



C O M P U T E R F O R E N S I C S L A B

London digital forensics & e-discovery · Established 2007

Metadata for Lawyers: What It Can and Cannot Prove

Reading the Data About the Data Without Being Misled by It

Metadata decides cases: the contract "signed in 2019" whose file was created in 2023; the photograph whose coordinates place the photographer forty miles from the claimed scene; the author field naming the person who swore they never touched the document. And metadata misleads exactly as powerfully: created dates that record a copy, not a creation; author fields inherited from templates; timestamps rewritten by every migration the estate ever suffered. The lawyer who reads metadata as testimony wins cases the documents alone cannot win; the lawyer who reads it as gospel pleads confidently into a trap. This guide is the field manual: the metadata families and what each field genuinely records, the innocent processes that rewrite them, corroboration before conclusion, preservation through collection and production, and the questions that separate a finding from a guess.

⌚ Estimated reading time: 25 min read

§ ABOUT THE AUTHOR

PREPARED BY COMPUTER FORENSICS LAB E-DISCOVERY TEAM

Computer Forensics Lab is a London-based digital forensics and electronic disclosure practice established in 2007. Metadata examination runs through nearly all of its casework: document dating and authorship analysis, backdating and manipulation findings, photo and communication metadata, and CPR Part 35 reports that state precisely what a field proves and what it does not. Its eDiscovery arm, **e-discovery.uk**, preserves and produces metadata to the standards this guide describes.

ESTABLISHED 2007 · LONDON

ISO 17025-ALIGNED PROCEDURES

METADATA & DOCUMENT DATING

MANIPULATION EXAMINATION

CPR PART 35 EXPERT REPORTS

FULL CHAIN-OF-CUSTODY DOCUMENTATION

§ CONTENTS

In this guide

- 01 Executive summary
- 02 The problem in plain English: the data about the data
- 03 The metadata families: file system, application, communication and media
- 04 What rewrites metadata innocently: copies, migrations and sync
- 05 Reading fields safely: corroboration before conclusion
- 06 Metadata as manipulation evidence: backdating and its signatures
- 07 Preserving and producing metadata: collection to load file
- 08 Source architecture: where else the evidence lives
- 09 Worked examples
- 10 Common mistakes and technical limitations
- 11 Questions to ask · Suggested wording
- 12 Checklist and red flags · When to involve a digital forensic expert
- 13 Frequently asked questions
- 14 Glossary · References · Disclaimer · How a specialist laboratory can assist

Executive summary

THE HEADLINE POINT: EVERY METADATA FIELD RECORDS A SPECIFIC TECHNICAL EVENT: KNOW WHICH EVENT, CORROBORATE ACROSS SOURCES, AND PLEAD EXACTLY THAT

Metadata is powerful because it is machine testimony: written by systems with no stake in the case: and dangerous because each field answers a narrower question than its name suggests. A file-system "created" date records when this copy arrived on this volume, not when the document was authored; an Office "author" records the profile of whoever's template began the file; an email "sent" time is the sender's clock at the client, while the header hops are the servers' story. Three disciplines convert this into safe practice. **Field literacy:** know the families (§3) and the precise event each field records: most "smoking gun" errors are category errors. **Innocent-rewrite awareness:** copies, migrations, sync clients, upgrades and even the wrong collection method rewrite fields wholesale (§4): a 2023 created date on a 2019 document is a copy event until proven otherwise. **Corroboration:** conclusions rest on multiple independent sources agreeing (file system plus application internals plus server logs plus counterpart copies): the standing lattice method: and are pleaded to what the fields actually establish. Read this way, metadata dates documents, places devices, attributes actions and exposes forgeries: §6's signatures: with the confidence that survives cross-examination.

- **Name the event, not the label:** "created" means this copy, here: the field's technical meaning governs, never its column header.
- **Suspect the innocent rewrite first:** migrations and copies explain most anomalies: eliminated before manipulation is alleged.
- **Corroborate across independent sources:** one field is a lead; agreement across systems is a finding.
- **Preserve before you touch:** opening, copying and emailing documents rewrites the evidence: forensic collection per guides 1 and 81.
- **Produce it properly:** agreed metadata fields travel with productions: flattening per guide 75 destroys the record.

The problem in plain English: the data about the data

Every digital document carries two texts: the one humans wrote, and the one machines wrote about it: who, when, where, with what, how many times. The machine text is the more candid witness: nobody drafts their file-system timestamps for effect: and litigation has learned to read it: the backdated agreement betrayed by its own creation sequence, the "independent" report authored in the counterparty's template, the photo that testifies against its photographer's account.

But the machine is a literal-minded witness answering literal questions. It says "a file with this content arrived on this disk at this time": humans hear "the document was written then". It says "this profile name was in the authoring application's settings": humans hear "she wrote it". The gap between the answered question and the heard one is where cases are lost: confident pleadings built on fields that never said what counsel thought, dismantled by an opposing expert in a paragraph. The cure is not scepticism about metadata: it is precision about it, which is what the next five sections install.

The metadata families: file system, application, communication and media

- **File-system metadata:** created, modified and accessed times, size, name and path: properties of this copy on this volume, rewritten by copying and migration: the most quoted family and the most misread.
- **Application (embedded) metadata:** what the document carries inside itself: Office authors, company fields, total editing time, revision counts and internal timestamps; PDF producers and creation chains; the internal record that travels with the file and often contradicts the wrapper: guide 75's ladders and §6's signatures both live here.
- **Communication metadata:** email headers and their server hops (guide 53), message timestamps and delivery states in chat platforms (guides 63-66): the transmission record, strongest at the server side.
- **Media metadata:** EXIF in photographs and video: device make and model, capture time, GPS coordinates, editing software traces: the location-and-moment family, with its own §4-style caveats (stripped by platforms, settable in-camera).
- **System and platform metadata:** the audit and activity layers around documents: DMS histories (guide 75), cloud version and access logs (guides 51-58), the registry and artefact strata (guide 41): not fields on the file but records about it, and frequently the corroboration §5 demands.

What rewrites metadata innocently: copies, migrations and sync

- **Copying and moving:** a copy to a new volume typically takes a new created date (younger than its modified date: the classic "anomaly" that is merely a copy); moves within a volume behave differently: the operation, not the document, wrote the field.
- **Migrations and upgrades:** server moves, laptop refreshes, archive restorations and platform migrations re-stamp estates wholesale: one estate-wide date cliff is a migration's signature, not a conspiracy's.
- **Sync clients and cloud platforms:** OneDrive, Dropbox and kin negotiate timestamps between local and cloud with per-platform rules; downloads from web interfaces are new local files with today's dates: the guide 51-58 estates carrying their own conventions.
- **Email and messaging transit:** attachments saved from mail take save-time stamps; platforms re-encode media and strip EXIF: the §3 media family arriving pre-laundered.
- **The examiner's own footprint:** opening a document updates access times and can touch more: why guide 1's imaging and guide 81's pipeline exist, and why "IT had a look first" is a sentence every examiner dreads.

Reading fields safely: corroboration before conclusion

- **State the event:** every finding phrased as what technically occurred ("a file with this content was written to this volume on...") before any inference about authorship or timing is layered on: the discipline that survives cross-examination.
- **Cross-family agreement:** file system against embedded internals against platform logs against communication records: independent writers agreeing is the finding; a single field alone is a lead.
- **The counterpart check:** the other party's copy of the same document, hash-compared and field-compared (guide 82 Example 3's method): agreement authenticates; divergence localises the question.
- **Baseline the estate:** the anomaly only means something against the estate's normal: the migration cliffs, the sync conventions, the template habits: mapped before any one document's fields are read as exceptional: the standing fairness rule.
- **Plead to the rung:** "the document's embedded metadata records X" versus "the defendant did X": the ladder stated, per the series' attribution discipline: metadata places files, profiles and devices; humans are reached by corroboration.

Metadata as manipulation evidence: backdating and its signatures

- **The internal contradiction:** a "2019" agreement whose embedded creation chain, font availability or producer string post-dates its own execution date: the document testifying against itself: guide 75's Example 3 and guide 82's Example 3 both being members of this class.
- **The sequence impossibility:** created after modified is a copy; but authored-after-signed, printed-before-created, or a revision history beginning years after the document's claimed life need explanations the innocent rewrites of §4 cannot supply.
- **The too-clean estate:** documents whose fields align perfectly with the pleaded case while the estate around them carries the usual migration mess: manufactured tidiness as its own signal, weighed carefully and never alone.
- **Clock-manipulation traces:** system-clock rollbacks leave collateral artefacts (event-log discontinuities, out-of-order identifiers per guide 76's anchors): the forger must fool every witness, and rarely does.
- **The honest both-ways rule:** the same examination that convicts a backdated document clears a genuine one whose fields were mangled by migration: findings state which, with the §4 alternatives expressly considered: the fairness discipline that makes the conviction findings believed.

Preserving and producing metadata: collection to load file

- **Collect forensically:** imaging and metadata-preserving acquisition per guides 1 and 81: the fields captured before anyone's curiosity rewrites them: with the §4 examiner's-footprint warning circulated to clients on day one.
- **The load-file fields:** the agreed metadata set (dates, authors, paths, hashes, family links) extracted at processing and travelling with every document into review and production: PD 57AD's DRD expects the conversation; guide 81's pipeline implements it.
- **Native production preserves, imaging flattens:** the guide 75 lesson generalised: producing PDFs of native documents strips the embedded family: native-or-fielded production per the protocol, with flattening a choice both sides made knowingly or a defect one side answers for.
- **Redaction interactions:** redaction tools can alter or leak metadata (guide 93's territory): production workflows check what the redacted file still carries.
- **Challenge rights preserved:** where a produced document's metadata matters, the native original and its forensic provenance are obtainable: the §11 requests, drafted before the argument needs them.

Source architecture: where else the evidence lives

Per this series' source-architecture discipline: a document's fields are one witness to its life, and the questions metadata is asked: when was this made, who made it, has it changed: are answered by the estate around the file as much as the file. The answers also live in: the platform layers (DMS version ladders and audit trails per guide 75, cloud version histories and access logs per guides 51-58: server-side records the desktop forger cannot reach); the communication record (the document as attached, dated by mail servers per guide 53: every sending a timestamped, third-party-witnessed copy); counterpart holdings (their copy's fields against yours); the artefact strata on devices (LNK files, registry traces, print spools per guide 41: the file's local biography); backups and archives holding the document's earlier states (the time machine per the backup guides); and the §3 system family generally: logs about the file outliving and out-testifying the fields on it. A manipulation question is answered from this lattice: the forged field is one witness changed; the lattice is many witnesses who were not.

QUESTION	FILE'S OWN FIELDS	DMS / CLOUD PLATFORMS	COMMUNICATION RECORD	COUNTERPART COPIES	DEVICE ARTEFACTS	BACKUPS / ARCHIVES
When was it authored?	Embedded internals	Y: version ladders	First sending dates it	Their receipt date	Local creation traces	Earliest captured state
Who made / edited it?	Author fields: profile-grade	Y: account-attributed saves	Sender context	n/a	Logged-in user strata	n/a
Has it changed?	Revision internals	Y: the ladder itself	Versions as sent over time	Y: the §5 check	n/a	Y: state comparison
Was it backdated?	§6's contradictions	Server dates vs claimed dates	Absence before claimed date	Their copy's honest fields	Clock-rollback traces	Absent from period backups
Where has it been?	Path fields per copy	Sharing and access logs	Recipients list	Their paths	USB / folder traces	Historic locations

Fallback strategy in one line: when the fields are contested, mangled or gone, date and attribute the document from the lattice: platforms, transmissions, counterparts, artefacts and backups: the file is one witness; the estate is the panel.

Worked examples

EXAMPLE 1 · THE ANOMALY THAT WAS A LAPTOP REFRESH

A claimant's team believed it had found the case: the defendant's key board minute, dated 2020, sat on the imaged laptop with a created date of March 2023: "plainly fabricated after proceedings were intimated". The §4 pass deflated it in a day: every document on the volume predating March 2023 carried the same created date cliff; IT records showed a laptop refresh that month with a copy-based migration; the minute's embedded internals recorded a 2020 creation, a plausible revision count and the company secretary's profile; and the 2020 backup set contained the identical hash. The finding, honestly reported both ways, cost the claimant a bad pleading and bought its expert the credibility that carried the genuinely anomalous document found two months later. The estate's normal, read first, is the whole of §5 in one story.

EXAMPLE 2 · THE 2019 AGREEMENT WRITTEN BY 2022 SOFTWARE

A loan dispute produced a signed PDF agreement "executed March 2019". Its embedded chain disagreed: the producer string named a PDF library first released in 2022; the document's fonts included a variant unavailable until 2021; the creation timestamp, in the file's internals, sat eleven days after the letter before action; and no copy of the agreement appeared in either party's mailboxes, DMS or backups before that date: the §8 lattice returning silence across five independent sources for the whole of the document's claimed three-year life. The forger had set the visible date and the file-system clock and could not reach the library string, the font record or the negative lattice. The Example 1 defendant's honest anomaly and this one differ exactly as §6 describes: one had an innocent explanation that survived testing; the other had none.

EXAMPLE 3 · THE PHOTO THAT PLACED THE PHOTOGRAPHER

An insurance claimant submitted photographs of flood damage "taken the morning after the storm" at the insured warehouse. The EXIF told a layered story: capture timestamps eight days before the storm; GPS coordinates at a different unit on the same estate; and a device model matching not the claimant's phone but his brother's, whose own social media had posted one of the same images months earlier: recovered with intact metadata via the platform's original-quality download. The claim was withdrawn before the insurer's Part 35 report was served. The example's discipline point: every EXIF finding was corroborated (platform copy, device records, weather data): because stripped, edited and re-stamped EXIF is common enough that media metadata alone is a lead, and the lattice made it a finding.

Common mistakes and technical limitations

Common mistakes

- **Field labels read literally:** "created" pleaded as authorship: Example 1's near-miss, made in filings daily.
- **Manipulation alleged before migration is excluded:** §4 skipped, the innocent rewrite discovered by the opposing expert instead.
- **Single-field findings:** conclusions balanced on one writable value: the lattice never assembled.
- **Author fields as identification:** profile names pleaded as hands on keyboards: the attribution ladder skipped at its first rung.
- **Evidence handled before collection:** the client's team opening, copying and forwarding the key documents: §4's examiner footprint, self-inflicted.
- **Flattened productions accepted silently:** PDFs where natives were owed: the embedded family destroyed without objection: guide 75's lesson unlearned.
- **EXIF trusted raw:** media metadata pleaded without platform-stripping and re-stamping considered: Example 3's corroboration omitted.
- **Baselines unread:** anomalies declared against estates whose normal nobody mapped.

Technical limitations

- **Most fields are writable:** tooling can set nearly anything file-side: which is why server-side and lattice evidence outranks fields, and findings say so.
- **Absence is weak:** stripped or empty metadata proves handling, not guilt: read with §4's launderers in mind.
- **Precision varies by field and system:** some timestamps are seconds, some days, some time-zone ambiguous: stated per field, per guide 70's clock disciplines.
- **Platforms change conventions:** sync and export behaviours move with versions: per-matter verification over remembered rules.
- **Interpretation is expert territory at the margins:** the deep internals (OLE streams, XMP chains, filesystem journals) reward specialist reading: the §12 threshold, stated plainly.

§ 11 · INTERROGATORIES & DRAFTING AIDS

Questions to ask · Suggested wording

Ask your client

- Who has touched the key documents since the dispute arose: opened, copied, emailed, "tidied": before we collected?
- What migrations, refreshes and platform moves has the estate had: the §4 events our baseline must know?
- Where else do the key documents live: DMS, cloud, backups, counterparties: for the §8 lattice?

Ask your opponent

- Please produce [documents] in native format with all embedded metadata intact, together with the agreed load-file fields, and identify the collection method used and whether it preserved file-system metadata.
- Please state the provenance of [document]: the system it was collected from, its original path, and the custodian: and produce any platform version history or audit records for it.
- Please confirm whether any migration, restoration or bulk-copy events have affected the relevant sources in [period], with dates and records.

Ask your eDiscovery / forensic provider

- What does each load-bearing field technically record: and which of our intended inferences does it actually support?
- What is the estate's baseline: migration cliffs, sync conventions, template habits: before we call anything anomalous?
- What does the §8 lattice say about the contested document: platforms, transmissions, counterparts, backups?

SUGGESTED WORDING · METADATA PARAGRAPH FOR THE DISCLOSURE PROTOCOL

"Documents shall be collected by methods preserving file-system and embedded metadata, and the collection method shall be recorded. Productions shall include the metadata fields listed at Appendix [A] for each document, and documents whose embedded metadata may be material shall be produced in native format. No party shall convert native documents to image or PDF formats in a manner that removes embedded metadata except as required for redaction, in which case the fact of conversion shall be indicated. Where the authenticity, date or authorship of a document is challenged, the producing party shall on request make available the native original, its forensic provenance, and any platform version or audit records relating to it."

Checklist and red flags · When to involve a digital forensic expert

The metadata checklist

- ☐ Key documents forensically collected before anyone's curiosity
- ☐ Load-file fields agreed and travelling with productions
- ☐ Estate baseline mapped: migrations, sync, templates
- ☐ Each finding phrased as its technical event first
- ☐ Innocent rewrites excluded before manipulation alleged
- ☐ Cross-family and counterpart corroboration assembled
- ☐ Attribution pleaded to its rung: file, profile, account, human
- ☐ Media metadata corroborated past platform laundering
- ☐ Native production preserved for metadata-material documents
- ☐ Both-ways findings reported honestly, Example 1 style

Red flags

- Created dates younger than modified dates on "key" documents: usually copies, occasionally §6.
- Documents whose internals post-date their claimed execution: Example 2's class.
- Estates with one suspiciously tidy document amid normal mess.
- Key documents absent from every period backup and mailbox.
- Productions flattened to PDF where natives were owed.
- Media evidence whose EXIF is intact, convenient and uncorroborated: or stripped precisely where it matters.
- Opposing metadata claims pleaded from single fields.

When to involve a digital forensic expert

Before anyone opens the key documents, so collection preserves what argument will need; at analysis, where field literacy, baselining and the §8 lattice turn leads into findings (Examples 1-3's machinery); at challenge, in either direction, where manipulation is alleged, tested and reported with the §6 both-ways discipline; and at production, where the protocol's metadata terms are drafted, policed and enforced. Under CPR Part 35, metadata findings state their events, their corroboration and their limits: which is precisely why they get believed. Through **e-discovery.uk**, the fields travel intact from collection to production as standard.

§ 13 · COMMON QUESTIONS

Frequently asked questions

Can metadata prove when a document was really created?

Often, but never from one field: authorship dating rests on the lattice: embedded internals, platform version ladders, first transmission, presence in period backups: agreeing. A file-system created date alone dates a copy event; Example 2 shows what real dating evidence looks like, and Example 1 shows why the single field misleads.

The author field names our opponent. Is that proof they wrote it?

It is proof a profile bearing that name was configured in the authoring application: templates, shared machines and IT-imaged laptops write other people's names daily. Treat it as the ladder's first rung and corroborate: platform saves under their account, drafting sessions on their device, the document travelling from their mailbox: per §5's rung discipline.

A key document's dates look wrong. Manipulation?

Statistically, probably a copy or migration: §4's innocent rewrites explain most anomalies, which is why the baseline and the IT history are read first: Example 1's sequence. What survives that filter: internal contradictions, impossible sequences, lattice silence: is §6's territory, and worth the full examination.

Does emailing or copying a document really destroy metadata?

It rewrites the file-system family (new dates, new paths) while embedded internals usually survive: "destroy" overstates it, but every handling event overwrites part of the record and adds noise the examination must later subtract. The practical rule stands: preserve first, then investigate: and tell the client's helpful IT team today.

Our documents were produced as PDFs and the other side now questions their dates. Are we stuck?

No: the natives still exist (production copies are derivatives), and the dating case runs from the §8 lattice regardless: DMS ladders, transmissions, backups: which no flattening touched. Produce the natives with provenance under the protocol, and let the lattice answer: it is usually stronger than the fields anyway.

Is metadata admissible, and how much weight does it get?

Routinely admissible as real evidence, with weight tracking exactly the disciplines in this guide: preserved provenance, stated events, excluded innocents, corroborating lattice: the difference between a finding a court adopts and a field an opposing expert re-explains. Weight is built at collection time, which is the quiet moral of the whole guide.

§ 1 4 · REFERENCE

Glossary

Baseline

The estate's normal metadata behaviour, mapped before anomalies are read.

Created / modified / accessed

File-system timestamps of this copy on this volume.

Embedded metadata

Fields inside the file: authors, internals, producer chains.

EXIF

Media metadata: device, capture time, GPS.

Lattice

Independent corroborating sources around a document.

Load-file fields

The agreed metadata set travelling with productions.

Migration cliff

An estate-wide date rewrite from a bulk copy event.

Producer string

The software chain recorded inside a PDF.

Profile author

The application-configured name written to author fields.

XMP

An embedded metadata standard carrying edit and history chains.

Sources and authoritative references

The metadata disciplines in this guide draw on materials published by our laboratory and e-discovery practice, referenced here as sources:

- cflab.uk, digital forensics services (metadata examination, document dating, manipulation findings, CPR Part 35 reporting): cflab.uk/digital-forensics-services · guides library: cflab.uk/guides
- e-discovery.uk, EDM Phase 03 · Collection and Phase 04 · Processing (metadata-preserving acquisition; load-file fields; native production): e-discovery.uk/edrm/collection · e-discovery.uk/edrm/processing
- e-discovery.uk, practice method and Knowledge Centre: e-discovery.uk · e-discovery.uk/knowledge
- PD 57AD and the Disclosure Review Document (metadata field agreement): justice.gov.uk, [PD 57AD](https://justice.gov.uk/pd-57ad) · CPR Part 35: [justice.gov.uk, Part 35](https://justice.gov.uk/part-35)
- Forensic Science Regulator, Code of Practice: gov.uk, [FSR Code](https://gov.uk/fsr-code) · ISO/IEC 27037: iso.org, [27037](https://iso.org/27037)
- ICO, UK GDPR guidance: ico.org.uk

DISCLAIMER

This publication provides general information about metadata as at August 2026. Field behaviour varies by system, platform and version and changes on vendors' schedules; verify the current position for the estates in the matter at hand. The worked examples are fictional. This guide is not legal advice, does not create a professional relationship, and should not be relied upon in place of advice on the facts of a specific matter from a qualified lawyer or a suitably qualified forensic practitioner. Links were live at the date of publication.

§ HOW A SPECIALIST LABORATORY CAN ASSIST

Working with Computer Forensics Lab

Metadata work at the laboratory runs the disciplines of this guide end to end: preservation-first collection before the client's estate rewrites itself; field-literate analysis phrased as technical events; the estate baseline and the §4 innocents read before anything is called anomalous; the §8 lattice assembled from platforms, transmissions, counterparts and backups; and manipulation examinations reported both ways under CPR Part 35, Example 1's honesty funding Example 2's conviction. Production-side, **e-discovery.uk** keeps the fields travelling: agreed load files, native handling, and the protocol terms of §11 drafted and policed. Conflicts checks, confidential scoping discussions and fixed-fee estimates are routine; and if a document's dates already smell wrong in your matter, the first instruction is the easiest: nobody opens it again until it is imaged.



I N S T R U C T T H E L A B

Speak to a forensic examiner, not a salesperson.

Describe your matter in confidence and an examiner will respond the same working day. Conflicts checks and scoping calls with US counsel are routine.

NEW ENQUIRIES

+44 (0)20 7164 6971

EMAIL

info@cflab.uk

E-DISCOVERY

e-discovery.uk

Computer Forensics Lab Ltd · Euro House, 133 Ballards Lane, Finchley, London N3 1LJ, United Kingdom

cflab.uk · Customer service +44 (0)20 7164 6915 · ICO ZB395023 · NPCC · ISO 17025 aligned practice · UK Gov Digital Marketplace