Cullen | Dykman

What is Metadata?

July 15, 2011

As part of Cullen and Dykman's new E-Discovery blog, we are happy to introduce our "What is…?" series. The series is intended to demystify some of the terminology, practices, and technology used to assess Electronic Discovery ("E-Discovery") issues that arise during litigation. The series aims to provide the reader with essential information to increase their understanding of this ever growing and complex area of law. Each short post communicates a key concept and issues related to that concept, which continue to be fleshed out by our courts.

Our second article of the series describes an intricate part of ESI known as "metadata." Metadata is often referred to as "data within data" or "data about data". It contains information about who created the data, when it was created, accessed, or modified, as well as other information related to a particular piece of data.

In order to view the metadata of ESI, you must have the data in its native format. This means that the data is in the format in which it was originally created (i.e. a Microsoft Word document is sent to the opposing party as a ".doc" rather than being converted into a portable document format [PDF]). Data in its native format typically keeps its metadata intact, unless it has been altered, which gives a Court discretion to impose severe punishments on the party who made the altercations. It is important to note that every action taken in a document, even opening the document, alters the metadata.

Metadata is not a separate piece of data, but rather embedded or "attached" to the piece of data in which it was retrieved from. For the most part, courts have recognized three separate types of metadata:

1. **Application metadata** is associated with a specific file and can consist of information about a file's display properties, including fonts, margins, color, and document edits.
2. **System metadata** is descriptive file information such as the file's title, the file author and owner, and other "profile" information.
3. **Embedded metadata** contains "text, numbers, content, data or other information that is directly or indirectly inputted into a [n]ative [f]ile by a user and which is not typically visible to the user viewing the output display' of the native file." Aguilar v. Immigration and Customs Enforcement Div., 255 F.R.D. 350, 356 (S.D.N.Y. 2008).

Both application and system metadata are straight forward concepts. Typically, you can view these forms of metadata by looking at the file in whatever software created it or by "right-clicking" on the file and viewing its properties. Viewing embedded metadata, however, requires third-party software or further investigation into the property functions of revealed by your Operating System. For a clearer understanding of embedded metadata, let's look at what information a normal picture has embedded in it:

(*Click the image to enlarge.*)

The above image is a digital picture taken in Alaska and transferred to a PC's hard drive. Simply by looking at the image, you cannot tell much about the file itself, except maybe the location – if you are familiar with Alaska scenery, that is. Now, let's take a look at what the embedded metadata reveals:

(*Click the image to enlarge.*)

As you can see, the metadata reveals information that is unavailable on the face of the image. Although metadata may differ depending on the type of camera used, a majority of the modern-day cameras store the following information:

- The make and model of the camera;
- The resolution of the image;
- The camera settings at the time of taking the picture;
- The compression format of the image; and
- The date the image was taken.

The image's metadata may be helpful in a lawsuit by establishing a material fact of the dispute. Unlike a printed version of the image, the electronically stored picture carries the picture's history. If the picture was simply printed, the metadata would be lost. Here, because the image is stored on a storage device, the metadata remains intact.

Metadata is not only embedded in pictures, but can be found in an array of other files, including Microsoft Word, Excel, and PowerPoint files, as well as emails, internet browsers, text messages, and pretty much any type of ESI. The metadata embedded into a file is different depending on what type of file the data is. For example, an email's metadata may reveal the sender, recipient, date and time sent or received, carbon copy recipients, attachments, attachment file locations, etc.

In addition to providing valuable information during factual disputes, metadata can also be helpful during the document review process. Metadata allows attorneys to use search functions and formulas to find documents more quickly than if they had to review each piece of data individually. By using keywords and filtered searches, the attorney may be able to avoid wasting time searching unnecessary documents during the review process.

As the litigation process continues to become more complex, attorneys must be aware of the importance of requesting metadata from an opposing party. When a case involves ESI, the use of data's metadata may save your case. Conversely, failure to account for metadata may cause you to lose an otherwise winnable case.

A special thanks to Sean Gajewski for helping with this post. Sean is a third-year law student at Hofstra University School of Law. You can reach him by email at srgajewski [at] gmail dot com. Bio:www.sgajewski.com.