What Is A Forensic Hash And Why Does It Matter

A forensic hash is a digital fingerprint proving a file's identity and integrity. This guide explains how hashing works, its practical applications, and its role in legal proceedings, including common misconceptions and how to present hash evidence in court.

e-discovery.uk, Computer Forensics Lab. Published 2026-08-30.
https://e-discovery.uk/library/what-is-a-forensic-hash-and-why-does-it-matter

HASHVALUES · A GUIDE FOR UK LAWYERS What Is a Forensic Hash and Why Does It Matter? MD5, SHA and Integrity, in Language a Court Will Follow COMPUTER FORENSICS LAB DISCOVERY. UK

§ ABOUT THE AUTHOR PREPARED BY COMPUTER FORENSICS LAB E-DISCOVERY TEAM COURT-EXPERIENCED EXPERT WITNESSES FULL CHAIN-OF-CUSTODY DOCUMENTATION

§ CONTENTS In this guide 01 Executive summary 02 The problem in plain English: "how do we know it is the same file?" 03 How hashing works, for non-mathematicians 04 The algorithm families: MD5, SHA-1, SHA-256 05 The collision question, answered honestly 06 What hashes are used for in practice 07 What a hash proves, and what it does not 08 The misconception gallery 09 Hashes in court: the exact language 10 Worked examples 11 Common mistakes and technical limitations 12 Questions to ask · Suggested wording 13 Checklist and red flags · When to involve a digital forensic expert 14 Frequently asked questions 15 Glossary · References · Disclaimer · How a specialist laboratory can assist

§ 01 · ORIENTATION Executive summary THE HEADLINE POINT: AHASHPROVESIDENTITYANDINTEGRITY, AND ONLY THAT

§ 02 · FIRST PRINCIPLES The problem in plain English: "how do we know it is the same file?"

§ 03 · THEMECHANISM How hashing works, for non-mathematicians

§ 04 · THENAMES The algorithm families: MD5, SHA-1, SHA-256 ALGORITHM OUTPUT LENGTH STATUS AND ROLE IN PRACTICE

§ 05 · THEHARDQUESTION The collision question, answered honestly

§ 06 · THEWORKHORSE What hashes are used for in practice

§ 07 · THEBOUNDARY What a hash proves, and what it does not A MATCHING HASH PROVES A MATCHING HASH DOES NOT PROVE

§ 08 · THEGALLERY The misconception gallery

§ 09 · THELANGUAGE Hashes in court: the exact language Presenting your own evidence Challenging the other side's

§ 10 · IN THE WILD Worked examples EXAMPLE1 · THEPLANTINGALLEGATION,

DISSOLVEDINFOURQUESTIONS EXAMPLE2 · " MD5ISBROKEN " MEETSTHEDISTINCTION EXAMPLE3 · THELATEHASHTHATCERTIFIEDAWEEK

§ 11 · WHEREITGOESWRONG Common mistakes and technical limitations Common mistakes Technical limitations

§ 12 · INTERROGATORIES & DRAFTING AIDS Questions to ask · Suggested wording Ask your client Ask your opponent Ask your e Discovery / forensic provider SUGGESTED WORDING · REQUESTFORVERIFIC AT IONM AT ERIAL

§ 13 · QUICK CONTROL Checklist and red flags · When to involve a digital forensic expert The hash-evidence checklist Red flags When to involve a digital forensic expert

§ 14 · COMMON QUESTIONS Frequently asked questions Can two different files ever share the same hash? Can a hash be reversed to reveal the document's contents? Should we insist on SHA-256 every where? What can be done if hashes were never taken? What is dual hashing, and is it necessary? Do hashes help us find our document on the other side's systems?

§ 15 · REFERENCE Glossary NSRL SHA-256 Sources and authoritative references 19 reporting of integrity evidence): cflab.uk/digital-forensics-services DISCLAIMER

§ HOW A SPECIALIST LABORATORY CAN ASSIST Working with Computer Forensics Lab Speak to a forensic examiner, not a salesperson. INSTRUCTTHELAB NEWENQUIRIESEMAILE - DISCOVERY

Questions and answers

Can two different files ever share the same hash?

In theory yes (the outputs are finite and the inputs infinite), but for a court's purposes the honest framing is this: no two meaningfully different files have ever been found to share a SHA-256 value; the engineered MD5 and SHA-1 collision pairs are laboratory constructions built together on purpose; and the accidental collision probabilities are so small that every hash comparison ever run in litigation, worldwide, has not produced one. Dual hashing takes even the theoretical residue off the table.

Can a hash be reversed to reveal the document's contents?

No: the algorithms are one-way, and the 64 characters of a SHA-256 value contain no recoverable content: a two-terabyte image and a birthday card produce values of identical length and equal opacity. This is why hash schedules can be exchanged freely as verification anchors, even where the underlying material is privileged, without disclosing anything of what it says.

Should we insist on SHA-256 every where?

For new work, expect it as the default (usually alongside MD5); for existing evidence, no: an MD5 computed at acquisition and verified through the chain is sound integrity evidence, per §5, and re-hashing the preserved masters with SHA-256 now is a cheap supplement any laboratory will perform. The instruction worth giving is not "use this algorithm" but "hash at acquisition, dual where practicable, verify at every step, and record everything".

What can be done if hashes were never taken?

Compute them now, and be honest about what they mean: integrity is established from today forward, and the earlier period rests on what ever else exists: custody records, storage access logs, consistency with other sources, witness evidence of handling. Sometimes that composite is persuasive; some time s Example 3 happens. Either way, the material is rarely excluded for the gap alone in civil proceedings; it is weighed with it, which is precisely the risk assessment your case plan should make before the other side makes it for you.

What is dual hashing, and is it necessary?

Computing two algorithms' values (conventionally MD5 and SHA-256) over the same acquisition, so verification can proceed against either and any theoretical attack must defeat both at once. Necessary, strictly, no: a verified SHA-256 alone is robust. But it is effectively free at acquisition, forecloses the entire "MD5 is broken" theatre in advance, and keeps evidence comparable across old and new tooling, which is why it is standard laboratory practice and a reasonable expectation of any provider.

Do hashes help us find our document on the other side's systems?

Yes, precisely: give the examiner the file and its hash, and any bit-identical copy any where in an estate is findable by fingerprint, regardless of filename or location: the standard technique in leak and exfiltration cases. Its limit is the avalanche effect: a copy that was edited, re-saved or format-shifted no longer matches, and the hunt shifts to similarity hashing and content analysis, which find the family rather than the twin. cflab. u k · e-disc ove r y. u k