications, which are these analyses that were prepared, if they are on those documents, we'll redact those and point out that they have been redacted, but still give them the electronic version of it, with the notation that that's missing.¹⁵

At the June 9, 2005 discovery conference, Plaintiffs again raised the issue of Defendant producing the RIF spreadsheets in their electronic Excel format. The Court renewed its show cause order, stating, "for today on this issue we'll leave the show cause that you're going to do the electronic spreadsheets by [June] 24th."¹⁶ This ruling was memorialized in paragraph 1 of the Court's June 16, 2005 Order, which stated: "The Court's previous Show Cause Order to Defendant remains on the following three categories of discovery: (a) Electronic versions of Excel and other spreadsheets, (b) Other documents (other than Minutes) relating to the RIF meetings, and (c) E-mails accompanying the spreadsheets."¹⁷

On June 23, 2005, Defendant tendered to Plaintiffs' counsel 3083 Excel spreadsheets in electronic form and indicated that there were 983 additional spreadsheets identified that had not been fully processed for production and would be produced no later than June 27, 2005.¹⁸

At the July 7, 2005 discovery conference, Plaintiffs' counsel advised the Court that Defendant, prior to producing the electronic versions of the Excel spreadsheets, had utilized software to scrub the spreadsheet files to remove the metadata. Plaintiffs claim this metadata would have contained information such as file names, dates of the file, authors of the file, recipients of the file, print-out dates, changes and modification dates, and other information. Plaintiffs' counsel stated that Defendant did not provide them

¹⁵*Id.* at p. 57-58, l. 15-25; 1-8.

¹⁶June 9, 2005 Tr. (doc. 3146) p. 17, l. 22-24.

¹⁷Doc. 2953.

¹⁸*See* Defendant's Response to Show Cause Order (doc. 2968).

with any type of log of what information was scrubbed. Plaintiffs' counsel also advised the Court that Defendant had locked certain cells and data on the Excel spreadsheets prior to producing them so that Plaintiffs could not access those cells.

Defendant admitted that it had scrubbed the metadata from and locked certain data on the spreadsheets prior to producing them. It argued that the spreadsheets' metadata is irrelevant and contains privileged information. Defendant further argued that Plaintiffs never requested the metadata be included in the electronic Excel spreadsheets it produced and that metadata was never discussed at any of the discovery conferences.

After hearing the respective arguments of counsel, the Court ordered Defendant to show cause why it should not be sanctioned for not complying with "what at least I understood my Order to be, which was that electronic data be produced in the manner in which it was maintained, and to me that did not allow for the scrubbing of metadata because when I talk about electronic data, that includes the metadata."¹⁹ 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."²⁰ 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

¹⁹July 7, 2005 Tr. (doc. 3147) p. 60, l. 20-25.

²⁰*Id.* at p. 61, l. 4-6.

“it’s going to take some clear showing or otherwise there are going to be appropriate sanctions, which at least will be the production of the information in the format it was maintained.”²¹

The Court’s corresponding written Order dated July 12, 2005 (doc. 3037), required Defendant to show cause in writing: (1) “why it should not produce the electronic spreadsheets in the manner in which they were maintained” and (2) “why it should not be sanctioned for its failure to comply with the Court’s ruling from the June 9, 2005 discovery conference, memorialized in the Court’s June 16, 2005 Order (doc. 2953), directing Defendant to produce electronic spreadsheets in the manner in which they were maintained.”²² The Order more specifically required Defendant “to show cause why it scrubbed the metadata and locked certain data on these electronic spreadsheets prior to producing them to Plaintiffs without either the agreement of the parties or the approval of the Court.”²³

II. Discussion

In its response to the Court’s Show Cause Order, Defendant states that it provided the spreadsheets as requested by Plaintiffs in native Excel format, with the following four modifications, none of which affected the discoverable data regarding the RIFs at issue this lawsuit: (1) Defendant deleted the adverse impact analyses; (2) Defendant deleted the social security numbers of employees referenced in the spreadsheets; (3) Defendant deleted metadata from the electronic files that included the spreadsheets; and (4) Defendant locked the value of the cells in the spreadsheets.

²¹*Id.* at p. 61, l. 9-13.

²²Doc. 3037.

²³*Id.*

Defendant asserts that these four modifications to the RIF spreadsheets were made in good faith and for legitimate purposes, namely to protect from disclosure information that Judge Lungstrum determined is not discoverable, to ensure the Court's rulings could not be circumvented, and to maintain the integrity of the data. Defendant maintains that under these circumstances its actions were appropriate and do not warrant the imposition of sanctions.

A. Adverse Impact Analyses and Social Security Numbers

Although Defendant's response to the Show Cause Order explains why it redacted or removed adverse impact analyses information and social security numbers from the electronic spreadsheets prior to producing them, the Court finds this explanation is not necessary, as Defendant previously indicated that it intended to redact this information. At the June 2, 2005 discovery conference, Defendant indicated to the Court that it intended to redact social security numbers and privileged communications consisting of the adverse impact analyses that were prepared.²⁴ Plaintiffs did not object at the June 2, 2005 discovery conference to Defendant's stated intention to redact this information prior to producing the spreadsheets. Plaintiffs also never raised Defendant's redaction of the adverse impact analyses information or social security numbers at the July 7, 2005 discovery conference when they reported that Defendant had scrubbed the spreadsheets' metadata and locked spreadsheet data. Moreover, the Court's Show Cause Order did not require Defendant to show cause why it deleted the adverse impact analyses and social security numbers. The Court therefore finds that Defendant need not show cause for the redaction or removal of the adverse impact analyses and social security numbers from the RIF spreadsheets.

B. Metadata

²⁴See June 2, 2005 Tr. p. 57-58.

The Court's Show Cause Order, however, does require Defendant to show cause for its actions in scrubbing the metadata from the electronic spreadsheets prior to producing them to Plaintiffs. Defendant claims that it scrubbed the metadata from the spreadsheets to preclude the possibility that Plaintiffs could "undelete" or recover privileged and protected information properly deleted from the spreadsheets and to limit the information in the spreadsheets to those pools from which it made the RIF decisions currently being litigated. In an attempt to justify its actions, Defendant contends that emerging standards of electronic discovery articulate a presumption against the production of metadata, which is not considered part of a document, unless it is both specifically requested and relevant. Defendant next argues that Plaintiffs never sought the production of metadata. Finally, Defendant argues that its removal of metadata was consistent with, if not compelled by, Judge Lungstrum's prior orders. Defendant asserts that these reasons support a determination that it has shown cause for its removal of the metadata from the Excel spreadsheets prior to producing them to Plaintiffs.

1. Emerging standards of electronic discovery with regard to metadata

a. What is metadata?

Before addressing whether Defendant was justified in removing the metadata from the Excel spreadsheets prior to producing them to Plaintiffs, a general discussion of metadata and its implications for electronic document production in discovery is instructive.

Metadata, commonly described as "data about data," is defined as "information describing the history, tracking, or management of an electronic document."²⁵ Appendix F to *The Sedona Guidelines*:

²⁵Proposed advisory committee note to Federal Rule of Civil Procedure 26(f). The pending rule amendments and notes can be viewed at:
(continued...)

Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age 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 (including data demographics such as size, location, storage requirements and media information.)”²⁶ Technical Appendix E to the *Sedona Guidelines* provides an extended description of metadata. It further defines metadata to include “all of the contextual, processing, and use information needed to identify and certify the scope, authenticity, and integrity of active or archival electronic information or records.”²⁷ Some examples of metadata for electronic documents include: a file’s name, a file’s location (e.g., directory structure or pathname), file format or file type, file size, file dates (e.g., creation date, date of last data modification, date of last data access, and date of last metadata modification), and file permissions (e.g., who can read the data, who can write to it, who can run it).²⁸ Some metadata, such as file dates and sizes, can easily be seen by users; other metadata can be hidden or embedded and unavailable to computer users who are not technically adept.²⁹

²⁵(...continued)

<http://www.uscourts.gov/rules/comment2005/CVAug04.pdf#page=24>. On September 20, 2005, the Judicial Conference approved the proposed rule amendments and they were transmitted to the Supreme Court for approval. The proposed rule amendments are set to become effective December 1, 2006.

²⁶*The Sedona Guidelines: Best Practice Guidelines & Commentary for Managing Information & Records in the Electronic Age*, App. F (The Sedona Conference Working Group Series, Sept. 2005 Version), available generally at <http://www.thesedonaconference.org> and more specifically at http://www.thesedonaconference.org/content/miscFiles/TSG9_05.pdf.

²⁷*Id.* at App. E.

²⁸*Id.* at App. E. n.1.

²⁹*Id.* at App. E.

Most metadata is generally not visible when a document is printed or when the document is converted to an image file.³⁰ Metadata can be altered intentionally or inadvertently and can be extracted when native files are converted to image files.³¹ Sometimes the metadata can be inaccurate, as when a form document reflects the author as the person who created the template but who did not draft the document.³² In addition, metadata can come from a variety of sources; it can be created automatically by a computer, supplied by a user, or inferred through a relationship to another document.³³

Appendix E to *The Sedona Guidelines* further explains the importance of metadata:

Certain metadata is critical in information management and for ensuring effective retrieval and accountability in record-keeping. Metadata can assist in proving the authenticity of the content of electronic documents, as well as establish the context of the content. Metadata can also identify and exploit the structural relationships that exist between and within electronic documents, such as versions and drafts. Metadata allows organizations to track the many layers of rights and reproduction information that exist for records and their multiple versions. Metadata may also document other legal or security requirements that have been imposed on records; for example, privacy concerns, privileged communications or work product, or proprietary interests.³⁴

The Microsoft Office Online website lists several examples of metadata that may be stored in Microsoft Excel spreadsheets, as well as other Microsoft applications such as Word or PowerPoint: author

³⁰*Id.*

³¹*The Sedona Conference Glossary*, p. 28-29.

³²*The Sedona Principles: Best Practices, Recommendations & Principles for Addressing Electronic Document Discovery*, Cmt. 12.a. (The Sedona Conference Working Group Series, July 2005 Version), available generally at <http://www.thesedonaconference.org>, and more specifically at available at http://www.thesedonaconference.org/content/miscFiles/7_05TSP.pdf.

³³*The Sedona Guidelines*, App. E.

³⁴*Id.*

name or initials, company or organization name, identification of computer or network server or hard disk where document is saved, names of previous document authors, document revisions and versions, hidden text or cells, template information, other file properties and summary information, non-visible portions or embedded objects, personalized views, and comments.³⁵

It is important to note that metadata varies with different applications. As a general rule of thumb, the more interactive the application, the more important the metadata is to understanding the application's output. At one end of the spectrum is a word processing application where the metadata is usually not critical to understanding the substance of the document. The information can be conveyed without the need for the metadata. At the other end of the spectrum is a database application where the database is a completely undifferentiated mass of tables of data. The metadata is the key to showing the relationships between the data; without such metadata, the tables of data would have little meaning. A spreadsheet application lies somewhere in the middle. While metadata is not as crucial to understanding a spreadsheet as it is to a database application, a spreadsheet's metadata may be necessary to understand the spreadsheet because the cells containing formulas, which arguably are metadata themselves, often display a value rather than the formula itself. To understand the spreadsheet, the user must be able to ascertain the formula within the cell.

Due to the hidden, or not readily visible, nature of metadata, commentators note that metadata created by any software application has the potential for inadvertent disclosure of confidential or privileged

³⁵Microsoft Office Online: Find and Remove Metadata (Hidden Information) in Your Legal Documents, <http://office.microsoft.com/en-us/assistance/HA010776461033.aspx>.

information in both a litigation and non-litigation setting, which could give rise to an ethical violation.³⁶ One method commonly recommended to avoid this inadvertent disclosure is to utilize software that removes metadata from electronic documents.³⁷ The process of removing metadata is commonly called “scrubbing” the electronic documents.³⁸ In a litigation setting, the issue arises of whether this can be done without either the agreement of the parties or the producing party providing notice through an objection or motion for protective order.

b. Whether emerging standards of electronic discovery articulate a presumption against the production of metadata

With the increasing usage of electronic document production in discovery, metadata presents unique challenges regarding the production of documents in litigation and raises many new discovery questions. The group of judges and attorneys comprising the Sedona Conference Working Group on Best Practices for Electronic Document Retention and Production (Sedona Electronic Document Working Group) identified metadata as one of the primary ways in which producing electronic documents differs from producing paper documents.³⁹ The Sedona Electronic Document Working Group also recognized that

³⁶See e.g., Brian D. Zall, *Metadata: Hidden Information in Microsoft Word Documents and its Ethical Implications*, 33-Oct. Colo. Law. 53 (2004).

³⁷*Id.* at 58.

³⁸*Id.*

³⁹*The Sedona Principles*, p. 5-6.

understanding when metadata should be specifically preserved and produced represents one of the biggest challenges in electronic document production.⁴⁰

Defendant contends that emerging standards of electronic discovery articulate a presumption against the production of metadata. To determine whether Defendant's contention is accurate, the Court must first identify the emerging standards for the production of metadata. Then the Court must determine whether these emerging standards provide any guidance on the issue before the Court, i.e., whether a court order directing a party to produce electronic documents as they are maintained in the ordinary course of business requires the producing party to produce those documents with the metadata intact. A related issue is determining which party has the initial burden with regard to the disclosure of metadata. Does the requesting party have the burden to specifically request metadata and demonstrate its relevance? Or does the party ordered to produce electronic documents have an obligation to produce the metadata unless that party timely objects to production of the metadata?

The Court starts with the current version of Federal Rule of Civil Procedure 34. This rule provides that "[a]ny party may serve on any other party a request (1) to produce and permit the party making the request, or someone acting on the requestor's behalf, to inspect and copy, any designated documents (including writings, drawings, graphs, charts, photographs, phonorecords, and other data compilations from which information can be obtained, translated, if necessary, by the respondent through detection devices into reasonably usable form)."⁴¹ "A party who produces documents for inspection shall produce them as

⁴⁰*Id.*

⁴¹Fed. R. Civ. P. 34(a).

they are kept in the usual course of business or shall organize and label them to correspond with the categories in the request.”⁴²

Federal Rule of Civil Procedure 34 includes “data compilations” in the listing of items that constitute a “document.” The 1970 amendment advisory committee note to Rule 34 states that “Rule 34 applies to electronics data compilations from which information can be obtained only with the use of detection devices, and that when the data can as a practical matter be made usable by the discovering party only through respondent’s devices, respondent may be required to use his devices to translate the data into usable form.”⁴³ Although neither Rule 34 nor its advisory committee notes defines “data compilations,” the 1972 proposed rules advisory committee note to Federal Rule of Evidence 803(6) explains that “the expression ‘data compilation’ is used as broadly descriptive of any means of storing information other than the conventional words and figures in written or documentary form. It includes, but is by no means limited to, electronic computer storage. The term is borrowed from revised Rule 34(a) of the Rules of Civil Procedure.”⁴⁴ Using this broad definition of a “data compilation,” an Excel spreadsheet, and perhaps the underlying metadata itself, would be considered a “data compilation” under Rule 34. The current version of Rule 34, however, provides limited guidance with respect to when “data compilations” or other types of electronic documents have to be produced and in what form they should be produced.

In the past year, the Civil Rules Advisory Committee has proposed to the Judicial Conference several amendments to the Federal Rules of Civil Procedure addressing the discovery of electronically

⁴²Fed. R. Civ. P. 34(b).

⁴³Fed. R. Civ. P. 34(a) advisory committee’s note.

⁴⁴Fed. R. Evid. 803(6) advisory committee’s note.

stored information.⁴⁵ One of the proposed amendments to Rule 34(a) adds “electronically stored information” as a separate category along with “any designated documents.” In addition, the proposed amendments to Rule 34(b) add the following language about the production of electronically stored information:

Unless the parties otherwise agree, or the court otherwise orders,

* * *

(ii) if a request for electronically stored information does not specify the form or forms of production, a responding party must produce the information in a form or forms in which it is ordinarily maintained, or in a form or forms that are reasonably usable.⁴⁶

The proposed committee note to Rule 34(b) provides the following guidance on the form of production:

The amendment to Rule 34(b) permits the requesting party to designate the form or forms in which it wants electronically stored information produced. The form of production is more important to the exchange of electronically stored information than of hard-copy materials, although a party might specify hard copy as the requested form. Specification of the desired form or forms may facilitate the orderly, efficient, and cost-effective discovery of electronically stored information. The rule recognizes that different forms of production may be appropriate for different types of electronically stored information. Using current technology, for example, a party might be called upon to produce word processing documents, e-mail messages, electronic spreadsheets, different image or sound files, and material from databases. Requiring that such diverse types of electronically stored information all be produced in the same form could prove impossible, and even if possible could increase the cost and burdens of producing and using the information. The rule therefore provides that the requesting party may ask for different forms of production for different types of electronically stored information.⁴⁷

⁴⁵ The pending rule amendments can be viewed at <http://www.uscourts.gov/rules/Reports/ST09-2005.pdf>. On September 20, 2005, the Judicial Conference approved the proposed rules amendments and they were transmitted to the Supreme Court for approval. The proposed rule amendments are set to become effective December 1, 2006.

⁴⁶ *Id.*

⁴⁷ *Id.*

Although the proposed amendments to Rule 34 use the phrase “in a form or forms in which it is ordinarily maintained,” they provide no further guidance as to whether a party’s production of electronically stored information “in the form or forms in which it is ordinarily maintained” would encompass the electronic document’s metadata.

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. In the case *In re Verisign, Inc. Securities Litigation*,⁴⁸ Judge Ware of the Northern District of California generally upheld the ruling of Magistrate Judge Trumbell that requiring the production of electronic documents in electronic format was not contrary to law but was supported by the federal rules. At issue in the *Verisign* case was the court’s prior order directing the defendants to produce responsive electronic documents in their native format.⁴⁹ The order expressly stated that “[p]roduction of TIFF version alone is not sufficient,” and that “[t]he electronic version must include metadata as well as be searchable.”⁵⁰ 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.

In another case, which involved the imposition of sanctions against a corporate defendant who repeatedly engaged in a number of discovery abuses, Magistrate Judge Hemann recommended that default

⁴⁸No. C 02-2270 JW, 2004 WL 2445243 (N.D. Cal. Mar. 10, 2004).

⁴⁹2004 WL 2445243, at *2.

⁵⁰*Id.* at *1.

judgment be entered against the defendant as sanctions for its discovery abuses.⁵¹ One of the discovery abuses involved the defendant's production of electronic files with metadata missing. Telxon and the plaintiffs argued that the missing documents, missing attachments, missing metadata, and hard copies of documents in a version different from the versions on any of the electronic databases produced suggested that the defendant was withholding or had improperly destroyed discoverable information. The magistrate judge was skeptical of the defendant's explanations for missing documents and metadata and for differences between hardcopy versions of documents and those on the electronic databases.⁵² Although the case does not directly state that metadata should have been produced, that conclusion can be inferred from the court's holding.

Having concluded that neither the federal rules nor case law provides sufficient guidance on the production of metadata, the Court next turns to materials issued by the Sedona Conference Working Group on Electronic Document Production. The Court finds two of the *Sedona Principles for Electronic Document Production* particularly helpful in determining whether Defendant was justified in scrubbing the metadata from the electronic spreadsheets. Principle 9 states that "[a]bsent a showing of special need and relevance a responding party should not be required to preserve, review, or produce deleted, shadowed,

⁵¹*In re Telxon Corp. Sec. Litig.*, No. 5:98CV2876, 1:01CV1078, 2004 WL 3192729 (N.D. Ohio July 16, 2004), reported in LEXIS as *Hayman v. Price Waterhouse Coopers, LLP*, 2004 U.S. LEXIS 27296. Note that the litigation concluded before the district judge issued any ruling on the magistrate judge's Amended Report and Recommendation.

⁵²*Id.* at *34 (The magistrate judge found the defendant's explanations for the missing documents and metadata and for differences between hardcopy versions of documents and those on any of the electronic databases less than convincing. She stated the defendant's explanations "may explain these phenomena in whole or in part, but 'may' is the operative conditioner here.")

fragmented, or residual data or documents.”⁵³ Principle 12 provides that “[u]nless it is material to resolving the dispute, there is no obligation to preserve and produce metadata absent agreement of the parties or order of the court.”⁵⁴

Comment 9.a. to the *Sedona Principles for Electronic Document Production* focuses on the scope of a “document” under Fed. R. Civ. P. 34. It notes that although Rule 34 was amended in 1970 to add “data compilations” to the list of discoverable documents, there was no suggestion that “data compilations” was intended to turn all forms of “data” into a Rule 34 “document.”⁵⁵ The comment suggests that the best approach to understanding what constitutes a “document” is to examine what information is readily available to the computer user in the ordinary course of business.⁵⁶ If the information is in view, it should be treated as the equivalent of a paper “document.”⁵⁷ Data that can be readily compiled into viewable information, whether presented on the screen or printed on paper, is also a “document” under Rule 34.⁵⁸ The comment, however, cautions that data hidden and never revealed to the user in the ordinary course of business should not be presumptively treated as a part of the “document,” although there are circumstances in which the data may be relevant and should be preserved and produced.⁵⁹ The comment

⁵³*The Sedona Principles*, Principle 9.

⁵⁴*The Sedona Principles*, Principle 12.

⁵⁵*The Sedona Principles*, Cmt. 9.a.

⁵⁶*Id.*

⁵⁷*Id.*

⁵⁸*Id.*

⁵⁹*Id.*

concludes that such data may be discoverable under Rule 34, but the evaluation of the need for and relevance of such discovery should be separately analyzed on a case-by-case basis.⁶⁰ Comment 9.a. provides the following illustration:

A party demands that responsive documents, “whether in hard copy or electronic format,” be produced. The producing party objects to producing the documents in electronic format and states that production will be made through PDF or TIF images on CD-ROMs. The producing party assembles copies of the relevant hard copy memoranda, prints out copies of relevant e-mails and electronic memoranda, and produces them in a PDF or TIF format that does not include metadata. Absent a special request for metadata (or any reasonable basis to conclude the metadata was relevant to the claims and defenses in the litigation), and a prior order of the court based on a showing of need, this production of documents complies with the ordinary meaning of Rule 34.⁶¹

Metadata is specifically discussed in depth in Comment 12.a. to the Sedona Principles. The comment states that “[a]lthough 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.”⁶² The comment further notes that it is likely to remain the exceptional situation in which metadata must be produced.⁶³

⁶⁰*The Sedona Principles*, Cmt. 9.a. (citing *McPeck v. Ashcroft*, 202 F.R.D. 31, 33 (D. D.C. 2001); *McPeck v. Ashcroft*, 212 F.R.D. 33, 35-37 (D.D.C. 2003)).

⁶¹*The Sedona Principles*, Cmt. 9.a.

⁶²*The Sedona Principles*, Cmt. 12.a. (citing *Pub. Citizen v. Carlin*, 184 F.3d 900, 908-11 (D.C. Cir. 1999), cert. denied, 529 U.S. 1003 (2000)).

⁶³*The Sedona Principles*, Cmt. 12.a.

The comment lists several ways in which routine preservation and production of metadata may be beneficial.⁶⁴ The comment balances these potential benefits against the “reality that most of the metadata has no evidentiary value, and any time (and money) spent reviewing it is a waste of resources.”⁶⁵ The comment concludes that a reasonable balance is that, unless the producing party is aware or should be reasonably aware that particular metadata is relevant, the producing party should have the option of producing all, some, or none of the metadata.⁶⁶ The comment sets forth one important caveat to giving the option of producing metadata to the producing party: “Of course, if the producing party knows or should reasonably know that particular metadata is relevant to the dispute, it should be produced.”⁶⁷

c. Application to this case

The narrow issue currently before the Court is whether, under emerging standards of electronic discovery, the Court’s Order directing Defendant to produce electronic spreadsheets as they are kept in the ordinary course of business requires Defendant to produce those documents with the metadata intact. As noted above, the Court finds insufficient guidance in either the federal rules or case law, and thus relies primarily on the Sedona Conference Principles and comments for guidance on the emerging standards of electronic document production, specifically with regard to metadata. While recognizing that the Sedona Principles and comments are only persuasive authority and are not binding, the Court finds the Sedona

⁶⁴*Id.*

⁶⁵*Id.*

⁶⁶*Id.*

⁶⁷*Id.*

Principles and comments particularly instructive in how the Court should address the electronic discovery issue currently before it.

Comment 9.a. to the *Sedona Principles for Electronic Document Production* approaches discoverability based on what constitutes a “document” under Rule 34. This comment uses viewability as the determining factor in whether something should be presumptively treated as a part of a “document.” Using viewability as the standard, all metadata ordinarily visible to the user of the Excel spreadsheet application should presumptively be treated as part of the “document” and should thus be discoverable. For spreadsheet applications, the user ordinarily would be able to view the contents of the cells on the spreadsheets, and thus the contents of those cells would be discoverable.

In light of the proposed amendment to Rule 34, which adds “electronically stored information” as its own separate category, it is no longer necessary to focus on what constitutes a “document.” With regard to metadata in general, the Court looks to Principle 12 and Comment 12.a. to the *Sedona Principles*. Based upon this Principle and Comment, emerging standards of electronic discovery appear to articulate a general presumption against the production of metadata, but provide a clear caveat when the producing party is aware or should be reasonably aware that particular metadata is relevant to the dispute.

Based on these emerging standards, the Court holds that when a party is ordered to produce electronic documents as they are maintained in the ordinary course of business,⁶⁸ the producing party should produce the electronic documents with their metadata intact, unless that party timely objects to production of metadata, the parties agree that the metadata should not be produced, or the producing party

⁶⁸This same reasoning would apply if the court ordered a party to produce the electronic documents as an “active file” or in their “native format.”

requests a protective order.⁶⁹ The initial burden with regard to the disclosure of the metadata would therefore be placed on the party to whom the request or order to produce is directed. 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. 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.

i. Relevancy

Defendant maintains that the metadata it removed from its electronic spreadsheets has absolutely no evidentiary value and is completely irrelevant. It argues that Plaintiffs' suggestion that the metadata may identify the computers used to create or modify the spreadsheets or reveal titles of documents that may assist in efforts to piece together the facts of the RIFs at issue in this case has no relevance to Plaintiffs' claim that Defendant maintained discriminatory policies or practices used to effectuate a pattern and practice of age discrimination. Defendant likewise argues that the metadata is not necessary because the titles of documents can be gleaned from the subject spreadsheets, and these titles adequately describe the data included in such spreadsheets.

The Court agrees with Defendant that certain metadata from the spreadsheets may be irrelevant to the claims and defenses in this case. The Court, however, does not find that all of the spreadsheets' metadata is irrelevant. In light of Plaintiffs' allegations that Defendant reworked pools of employees in

⁶⁹The same principle may apply when a party *requests* electronic documents be produced as they are maintained in the ordinary course of business, as an "active file," or in their "native format."

order to improve distribution to pass its adverse impact analysis, the Court finds that some of the metadata is relevant and likely to lead to the discovery of admissible evidence. While the Court cannot fashion an exhaustive list of the spreadsheet metadata that may be relevant, the Court does find that metadata associated with any changes to the spreadsheets, the dates of any changes, the identification of the individuals making any changes, and other metadata from which Plaintiffs could determine the final versus draft version of the spreadsheets appear relevant. Plaintiffs' allegation that Defendant reworked the pools is not a new allegation. Thus, Defendant should reasonably have known that Plaintiffs were expecting the electronic spreadsheets to contain their metadata intact. Furthermore, if 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.

ii. Reliability

Defendant also argues that the metadata removed from the electronic spreadsheets may be inaccurate and therefore has no evidentiary value. The Court finds that this is not sufficient justification for removing the metadata absent agreement of the parties or the Court's approval. If Defendant had any concerns regarding the accuracy or reliability of the metadata, it should have communicated those concerns to the Court before it scrubbed the metadata.

iii. Privilege

Defendant also argues that production of certain metadata removed by Defendant would facilitate the revelation of information that is attorney-client privileged and/or attorney work product. Defendant claims that through the use of easily accessible technology, metadata may reveal information extracted from

a document, such as the items redacted by Defendant's counsel, as well as other protected or privileged matters. It further claims that metadata may create a data trail that reveals changes to prior drafts or edits.

The Court agrees with Defendant that it should not be required to produce metadata directly corresponding to the information that it was permitted to redact, namely the adverse impact analyses and social security numbers. The Court is cognizant that all or some of the metadata may reveal the redacted information. The Court will therefore permit Defendant to remove metadata directly corresponding to Defendant's adverse impact analyses and social security number information.

For any other metadata Defendant claims is protected by the attorney-client privilege or as attorney work product, the Court finds that Defendant should have raised this issue prior to its unilateral decision to produce the spreadsheets with the metadata removed. Fed. R. Civ. P. 26(b)(5) requires a party withholding otherwise discoverable information on the basis of privilege to make the claim expressly and to describe the nature of the documents, communications, or things not produced or disclosed in a manner that, without revealing the privileged information, will enable the other parties to assess the applicability of the privilege. Normally, this is accomplished by objecting and providing a privilege log for "documents, communications, or things"⁷⁰ not produced.

In this case, Defendant has failed to object and has not provided a privilege log identifying the electronic documents that it claims contain privileged metadata. Defendant has not provided the Court with even a general description of the purportedly privileged metadata that was scrubbed from the spreadsheets. As Defendant has failed to provide any privilege log for the electronic documents it claims contain metadata that will reveal privileged communications or attorney work product, the Court holds that Defendant has

⁷⁰Fed. R. Civ. P. 26(b)(5).

waived⁷¹ any attorney-client privilege or work product protection with regard to the spreadsheets' metadata except for metadata directly corresponding to the adverse impact analyses and social security number information, which the Court has permitted Defendant to remove from the spreadsheets.⁷²

2. Plaintiffs never requested the production of metadata

Defendant also argues that Plaintiffs never requested the metadata and that metadata was never mentioned during any of the discovery conferences. While metadata was never mentioned during any of the discovery conferences or in any of the Court's orders, the Court finds that Defendant should reasonably have been aware that the spreadsheets' metadata was encompassed within the Court's directive that it produce the electronic Excel spreadsheets as they are maintained in the regular course of business. Defendant is correct in asserting that Plaintiffs never expressly requested metadata and that the Court never expressly ordered Defendant to produce the electronic spreadsheets' metadata. However, taken in the context of Plaintiffs' stated reasons for requesting the Excel spreadsheets in their native electronic format and the Court's repeated statements that the spreadsheets should be produced in the electronic form in which they are maintained, the Court finds that Defendant should have reasonably understood that the

⁷¹See *Employer's Reinsurance Corp. v. Clarendon Nat'l Ins. Co.*, 213 F.R.D. 422, 428 (D. Kan. 2003) (stating that it is well-established law in this district that failure to produce a privilege log can result in a waiver of any protection afforded to those documents); *Haid v. Wal-Mart Stores, Inc.*, No. 99-4186-RDR, 2001 WL 964102, at *1 (D. Kan. June 25, 2001) (failure to produce a privilege log or production of an inadequate privilege log may be deemed a waiver of the privilege asserted); *Starlight Int'l, Inc. v. Herlihy*, No. 97-2329-GTV, 1998 WL 329268, at *3 (D. Kan. June 16, 1998) (same).

⁷²The Court recognizes that this ruling impacts the relief requested by Plaintiffs in their pending Motion to Declare Defendant Has Waived Any Asserted Privilege With Regard to Certain Documents and Electronic Spreadsheets (doc. 3192). The Court intends its ruling herein to resolve the issue of waiver with regard to metadata only. The remainder of Plaintiffs' Motion to Declare Defendant Has Waived Any Asserted Privilege With Regard to Certain Documents and Electronic Spreadsheets (doc. 3192) is still pending.

Court expected and intended for Defendant to produce the spreadsheets' metadata along with the Excel spreadsheets. If Defendant did not understand the Court's ruling, it should have requested clarification of the Court's order. As the Sedona Working Group on Electronic Document Production observed: "Of course, if the producing party knows or should reasonably know that particular metadata is relevant to the dispute, it should be produced."⁷³ Here, the Court finds that Defendant should have reasonably known that the metadata was relevant to the dispute and therefore should have either been produced or an appropriate objection made or motion filed.

3. Whether Judge Lungstrum's prior orders compel the removal of metadata

Defendant next argues that its removal of metadata from the RIF spreadsheets was consistent with, if not compelled by, Judge Lungstrum's prior orders. Defendant appears to be arguing that Judge Lungstrum's prior rulings with regard to the adverse impact analyses compel the removal of all metadata. As discussed in section II.A. above, Defendant may redact its adverse impact analyses and social security numbers from the spreadsheets prior to producing them to Plaintiffs. Defendant may also remove the metadata directly corresponding to the redacted information. The undersigned Magistrate Judge, however, does not find that Judge Lungstrum's prior rulings compel the removal of all other metadata. But in any event, if Defendant believed that Judge Lungstrum's rulings compelled the removal of all metadata, then Defendant should have either objected or requested a protective order before it produced the spreadsheets with all metadata removed.

C. Locked Spreadsheet Cells and Data

⁷³*The Sedona Principles*, Cmt. 12.a.

The Court next addresses whether Defendant has shown cause for the locking of certain data and cells on the Excel spreadsheets produced to Plaintiffs. Defendant states that it locked the value of the cells in the spreadsheets to ensure the integrity of the data regarding RIFs, i.e., to ensure that the data could not be accidentally or intentionally altered. Defendant claims its purpose was not to preclude Plaintiffs from sorting or filtering the data, a task it claims Plaintiffs could easily accomplish by copying the data to another spreadsheet. Instead, Defendant claims it locked the data to preclude inadvertent or intentional modification of the data it produced. It argues that because electronic data is not ordinarily static, locking the data was essential to ensure that Defendant could demonstrate data subsequently used in the case was identical to data it produced electronically. It asserts that no malicious intent was associated with Defendant's efforts to preserve the integrity of the data it produced.

The Court finds that Defendant has failed to show sufficient cause for its unannounced and unilateral actions in locking certain data and cells on the Excel spreadsheets the Court ordered it to produce to Plaintiffs in the manner in which they were maintained. None of the reasons asserted by Defendant justifies its decision to lock the spreadsheet cells and data prior to producing them to Plaintiffs. While the Court's Order did not expressly state that the spreadsheets should be produced "unlocked," Defendant should have been reasonably aware that locking the spreadsheets' cells and data was not complying with the spirit of the Court's directive that the spreadsheets be produced as they are kept in the ordinary course of business. Moreover, at the June 2, 2005 discovery conference, Plaintiffs specifically detailed their difficulties with the hard copy versions of the spreadsheets produced by Defendant, including their complaints that the hard copy versions cut off the information contained in the spreadsheet columns and cells.

Defendant's concerns regarding maintaining the integrity of the spreadsheet's values and data could have been addressed by the less intrusive and more efficient use of "hash marks." For example, 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.⁷⁴ 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.⁷⁵ 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.⁷⁶ This method allows a large amount of data to be self-authenticating with a rather small hash mark, efficiently assuring that the original image has not been manipulated.⁷⁷

Defendant also states that despite its ongoing concerns with the integrity of the data in the electronic spreadsheets, it has voluntarily reproduced "unlocked" copies of the spreadsheets that were produced before July 7, 2005. It also represents that all successive productions of spreadsheets have not been locked. The Court concludes from Defendant's statement that it has already produced "unlocked" spreadsheets to Plaintiffs. To the extent that Defendant has not produced "unlocked" versions of all of the spreadsheets previously produced with locked cells or data, Defendant shall produce "unlocked" versions

⁷⁴Dean M. Harts, *Reel to Real: Should You Believe What You See? Keeping the good and eliminating the bad of computer-generated evidence will be accomplished through methods of self-authentication and vigilance*, 66 Def. Couns. J. 514, 522 (1999).

⁷⁵*Id.*

⁷⁶*Id.*

⁷⁷*Id.*

of those spreadsheets by the next discovery conference on **October 6, 2005**, or advise the Court at the discovery conference of the date when this information will be produced.

III. Sanctions

The Court's Show Cause Order also required Defendant to show cause why it should not be sanctioned for its failure to comply with the Court's ruling directing Defendant to produce electronic spreadsheets in the manner in which they were maintained. Defendant states that it did not understand the Court's direction to produce electronic spreadsheets included the production of metadata and that its actions were not made in bad faith. It points out that it has already produced hundreds of documents in response to formal and informal requests for production, answered hundreds of interrogatories, and produced and scheduled scores of witnesses for deposition in support of its assertion that it has acted in good faith throughout this litigation.

The Court concludes that Defendant has shown cause why it should not be sanctioned for its actions in scrubbing the metadata and locking spreadsheet cells and data. Although the Court intended its ruling requiring Defendant to produce the electronic RIF-related spreadsheets in the manner in which they were ordinarily maintained to include the metadata, the Court recognizes that the production of metadata is a new and largely undeveloped area of the law. This lack of clear law on production of metadata, combined with the arguable ambiguity in the Court's prior rulings, compels the Court to conclude that sanctions are not appropriate here.

The Court, however, wants to clarify the law regarding the production of metadata in this case. When the Court orders a party to produce an electronic document in the form in which it is regularly maintained, i.e., in its native format or as an active file, that production must include all metadata unless that

party timely objects to production of the metadata, the parties agree that the metadata should not be produced, or the producing party requests a protective order.

Defendant avoids sanctions with regard to its locking of the spreadsheet cells and data by its decision to voluntarily reproduce “unlocked” versions of these spreadsheets to Plaintiffs. As directed above, to the extent that Defendant has not produced “unlocked” versions of all the spreadsheets previously produced with locked cells or data, it shall produce “unlocked” versions of those spreadsheets by the next discovery conference on **October 6, 2005**, or advise the Court at the discovery conference of the date when this information will be produced.

IT IS THEREFORE ORDERED that Defendant has failed to show cause why it should not produce the electronic spreadsheets in the manner in which they were maintained. Defendant therefore shall produce the electronic spreadsheets in the manner in which they were maintained, which includes the spreadsheets’ metadata. Defendant may, however, redact its adverse impact analyses and any social security numbers from the spreadsheets prior to producing them to Plaintiffs. Defendant may also remove the metadata directly corresponding to the redacted information. Defendant shall produce the electronic spreadsheets by the next discovery conference on **October 6, 2005**, or advise the Court at the October 6, 2005 discovery conference of the date when these spreadsheets will be produced. Any assertion of attorney-client privilege or work product protection with regard to the metadata contained within these Excel spreadsheets, other than the metadata directly corresponding to the adverse impact analyses and social security number information, is deemed waived due to Defendant’s failure to object and produce a privilege log regarding the metadata.

IT IS FURTHER ORDERED that to the extent Defendant has not reproduced “unlocked” versions of the spreadsheets previously produced with locked cells or data, Defendant shall produce “unlocked” versions of those spreadsheets by the next discovery conference on **October 6, 2005** or advise the Court at the discovery conference of the date when these spreadsheets will be produced.

IT IS FURTHER ORDERED that Defendant has shown sufficient cause why it should not be sanctioned for its failure to comply with the Court’s ruling directing it to produce electronic spreadsheets in the manner in which they were maintained.

IT IS SO ORDERED.

Dated in Kansas City, Kansas on this 29th day of September, 2005.

s/ David J. Waxse
David J. Waxse
United States Magistrate Judge

cc: All counsel