Coates' Canons Blog: Is Metadata a Public Record?: Part 1

By Kara Millonzi

Article: https://canons.sog.unc.edu/is-metadata-a-public-record-part-1/

This entry was posted on March 04, 2010 and is filed under Open Government, Public Records (Electronic)

(Updated March 15, 2010) Is metadata a public record? There is not a clear answer to this question, at least in North Carolina. But that does not mean that public entities should ignore the question. In fact, public officials, particularly attorneys and information technology professionals, are well advised to stay abreast of this emerging issue, as it likely will impact future responses to public records requests for electronic information and may even alter retention practices with respect to that information. This post defines metadata and summarizes the few cases (in other states) that have squarely addressed whether, and to what extent, metadata is a public record. In my next post, I will discuss the potential treatment of metadata under North Carolina's public records laws, and flag additional issues posed by metadata in the public records context.

What is metadata?

Metadata, commonly referred to as "data about data" is information describing the history, tracking, or management of an electronic document. There is no single, "dictionary" definition of metadata. The term has come to be understood, at least in the legal community, to describe a variety of information associated with electronic documents or files. Examples of metadata include a document's or file's designation, dates it was created, last accessed, or modified, its author, and its edit history. Metadata also may include information that is recorded by a computer to assist in storing and retrieving a file, or aid in its searchability. Finally, metadata allows the functioning of routines within a document or file, such as cell formulae in spreadsheets. Some metadata is supplied by the creator or author of electronic information; other metadata is created automatically by a computer. Metadata often is not static—at least some of the metadata associated with a particular electronic record may be modified multiple times and even deleted.

My colleague, Shannon Tufts, has created helpful tutorials illustrating the metadata associated with **Microsoft Outlook e-mails, Microsoft Word, Excel, and PowerPoint documents and portable document format (.pdf) documents.**

There is a growing body of case law addressing the treatment of metadata in the context of complying with state and federal litigation discovery requirements. Several courts have found it useful to group the various types of metadata into three categories—substantive (or application) metadata; system metadata; and embedded metadata. *See, e.g.*, Aguilar v. Immigration and Customs Enforcement Division of the United States Department of Homeland Security, 255 F.R.D. 350 (2008) (citing, among other sources, a **publication by the Sedona Conference**, the leading commentator on issues related to the discovery of electronic information). And, what category a particular type of metadata falls in may prove legally significant in analyzing whether or not it constitutes a public record.

Substantive (Application) Metadata

Substantive metadata reflects modifications to an electronic document, such as edits or editorial comments (for example, Track Changes in Microsoft Word documents). Thus, it is useful in showing the genesis of a document, including its history of revisions. It also includes data that instructs the computer how to display information—including fonts and spacing. Substantive metadata is created as a function of the software application used to create the document. It is embedded in the document it describes and remains with the document when the document is moved or copied.

System Metadata

System metadata encompasses information that is created automatically by a computer system. System metadata often is not embedded within the document or file it describes. Examples include the author of a document (based on information that is pre-set in the computer), and the date and time a document was created, last modified, and last accessed. This

type of metadata also allows a computer "user" to search and sort through multiple documents and files efficiently. System metadata may be relevant if a document's authenticity is at issue, or there are questions as to who received a document or when it was received.

Embedded Metadata

Embedded metadata is information that is inputted into a document by a user but is not typically visible on the "face" of the document. Embedded metadata includes the cell formulae in spreadsheets, externally linked files, hyperlinks, and certain database information. This type of information often is crucial to fully understanding the visible information in an electronic document. Some courts have held that embedded metadata is a subset of substantive (application) metadata.

(For a good discussion of the evolving treatment of metadata in the civil litigation context, click **here**.)

Case Law on Metadata as a Public Record

To date, there are three reported appellate cases addressing the application of public records laws to metadata. The courts in all three cases have held that at least some metadata constitutes a public record. Before drawing conclusions about the impact of these holdings, it is important to understand the context in which these decisions occurred. The following is a summary of the three cases:

O'Neill v. City of Shoreline, 187 P.3d 822 (Wash.App. 2008)

In *O'Neill*, the city's deputy mayor announced at a city council meeting that she had received an e-mail from two individuals that alleged improper influence by council members over a zoning issue. One of the individuals that the deputy mayor claimed she had received the e-mail from, Ms. O'Neill, subsequently requested to see the e-mail. The deputy mayor had received the e-mail on her personal e-mail account. In response to the request, the deputy mayor forwarded the e-mail to herself, and in the process deleted the first four lines of the header (including the "To," "From" and "Internet protocol address" information). The e-mail was produced to Ms. O'Neill without this header information. Ms. O'Neill made several subsequent oral and written requests to view the "entire e-mail string," "all information relating to this e-mail: how it was received . . . from whom it was received, and the forwarding chain of the e-mail," and "a copy of the e-mail . . . including all metadata, memos, and any other correspondence relating to this document." Some time after the deputy mayor forwarded the altered e-mail to herself, she deleted the original e-mail. The deputy mayor eventually asked the original sender of the e-mail to resend it to her and the city produced a paper copy of the re-sent e-mail to Ms. O'Neill with the header information displayed.

Ms. O'Neill brought an action against the city under the state's Public Records Act (PRA) for disclosure of the original e-mail and the metadata associated with the e-mail. The superior court dismissed the action. The Washington Court of Appeals vacated a portion of the superior court's judgment, holding, in relevant part, that at least some of the metadata associated with the e-mail was a public record.

The Washington PRA defines a public record as "any writing containing information relating to the conduct of government or the performance of any governmental or proprietary function prepared, owned, used, or retained by any state or local agency regardless of physical form or characteristics." The court of appeals indicated that the metadata at issue (e-mail header information) itself constituted a public record because it was "owned" by the deputy mayor in her official capacity, and it contained information that related to city business—specifically "the e-mail addresses of a person who may have knowledge of alleged government improprieties in dealing with a zoning matter." It is hard to imagine, however, that the Washington Court of Appeals would have held that the e-mail header information "relate[d] to city business" if it was completely divorced from the e-mail text which described the purported improprieties. Thus, the holding might better be understood as standing for the proposition that if an e-mail constitutes a public record, the e-mail's header information also constitutes a public record. The court further held that providing the re-sent e-mail to Ms. O'Neill did not necessarily satisfy the city's requirement under the PRA because the city had failed to demonstrate that it had provided the exact metadata requested by Ms. O'Neill. The court of appeals remanded to the trial court to determine if the requested metadata still existed on the deputy mayor's hard drive or elsewhere.

(Note that the Supreme Court of Washington granted a petition for discretionary review of this decision in the spring of 2009.)

Lake v. City of Phoenix, 218 P.3d 1004 (2009)

In *Lake*, a city police officer, David Lake, filed a lawsuit against the city alleging employment discrimination. Mr. Lake also submitted a public records request to the city seeking certain notes kept by Mr. Lake's supervisor documenting the officer's work performance. After reviewing paper copies of the notes, Mr. Lake suspected that they had been backdated. He subsequently requested the production of "metadata" or "specific file information contained inside" the electronic file containing the supervisor's notes. Mr. Lake specified that he wanted information reflecting the true creation date of the file, the access dates for each time the file was accessed, who accessed the file, and dates the file was printed. The city denied the request, claiming that metadata is not a public record.

The Arizona Court of Appeals affirmed the superior court's denial of production of the metadata embedded in the supervisor's notes. The court of appeals determined that metadata is not embraced by the common law definition of public records. (Public records are not defined under Arizona statutes. Instead, Arizona courts have adopted three definitions of the term "public record." *See* Mathews v. Pyle, 251 P.2d 893 (1952).) The Arizona Supreme Court reversed, holding that "when a public entity maintains a public record in an electronic format, the electronic version of the record, including any embedded metadata, is subject to disclosure" under the state's public records laws. In so holding, the court noted that Arizona law defines public records broadly and requires disclosure of documents with a "'substantial nexus' to government activities. . . ."

With respect to metadata, the court determined that the pertinent inquiry was not whether metadata considered by itself is a public record, but whether a "public record" maintained in electronic format includes the record's embedded metadata. The court concluded that embedded metadata is a part of an electronic document, stating that

> [i]t would be illogical, and contrary to the policy of openness underlying the public records laws, to conclude that public entities can withhold information embedded in an electronic document, such as the date of creation, while they would be required to produce the same information if it were written manually on a paper public record.

Somewhat confusingly, the court stated that its holding was limited to "application metadata" which is "embedded in the file it describes and moves with the file when it is moved or copied," and that it's analysis did "not encompass external or 'system metadata' . . . ." The facts of the case indicate, however, that the metadata at issue in the case was "system metadata," as it is defined by the case law and treatise cited by the court as precedent for its holding. It is unclear if the court misunderstood the categorizations of metadata set forth in those cases and treatise or if the court was re-categorizing certain types of metadata for purposes of analyzing it in the public records context.

Irwin v. Onondaga County Resource Recovery Agency, 2010 WL 462948 (N.Y.A.D. 4 Dept. Feb. 11, 2010)

In *Irwin*, a public agency that handles solid waste and recycling for Onondaga County, sent an e-mail newsletter to subscribers that included a picture of John Irwin engaged in the process of disposing of leaves at the agency's compost site. After unsuccessfully extracting "payment" in the form of two free annual compost site passes in exchange for the use of his picture in the newsletter, Mr. Irwin made a request under the state's Freedom of Information Law (FOIL) seeking, among other things, any and all records involving the photograph that was used in the e-mail newsletter, including the image file itself and any associated metadata. (Mr. Irwin also made extensive requests for all photographs available for use in any of the agency's publications, except those of agency employees, and all computer records associated with the photographs.) In response to the request, the agency provided Mr. Irwin with digital copies of the multiple photographs, include two of Mr. Irwin. The photographs were produced digitally, but they had "been reduced in quality and resolution and were bereft of metadata."

Mr. Irwin commenced an action against the agency, claiming that it unlawfully had denied part of this FOIL request by, among other things, failing to supply the requested metadata associated with the photographs. The trial court denied Mr. Irwin's request to compel disclosure of the requested metadata. The appellate court reversed, holding that the trial court should have ordered the agency to disclose the metadata associated with the photographs were subject to disclosure

under FOIL and requested by Mr. Irwin.

The appellate court identified the metadata at issue as system metadata—encompassing “file names and extensions, sizes, creation dates, and latest modification dates of digitally-stored photographs . . . .” According to the court because “[r]ecords stored in electronic format are subject to FOIL . . . .[the court was] constrained to conclude that the subject ‘system’ metadata, which is at its core the electronic equivalent of notes on a file folder indicating when the documents stored therein were created or filed” was subject to disclosure. (Under FOIL, a record is defined as “any information kept, held, filed, produced or reproduced by, with or for an agency or the state legislature, in any physical form whatsoever including, but not limited to, reports, statements, examinations, memoranda, opinions, folders, files, books, manuals, pamphlets, forms, papers, designs, drawings, maps, photos, letters, microfilms, computer tapes or discs, rules, regulations or codes.”) The court indicated that it was not addressing whether “substantive” or “embedded” metadata would similarly constitute a public record, choosing instead to forge a limited path in this emerging area.

What conclusions or inferences can we draw from these cases about whether, to what extent, and in what context, metadata constitutes a public record in North Carolina? That will be the subject of my next post. In the meantime, I would love to hear how local governments currently are addressing this issue.

Links

- www.thesedonaconference.org/content/miscFiles/publications_html?grp=wgs110
- jolt.richmond.edu/v14i3/article10.pdf