ARTICLES

SHOW ME THE METADATA: THE CASE TO STANDARDIZE AUTOMATIC METADATA PRODUCTION FOR E-MAIL COMMUNICATIONS

Cory Simmons-Edler[*]

---

* Cory Simmons-Edler, Senior Editor, *Rutgers Computer and Technology Law Journal*, Rutgers School of Law—Newark, J.D. anticipated May 2015; B.A., English and American Literature, New York University. The Author would like to thank the staff of the University of Florida Levin College of Law, *Journal of Technology Law and Policy* for all of their hard work in the preparation of this Article for publication, particularly Editor-in-Chief Kristen Vogl and Articles Editor Nick Klimas. The Author would also like to thank Changi Wu for his feedback and guidance in the writing of this Article.

I. INTRODUCTION

In the last thirty-five years, our world has rapidly made the transformative shift from analog to digital. Where once reams of paper records were stored in rows of file cabinets, the same information can now potentially fit in a storage unit no larger than a fingernail.[1] Common practice in document maintenance has shifted so much in fact, that "[e]lectronically stored information has become the dominant form of discovery in the litigation process."[2]

Concurrent with this transition from physical paper to electronic document is the shift of many of our oral conversations to written format. What once might have been discussed orally between colleagues or friends at the water cooler or over coffee now takes place via Short Message System (SMS) text or e-mail. Both the importance and volume of the written word in litigation has never been greater. Of the many forms that the written word can take, e-mail is probably the most prevalent in the business and litigation context.[3] In 2012, 2.2 billion people worldwide used e-mail and 144 billion e-mails were sent each day.[4]

Both the sheer number of individual communications, as well as the difficulties in processing and categorizing this exploding volume of information presents new challenges to the legal community. Within the past ten to fifteen years a number of treatises and scholarly works[5] have been produced, and some have been enormously influential on the evolution of law surrounding the discovery of Electronically Stored Information (ESI). Chief among these is the publication of the Sedona

1. *See Secure Digital*, WIKIPEDIA, http://en.wikipedia.org/wiki/Secure_Digital (last visited Dec. 9, 2014).

2. Burke T. Ward et al., *Electronic Discovery: Rules for a Digital Age*, 18 B.U. J. SCI. & TECH. L. 150, 151 (2012).

3. *See, e.g.*, FED. R. CIV. P. 34 advisory committee's note, amend. 2006, subdiv. (a).

4. *Internet 2012 in Numbers*, ROYAL PINGDOM (Jan. 16, 2003), http://royal.pingdom.com/2013/01/16/internet-2012-in-numbers/.

5. *See, e.g.*, Philip J. Favro, *A New Frontier in Electronic Discovery: Preserving and Obtaining Metadata*, 13 B.U. J. SCI. & TECH. L. 1 (2007); *see also* W. Lawrence Wescott II, *The Increasing Importance of Metadata in Electronic Discovery*, 14 RICH. J.L. & TECH. 10 (2008).

Conference Principles in 2004,[6] and the revised principles in 2007.[7] One of the main differences between ESI and conventionally stored information is the fact that ESI includes certain data points demonstrably outside of the text or document itself.[8] These data points are called metadata.[9]

This Article examines the way that electronically stored documents are treated in the discovery process, and particularly focuses on the metadata attached to email, the most prevalent form of communication in the litigation context. This Article also surveys how the Federal Rules of Civil Procedure treat metadata, as well as case law ruling on the discovery of metadata, and the evolution of law surrounding it. The Article finishes by interpreting this evolution, and suggests a modest standardization and automatic disclosure during discovery of specific e-mail metadata fields. These changes will limit the cost and streamline the process of e-discovery, particularly for unsophisticated litigants.

II. WHAT IS DISCOVERY?

Broadly speaking, discovery is a fact-finding process that occurs after a lawsuit is filed, whereby litigants manufacture evidence in support of their position, and learn about information detrimental to their case.[10] The process is "based on the belief that a free exchange of information is more likely to help uncover the truth regarding the facts in issue."[11] The discovery process is generally "designed to clarify issues in litigation, obtain evidence not readily accessible to opposing counsel, and to ascertain information that may be used at trial."[12]

III. ELECTRONICALLY STORED INFORMATION (ESI)

ESI is a wide-ranging category, but information is generally "considered 'electronic' if it exists in a medium that can only be read by

---

6. *See generally* THE SEDONA PRINCIPLES: BEST PRACTICES RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION (Jonathan M. Redgrave et al. eds., 1st ed. 2004) [hereinafter SEDONA PRINCIPLES 1st ed.].

7. *See generally* THE SEDONA CONFERENCE, THE SEDONA PRINCIPLES: BEST PRACTICES RECOMMENDATIONS & PRINCIPLES FOR ADDRESSING ELECTRONIC DOCUMENT PRODUCTION (Jonathan M. Redgrave et al. eds., 2d ed. 2007) [hereinafter SEDONA PRINCIPLES 2d ed.].

8. SEDONA PRINCIPLES 1st ed., *supra* note 6, at 5.

9. *Id.*

10. *What is Discovery?*, THE COCHRAN FIRM, http://www.cochranfirm.com/resources/Ask-our-Lawyers/whatisdiscovery.html (last visited Dec. 9, 2014).

11. Ward et al., *supra* note 2, at 153.

12. *Id.* at 154.

a computer, including email, web pages, word processing files, audio and video files, images, computer databases, spreadsheets and virtually anything else that is stored on a computing device."[13] By this definition, e-mail is a subset of ESI. While the common definition of e-mail may seem obvious (you know it when you see it), the legal definition of what constitutes e-mail has proven to be somewhat difficult.[14] E-mail can vary greatly in form, function, and content.[15] The Alliance for Telecommunications Industry Solutions defines an e-mail as

> [a]n electronic means for communication in which (a) usually text is transmitted (but sometimes also graphics and/or audio information), (b) operations include sending, storing, processing, and receiving information, (c) users are allowed to communicate under specified conditions, and (d) messages are held in storage until called for by the addressee. Some e-mail software permits the attachment of separate electronic files, e.g., word-processor files, graphics files, audio files.[16]

Similarly, another well-established U.S. standards bureau, defines "e-mail," and "e-mail system" thusly:

> *Electronic mail message*. A document created or received on an electronic mail system including brief notes, more formal or substantive narrative documents, and any attachments, such as word processing and other electronic documents, which may be transmitted with the message.[17]
>
> *Electronic mail system*. A computer application used to create, receive, and transmit messages and other documents. Excluded

---

13. *Id.* at 155.

14. *See* John C. Montaña, *Legal Obstacles to E-Mail Message Destruction*, THE ARMA INT'L EDUC. FOUND., at 8.

15. *See id.*

> Everyone who gets e-mail gets a great deal of automated e-mail – advertisements and order acknowledgements are the commonest examples – that cannot meaningfully be said to have a human sender at the other end; and most e-mail users have, at one point or another sent e-mail to an auto-receipt address at a business or government agency that either deals with the response automatically or directs it to some unknown person. Thus, the human-to-human element of e-mail that we often associate with it is clearly not a necessary prerequisite.

Id.

16. Alliance for Telecommunications Industry Solutions (ATIS), ATIS TELECOM GLOSSARY, http://www.atis.org/glossary/definition.aspx?id=7643 (last visited Dec. 9, 2014).

17. Montaña, *supra* note 14, at 10.

> from this definition are file transfer utilities (software that transmits files between users but does not retain any transmission data), data systems used to collect and process data that have been organized into data files or data bases on either personal computers or mainframe computers, and word processing documents not transmitted on an e-mail system.[18]

E-mail communications are not altogether different in kind from other sorts of ESI, but do commonly have several unique characteristics that are not often found in other forms of ESI, most notably as a vehicle to send information (including other forms of ESI such as photos, text documents, and spreadsheets) electronically between one or more senders and receivers.[19] Because of this, e-mail and particularly the metadata of e-mail relating to its unique characteristics, the habitual sending and receiving, warrants special treatment within the context of discovery.

IV. DIFFERENCES BETWEEN CONVENTIONAL AND ELECTRONIC DISCOVERY

E-discovery as a formally distinct concept was first addressed by the Judicial Conference Advisory Committee on Civil Rules ("Advisory Committee") during the 1996 discovery project.[20] During the project, several lawyers brought up the concept of electronic discovery, and the potentially massive change that could accompany it.[21]

At first the subject was approached with timidity; members of the Advisory Committee and the legal community at large did not necessarily understand the substantive differences between ESI and traditional paper documents or the need for special treatment.[22] After all, "[T]here were no special rules added to deal with the discovery challenges produced by the introduction of photocopiers. . ."[23] However, in the years since the topic of e-discovery was first breached, it has become evident that the field is here to stay, and presents challenges not before encountered in the pre-digital era.

The Sedona Conference outlines six broad differences between

---

18. *Id.* (citing 36 C.F.R. § 1234.2).

19. Even this characteristic is not absolutely essential to a medium's definition as e-mail. *See, e.g.*, Max Fisher, *Here's the E-mail Trick Petraeus and Broadwell Used to Communicate*, WASH. POST (Nov. 12, 2012), http://www.washingtonpost.com/blogs/worldviews/wp/2012/11/12/heres-the-e-mail-trick-petraeus-and-broadwell-used-to-communicate/.

20. Richard Marcus, *Only Yesterday: Reflections on Rulemaking Responses to E-discovery*, 73 FORDHAM L. REV. 1, 7 (2004).

21. *Id.*

22. *See id*.

23. *Id.* at 8.

producing paper documents and ESI.[24] These are (A) volume and duplicability, (B) persistence, (C) dynamic, changeable content, (D) metadata, (E) environment-dependence and obsolescence, and (F) dispersion and searchability.[25]

A. *Volume and Duplicability*

Perhaps the most notable difference in E-discovery is the "staggering" quantity of ESI that it is now not only possible, but relatively easy to retain.[26] Additionally, "[e]mails have replaced other forms of communication besides just paper-based communication. Many informal messages that were previously relayed by telephone or at the water cooler are now sent via email."[27] "[A] complex litigation between two large corporate parties can generate the equivalent of more than one hundred million pages of discovery documents, requiring over twenty terabytes of server storage space."[28] When it comes to the reviewing this massive volume of information, new issues have to be addressed regarding the cost of review, and the most efficient way to go about doing it.

B. *Persistence*

The persistence of ESI distinguishes it as well. Generally, ESI is more difficult to dispose of than a paper document. While a shredded paper document is generally beyond repair, and a paper document in the garbage is usually beyond reach once it has been removed from the premises for waste processing,[29] a "deleted" electronic document is seldom beyond the reach of a trained professional.[30]

---

24. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 2–5.

25. *Id.*

26. Marcus, *supra* note 20, at 12.

27. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 2 (quoting *Byers v. Ill. State Police*, 53 Fed. R. Serv. 3d 740 (N.D. Ill. 2002) ("Additionally, computers have the ability to capture several copies (or drafts) of the same email, thus multiplying the volume of documents."). E-mail systems also have the tendency to replicate documents unnecessarily, which has a multiplying effect, especially when e-mails are internal, as much corporate correspondence is. *Id.*

28. Marcus, *supra* note 20, at 12 (quoting Robert Douglas Brownstone, *Collaborative Navigation of the Stormy e-Discovery Seas*, 10 RICH. J.L. & TECH. 53 (2004)) ("Assuming a review rate of one box of paper documents per weekday, per reviewer, a one hundred million page volume corresponds to over thirty person-years of review for each party. In ecological terms, each side would require approximately 6,250 trees just to print one copy of each of the documents it produced and of each of the documents it received.").

29. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 3.

30. *See, e.g.*, Ceglia v. Zuckerberg, 2013 WL 1208558, at *13 (W.D.N.Y. Mar. 26, 2013).

C. *Dynamic, Changeable Content*

Another characteristic of ESI that distinguishes it from traditionally stored data is its dynamic nature. Often, information is stored in such a manner that it is subject to change without any sort of direct human interference.[31] By way of example, simply think of a backup system for a household computer. Once a computer is programmed to back up data at periodic points (say once a week), the computer needs no further direction in order to accomplish that task, it does so automatically.[32] Thus, the contents of a backup drive are dynamic—they can change without direct action by the user. With a traditional paper file, you would have to take some sort of direct action such as photocopying that file in order to have a backup.[33] In a business context, computerized records "often . . . consist of dynamic databases that 'exist' only in the sense that they will provide responsive information when queried."[34] Failure to modify automated ESI retention protocols can be grounds for sanctions[35] when a litigation hold is put in place, or when that party can anticipate a reasonable likelihood of future litigation.[36]

D. *Metadata*

Quite notably, unlike paper documents, electronic documents generally have metadata attached to them. Often, this information is not readily apparent to the viewer.[37] Numerous characteristics are encompassed within the meaning of the word metadata, and the term is often misunderstood.[38] By way of an example, say an individual hires a private detective to eavesdrop on their spouse. That detective might tap the spouse's phone, bug his office, or open his mail. The result of these activities would be the data—analogous to the meaning of the text of an e-mail.[39] Alternately, imagine if that same individual hires that detective again to surveil her spouse. The result would be the details of where he

31. *See generally* Marcus, *supra* note 20, at 13.

32. *Id.*

33. *See id.*

34. *Id.* ("Such databases are difficult to conceptualize as 'documents' in the traditional way, and discovery about or from them blurs the distinction between Rule 33 interrogatories and Rule 34 document requests.").

35. *See* FED. R. CIV. P. 37 (providing various options to address failures to produce requested documentation).

36. Apple v. Samsung Electronics*: The Perils of Email Auto Deletion*, NUTTER (July 27, 2012), http://www.nutter.com/Apple-v-Samsung-Electronics-The-Perils-of-Email-Auto-Deletion-07-27-2012/#.Uw0DsEJdWLM.

37. *See*, *e.g.*, Wescott, *supra* note 5, at 3.

38. *Id.* at 4.

39. Bruce Schneier, METADATA EQUALS SURVEILLANCE, SCHNEIER ON SECURITY (Sept. 23, 2013, 6:21 AM), https://www.schneier.com/blog/archives/2013/09/metadata_equals.html.

went, who he talked to, what he looked at, and how he spent his day. All of this information would be considered metadata.[40] Metadata includes

> [s]uch information [as] file designation, create and edit dates, authorship, comments, and edit history . . . . [E]mail has its own metadata elements that include, among about 1,200 or more properties, such information as the dates that mail was sent, received, replied to or forwarded, blind carbon copy ("bcc") information, and sender address book information.[41]

Metadata can be separated into three basic types: system, substantive, and embedded metadata.[42] System metadata is simply, "data that is automatically generated by a computer system."[43] Examples of system metadata include "[T]he author, date and time of creation, and the date a document was modified."[44] More often than not when people mention metadata, they are referring to systems metadata.

"Substantive Meta-Data is data that reflects the substantive changes made to the document by the user. For example, it may include the text of actual changes to a document."[45] Substantive metadata poses perhaps the biggest risk to attorneys and others who routinely handle sensitive information in practice. An action as simple as sending a word doc without disabling the "undo changes" function can allow for the inadvertent disclosure of sensitive information.[46]

Embedded metadata is defined as "the text, numbers, content, data, or other information that is directly or indirectly inputted into a Native File by a user and which is not typically visible to the user viewing the output display of the Native File on screen or as a print out."[47]

---

40. *Id.*

41. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 3.

42. Wescott, *supra* note 5, at 2 (citing U.S. DISTRICT COURT OF MARYLAND, SUGGESTED PROTOCOL FOR DISCOVERY OF ELECTRONICALLY STORED INFORMATION 25 (2006) [hereinafter SUGGESTED PROTOCOL], *available at* http://www.mdd.uscourts.gov/news/news/ESIProtocol.pdf (last visited Dec. 10, 2014)).

43. SUGGESTED PROTOCOL, *supra* note 42, at 25–26.

44. *Id.* at 26.

45. *Id.*

46. *See, e.g*, Wescott, *supra* note 5, at 3.

47. SUGGESTED PROTOCOL, *supra* note 42, at 27.

> Examples of Embedded Meta-Data include, but are not limited to, spreadsheet formulas (which display as the result of the formula operation), hidden columns, externally or internally linked files (*e.g.*, sound files in PowerPoint presentations), references to external files and content (*e.g.*, hyperlinks to HTML files or URLs), references and fields (*e.g.*, the field codes for an auto-numbered document), and certain database information if the data is part of a database (*e.g.*, a date field in a database will display as a formatted date, but its actual value is

E. *Environment-Dependence and Obsolescence*

Unlike data stored in a paper medium, electronically stored data can be very much dependent on its electronic environment- which programs open and operate upon the file.[48] Often, electronically stored data can be incomprehensible absent the proper program.[49] "If the raw data (without the underlying structure) in a database is produced, it will appear as merely a long list of undefined numbers. To make sense of the data, a viewer needs the context, including labels, columns, report formats, and similar information."[50]

Additionally, the scope of what technology can do is expanding not only quickly, but at a relatively steady rate.[51] If one were to leave a stack of papers in the corner for 20 years, they might be a bit brittle and yellowed, but still would be perfectly usable. However, if one were to leave a stack of floppy disks in the corner in 1994, the ability to use the information contained on those floppy disks is much more questionable. Change in computer systems and methods of storing information are inevitable, and it is only a matter of time before any given method of storing information becomes obsolete. In fact, "it is not unusual for an organization to undergo several migrations of data to different platforms within a few years."[52] Often too, metadata can be lost when converting files from one format to another.[53]

F. *Dispersion and Searchability*

While paper documents tend to be consolidated in a central location, electronically stored documents are often found in a variety of networked drives, files, and directories, and the physical location that data storage devices are located in can vary greatly.[54] Additionally, many computer systems have auto saving and recovery functions that can produce

---

typically a long integer).

Id.

48. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 4.

49. *Id.*

50. *Id.*

51. *See e.g., Moore's Law*, WIKIPEDIA, http://en.wikipedia.org/wiki/Moore's_law (last visited Dec. 10, 2014).

52. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 4.

53. *See id.*

54. Off-site storage, dubbed "cloud computing," can pose a particularly tough challenge, as it is technically possible for information to be stored at a facility in another country, potentially one that has much stricter privacy protections, posing an issue for American style discovery. *See, e.g.*, Danny Hakim, *Europe Aims to Regulate the Cloud*, N.Y. TIMES (Oct. 6, 2013), http://www.nytimes.com/2013/10/07/business/international/europe-aims-to-regulate-the-cloud.html?_r=0.

multiple, sometimes only vaguely different copies of documents in diverse networked locations, where due to drive organization, ownership can be difficult to establish.[55]

V. FEDERAL RULES AND ELECTRONICALLY STORED INFORMATION: THE 2006 AMENDMENTS

Prior to the 2006 amendments to the Federal Rules of Civil Procedure, there was a great deal of uncertainty as to how ESI should be treated under the rules.[56] While some documents were substantially the same in electronic form as they are in paper form (a Word document for example), other ESI is distinctly different, not just in content, but in kind. With information increasingly being stored in dynamic databases that are constantly changing and updating, it became increasingly difficult to shoehorn these conceptually new methods of storing information into the traditional meaning of "document."[57] "As originally adopted, Rule 34 focused on discovery of "documents" and "things.""[58] Expanded and clarified in 2006, the intent and effect of the addition was to explicitly include electronically stored data, and ensure that its discovery was regarded on an equal basis with that of traditional documents.[59] The 2006 amendments were in many ways kept intentionally broad, due to the rapid pace at which technology and information management systems are evolving.[60]

Generally, under Rule 34, documents and ESI must be produced "as they are kept in the usual course of business."[61] Unless a specific form of production is stipulated in the discovery request, this information must be produced in the "form or forms in which it is ordinarily maintained or in a reasonably usable form or forms[,]"[62] and the producing party must

---

55. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 5. "[W]hile electronically stored information may be stored on a single location, such as a local hard drive, it is likely that such documents may also be found on high-capacity, undifferentiated backup tapes, or on network servers— not under the custodianship of an individual who may have 'created' the document." *Id.*

56. *See, e.g.*, Bennett B. Borden et al., *Four Years Later: How the 2006 Amendments to the Federal Rules Have Reshaped the E-Discovery Landscape and are Revitalizing the Civil Justice System*, 17 RICH. J.L. & TECH. 10 (2011).

57. *See* FED R. CIV. P. 34 advisory committee's note, amend. 2006.

58. *Id.*

59. *See id.*

60. *See id.* ("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.").

61. FED. R. CIV. P. 34(b)(2)(E)(i).

62. FED. R. CIV. P. 34(b)(2)(E)(ii); *see also* FED. R. CIV. P. 34(a)(1)(A).

state in their response "the form or forms it intends to use."[63] The 2006 edits to include ESI in Rule 34 also impacted other Rules relating to discovery generally. Most notable among these are Rule 26— specifically Rule 26(b)(2)(B), which outlines "Specific Limitations on Electronically Stored Information."[64]

The Rules Advisory Committee was concerned that while ESI is frequently easier to uncover and produce than traditional documents, it is possible for a system to retain specific data in a manner than to make it "access[ible] only with substantial burden and cost."[65] Nonetheless, the Rules provide that the requesting party can make a motion to request this information if there is good cause and the considerations for limitation of discovery in Rule 26(b)(2)(C) are not implicated.[66]

VI. FEDERAL RULES TREATMENT OF METADATA

The word metadata does not appear in the Federal Rules of Civil Procedure. However, metadata is easily encompassed within the scope of ESI as defined in Rule 34.[67]The Federal Rules of Civil Procedure provide a discovery structure that incorporates ESI throughout the process, and would allow a reasonably informed requesting party to specify the form of information for production, including metadata. Litigants also have an opportunity to make specific requests relatively early in the process through the Rule 26(f)(3) conference.[68] If the parties fail to reach agreement on the scope and form of the discovery (whether or what ESI metadata should be included), the court will be notified through the Rule

63. FED. R. CIV. P. 34(b)(2)(D).

64. FED. R. CIV. P. 26(b)(2)(B).

65. FED. R. CIV. P. 26 advisory committee's note, amend. 2006, subdiv. (b)(2)("In a particular case, these burdens and costs may make the information on such sources not reasonably accessible.").

66. FED. R. CIV. P. 26(b)(2)(B–C). The Rules Advisory Committee notes list several considerations for compelling production of ESI not reasonably accessible. "Appropriate considerations may include: (1) the specificity of the discovery request; (2) the quantity of information available from other and more easily accessed sources; (3) the failure to produce relevant information that seems likely to have existed but is no longer available on more easily accessed sources; (4) the likelihood of finding relevant, responsive information that cannot be obtained from other more easily accessible sources; (5) predictions as to the importance and usefulness of the further information; (6) the importance of the issues at stake in the litigation; and (7) the parties' resources. *Id.*

67. *See* FED. R. CIV. P. 34.

68. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 65 ("Specifically, Rule 26(f)(3) mandates that the parties meet, confer, and develop a proposed discovery plan that includes the parties' views and proposals regarding, among other topics, "any issues relating to disclosure or discovery of electronically stored information, including the form or forms in which it should be produced.").

26(f) report and the issue can be addressed in the Rule 16(b) conference. Notwithstanding, parties need to agree to a form of production early in the discovery process, as under FRCP 34(b)(2)(E)(iii), "a party need not produce the same electronically stored information in more than one form."[69]

When read expansively, Rule 34(b) can be interpreted to require the production of metadata, if the metadata is integral to the way that the ESI is "kept in the usual course of business," or to the "form . . . in which it is ordinarily maintained[70] or in a reasonably usable form."[71] However, case law has not always reflected or endorsed this interpretation.[72] One practical obstacle to the production of metadata is the difficulty of Bates stamping documents. When metadata can be produced as a printout or Portable Document Format (PDF), it is relatively easy to Bates stamp the document. When metadata involves functions that do not easily translate to printout or PDF (for example, the formula or information in cells of an Excel spreadsheet), the task of producing this ESI absent potentially privileged information can become much more difficult and complicated. Luckily, one solution that can work in some cases is the use of a "hash" mark.[73] A hash is in essence an algorithm that is applied to a document that uses specific values in the document to produce a number that will stay the same so long as the document is not altered.[74] Thus, the authenticity of the document is ensured as the values cannot have been tampered with or changed without altering the hash value.[75]

As e-discovery practice has generally evolved, the custom has simply become to request metadata specifically in the discovery conference. Usually, if reasonable and not at a great expense, the metadata information requested is produced without difficulty. However, confusion still exists among unsophisticated litigants as to the scope of what metadata is appropriate to disclose, and what is commonly thought of as not useful.

---

69. FED. R. CIV. P. 34(b)(2)(E)(iii).

70. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 63 ("The form in which electronically stored information is 'ordinarily maintained' is not necessarily synonymous with the form in which it was created. There are occasions when business considerations involve the migration or transfer of electronically stored information to other applications or systems.").

71. FED. R. CIV. P. 34(b)(2)(E)(i),(ii).

72. *See, e.g.*, Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 649–50 (D. Kan. 2005).

73. *See, e.g.*, Ralph C. Losey, *Hash: The New Bates Stamp*, 12 J. TECH. L. & POL'Y 1, 2 (2007).

74. *Id.*

75. *Id.*

VII. CURRENT CASE LAW AND SCHOLARSHIP'S TREATMENT OF METADATA

E-discovery generally has been a hot topic within the litigation and discovery community since the mid-1990s.[76] Notwithstanding, the level of scholarly production and the number of cases discussing e-discovery and metadata more specifically has somewhat died down in the past three to four years. This allows us to examine the record and see if any kind of consensus has been reached, either legally or in best practice.

Most influential over the evolution of e-discovery protocols and best practices is perhaps the Sedona Principles. First published in 2004, and revised to a second edition in 2007,[77] the import of this treatise cannot be overstated. In the first edition, the drafters of The Sedona Principles were of the opinion that "there should be a modest legal presumption in most cases that the producing party need not take special efforts to preserve or produce metadata."[78] The second edition of the Sedona Principles, released in 2007, stepped away from this presumption, instead stating "[t]he extent to which metadata should be preserved and produced in a particular case will depend on the needs of the case."[79] This is a more reasonable position to take on the production of metadata. Several cases subsequent to the publication of The Sedona Principles in 2004 have cited the principles presumption against the production of metadata, yet after the revised principles were published in 2007, there was a subtle, but not drastic, shift in case law.[80]

Perhaps the earliest case to address metadata within what we now consider the bounds of e-discovery is *Armstrong v. Exec. Office of the President.*[81] In *Armstrong*, the court struggled with the concept of e-mail and the problems of categorization.[82] The Court found that while the main text of the e-mail when printed out is substantially the same as a comparable typed document on paper, other qualities such as the times sent and received as well as recipients were important, and that their exclusion would be analogous to cutting the header off of a conventional memo.[83]

The *Armstrong* Court also saw fit to establish that e-mail communications did satisfy the Federal Records Act definition of a

76. *See, e.g.*, Marcus, *supra* note 20, at 7.

77. *See* SEDONA PRINCIPLES 1st ed., *supra* note 6; *see also* SEDONA PRINCIPLES 2d ed., *supra* note 7.

78. SEDONA PRINCIPLES 1st ed., *supra* note 6, at 41.

79. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 61.

80. *See infra* Part VII.

81. *See* Armstrong v. Exec. Office of President, 1 F.3d 1274 (D.C. Cir. 1993) (challenging administrative guidelines for electronic document destruction and retention).

82. *See id*. at 1279–80.

83. *Id.* at 1280.

record, and that because there were often meaningful differences between electronic and paper copies, the electronic versions did not lose their status as records, and so had to be preserved.[84]

In *Public Citizen v. Carlin*, the Court allowed that while there were certain benefits to retaining documents in electronic form, these benefits could be overridden by the necessity of working within budgetary and organizational restraints.[85] Ultimately, the Court decided "a record in electronic form lacks sufficient value to warrant preservation once it is transferred intact to a paper recordkeeping system."[86] However, the Court did not view this as a departure from *Armstrong*, but rather a clarification, indicating that so long as reasonable steps were taken to "'preserve[] the[] content, structure, and context' of a record[,]" records created electronically could be archived as a paper copy.[87] This deference to the practical reality of recordkeeping systems has largely continued to be the norm in discovery, despite technological advances in the past fifteen years that make both storage and export of data far cheaper, and review for privilege much less onerous.[88]

Zubulake v. UBS Warburg LLC followed the established line of reasoning, recognizing that as "there are many ways to manage electronic data, litigants are free to choose how this task is accomplished."[89] However, the *Zubulake* Court further elaborated, and indicated that once the duty to preserve attaches, that duty would dictate that the documents are preserved "in the state they existed at that time[.]"[90]

Shortly following *Zublake*, the first edition of the *Sedona Principles* were published in 2004.[91] What the Sedona Principles did was record what the best practices and emerging trends in e-discovery throughout all U.S. jurisdictions were at that point.[92] As a result, many cases that

---

84. *Id.* at 1287.

85. Pub. Citizen v. Carlin, 184 F.3d 900, 910 (D.C. Cir. 1999).

86. *Id.*

87. *See id.* at 910–11 (citing 60 Fed. Reg. at 44,646/3, 44,644/1.).

88. *See, e.g.*, CONCORDANCE, http://www.lexisnexis.com/en-us/litigation/products/concordance.page (last visited Dec. 10, 2014); RELATIVITY, http://kcura.com/relativity/ (last visited Dec. 10, 2014). These are two leading providers of document review software, enabling fully integrated electronic document review and data processing.

89. Zubulake v. UBS Warburg LLC, 220 F.R.D. 212, 218 (S.D.N.Y. 2003).

90. *See id.* This is also in line with the spirit of FED. R. CIV. P. 34(b)(2)(E)(i),(ii).

91. SEDONA PRINCIPLES 1st ed., *supra* note 6.

92. *About Us*, THE SEDONA CONFERENCE, https://thesedonaconference.org/aboutus (last visited Dec. 10, 2014).

> The Sedona Conference (TSC) is a nonprofit, 501(c)(3) research and educational institute dedicated to the advanced study of law and policy in the areas of antitrust law, complex litigation, and intellectual property rights. The mission of TSC is to drive the reasoned and just advancement of law and policy by stimulating ongoing dialogue amongst leaders of the bench and bar to achieve

followed cite the *Sedona Principles*, particularly its presumption against the production of metadata.[93] In 2007, the second edition of the *Sedona Principles* was published, but included no such presumption against the production of metadata.[94] Rather, the second edition noted that in the absence of a specified form of production, "production should be made in the form or forms in which the information is ordinarily maintained or in a reasonably usable form, taking into account the need to produce reasonably accessible metadata . . ."[95] However, even after publication of the second edition of the *Sedona Principles*, courts continued to cite the first edition, particularly its use in *Williams* with regard to the presumption against the production of metadata.[96]

Currently, judges seem to recognize the necessity of metadata for giving context to communications, but are still generally wary that broad requests for metadata will be a waste of parties' time and money.[97] In response to this threat, courts have adopted the approach that parties should focus their requests on specific documents or sets of data, and specify precisely which fields of metadata should be produced.[98] The consensus seems to be that "[t]he safest practice for parties seeking metadata is likely to request ESI in native format to preserve metadata."[99] As the law exists currently, most courts follow the "Default Standard under which the need for metadata must be shown[,]"[100] and "[t]he issue of whether metadata is relevant or should be produced . . . . ordinarily should be addressed by the parties in a Rule 26(f) conference."[101]

> consensus on critical issues. TSC brings together the brightest minds in a dialogue-based, think-tank setting with the goal of creating practical solutions and recommendations of immediate benefit to the bench and bar.

Id.

93. *See, e.g.*, Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 651 (D. Kan. 2005); Kentucky Speedway LLC v. NASCAR, Inc., 2008 WL 7427284 (E.D. Ky. 2006); Wyeth v. Impax Labs., Inc., 248 F.R.D. 169, 171 (D. Del. 2006).

94. *Compare* SEDONA PRINCIPLES 2d ed., *supra* note 7, at 3, *with* SEDONA PRINCIPLES 1st ed., *supra* note 6, at 41.

95. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 60. Additionally, "[t]he 2006 amendment to Rule 34(a) no longer requires production of ESI in its native format that would include metadata." John B. v. Goetz, 879 F. Supp. 2d 787, 879 (M.D. Tenn. 2010).

96. *See, e.g.*, Autotech Tech's., Ltd. v. AutomationDirect.com, Inc., 248 F.R.D. 556, 560 (N.D. Ill. 2008); Pace v. Int'l Mill Serv., 2007 WL 1385385 (N.D. Ind. 2007).

97. *See, e.g.*, Dahl v. Bain Capital Partners LLC, 655 F. Supp. 2d 146, 149–50 (D. Mass. 2009).

98. *Id.* at 150. The Court goes on to state that "[t]his more focused approach will, the court hopes, reduce the parties' costs and work." *Id.*

99. S2 Automation LLC v. Micron Tech., Inc., 2012 WL 3656454, at *89 (D.N.M. 2012).

100. *Goetz*, 879 F. Supp. 2d at 879.

101. *Id.* (citations omitted).

VIII. ANALYSIS

While the current best practice allows for generally adequate electronic discovery and allows parties to request metadata if the request is made with sufficient specificity,[102] it is lacking in several important characteristics.

A. *Producing Party Protections*

Courts have developed a series of protections from potentially abusive discovery for producing parties in litigation. Among these are the abovementioned limits on the time that a request for documents can be made, and the breadth of the request.[103] Additionally, producing parties must review all of their own information for privilege before submitting it as part of a discovery request. Based on this review, a producing party can then state their objections based on a fuller knowledge of what is being sought. This is not a bad thing in the abstract, however, it does allow for potentially vital contextual information to be excluded if discovery requests are made improperly. While a party in federal court is required to produce electronic documents "as they are kept in the usual course of business,"[104] if the requesting party does not specify the form of production, the producing party need only produce the requested information in a "reasonably usable form or forms."[105] Further, once electronically stored information is produced in one form, a party's ability to require re-production in a more usable and data rich form is limited.[106] When discovery files have been produced on paper or in the somewhat limited PDF form, or alternately in the slightly more manipulable Tagged Image File Format (TIFF),[107] requests for the equivalent information in electronic form can be grounds for shifting the cost to the requesting party.[108]

A parallel concern is the specificity of the original discovery request. "The less specific the requesting party's discovery demands, the more appropriate it is to shift the costs of production to that party."[109] Further, if a request is not made in a limited or confined fashion, a court may

---

102. *See, e.g.*, Rowe Entm't, Inc. v. William Morris Agency, Inc., 205 F.R.D. 421, 430 (S.D.N.Y. 2002).

103. *See, e.g.*, *supra* notes 92, 93.

104. FED. R. CIV. P. 34(b)(2)(E)(i).

105. FED. R. CIV. P. 34(b)(2)(E)(ii).

106. FED. R. CIV. P. 34(b)(2)(E)(iii).

107. *Tagged Image File Format*, WIKIPEDIA, http://en.wikipedia.org/wiki/Tagged_Image_File_Format (last visited Dec. 11, 2014).

108. *See* Zubulake v. UBS Warburg LLC, 217 F.R.D. 309, 318 (S.D.N.Y. 2003).

109. Rowe Entm't, Inc. v. William Morris Agency, Inc., 205 F.R.D. 421, 430 (S.D.N.Y. 2002).

consider it to impose an "undue burden or expense on the responding party."[110] Again, this is beneficial when one considers the need to streamline the discovery process and speed up litigation (or settlement as the case may be), but can inhibit an unsophisticated litigant's ability to learn information and construct a full timeline of events.[111]

B. *The Pro Se Problem*

As it stands now, the discovery procedure also requires a certain level of legal sophistication that many *pro se* litigants simply do not possess, along with a degree of computer savvy that *pro se* litigants and lawyers, not just elderly lawyers, generally are not accustomed to.[112] When a layman speaks of an e-mail communication, no differentiation is generally made between the text of the communication and such basic information as the recipients, time sent, time received. He probably assumes that they are one and the same. If a defendant retains counsel proficient in e-discovery, it is quite possible for them to thwart even a basic e-discovery request of an unsophisticated litigant, particularly one whose complaint is based upon the timing of communications and the possession of knowledge.

A rule requiring the production of specific fields of e-mail metadata in electronic form, or alternately requiring the production of e-mails in as close to native format as technology allows[113] might mitigate the impact of procedural chicanery. Additionally, if lawyers are habitually presented

---

110. *Zubulake*, 217 F.R.D. at 318. The *Zubulake* Court listed seven factors to be considered when conducting a cost shifting analysis:

> 1. the extent to which the request is specifically tailored to discover relevant information;
> 2. the availability of such information from other sources;
> 3. the total cost of production, compared to the amount in controversy;
> 4. the total cost of production, compared to the resources available to each party;
> 5. the relative ability of each party to control costs and its incentive to do so;
> 6. the importance of the issues at stake in the litigation; and
> 7. the relative benefits to the parties of obtaining the information.

Id. at 316.

111. *See* Hickman v. Taylor, 329 U.S. 495, 501 (1947). The modern U.S. discovery process is intended "for the parties to obtain the fullest possible knowledge of the issues and facts before trial." *Id*.

112. *E.g*., Martha Neil, *Asked to Demonstrate Computer Skills, 0 of 9 Law Firms Passed in-House Hiring Test*, ABA J. (May 23, 2013, 12:40 PM), http://www.abajournal.com/news/article/in-house_lawyer_tests_biglaw_firms_for_computer_skills_before_hiring_them/.

113. The way that data is stored in some e-mail systems makes it virtually impossible to produce a 'true native' version of a specific e-mail. Often one will have to settle for an exported file format type, such as .eml, .pst, or .msg that for all intensive purposes is the functional equivalent of a native file.

with metadata data sets in electronic format that can potentially be mined for information, lawyers who are not specialists in the e-discovery field will be incentivized to become more familiar with e-discovery techniques and procedures.

C. *Neglected Value of Relational Data*

Historically, discovery has been paper-based, and only as technology has shifted have the rules changed and lawyers have opened up to new ways of reviewing documents and data. One function which must have been nearly impossible to accomplish in the pre-ESI era is relational data visualization. With the aid of software, an individual who has access to simple e-mail metadata can use discrete data fields such as the "From," "To," subject, and time sent to construct a graphical representation of communication patterns between individuals.[114] Even with a digital image based file like a PDF or TIFF this sort of analysis would be extremely difficult but not impossible. One can easily imagine how important such a relational "map" could be in litigation involving the timeline of relationships or the possession of "inside" or "proprietary" knowledge. Courts have established a propensity for suspicion of broad requests for data,[115] and a hesitancy for allowing multiple rounds of discovery for the same or similar information.[116]

D. *Informational Integrity*

The metadata information contained within an e-mail communication does not reveal information outside the scope of request or inquiry when e-mails are requested generally. While the type of information classified as metadata is surely different in kind than that of the textual body, it is not altogether unrelated or distant enough to be considered as a basis for exclusion. For example, while the metadata information contained in a popup cell of a Microsoft Excel spreadsheet may have to be screened for privilege,[117] it is difficult to imagine that metadata for an otherwise

---

114. *See Immersion: A People-Centric View of Your Email Life*, IMMERSION, https://immersion.media.mit.edu/ (last visited Dec. 11, 2014). Even a relatively streamlined and user-friendly consumer oriented program can demonstrate the value of relational metadata information. One can easily see the potential here for "Big Data" style relational and key word data mining in discovery.

115. Dahl v. Bain Capital Partners LLC, 655 F. Supp. 2d 146, 149–50 (D. Mass. 2009).

116. *See, e.g.*, Rowe Entm't, Inc. v. William Morris Agency, Inc., 205 F.R.D. 42 1, 430 (S.D.N.Y. 2002).

117. Metadata contained in a Microsoft Excel spreadsheet popup is not different in kind from textual information visible in a cell when a spreadsheet is printed out. In contrast, once the text of an e-mail is screened for privilege, the likelihood that any metadata accompanying that text is privileged is probably very low.

unprivileged e-mail text would be. Additionally, it is difficult to imagine a circumstance where such essential metadata information as sender, receiver(s), and times sent and received would not be essential to a full and complete discovery.

While the defense bar might point to the data visualization tools mentioned in Part VIII.C, as an indication that mandatory disclosure of specific e-mail metadata fields alters the scope of the discovery request, it is not the data that is wholly different than its antecedents (for example, a header on a memorandum) but that we now have tools that enable one to process that data. The metadata in an e-mail simply gives context to the textual content of the e-mail, and thus provides a much fuller picture of the meaning of the communication.[118] No reason exists from the standpoint of informational integrity to exclude such basic and helpful data as the "From," "To," "CC (and BCC)," "Timestamp," "Subject," and the "Internet Message-ID."[119]

Even the way that metadata for a particular e-mail system is formatted can provide valuable evidence to prove the veracity of an alleged communication at a specific time. This issue was addressed at length in *Ceglia v. Zuckerberg*.[120] Even when all e-mails in controversy are produced or reproduced in a supposedly uniform manner, information fields like e-mail headers are "'automatically generated when an e-mail is created, not typed by the user,' [and] the inconsistent formatting indicates [alteration of the original e-mail metadata or content]."[121] In *Ceglia*, the defendant's forensic expert noted specifically both that internal inconsistencies would not be present if the e-mail files in question "were actually copied-and-pasted from an authentic source," and the manner of date abbreviation automatically generated by the e-mail system would not have varied.[122]

E. *Technological Limits*

The boundary of what technology can do is continuously being pushed outward, and it is with this inexorable fact in mind that rule-making and advisory committees should plan for the future. While there certainly are legacy systems currently in place, they are only likely to remain entrenched in the near term. Already long accepted by consumers at

118. This is in keeping with the spirit of the Federal Rules of Civil Procedure. *See* Ward et al., *supra* note 2, at 152–53.

119. *See Message-ID*, WIKIPEDIA, http://en.wikipedia.org/wiki/Message-ID (last visited Dec. 11, 2014).

120. 2013 WL 1208558, at *172–76 (W.D.N.Y. Mar. 26, 2013).

121. *Id.* at *172–73 (citations omitted).

122. *Id.* at *173 (citations omitted).

large,[123] cloud based storage is increasingly gaining acceptance in business contexts.[124] Because these new technologies do not rely on the magnetic backup tapes of old, automatically requiring production of defined fields of metadata is no more difficult than producing the e-mail messages themselves.

Currently, best practices in litigation where extensive discovery of ESI is necessary dictates that the requested documents are provided in native format or separated into a text file and a "Load File."[125] A load file is used to import the data into a document review management program.[126] A document review management program can process metadata information just as easily as the body text of a document. The key is to have the load file provided in such a manner that the data remains in native format, or the export file type represents as close as possible the original content and metadata of the communication under review. Even small changes in metadata format can be indicative of inappropriate alteration by the producing party,[127] and it can be difficult to tell if slight inconsistencies in metadata format are a result of tampering or inadvertent alteration once the information is simply cut and pasted or otherwise copied manually to another document type for printing and review.[128]

Far from increasing the costs associated with production, automatically requiring production of specific fields of e-mail metadata will bring down discovery costs if it has a significant effect on costs at all. Since best practice dictates that ESI should be requested and produced in native format if possible,[129] it will generally be unnecessary for producing parties to convert the information into any other form. Aligning requirements with what is already best practice simply serves to further the efficient and fair administration of justice. If production of fields of simple metadata is required, it may also serve to expedite the

123. Sean Ludwig, *Gmail Finally Blows Past Hotmail to Become the World's Largest Email Service*, VENTURE BEAT (June 28, 2012, 5:05 PM), http://venturebeat.com/2012/06/28/gmail-hotmail-yahoo-email-users/. Gmail is an Internet based e-mail program that relies on cloud storage for users e-mails.

124. *Cloud Infographic: Worldwide Big Data Ecosystem*, CLOUDTWEAKS (Sept. 4, 2013, 6:59 AM), http://cloudtweaks.com/2013/09/cloud-infographic-worldwide-big-data-ecosystem/; Michael Singer, *Dell's Business Model Shifts to the Cloud in Pact with Dropbox*, READWRITE (Dec. 17, 2013), http://readwrite.com/2013/12/17/dell-dropbox-pact-perks-up-business-argument-for-online-storage (the shift in strategy by Dell is particularly important here because the company is a leading technology provider to businesses).

125. *Load File*, WIKIPEDIA, http://en.wikipedia.org/wiki/Load_file (last visited Dec. 11, 2014).

126. *Id.*

127. *See* Ceglia v. Zuckerberg, 2013 WL 1208558, at *172–76 (W.D.N.Y. Mar. 26, 2013).

128. *See id.* at 166.

129. *See, e.g.*, John B. v. Goetz, 879 F. Supp. 2d 787, 879 (M.D. Tenn. 2010).

litigation process, as producing parties will have less of an incentive to challenge or otherwise obstruct production. Ultimately, this may serve to lower costs for producing parties, as there will be less of a temptation to strip metadata from individual files, which can be an expensive and time consuming process.

The Federal Rules preference for widely defined ESI production definitions and guidelines is not incompatible with standardizing production for a specific and largely mature document type. Overly general production requests can even be detrimental to the flow of litigation by resulting in unnecessary complication and even cost shifting for document processing.[130] While postal services have been a part of human society at least since Roman times,[131] e-mail has largely taken the place of conventional paper mail for most day-to-day business applications. Given this, e-mail, as a method for sending information, is unlikely to drastically shift or disappear within the foreseeable future. Once standards for metadata production are established, software providers will have an incentive to further streamline preservation tools to allow for later production (to the extent that there may have been any technical difficulties in producing this information at all). Hence, moderate procedural standardization is unlikely to be detrimental to the cost or procedural complexity of litigation.

Some of the issues that detractors bring up regarding the difficulty of working with information in native format do not apply to e-mail communications.[132] E-mails are often relatively easy to produce in either native format or in a searchable image format accompanied by a load file. This is not the case "for certain types of electronically stored information such as spreadsheets, [audio or video files,] and dynamic databases."[133] Not only is an e-mail two sided,[134] but in addition to the ability to create a unique 'hash' value for the document, each e-mail has its own unique message-ID.[135]

The issue of software compatibility for native format production is quickly fading into the past as enterprising technologists create solutions

---

130. John Hopkins, *Beware of Too General Production Agreements in E-Discovery*, SEARCY LAW (Aug. 9, 2013), http://www.searcylaw.com/beware-of-too-general-production-agreements-in-e-discovery/.

131. Joan Brown Wettingfeld, *Sophisticated Postal Service Existed in Ancient Rome*, TIMES LEDGER (July 20, 2012), http://www.timesledger.com/stories/2012/29/wettingfeld_all_2012_07_19_q.html.

132. *See, e.g.*, Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 640 (D. Kan. 2005); SEDONA PRINCIPLES 1st ed., *supra* note 6, at 4–5; SEDONA PRINCIPLES 2d ed., *supra* note 7, at 62.

133. SEDONA PRINCIPLES 2d ed., *supra* note 7, at 62.

134. This feature potentially allows for spot-checking specific communications where the authenticity or integrity of the document is in doubt.

135. *See supra* text accompanying note 119. Message-ID can be valuable when the authenticity of a document is in doubt, as no two are ever alike.

to the problem of compatibility in document review, lured by the money available in a high stakes litigation context. Even if information has to be recovered from legacy magnetic backup tapes, producing the metadata along with the body text already required to be produced should add little technical complication. Even then, this document recovery process will only become easier as time progresses and legacy systems are phased out.

F. *Consistency with the Federal Rules of Civil Procedure*

The Advisory Committee for the 2006 Amendment to the Federal Rules of Civil Procedure was very much aware of the lightning pace at which technology evolves, particularly in comparison to procedural rules.[136] With this in mind, the committee drafted Rule 34(a)(1) "to be broad enough to cover all current types of computer-based information, and flexible enough to encompass future changes and developments."[137] Even then however, almost eight years ago, the advisory committee makes note that "[a] common example often sought in discovery is electronic communications, such as email."[138] The Advisory Committee's seeming skepticism of their mastery over technology and wise caution is again demonstrated when the committee specifically notes that standardizing the required form of ESI production "could prove impossible, and even if possible could increase the cost and burdens of producing and using the information."[139] Nonetheless, Rule 34(b) requires ESI production "in a form or forms in which it is ordinarily maintained or in a reasonably usable form or forms[.]"[140] Simple metadata is easily viewable for e-mail as it is ordinarily maintained in the course of business, and arguably an e-mail divorced from identifying information such as sender and subject is not reasonably usable.[141]

While this hesitance to impose costs on parties in litigation is admirable, the fact of the matter is that responding parties can sometimes easily spend more money in stripping metadata information from unprivileged documents than it would cost to simply produce that same document in its unstripped and unredacted form.[142] An amendment

136. *See* FED. R. CIV. P. 34 advisory committee's note, amend. 2006, subdiv. (a).

137. *Id.*

138. *Id.*

139. *Id*. subdiv. (b).

140. FED. R. CIV. P. 34(b)(2)(E)(ii).

141. Compare to a handwritten letter where the sender and the recipient have been redacted. The evidentiary quality of that letter has been severely impaired.

142. *See, e.g.*, John Hopkins, *Electronically Stored Information (ESI)- Search and Identify*, SEARCY LAW (Dec. 6, 2012), http://www.searcylaw.com/electronically-stored-information-esi-search-and-identify/ ("Producing parties want to limit their time and expense to the smallest possible number[]" yet "[t]hey want to produce only the bare minimum required by the law."); Mike Breen, *Nothing to Hide: Why Metadata Should be Presumed Relevant*, 56 KAN. L. REV. 439,

requiring specific e-mail metadata disclosure could very likely counteract this perverse incentive,[143] and ultimately lead to lower overall costs, despite protests from the defense bar.

The proposition that e-mail metadata production is standardized and automatically required is not a drastic departure from the larger trends in ESI discovery. The shift in opinion is evident in the changes between the first and second editions of The Sedona Principles.[144] Slowly but steadily the law moves forward, as judges become more knowledgeable and courts become more sympathetic to production requests for metadata. Few if any judges are likely to deny a request for simple e-mail metadata, as the value of such information to understanding the context of the communication is fairly self-evident. Why make parties specifically ask for it? In sum, the time has come to amend the Federal Rules of Civil Procedure to standardize some aspects of the law surrounding ESI discovery, and require more expansive metadata disclosure.

IX. CONCLUSION

For the most part, the law governing e-discovery, metadata, and e-mail, has developed as it should. Courts and legal scholars take incremental steps toward aligning the law with our digital reality. Most notable here is that this alignment process has been done with a certain level of humility. As the Federal Rules Advisory Committee notes, the technological sands are constantly shifting at a much faster rate than the law can adapt.[145] With this in mind, the framers built in a degree of flexibility as to what the rules require, and purposefully left definitions broad to allow for future developments.[146] Within this space, organizations like the Sedona Conference promulgated guidance on best practices.[147] In turn, state and federal courts have adopted these suggestions[148] with courts going on to promulgate their own guidelines for electronic discovery,[149] some of which have in turn become influential

461–62 (2008) (citations omitted); *see generally* Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640 (D. Kan. 2005).

143. *See, e.g.*, Breen, *supra* note 142, at 465.

144. Aguilar v. Immigration & Customs Enforcement Div., 255 F.R.D. 350, 355–56 (S.D.N.Y. 2008).

145. FED. R. CIV. P. 34 advisory committee's note, amend. 2006, subdiv. (a).

146. *Id.*

147. *See generally* SEDONA PRINCIPLES 1st ed., *supra* note 6; *see also* SEDONA PRINCIPLES 2d ed., *supra* note 7.

148. *See, e.g.*, Williams v. Sprint/United Mgmt. Co., 230 F.R.D. 640, 651–52 (D. Kan. 2005).

149. *See, e.g.*, U.S. District Court of Maryland, *Suggested Protocol for Discovery of Electronically Stored Information*, at 25–26, *available at* http://www.mdd.uscourts.gov/news/

over the field in their own right.

Despite all of the advancement made in the field, and the general prudence of keeping definitions and requirements broad to encompass diverse existing systems and future developments, it is time for a refinement. That is, defined "essential" fields of e-mail metadata should be automatically included in a document request that includes e-mail communication, regardless of the format or system. The burden of requesting additional metadata information would still be on the requesting party.[150] Similarly, if the producing party feels that production of this simple metadata information is too burdensome, they can object in their response under Rule 34(b) as it now exists to challenge the necessity of production,[151] or if that fails, request cost shifting as outlined in *Zubulake*.[152]

While some might argue that an additional requirement as outlined above is contra to the intent of the framers of the 2006 e-discovery requirements, that opinion would belie a lack of depth in understanding. The FRCP simply exemplifies a wariness for locking the legal and business community into too strict of a regime where the practical reality is quickly shifting. While it is certainly true that the development of digital technology, both hardware and s

oftware is moving at a comparatively lightning pace, one can safely believe that the concept of e-mail is here to stay. The program with which a user accesses e-mail, and the method by which it is sent from one user to the other may change of course, but essential elements that go to the root of how individuals use the medium will not. These qualities are agnostic to the underlying software and hardware platforms.

Standardizing e-mail metadata production will reduce inequity in legal representation and level the playing field for *pro se* litigants and lawyers who do not concentrate their practice on e-discovery issues. Complexity of disputes during the discovery phase will be reduced, and overall the process will be expedited. All of these advantages serve to further the goal of the Federal Rules of Civil Procedure.[153]

---

news/ESIProtocol.pdf.

150. A producing party will have an incentive to simply provide all metadata fields when some are required in native format. It is unlikely to be worth the cost to go through the trouble of stripping that metadata as it is generally of little evidentiary value in most situations, even when the producing party intends to be obstructive.

151. FED. R. CIV. P. 34(b)(2). Scope of discovery is also limited under FED. R. CIV. P. 26(b)(2)(B).

152. Zubulake v. UBS Warburg LLC, 217 F.R.D. 309, 317–18 (S.D.N.Y. 2003).

153. *See* Ward et al., *supra* note 2, at 153 n.5.