June 2015

The Legal Ethics of Metadata Mining

Andrew M. Perlman

THE LEGAL ETHICS OF METADATA MINING

Andrew M. Perlman[*]

* Professor of Law, Suffolk University Law School. Thanks to Monroe Freedman and Gabriel Teninbaum for their feedback on an earlier draft of this article and to librarian Rick Buckingham for his research assistance.

I. INTRODUCTION

Bar associations have produced a number of legal ethics opinions that address the practice of metadata mining.[1] These opinions examine whether a lawyer is permitted to extract an electronic document's embedded information, such as the document's author history, without first seeking the permission of either a court or the lawyer's adversary.

This essay explains the nature of the problem, reviews the ethics opinions that have addressed it, and contends that the issue is simply a variation of the oft-examined problem of inadvertently disclosed documents. The essay concludes that flat bans on metadata mining are misguided and that metadata mining should be treated in the same manner as inadvertent disclosures more generally. Under this approach, if a state permits lawyers to review inadvertently disclosed privileged documents, the jurisdiction should also permit lawyers to review the metadata contained in electronic documents. In contrast, if a jurisdiction prohibits the review of misdirected privileged documents, the state should ban metadata mining, but only when recipients have reason to believe that the metadata contains protected information.

II. LEGAL ETHICS ISSUES CONCERNING METADATA MINING

Metadata is essentially information that is embedded in—not apparent on the face of—electronic documents, such as word processing files or spreadsheets.[2] Metadata can contain a wide range of information, including the name of the person who originally authored the document, the date the document was created, the dates it was edited, the names of other people who edited it, and even the contents of previous edits.[3]

The existence of metadata, and the potential to extract it surreptitiously, has prompted bar associations to address several legal

1. ABA Comm. on Ethics and Prof'l Responsibility, Formal Op. 442 (2006); Ala. St. Bar Office of Gen. Counsel, Formal Op. 02 (2007); St. Bar of Ariz. Ethics Comm., Ethics Op. 03 (2007).; Colo. Bar Ass'n Ethics Comm., Ethics Op. 119 (2007); D.C. Bar Legal Ethics Comm., Op. 341 (2007); Fla. Bar Ethics Dep't, Ethics Op. 02 (2006); Me. Bd. of Overseers of the Bar, Prof'l Ethics Comm. Op. 196 (2007).; Md. St. Bar Ass'n, Comm. on Ethics, Ethics Docket 09 (2007); N.H. Bar Ass'n, Ethics Comm. Op. 4 (2008-2009); N.Y. St. Bar Ass'n Comm. on Prof'l Ethics, Op. 749 (2001); NYCLA Comm. on Prof'l Ethics, Op. 738 (2008); Pa. Bar Ass'n, Comm. on Legal Ethics and Prof'l Responsibility, Formal Op. 100 (2009); Vt. State Bar Ass'n Ethics Op. 01 (2009); W. Va. Bar Ass'n, Lawyer Disciplinary Bd., L.E.O. 01 (2009).

2. Elizabeth W. King, *The Ethics of Mining for Metadata Outside of Formal Discovery*, 113 PENN ST. L. REV. 801, 805-807 (2009) (offering an extensive definition).

3. *Id.*

ethics questions. The most controversial question, which can arise in both the transactional and litigation contexts, is whether the recipient of electronic documents can look at the metadata without first getting permission to do so.[4]

A. The Transactional Context

Imagine that a corporation produces a spreadsheet during the negotiation of a business deal, and the receiving lawyer wants to know how the corporation generated its data. Specifically, the lawyer wants to look at various formulas embedded in the spreadsheet to ensure that the corporation has produced its figures accurately. Moreover, the lawyer wants to know who created the spreadsheet in order to identify the person who was most likely responsible for collecting the data. Does the lawyer have to ask the corporation for permission to look at this information?

Alternatively, imagine that a lawyer is negotiating a contract through the exchange of an electronic document created in WordPerfect. During the negotiations, the client instructs the attorney to make an important concession in one of the contract's provisions. The lawyer makes the change in the electronic version of the document, but before emailing the proposed change to opposing counsel, the client decides not to offer the concession. The lawyer edits the document back to its original state and sends it to the other party's attorney. If the sending lawyer left the "undo" command enabled and saved the document with past edits intact, the receiving attorney could click the "undo" button to see the past changes and discover that the sending attorney's client was considering an important concession. Can the receiving lawyer look for this hidden information?

B. Litigation Examples

Consider a case in which a large volume of electronic documents are produced in response to a discovery request. The parties did not agree in advance whether metadata was discoverable, and the recipient wants to review the metadata of the produced documents to determine who authored them and on what dates.[5] In some cases, the discovery

4. Some ethics opinions also address the ethical obligation of the sending attorney, but the recipient's obligations are the subject of greater disagreement. *See* King, *supra* note 2, at 817-19 (making a similar observation).

5. The Federal Rules of Civil Procedure address the discoverability of electronic documents generally and imply that lawyers should be permitted to review metadata. *See* King, *supra* note 2,

documents were produced in their native format (e.g., Microsoft Word's ".doc" format), so the information is easily discoverable in the metadata.

Other electronic documents, however, were converted to Adobe's ".pdf" format before production. The sender digitally redacted (and asserted a privilege regarding) some of the text in those documents through the use of what is effectively a digital black magic marker that covers the visible text. The receiving lawyer, however, knows how to remove the digital "black out" and examine the text that lies underneath.[6] Is it ethically permissible for the lawyer to do so?

III. THE CURRENT STATE OF THE LAW

To date, fourteen bar associations have examined whether lawyers should be permitted to engage in the metadata mining described in the above examples.[7] The opinions fall into three categories: Some say metadata mining should always be impermissible (seven opinions); some say it should always be permissible (three opinions); and some say that it should usually be permissible, but with a few limitations (four opinions).

A. Opinions Prohibiting Metadata Mining

Seven bar associations have concluded that it is generally unethical to review a document's metadata unless the sending party has expressly permitted it. The New York State Bar Association's Committee on Professional Ethics was the first to adopt this view,[8] concluding that a "lawyer may not make use of computer software applications to surreptitiously 'get behind' visible documents . . . ."[9] Since then, bar associations in Alabama, Arizona, Florida, Maine, New Hampshire, and New York City have reached a similar conclusion.[10]

---

at 810-15. The Federal Rules, however, do not resolve all of the ethics issues, *id*., and most states do not yet address the problem in their rules of civil procedure. Although some ethics opinions suggest that their scope is limited to the non-litigation context, the ethical issues (such as the redacted PDF document) would appear to be the same in the litigation context as well.

6. Depending on how the data was concealed, revealing it can be as simple as cutting and pasting the blacked out text into a new document.

7. *See supra* note 1.

8. N.Y. St. Bar Ass'n Comm. on Prof'l Ethics, Op. 749 at *3 (Dec. 14, 2001).

9. *Id.* at *4.

10. Ala. St. Bar Office of Gen. Counsel, Formal Op. 02 (2007) (limiting its conclusion to the non-litigation context); St. Bar of Ariz. Ethics Comm., Ethics Op. 03 (2007); Fla. Bar Ethics Dep't, Ethics Op. 02 (2006); Me. Bd. of Overseers of the Bar, Prof'l Ethics Comm. Op. 196 (2007); N.H. Bar Ass'n, Ethics Comm. Op. 4 (2008-2009) (excluding from its analysis "electronic materials subject to discovery"); NYCLA Comm. on Prof'l Ethics Op. 738 (2008) (same).

These opinions all rely on similar rationales. Primarily, they argue that metadata mining would damage the attorney-client relationship because clients would be less willing to communicate with counsel out of fear that their communications could not be adequately safeguarded.[11] The opinions also assume that, when a lawyer intentionally transmits an electronic document, "counsel plainly does not intend the [opposing] lawyer to receive the 'hidden' material or information."[12] Accordingly, the recipient of an electronic document should assume that the metadata is not subject to review.

B. Opinions Permitting Metadata Mining

Three bar associations—the American Bar Association, the Maryland State Bar Association, and the Vermont State Bar Association—have rejected these arguments.[13] They emphasize that most metadata does not contain protected information and is thus unlikely to affect the attorney-client relationship.[14] They also note that the sending attorney can take measures to extract metadata, so if an attorney distributes an electronic document with the metadata intact, it is reasonable to conclude that the sending attorney intended to include the metadata and make it available for review.[15]

Relying on this reasoning, the American Bar Association has concluded that metadata mining should be handled in the same way as inadvertent disclosures more generally.[16] The ABA's opinion notes that Rule 4.4(b) gives a lawyer the discretion to review misdirected documents, so a lawyer should have the same discretion to review a document's metadata.[17] The Maryland and Vermont bar associations have reached the same conclusion using similar reasoning.[18]

---

11. *See* King, *supra* note 2, at 819.

12. N.Y. St. Bar Ass'n Comm. Prof'l Ethics, Op. 749 at *3 (2001).

13. ABA Comm. on Ethics and Prof'l Responsibility, Formal Op. 442 at 1 (2006); Md. St. Bar Ass'n, Comm. on Ethics, Ethics Docket 09 (2007); Vt. State Bar Ass'n Ethics Op. 01 (2009).

14. ABA Comm. on Ethics and Prof'l Responsibility, Formal Op. 442 at 3 (2006).

15. *Id.* at 4-5.

16. *Id.* at 4 (analogizing to newly adopted Rule 4.4(b) of the Model Rules of Professional Conduct).

17. *Id.*

18. Md. St. Bar Ass'n, Comm. on Ethics, Ethics Docket 09 (2007); Vt. State Bar Ass'n Ethics Op. 01 (2009).

C. Opinions Permitting Metadata Mining in Certain Circumstances

Four bar associations have concluded that metadata mining should be permissible, at least in some circumstances. For example, the Bar of the District of Columbia has concluded that "[a] receiving lawyer is prohibited from reviewing metadata sent by an adversary only where he has actual knowledge that the metadata was inadvertently sent."[19] Because lawyers usually will not have such knowledge prior to reviewing the metadata, the District of Columbia opinion will typically permit an initial review of a document's metadata. Similarly, Colorado permits metadata mining unless the receiving attorney is notified by the sender prior to the recipient's review of the metadata that the metadata contains confidential information.[20]

The West Virginia Bar is somewhat more restrictive but falls short of a flat ban. It explains that, "if a lawyer has received electronic documents and has actual knowledge that metadata was inadvertently sent, the lawyer should not review the metadata before consulting with the sending lawyer to determine whether the metadata includes work product or confidences."[21] If, however, the recipient is not sure whether the disclosure of metadata was inadvertent, the lawyer is encouraged (though apparently not required) to seek clarification from the sender before reviewing the metadata.[22]

Finally, the Pennsylvania Bar Association has determined that the answer should turn on a case-by-case inquiry and vary depending on a number of factors, including whether the lawyer could use the metadata as a matter of substantive law (e.g., whether the privilege would be waived), the potential effect on the client's matter if the lawyer reviews the metadata, and the client's views about metadata mining.[23] In sum, bar associations are divided as to whether metadata mining should be ethically permissible.[24]

---

19. D.C. Bar Legal Ethics Comm., D.C. Op. 341 (2007).

20. Colo. Bar Ass'n Ethics Comm., Ethics Op. 119 (2007).

21. W. Va. Bar Ass'n, Lawyer Disciplinary Bd., L.E.O. 01 (2009) (citing the New York State Bar opinion).

22. *Id.*

23. Pa. Bar Ass'n, Comm. on Legal Ethics and Prof'l Responsibility, Formal Op. 100 (2009). This opinion replaced an older opinion on the issue, which had offered even more ambiguous guidance. The Pennsylvania Bar Association had previously concluded that "each attorney must . . . determine for himself or herself whether to utilize the metadata contained in documents and other electronic files based upon the lawyer's judgment and the particular factual situation." Pa. Bar Ass'n Comm. on Legal Ethics and Prof'l Responsibility, Formal Op. 500 (2007).

24. Metadata Ethics Opinions Around the U.S., American Bar Association, http://www.abanet.org/tech/ltrc/fyidocs/metadatachart.html (last visited Apr. 13, 2010)

IV. THE PROBLEMATIC FLAT BAN APPROACH TO METADATA MINING

Ethics opinions that permit metadata mining are more persuasive than the opinions that have adopted an outright ban. In particular, flat bans are overly broad, incorrectly assume that metadata mining is intended to uncover protected information, make false analogies to clearly unethical behavior, and wrongly imply that metadata mining will increase the cost of legal services.

A. Flat Bans Are Overly Broad

Opinions that endorse a wholesale ban on metadata mining fail to acknowledge the disparate contexts in which the issue can arise. Specifically, they do not distinguish between electronic documents that contain obviously privileged or confidential metadata and electronic documents that contain metadata that is unlikely to receive such protection.

Consider the previously described litigation example, where the recipient of electronic discovery wants to know who originally authored Microsoft Word documents and when those documents were created. There is no reason to think that the metadata contained in these documents is privileged, covered by the work product doctrine, or is otherwise confidential. The same is true for the spreadsheet example in the transactional context, where the recipient wants to look at the formulas that were used or the name of the spreadsheet's original author. In each of these cases, the metadata is not likely to reveal any protected information and is simply going to disclose information that is relevant to the underlying legal matter.

In contrast, there are situations where metadata mining is more troubling, such as when the metadata is obviously subject to a claim of privilege. In the PDF digital redaction scenario, for instance, the sender is clearly communicating to the recipient that the underlying text may be protected, and the only purpose for metadata mining would be to uncover this protected information.[25] Metadata mining, therefore, can occur in many different contexts, only some of which give rise to a

25. Just because the sender believes the document is privileged does not mean that it actually is. Nevertheless, the parties can bring the issue to a court to resolve the privilege question. The question of whether the recipient should be permitted to look at the document is a distinct issue. Andrew M. Perlman, *Untangling Ethics Theory from Attorney Conduct Rules: The Case of Inadvertent Disclosure*, 13 GEO. MASON L. REV. 767, 777-80 (2005).

concern that a lawyer will uncover protected information.[26] By treating all metadata mining the same way, however, flat ban opinions are broader than they need to be given the stated goal of protecting attorney-client communications.

B. Metadata Mining Serves Legitimate Purposes

Many flat ban opinions imply that, even though metadata mining will not always reveal protected information, the practice should nevertheless be prohibited because the only conceivable purpose for metadata mining is to uncover confidential information. Professor Hricik makes this argument while endorsing the flat ban approach, stating that "[e]mbracing the proposition that embedded data is not always—*or at least not presumptively*—included unintentionally is startling . . . [I]t is hard to imagine a scenario where a lawyer would *intentionally* include confidential information in the form of embedded information . . . [A] lawyer at least *should know* that any embedded confidential information was sent inadvertently."[27]

This argument incorrectly and implicitly assumes that metadata mining is typically undertaken in an effort to reveal inadvertently sent confidential information. In reality, most electronic documents do *not* contain confidential metadata, and lawyers may have legitimate reasons for mining that non-confidential metadata. For example, a transactional lawyer who receives electronic documents as part of due diligence may have a legitimate interest in knowing who edited a company's memorandum regarding its financial status or future sales projections. That embedded information is relevant to the transaction and, because it is simply a business document (not created by or for attorneys), there is no reason to conclude that it is confidential or otherwise protected. Lawyers, therefore, will often have a sound strategic reason for looking at non-confidential metadata in both the litigation and non-litigation contexts.

Not only do lawyers have many sound reasons for mining non-confidential metadata, but the likelihood of uncovering confidential metadata is decreasing with time. Lawyers are becoming more aware of metadata's existence and the dangers associated with it, so they are

26. King, *supra* note 2, at 807 n.24 (noting that ordinarily "metadata will have no material evidentiary value") (quoting *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production*, at 5 (Jan. 2004)).

27. David Hricik, *Mining for Embedded Data: Is it Ethical to Take Advantage of Other People's Failures?*, 8 N.C. J.L. & TECH 231, 241 (2007).

taking more precautions to prevent the dissemination of confidential metadata.[28] Indeed, it is now more reasonable to expect attorneys to check for protected metadata before disseminating electronic documents than it would have been just a few years ago, when some bar associations announced prohibitions on metadata mining. Accordingly, the assumption (such as the New York State Bar Association's and Professor Hricik's)[29] that lawyers do not intend to send metadata when they distribute electronic documents is increasingly inaccurate. And the related assumption that the existence of protected metadata is sufficiently common to justify a flat ban on metadata mining is inaccurate as well.

Finally, some commentators argue that the occasional revelation of confidential metadata is so costly that we should ban the practice entirely.[30] This argument, however, proves far too much, because it would justify flat bans on electronic discovery, e-mail, fax machines, and other common practices that carry a risk of inadvertent disclosure. Indeed, these practices and methods of communication are much more likely to lead to the inadvertent disclosure of privileged information than metadata mining.[31]

Arguably, these practices differ from metadata mining because they produce many benefits (e.g., facilitating communication), but metadata mining has many benefits as well. For example, it could reveal an attempt to engage in fraud or other wrongdoing, and it could help a client to win a case or prove what really happened. Given that metadata mining will rarely cause problems and is likely to produce as many benefits as costs, an outright ban is not justifiable.

---

28. *See* N.Y. St. Bar Ass'n Comm. on Prof'l Ethics, Op. 782 (imposing on lawyers an obligation to use reasonable care to prevent the dissemination of confidential metadata). *See also* Jason Krause, *Hidden Agendas: Unlocking Invisible Electronic Codes Can Reveal Deleted Text, Revisions*, 90 A.B.A. J. 26, 27 (Jul. 2004) (quoting a technology specialist who believes that "[l]awyers can't plead ignorance when it comes to [metadata] anymore").

29. N.Y. St. Bar Ass'n Comm. Prof. Ethics, Op. 749 at *3 (Dec. 14, 2001); Hricik, *supra* note 27, at 241.

30. King, *supra* note 2, at 833-34.

31. The large increase in commentary, case law, and ethics opinions regarding inadvertent disclosure since the advent of fax machines offers ample evidence of how technology increases the frequency of inadvertent disclosures. *See* Perlman, *supra* note 25, at 772-73 (tracing the history of inadvertent disclosure law and demonstrating an increase in attention to the issue after lawyers started using fax machines).

C. Metadata Mining is Not Like Snooping in Someone's Briefcase

Even if electronic documents are unlikely to contain protected metadata, flat bans might still be justified on the grounds that the very practice of metadata mining is abhorrent. For example, some opinions have analogized the practice to looking through an opposing counsel's briefcase when she steps out of the room. Other opinions and commentators argue that the practice is simply too "sneaky" or "deceitful."[32]

The implicit assumption here is that an electronic document contains only what is visible on its face and that anything else in the document is private or should be assumed not to exist. This view of electronic documents is neither accurate nor reasonable. The increasingly widespread use of metadata scrubbers, which remove metadata from electronic documents before they are transmitted by e-mail, as well as innumerable continuing legal education programs have sensitized lawyers to metadata's existence. Lawyers know, or at least should know, that when they transmit an electronic document, the document contains more information than what is on the document's face.[33] Metadata mining, which is simply the process of examining the entirety of an electronic document, is thus unlike briefcase snooping, where a lawyer has every reason to believe and expect that her briefcase is free from snooping eyes. And it is not deceitful because lawyers should now be aware that metadata is ultimately an integral part of what an electronic document is.

Moreover, and more importantly, even if a lawyer sends a document without realizing that it contains metadata, that metadata is very unlikely to be privileged, protected by the work product doctrine, or otherwise confidential. As explained above, the vast majority of electronic documents do not contain protected metadata. Thus, if the receiving lawyer engages in metadata mining, she is either not going to get useful information, or if she gets useful information, it is unlikely to be confidential, privileged, or subject to work-product protection.

D. Metadata Mining Will Not Increase the Cost of Legal Services

Another possible objection to metadata mining is that it will add to the cost of legal services.[34] By permitting the practice, the concern is

32. King, *supra* note 2, at 836-37; N. H. Bar Ass'n Ethics Comm., Op. 4 at 6 (2008-2009).
33. *See supra* note 28.
34. King, *supra* note 2, at 830.

that sending lawyers will have to undertake costly efforts to extract metadata, or recipients might feel compelled to engage in an expensive review of metadata to ensure that they are uncovering all relevant information.

This concern is overstated. Today, litigation and business transactions often involve thousands, if not millions, of electronic documents. There is already considerable expense associated with conducting a privilege review of the visible portions of these documents. The additional cost of reviewing the metadata is often negligible.[35] Moreover, to the extent that parties want to avoid the potential disclosure of damaging metadata, they can agree in advance that metadata mining is impermissible.[36]

V. THE CONCEPTUAL IDENTITY OF METADATA MINING AND THE REVIEW OF INADVERTENT DISCLOSURES

The best approach to metadata mining is to analogize it to the review of inadvertently disclosed documents more generally. The two issues are conceptually indistinguishable.[37]

A. The Obviously Privileged Document

Consider the classic "errant fax,"[38] where a lawyer mistakenly sends a privileged communication by fax machine (or e-mail) to her opponent, and the cover page makes clear that the document was

35. Correspondence with electronic discovery service provider is on file with the author.

36. The Federal Rules of Civil Procedure now address some of these concerns. The newly adopted Rule 26 states as follows:

> If information is produced in discovery that is subject to a claim of privilege or of protection as trial-preparation material, the party making the claim may notify any party that received the information of the claim and the basis for it. After being notified, a party must promptly return, sequester, or destroy the specified information and any copies it has and may not use or disclose the information until the claim is resolved. A receiving party may promptly present the information to the court under seal for a determination of the claim. If the receiving party disclosed the information before being notified, it must take reasonable steps to retrieve it. The producing party must preserve the information until the claim is resolved.

Fed. R. Civ. P. 26(b)(5)(B).

37. This is the view of the American Bar Association, the Colorado Bar, the District of Columbia Bar, and the Pennsylvania Bar. *See* ABA Standing Comm. on Ethics and Prof'l Responsibility, Formal Op. 442 (2006); Colo. Bar Ass'n Ethics Comm., Ethics Op. 119 (2007); D.C. Bar Legal Ethics Comm., D.C. Op. 341 (2007); Pa. Bar Ass'n Comm. on Legal Ethics and Prof'l Responsibility, Formal Op. 100 (2009).

38. Monroe Freedman, *The Errant Fax*, LEGAL TIMES, Jan. 23, 1995, at 26.

inadvertently sent. This situation is no different from the PDF digital redaction. In both cases, the receiving lawyer knows that she has received information that the sender intended for her not to see, and in both cases the sending lawyer took protective measures to prevent the recipient from seeing the information. In the fax scenario, the sender included a fax cover sheet with a clear statement that the document contained privileged information. In the PDF example, the lawyer used a digital version of a black magic marker.

Moreover, in both cases, the recipient cannot see the allegedly privileged information without taking additional steps. In the fax example, the recipient has to turn the page and start reading the document. In the PDF example, the recipient has to cut and paste the blacked out text into a new document.[39]

One possible distinction is that the sending lawyer in the misdirected fax example is more culpable for her error because she could have prevented a misdirected fax more easily than she could have extracted privileged metadata. This culpability distinction, however, does not hold up under closer scrutiny. First, the difficulty of identifying and redacting privileged metadata is no greater than preventing the dissemination of privileged documents more generally. Given the sheer volume of documents that are involved in most cases and transactions, a reasonably diligent attorney will often be unable to prevent an inadvertent disclosure. Thus, many misdirected fax scenarios do not involve any greater degree of negligence than the failure to properly redact the metadata from an electronic document. In fact, in many cases, metadata scrubbing will be easier than preventing a misdirected fax.

Second, and more importantly, the degree of the sender's negligence is typically irrelevant to the ethics inquiry. In states that permit a lawyer to examine an adversary's inadvertently disclosed privileged documents, the recipient can look at the document regardless of the precautions that the sender took to prevent the disclosure.[40] Accordingly, if a state permits the review of an obviously privileged document, the state should also permit a lawyer to engage in metadata

39. For an example of how this can be done, see Terry Frieden, *Justice Dept. Defends Editing Charge on Diversity Study*, CNN, Oct. 31, 2003, *available at* http://www.cnn.com/2003/LAW/10/31/justice.diversity. The fully redacted document is available at http://thememoryhole.org/feds/diversityanalysis.pdf.

40. Many courts take into account the sender's culpability when considering whether the inadvertent disclosure waived the attorney client privilege. Perlman, *supra* note 25, at 776-77. But the reasonableness of the sender's culpability does not affect the recipient's ethical duties. *Id.* at 777-80 (explaining the difference between the ethical obligations and privilege waiver).

mining, even when it will reveal privileged information (such as in the PDF example). In contrast, if a state prohibits the review of documents that are likely to be privileged, the state should also prohibit a lawyer from engaging in metadata mining, but only when the metadata is likely to be privileged.[41]

B. The Document that Is Not Protected on Its Face

Now imagine that, instead of receiving an obviously misdirected document, a lawyer receives a document that is not protected on its face, perhaps because it does not contain a cover page or other identifying information. The recipient begins to read the document and only then discovers that it is confidential, work product, or privileged. Under these circumstances, no state would find that the lawyer has engaged in misconduct for having started to read the document. Indeed, even in states that prohibit a lawyer from reviewing inadvertently disclosed privileged information, the prohibition only applies once the recipient realizes that the document is, in fact, misdirected.

Using the same reasoning, if lawyers have no reason to believe that an electronic document contains protected metadata or that the metadata was sent by mistake, lawyers should be permitted to look at it. Again, the metadata is unlikely to be privileged or otherwise protected, so giving lawyers the permission to look at the metadata under these circumstances is unlikely to cause any harm. In sum, lawyers should be permitted to review the metadata in documents, such as the Microsoft Word and spreadsheet examples, at least until they have reason to believe that the metadata contained in those documents is protected.

C. The Inadequacy of Existing Analyses

Many ethics bar opinions cite to the law of inadvertent disclosures for guidance, but they often draw the wrong conclusions. For example, to justify its broad ban on metadata mining, the New York opinion refers to New York law on inadvertent disclosures, which prohibits lawyers from reviewing inadvertently disclosed privileged information.[42]

The problem is that the New York opinion incorrectly assumes that a document's metadata is both inadvertently disclosed and privileged. As explained above, it is becoming less reasonable to assume that the

41. My own view is that states should adopt the latter approach. *Id.* at 813-16. The point, however, is that there is no reason to treat metadata mining any differently than the jurisdiction treats misdirected documents of any kind.

42. N.Y. St. Bar Ass'n Comm. on Prof'l Ethics, Op. 749 at *3 (2001).

sender of an electronic document intends to send only the visible portion of the document. Moreover, even if lawyers are still regularly disseminating electronic documents without any awareness of metadata's existence, most electronic documents are quite unlikely to contain privileged metadata. So although the law of inadvertent disclosures is directly applicable to the metadata mining context, the law on that subject does not justify a flat ban on the practice.

West Virginia is somewhat more consistent with the law governing inadvertent disclosures because it prohibits metadata mining only in certain circumstances: when the lawyer knows that the metadata was inadvertently sent. The opinion, however, does not make clear why that should be the deciding factor as opposed to whether the metadata contains protected information. The law on inadvertent disclosures focuses on the latter, not the former, and for good reason. The primary concern is that the lawyer will uncover privileged information, and that concern does not arise simply because a document was misdirected (or contains metadata). It arises only when the document or the metadata is likely to contain protected information.

VI. CONCLUSION

The ABA's opinion correctly recognizes that the law of inadvertent disclosures provides the best framework for understanding the practice of metadata mining. Although states vary widely in their approach to inadvertent disclosures, they all recognize that a lawyer can examine a document until there is a reason to believe that the document contains protected information. Using the same logic, a lawyer should be permitted to engage in metadata mining of an opponent's documents, such as in the Microsoft Word and spreadsheet examples, as long as the lawyer has no reason to believe that the metadata contains privileged information.

In contrast, if a lawyer knows that the metadata is subject to a claim of privilege, the lawyer should abide by the jurisdiction's approach to inadvertently disclosed privileged documents more generally. If the state prohibits review of such documents, the lawyer should not review metadata that is subject to a claim of privilege, such as in the PDF document example. If the state permits the review of inadvertently disclosed privileged information, however, the state should also permit the lawyer to review metadata that contains privileged information.

Ultimately, ethics opinions concerning metadata mining assume that the practice is somehow different from the more common problem

of misdirected information. The situations are, in fact, analytically identical, and ethics opinions should treat them that way.