

AI Generated Documents Provenance And Detection

AI can generate plausible documents rapidly, but text analysis for detection is weak. This guide explains how provenance, examining a document's origin and history, provides the forensic solution for UK legal professionals.

e-discovery.uk, Computer Forensics Lab. Published 2026-08-31.

<https://e-discovery.uk/library/ai-generated-documents-provenance-and-detection>

AI - GENERATEDDOCUMENTS · A GUIDE FOR UK LAWYERS AI-Generated Documents:
Provenance and Detection Fabricated Evidence, Invented Citations and the Return to
Provenance COMPUTER FORENSICS LAB

§ ABOUT THE AUTHOR PREPARED BY COMPUTER FORENSICS LAB E-DISCOVERY TEAM
ESTABLISHED 2007 · LONDON ISO 17025-ALIGNED PROCEDURES DOCUMENT
PROVENANCE ANALYSIS

§ CONTENTS In this guide 01 Executive summary 02 The problem in plain English: plausible
text is cheap now 03 How AI-generated and AI-assisted documents arise 04 Why text-based
detection is weak 05 Provenance: the forensic answer 06 Metadata, corroboration and honest
limits 07 Deployment: fabrication, hallucinated citations and duties 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

§ 01 · ORIENTATION Executive summary The headline point: AI can now fabricate a
plausible document in seconds, but detecting AI from the text alone is unreliable, so the
forensic answer is not a detector but provenance: where the document actually came support
its claimed origin.

§ 02 · FIRST PRINCIPLES The problem in plain English: plausible text is cheap now

§ 03 · HOWTHEYARISE How AI-generated and AI-assisted documents arise

§ 04 · WHYTEXTDETECTIONISWEAK Why text-based detection is weak

§ 05 · PROVENANCE Provenance: the forensic answer

§ 06 · METADATAANDHONESTLIMITS Metadata, corroboration and honest limits

§ 07 · DEPLOYMENT Deployment: fabrication, hallucinated citations and duties

§ 08 · THEWIDERMAP Source architecture: where else the evidence lives CORROBORATING
RECIPIENT / SERVER / BACKUPS / APPLICATION / SOURCE COUNTERPART CLOUD
VERSIONS DATABASE THE SENDER / DOCUMENT AUTHOR ITSELF SYSTEM

§ 09 · IN THE WILD Worked examples EXAMPLE1 ·
THEEMAILTHATEXISTEDONLYASAFILE EXAMPLE2 ·

THE SUBMISSION WITH AUTHORITIES THAT DID NOT EXIST EXAMPLE 3 ·
THE GENUINE DOCUMENT DEFENDED BY ITS PROVENANCE

§ 10 · WHERE IT GOES WRONG Common mistakes and technical limitations Common mistakes
Technical limitations

§ 11 · INTERROGATORIES & DRAFTING AIDS Questions to ask · Suggested wording Ask
your client Ask your opponent Ask your e Discovery / forensic provider SUGGESTED
WORDING · INSTRUCTION FOR A DOCUMENT- PROVENANCE EXAMINATION

§ 12 · QUICK CONTROL Checklist and red flags · When to involve a digital forensic expert The
provenance checklist Red flags When to involve a digital forensic expert

§ 13 · COMMON QUESTIONS Frequently asked questions Can you tell if a document was
written by AI? How do you prove a document is fabricated? Why do we need the native file and
not a printout? Our opponent's submissions cite cases we cannot find. What does that mean?
Can metadata itself be faked? Is using AI to draft documents a problem in itself?

§ 14 · REFERENCE Glossary Sources and authoritative references 19 reporting):
cflab.uk/digital-forensics-services · guides library: cflab.uk/guides DISCLAIMER

§ HOW A SPECIALIST LABORATORY CAN ASSIST Working with Computer Forensics Lab
Speak to a forensic examiner, not a salesperson. INSTRUCT THE LAB
NEW ENQUIRIES EMAIL - DISCOVERY

Questions and answers

Can you tell if a document was written by AI?

Not reliably from the text, and no honest examiner will claim to (§4). AI-text detectors produce false positives and negatives and are easily defeated, so a "this was written by AI" opinion based on the prose does not survive scrutiny. What can be established, objectively, is provenance: whether the document's metadata and creation history support its claimed origin, and whether it exists where a genuine document would (§5). That, not text analysis, is the reliable answer.

How do you prove a document is fabricated?

By provenance and corroboration, not prose (§5, Example 1). A genuine email exists in mailboxes, on servers and in backups; a genuine file has a consistent creation and revision history and copies in the cloud and with others. A fabricated document typically exists only where the party produced it, contradicts its own creation metadata, and is absent from every system that should hold it (§8). That objective absence and inconsistency, not how the text reads, exposes the fake.

Why do we need the native file and not a printout?

Because provenance lives in the native file and its system context (§5). A printout or screenshot strips the metadata, the author in g records, the creation and revision history, the very things the analysis depends on, and flattens the document into an image that cannot be tested. To examine provenance you need the electronic original with its metadata intact, obtained from the system where it lives, with integrity preserved.

Our opponent's submissions cite cases we cannot find. What does that mean?

It may mean the authorities were hallucinated by an AI tool used in drafting and never verified (§3, Example 2), an increasingly common and serious problem. The check is the simplest one: verify each citation against a primary source. Authorities that exist nowhere are a real risk to the party relying on them, engaging duties to the court, and the safeguard is verification, not any attempt to "detect AI" in the writing.

Can metadata itself be faked?

Yes, metadata can be altered, which is why it is never taken alone (§6). Its power comes from corroboration: a creation date is tested against independent system records, and a document's claimed metadata is checked against the mailboxes, servers and backups that should independently reflect it. Consistent corroboration across systems supports authenticity; contradiction and absence expose fabrication. Metadata is a strong tool, used with the same corroboration discipline as any artefact.

Is using AI to draft documents a problem in itself?

No, honest, checked and, where required, disclosed AI-assisted drafting is legitimate, and this guide does not treat the tool as the wrong (§7). The problems are fabrication, using AI to manufacture false evidence, and negligence, relying on invented authorities or facts without verifying them. The line is between assistance that is checked and disclosed and deception or carelessness that is not, and provenance and verification, not hostility to the technology, are

how that line is policed. cflab. u k · e-disc ove r y. u k ©2026 Computer Forensics Lab Ltd
·cflab.uk ·e-discovery.uk ·info@cflab.uk ·+44 (0)20 7164 6915 Page 15 of 17

Source: <https://e-discovery.uk/library/ai-generated-documents-provenance-and-detection>. Free to read and share with attribution to Computer Forensics Lab, e-discovery.uk.