PUBLICATION

The Problem of Metadata

Metadata is data that is embedded or hidden in other data. It is sometimes called "data about data." A problem with metadata arises when one sends an email or a document unwittingly disclosing confidential information.

The term "metadata" has come into prominence lately in the context of electronic information. In this context, metadata describes the contents, location, physical attributes, type and form of the electronic information. Typically, metadata tracks the changes and developments within a document and may contain information such as the author's name, the name of the server or disk where the document was saved, properties or summary information about the document, the names of previous authors and revisions to the documents. In essence, metadata addresses the who, what, when, where and how of the underlying data.

Metadata is encoded in order to allow it to be processed by a computer program. It may be encoded by using different protocols or schemes. The basic purpose of metadata is to allow electronic information to be located, organized and used. Depending on its type, metadata may assist in identifying the location of the information by supplying call numbers or other identifying information much like those used in library identification systems, such as the Dewey Decimal System. Those identifying systems might be based on logical groupings of the elements of the data.

Metadata may be attached to the underlying document in a number of ways. It can be embedded in the document itself. It may be in a separate document that is linked to the underlying data. Finally, it may be in a separate database. Ordinarily, metadata is created automatically by a computer tool.

If emails or other electronic data are transmitted in native format (or, in other words, in original format), the recipient will be able to access this metadata. It may contain information that the sender does not want revealed. A typical example would be a draft of an agreement. The metadata might show previous drafts of the agreement and give the other side in the transaction a strategic advantage in the conduct of the negotiations. In the discovery process in litigation, this information can be crucial. Whether metadata is supplied or not needs to be considered by both the client and lawyer before electronic information is disclosed to the other side in negotiations or litigation.

Metadata can easily be removed in a number of different ways. The document can be printed or scanned into PDF; it can be transferred to another program in a way that removes the metadata; or it can be scrubbed by a software tool that removes the metadata. Whether to remove metadata is a decision that needs to be on the checklist of every client and attorney engaged in any significant business negotiation or any litigation.