INVESTIGATION · DIGITAL INVESTIGATION

Metadata as Evidence

Metadata, the data that describe other data, can show when a document was created, where a photograph was taken and who contacted whom. This article explains the main kinds of metadata, what they can prove, how they can mislead, and how to handle them as evidence.

ModQu research article · 1,298 words · 14 references (APA 7th edition)

When a document is saved, a photograph taken, an email sent or a call made, systems record information about the event as well as its content: when it happened, which device was used, who was involved and sometimes where. This information about information is called metadata. Riley (2017) describes metadata as structured information that describes, explains, locates or otherwise makes it easier to retrieve, use or manage a resource. In investigations, metadata often matter as much as content. A contract's text may be disputed, but its creation date, author and editing history can show whether it was written when claimed. A photograph may show a scene, but its embedded data can show when and where it was taken and with what camera.

Kinds of metadata

Common kinds of metadata in investigations

Each can be valuable evidence, and each has characteristic weaknesses

	What it can show	Cautions
File system	When files were created, modified and accessed on a device	Changed by copying, opening or system activity
Document properties	Author, organisation, creation date, editing time, revisions	Inherited from templates; easily edited
Image data (EXIF)	Camera model, date and time, settings, sometimes GPS location	Device clocks can be wrong; stripped by many platforms
Email headers	Sending servers, route, time stamps, message identifiers	Some fields can be forged; servers vary
Communications records	Who contacted whom, when, for how long, from where	Shows contact, not content or who held the device

Based on Casey (2011), Carrier (2005) and CIPA (2019) ModQu.com

Figure 1. Five kinds of metadata that commonly appear as evidence. Their value depends on understanding how the system that produced them works. Diagram by ModQu.

File system metadata records when files were created, last modified and last accessed on a storage device, along with their size and location. Carrier (2005) explains how different file systems record these values and what each actually means, which is often not what a lay reader would assume. Buchholz and Spafford (2004) examined how file system metadata can support forensic reconstruction of events, and the limits of what it can show.

Document properties are stored inside files by applications such as word processors and spreadsheets. They may include the author's name, the organisation, the creation date, the total editing time and the number of revisions. Some formats retain earlier versions or tracked changes.

Image metadata is recorded by cameras and phones in a standard format known as EXIF (Camera & Imaging Products Association [CIPA], 2019). It can include the camera model, the date and time, the exposure settings and, on many phones, the location from satellite positioning.

Email headers record the path a message took between servers, with time stamps at each step and unique message identifiers. They can help establish when and from where a message was sent and whether it is genuine.

Communications metadata held by telephone companies and internet providers shows who contacted whom, when, for how long and, for mobile phones, which cell towers were used.

What metadata can reveal

Metadata can be highly revealing even without content. Mayer et al. (2016) studied telephone metadata from volunteers and showed that it was possible to infer sensitive information, such as medical conditions and religious affiliation, from patterns of calls alone. Schneier (2015) argues that metadata about communications is often more revealing than the content, because it shows patterns of relationships and movement over time.

Courts have recognised this. In *Carpenter v. United States* (2018), the United States Supreme Court held that police generally need a warrant to obtain historical cell-site location records covering an extended period, because they reveal the "privacies of life". In *Digital Rights Ireland* (2014), the Court of Justice of the European Union struck down a directive requiring the general retention of communications metadata, finding that it allowed very precise conclusions to be drawn about people's private lives.

In investigations, metadata are used to establish timelines, test claims about when documents were created, link devices and accounts, and verify images. A document said to have been written in 2018 but containing metadata showing it was created with software released in 2021 is likely to have been fabricated or recreated. The documents presented by a television programme in 2004 as memos written in the 1970s were widely questioned because their typography appeared to reflect modern word processing rather than typewriters of the period; an independent review found that the network had failed to establish their authenticity (Thornburgh & Boccardi, 2005).

How metadata can mislead

Metadata are produced by systems, and they reflect how those systems work, not necessarily what a person did.

- **Clocks can be wrong.** Time stamps depend on the device's clock and time zone settings. A camera with its clock set incorrectly will record the wrong time on every photograph.
- **Ordinary activity changes metadata.** Copying a file, opening it, scanning it with antivirus software or moving it between systems can change file system dates. Investigators who open files on the original device may themselves alter the evidence.
- **Values are inherited.** A document created from a template may carry the template author's name. A file copied from another computer may keep its original creation date or receive a new one, depending on the system.
- **Metadata can be edited or removed.** Many tools allow document properties and image data to be changed, and many social media platforms strip location and other data from uploaded images.
- **Metadata show the device, not the person.** A record that a phone connected to a cell tower shows where the phone was, not who was carrying it.

Each of these limitations means that metadata should be interpreted with knowledge of the system that produced them and corroborated with other evidence.

Handling metadata as evidence

Because metadata are fragile, they must be preserved before anyone examines the original files. Standard digital forensic practice applies: create a verified forensic copy, record hash values, and work on the copy (International Organization for Standardization [ISO], 2012; Casey, 2011). When documents are collected for litigation, they should be collected in their native format with metadata intact rather than printed or converted, and guidance on electronic disclosure emphasises that parties should agree how metadata will be preserved and produced (The Sedona Conference, 2018).

Courts require that electronic evidence be authenticated, meaning that there is evidence that it is what it is claimed to be. In *Lorraine v. Markel American Insurance Co.* (2007), a United States federal court set out in detail how electronically stored information, including metadata, can be authenticated and the problems that arise when parties fail to do so. United States federal rules now allow data copied from a device to be authenticated by a certification that the copy was verified by a process of digital identification such as a hash value (Fed. R. Evid. 902(14)).

The investigator's report should explain what each metadata value is, how it was obtained, what it shows, and what alternative explanations exist. "The file was created on 3 March" is a weaker and more accurate statement when expressed as "the file system on the laptop records a creation time of 3 March at 14:02, according to the laptop's clock, which was checked and found to be accurate within two minutes".

A short example

The following hypothetical case illustrates the approach. In a contract dispute, one party produces a letter, dated two years earlier, in which the other party appears to agree to a price change. The other party says it never received the letter. Examination of the native file shows that its document properties record a creation date eight months ago, and that it was

created from a letterhead template that the company introduced only last year. The file system records on the sender's computer show the same recent creation date. The sender suggests that the letter was retyped from an earlier paper copy, but no paper copy, no earlier electronic version and no record of posting can be found. The metadata do not prove that the letter is false, but they remove its support as a contemporaneous document and shift the burden to the party relying on it.

Conclusion

Metadata turn ordinary files and records into evidence of when, where and how things happened. They can confirm or undermine accounts, expose fabrications and reveal patterns that content alone would not show. But they are produced by systems with clocks, defaults and quirks, and they can be changed by ordinary handling or deliberate editing. Used well, they are among the most valuable evidence available to investigators; used carelessly, they can support confident conclusions that are simply wrong.

References

Buchholz, F., & Spafford, E. (2004). On the role of file system metadata in digital forensics. *Digital Investigation, 1*(4), 298–309. https://doi.org/10.1016/j.diin.2004.10.002

Camera & Imaging Products Association. (2019). *Exchangeable image file format for digital still cameras: Exif version 2.32* (CIPA DC-008-2019).

Carpenter v. United States, 585 U.S. 296 (2018).

Carrier, B. (2005). *File system forensic analysis*. Addison-Wesley.

Casey, E. (2011). *Digital evidence and computer crime: Forensic science, computers and the Internet* (3rd ed.). Academic Press.

Digital Rights Ireland Ltd v Minister for Communications, Marine and Natural Resources (C-293/12 and C-594/12) EU:C:2014:238.

Fed. R. Evid. 902(14).

International Organization for Standardization. (2012). *Information technology — Security techniques — Guidelines for identification, collection, acquisition and preservation of digital evidence* (ISO/IEC Standard No. 27037:2012).

Lorraine v. Markel American Insurance Co., 241 F.R.D. 534 (D. Md. 2007).

Mayer, J., Mutchler, P., & Mitchell, J. C. (2016). Evaluating the privacy properties of telephone metadata. *Proceedings of the National Academy of Sciences, 113*(20), 5536–5541. https://doi.org/10.1073/pnas.1508081113

Riley, J. (2017). *Understanding metadata: What is metadata, and what is it for?* National Information Standards Organization.

Schneier, B. (2015). *Data and Goliath: The hidden battles to collect your data and control your world*. W. W. Norton.

The Sedona Conference. (2018). The Sedona Principles, third edition: Best practices, recommendations & principles for addressing electronic document production. *The Sedona Conference Journal*, *19*, 1–207.

Thornburgh, D., & Boccardi, L. D. (2005). *Report of the independent review panel on the September 8, 2004 60 Minutes Wednesday segment "For the Record" concerning President Bush's Texas Air National Guard service*. CBS News.