AVANSIC®
E-Discovery & Digital Forensics

When You Shouldn't Trust the Metadata:
The Truth Behind Creation, Modified, and Accessed Date Information

By Dr. Gavin W. Manes, President and CEO

Date filtering is a common method to reduce a data set during e-discovery, but creation, accessed, and modified dates may not be as trustworthy as you think. Computers change these dates under very specific circumstances that are not as intuitive as their titles imply. Knowing the conditions under which these dates are altered allows you to make informed decisions about their usefulness in e-discovery filtering. Note that the email metadata field "sent date" is a reliable field for use in filtering and provides insight into when an email attachment existed.

"Creation" Date

Creation date is a misleading term since the human definition and a computer's definition are very different. Most assume that it represents the first time the file was made, wherever that may have been. However, the computer defines creation date as the time that a change occurred to the file's location on a given file system. If you copy a file from another digital device (like a thumb drive), the creation date of the file on your system will be the date that you downloaded it.

For example, when you copy a file the computer's file system is creating what it recognizes as a new file. It does not reflect when the file was originally created (on that computer, or on any other computer that may have been used to generate the file); rather, it is the last time that the file system data was changed.

Indeed, creating a new file changes the location of that file since it did not exist before. In this case, the new file's creation date is the first time the file was made. If that file never moves within the file system, its creation date will represent the file's original "born date."

"Accessed" Date

The accessed date represents the last time that a program read the contents of a file. Although it sounds simple and intuitive, access dates are one of the least useful pieces of metadata for e-discovery and modern digital forensics purposes. In fact, newer versions of Windows (such as Vista and Windows 7) no longer update access times in order to increase file read times.

For example, a virus scanner program on a Windows XP computer goes through a computer and reads all of the documents looking for anything harmful and through those actions change the access dates of each scanned file. Any other program designed to scan the contents of the computer will have the same effect. Of course, a seasoned forensics investigator would be able to identify this activity using link file or plist examination or even noticing the large number of files with very similar access dates, e-discovery date filtering based on this access time would return erroneous hits.

"Modified" Date

The modified date represents the last time that new data was saved to the file. Therefore, if you open an Excel document, make content changes, and save it, the modified date will be changed. Some programs may prompt you to save even if you haven't made any obvious changes; Microsoft Word is a good example. If you open a document, print it, make no alterations to the text content, and then save, the modified date will change.

Examples

Below are examples of commonly encountered situations and their corresponding metadata changes for both an image file (such as a .JPG) and a Word document.

Graphic File (JPG)	Created	Accessed	Modified
View as a thumbnail		X*	
View in a program (such as Windows viewer, Photoshop) without saving		X*	
Saving as	X	X	X
Open, alter, and save		X	X
* Not updated by default on Windows 7 and Windows Vista			

Word Document	Created	Accessed	Modified
Open and save (whether or not changes are made)		X*	X
Open and don't save		X*	
Change filename without opening		X*	
Download and save an email attachment	X	X	X
* Not updated by default on Windows 7 and Windows Vista			

Deleted Documents

A file's timestamp changes when it is sent to the Recycle Bin because the computer changes the file's name and location. Therefore, the created and modified times may not be accurate for date filtering purposes once they have been placed in the Recycle Bin. The created and modified times do represent the time those files were placed in the Recycle Bin.

Conclusion

Although document metadata such as creation, accessed, and modified times may seem to be helpful during e-discovery date filtering, that information may be disingenuous. Internal file metadata may be more useful, but not all files have that information available. For example, Word captures its own metadata that is more comprehensive and more consistent that the file system's metadata. Image files have metadata also, called EXIF, that may store information about the camera itself or photo imaging software.

Therefore, legal professionals should take care when using date filtering as a search parameter during e-discovery projects.