

- Coates' Canons: NC Local Government Law Blog - <http://sogweb.sog.unc.edu/blogs/localgovt> -

Is Metadata a Public Record? An Analysis Under Federal FOIA

Posted By [Kara Millonzi](#) On February 18, 2011 @ 1:23 PM In [Open Government](#) | [6 Comments](#)



[1]

UPDATED: June 22, 2011. The court withdrew the opinion analyzed below on June 17, 2011. See *National Day Laborer Organizing Network v. United States Immigration and Customs Enforcement Agency* (S.D.N.Y. June 17, 2011). Although the opinion no longer has any precedential value, Judge Scheindlin's analysis of whether and to what extent metadata is part of a public record is likely to influence future developments in this area of law.

As detailed in previous posts (see [here](#) [2], [here](#) [3], and [here](#) [4]), analyzing the emerging issue of whether, and to what extent, the metadata associated with a public record is part of that public record for purposes of public access is akin to piecing together a puzzle on a case-by-case basis. There is now an additional piece to that puzzle—perhaps an important one.

Recently, a federal district court judge issued an opinion addressing the extent to which metadata is part of an electronic record pursuant to the federal Freedom of Information Act (FOIA). Specifically, in *National Day Laborer Organizing Network v. United States Immigration and Customs Enforcement Agency*, 2011 WL 381625 (S.D.N.Y. Feb. 7, 2011), Judge Shira Scheindlin held that “metadata maintained by an agency as part of an electronic record is presumptively producible under FOIA, unless the agency demonstrates that such metadata is not ‘readily reproducible.’” (emphasis in original).

Facts of *Nat'l Day Laborer Case*

The basic facts of *Nat'l Day Laborer* are as follows. Three advocacy groups submitted FOIA requests to each of four federal agencies pertaining to the operation of a particular federal program. The requested records existed in various formats, including electronic text records, e-mails, spreadsheets, and paper records. Although there is some dispute as to whether, and to what extent, the groups specifically requested the metadata associated with any of the responsive electronic records, at a minimum they asked the government agencies to “(1) produce the responsive records on a CD and, if possible, as an attachment to an email; (2) save each document on the CD as a separate file; (3) provide excel documents in excel file format and not as PDF screen shots; and (4) produce all documents with consecutively numbered bate[s] stamps” The agencies did not initially respond to the FOIA requests, and the advocacy groups filed suit to compel production. After extensive negotiations and some court intervention, the government agencies produced a subset of the requested records. Among the documents produced were five Portable Document Format (PDF) files, in which the government agencies combined a number of separate electronic and scanned paper documents. (There were between 2,000-3,000 pages in the five PDF files.) The PDF files were not electronically searchable. Furthermore, the electronic documents in the PDF files had been

stripped of all their metadata, the attachments associated with e-mails were not linked to the e-mails, and there was no way to identify individual documents without reading through the entire PDF files. The advocacy groups objected to the government agencies' production because (1) the data was produced in unsearchable PDF format and all the paper and electronic documents were indiscriminately merged together into a single file; and (2) the electronic records were stripped of all metadata.

District Court Opinion and Order

As discussed below, *Nat'l Day Laborer* is a district court opinion and order and, as such, it only applies to the case at issue. Several aspects of the court's order are unique to the facts and procedural posture of the case. The court's broader analysis of the issues presented in the case—namely whether and to what extent metadata is part of a government record for purposes of FOIA and, if so, whether it must be specifically requested—may provide a helpful framework for analyzing these issues in other contexts, though. The court's analysis can be broken down into four key parts, as set forth below.

Is metadata part of a record under FOIA?

As discussed in a previous [post](#) ^[2], the term metadata encompasses a broad array of information that relates to an electronic document—ranging from who authored the document and when it was last accessed, to the edit history of the document, to where the document resides within an information network. Some metadata is part of the electronic document, whereas other metadata is external to the document itself. One of the arguments raised by the government agencies in *Nat'l Day Laborer* was that metadata constitutes substantive information that must be explicitly requested and reviewed by an agency for possible exemptions from public disclosure. The court rejected this argument almost out of hand. Acknowledging that no federal court had yet to recognize that metadata is part of a public record as defined in FOIA, the court nonetheless held that "it is well accepted, if not indisputable, that metadata is generally considered to be an integral part of an electronic record." Thus, consistent with the state courts that have addressed this issue to date, the court determined that metadata is part of the underlying record with which it is associated; it does not constitute separate, substantive information.

The court then noted that the relevant FOIA provision requires an agency to "provide [a] record in any form or format requested by the person if the record is readily reproducible by the agency in that form or format," [5 U.S.C. § 552\(a\)\(3\)\(B\)](#) ^[5]. The court interpreted this provision to allow a requesting party to select a form or format that includes any metadata that "is an integral or intrinsic part of an electronic record. . . ," finding that such metadata "is 'readily reproducible' in the FOIA context."

What metadata is part of a record under FOIA?

The more difficult inquiry for the court was determining what types of metadata are "an integral or intrinsic part of an electronic record" for purposes of FOIA. The court recognized that the answer to this question will depend both on the type of electronic record at issue (*e.g.* text record, e-mail, or spreadsheet) and on how a government agency maintains its records (*e.g.* native format, imaged document, or hard-copy only). It resolved that the best approach is to create a default presumption that all the metadata maintained by an agency as part of an electronic record is producible under FOIA, unless the agency demonstrates that such metadata is not "readily reproducible." This appears to suggest that whatever metadata is associated with an electronic record, as that record is kept in the ordinary course by an agency, is presumptively producible under FOIA, at least pursuant to a specific request for the metadata. (More on whether or not the metadata must be requested below.) It imposes the burden on the responding agency to demonstrate that it is unable to produce one or more particular types of metadata.

[The court's analysis was informed heavily by the federal rules of civil procedure (FRCP) governing the discovery of electronic information in civil litigation. In fact, the FRCP provides a general

framework for the production of electronic information in civil litigation whereby all relevant information is discoverable unless the producing party demonstrates that it is not reasonably accessible because of undue burden or cost. (For detailed information on the required procedures under the Rule, click [here](#) [6].)]

The court went further, though, carving out specific “fields” of metadata that the government agencies were to include in all subsequent productions of records in the case at issue. Although the court indicated that these fields were not intended to encompass a standard production protocol in all cases, it clearly signaled the importance of these particular fields, going so far as to state that, in the court’s opinion, they “are the minimum fields of metadata that should accompany any production of a significant collection of ESI [Electronically Stored Information].”

The court identified the following metadata fields to be produced with respect to any electronic record:

- (1) Identifier—a unique production identifier of the item
- (2) File Name—the original name of the item or file
- (3) Custodian—the name of the custodian or source system from which the item was collected
- (4) Source Device—the device from which the item was collected
- (5) Source Path—the file path from the location from which the item was collected
- (6) Production Path—the file path to the item produced from the production media
- (7) Modified Date—the last date the item was modified before it was collected
- (8) Modified Time—the last time the item was modified before it was collected
- (9) Time Offset Value—the universal time offset of the item’s modified date and time based on the source system’s time zone and daylight savings time settings

Furthermore, the court specified the following additional fields to accompany any e-mail production:

- (10) To—addressee(s) of the message
- (11) From—e-mail address of the person sending the message
- (12) CC—person(s) copied on the message
- (13) BCC—person(s) blind copied on the message
- (14) Date Sent—date the message was sent
- (15) Time sent—time the message was sent
- (16) Subject—subject line of the message
- (17) Date Received—date the message was received
- (18) Time Received—time the message was received
- (19) Attachments—the unique identifier number(s) of any attachments to the e-mail

Finally, the court ordered that all spreadsheets be produced in their native format, presumably to ensure that all the embedded metadata, such as underlying formulas, are included.

The court rejected requests for additional metadata fields in this case—specifically Parent Folder; File

Size; File Extension; Record Type; Master Date; and Author—although it noted that the production of these and other additional metadata should be evaluated by courts on a case-by-case basis.

The court's identification of specific metadata fields represents a departure from the **category approach** [2] taken by other courts. In that respect, it appears to be a positive sign that courts are beginning to recognize that all metadata is not created equally and is not equally essential to the record with which it is associated. On the other hand, at least some of the metadata fields listed above are external to the electronic record itself, and, therefore, may be more difficult to capture and produce.

Is certain metadata exempt from public access under FOIA?

One of the main concerns expressed by the government agencies was that the disclosure of some of the metadata associated with a particular record might lead (directly or indirectly) to the disclosure of information that is exempted from public access. In particular, the agencies were concerned with the ability of a requesting party to use the metadata to engage in reverse-engineering that might lead to the disclosure of protected information from redacted documents. The court acknowledged this concern, but indicated that there were available means to guard against inappropriate disclosures. A record could be produced in its native format but with certain substantive information and appropriate metadata redacted. Alternatively, a record could be produced in static format with the protected information redacted, along with an accompanying load file. (A load file is a file that relates to a set of scanned images of electronically processed files and that indicates "where individual pages or files belong together as documents, to include attachments, and where each document begins and ends. A load file may also contain data relevant to the individual documents, such as selected metadata, coded data, and extracted texts." See [The Sedona Conference Glossary: E-Discovery & Digital Information Management \(3d ed. Sept. 2010](#) [7].)

Must metadata be specifically requested under FOIA?

The final issue that the court addressed is whether metadata must be specifically requested. The court's analysis of this issue is a bit muddled, in part due to the facts of the case at issue. The court noted on a couple of occasions that the requesting parties never specified that they wanted the "metadata" associated with the underlying records sought. Because of this, the court did not order the government agencies to re-produce all of the records that had already been produced, even though the records had been stripped of all metadata. (The court did find, however, that there was an appropriate request for the metadata associated with the spreadsheet records, because these were solicited in their native format.) And, in its concluding remarks, the court indicated that metadata should be specifically requested. This requirement is consistent with the other cases that have addressed this issue thus far.

However, the court found the lack of the specific request for metadata to be a "lame excuse for [the agencies' failure] to produce the records in an unusable format." Even if the agencies did not have to provide all the metadata fields, absent a specific request, it was under a duty to provide the electronic records in a format that did not significantly degrade the electronic records' searchability. Thus, according to the court, "it is no longer acceptable for any party, including the Government, to produce a significant collection of static images of ESI without accompanying load files" to make the production searchable and therefore reasonably usable. This means that at least some metadata (the metadata that enables searchability) must be produced even absent a specific request.

Application of National Day Laborer Opinion

As with the other cases addressing the application of public records laws to metadata to date, *Nat'l Day Laborer* has no direct precedential value in North Carolina. In fact, as a district court opinion, it only applies to the case at issue. Judge Scheindlin, however, is a very well respected jurist. She authored the seminal series of opinions on the application of the federal rules of civil procedure to the discovery of electronic information. See, e.g., *Zubulake v. UBS Warburg*, 216 F.R.D. 280 (S.D.N.Y. 2003). And her opinions in this area are among the most, if not the most, widely cited by

federal and state courts alike. They have provided the framework for amendments to the federal rules, as well as many state court rules of civil procedure, and continue to influence litigation practices. Unless her opinion in *Nat'l Day Laborer* is overturned on appeal, it is likely to be influential in the emerging area of the application of public records laws to electronic information.

More specific analysis of the potential application of the framework set forth by Judge Scheindlin to North Carolina public records requirements awaits a future post...

Article printed from Coates' Canons: NC Local Government Law Blog:
<http://sogweb.sog.unc.edu/blogs/localgovt>

URL to article: **<http://sogweb.sog.unc.edu/blogs/localgovt/?p=4023>**

URLs in this post:

[1] Image: **<http://sogweb.sog.unc.edu/blogs/localgovt/wp-content/uploads/2011/02/puzzle.jpg>**

[2] here: **<http://sogweb.sog.unc.edu/blogs/localgovt/?p=1984>**

[3] here: **<http://sogweb.sog.unc.edu/blogs/localgovt/?p=2064>**

[4] here: **<http://sogweb.sog.unc.edu/blogs/localgovt/?p=3417>**

[5] **5 U.S.C. § 552(a)(3)(B)**: **http://www.law.cornell.edu/uscode/5/usc_sec_05_00000552--000-.html**

[6] here: **<http://www.sog.unc.edu/pubs/electronicversions/pdfs/ediscovery09.pdf>**

[7] The Sedona Conference Glossary: E-Discovery & Digital Information Management (3d ed. Sept. 2010): **<http://sogweb.sog.unc.edu/blogs/localgovt/wp-content/uploads/2011/02/glossary2010.pdf>**

Copyright © 2011 School of Government at the University of North Carolina. All rights reserved.