COMPUTER FORENSICS LAB

London digital forensics & e-discovery · Established 2007

Website, Web Server and Application Logs

Request Trails, User Journeys and Proving What Happened on a Website

Disputes increasingly happen on websites: the order placed or not placed, the terms accepted or never shown, the defamatory post and who made it, the account taken over, the price displayed at the moment of contract. The record of all of it is in logs: web servers noting every request with IP, timestamp and user agent; applications recording logins, actions and errors; CDNs, load balancers and analytics platforms adding their own layers. This guide equips UK lawyers to use that estate: what each log layer records, reconstructing user journeys and click-through evidence, proving what a page showed at a moment in time, attribution from IP to account to human, retention realities and the preservation letter, and the join to the counterpart evidence in the user's own browser per this series' method.

⏱ 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. Web-log evidence runs through its e-commerce, defamation, account-compromise and online-contract casework: server and application log collection and analysis, user-journey reconstruction, page-state proof, and IP-to-account attribution under CPR Part 35. Its eDiscovery arm, **e-discovery.uk**, hosts log-derived chronologies beside the documentary estate.

ESTABLISHED 2007 · LONDON | ISO 17025-ALIGNED PROCEDURES | WEB & APPLICATION LOG FORENSICS | USER-JOURNEY RECONSTRUCTION | CPR PART 35 EXPERT REPORTS | FULL CHAIN-OF-CUSTODY DOCUMENTATION

§ CONTENTS

In this guide

01 Executive summary

02 The problem in plain English: the site remembers every request

03 The log estate: servers, applications, CDNs and analytics

04 Reconstructing journeys: sessions, clicks and transactions

05 Proving page state: what the site showed and when

06 Attribution: IP, account, device, human

07 Retention, collection and the preservation race

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

§ 0 1 · ORIENTATION

Executive summary

THE HEADLINE POINT: LAYERED LOGS RECONSTRUCT WHAT A USER DID; PAGE-STATE PROOF AND ATTRIBUTION DECIDE WHAT IT MEANS

A website writes several diaries at once: web-server access logs (every request: IP, time, URL, status, user agent), application logs (logins, actions, orders, errors: the business layer), and the surrounding infrastructure (CDNs, load balancers, payment gateways, analytics): so the disputed episode can usually be rebuilt as a sessionised journey: arrived, browsed, clicked, submitted, confirmed: with server-side timestamps. Two further disciplines convert the journey into a case. **Page state:** the logs prove requests; what the page displayed comes from deployment history, template versions, CMS records, archives and the user's own browser copy: essential in online-contract and pricing disputes, where "what was on the screen" is the issue. **Attribution:** the IP is the weakest identifier (shared, dynamic, proxied); the account, session and device signals climb toward the human, per this series' ladder. Retention is short and configured: the preservation letter names every layer, and collection follows guide 71's structured-data disciplines.

- **Collect the layers together:** server, application, CDN and gateway logs corroborate each other: one alone is challengeable.
- **Journeys, not rows:** sessionised reconstruction with corrected timestamps is the deliverable: the click-through sequence the court can follow.
- **Page state is separate work:** what was shown is proved from deployments, CMS history, archives and the counterpart browser: not assumed from the URL.
- **IPs identify connections, not people:** attribution climbs IP to account to device to human, stated to its rung.
- **Retention is days-to-months by default:** the layered preservation letter goes out in week one.

§ 0 2 · FIRST PRINCIPLES

The problem in plain English: the site remembers every request

Every page view is a request; every request is a log line. When a customer says the checkout never showed the surcharge, when a claimant says the defamatory comment was posted by the defendant's account at 23:12, when a business says its portal was raided by a competitor's scraper: the events are all request sequences, and the request sequences were written down as they happened, by machines with no stake in the outcome. Few evidence classes offer this density: a contested ninety seconds on a website can yield dozens of timestamped, attributable entries across independent layers.

The catch is threefold, and it structures this guide: the diaries rotate quickly (retention is an engineering setting, not an evidence policy); they record traffic, not screens (what the user saw needs separate proof); and they name network addresses, not people (attribution is built). A lawyer who moves fast on preservation, asks for page-state evidence by name, and pleads attribution to its proven rung gets from web logs what they actually offer: which is usually the case.

§ 0 3 · THE ESTATE

The log estate: servers, applications, CDNs and analytics

- **Web-server access and error logs:** the base layer: timestamp, client IP, method and URL, status code, response size, referrer, user agent: every request, including the failed and the automated: with error logs recording what broke and when.
- **Application and audit logs:** the business layer: registrations, logins and failures, password resets, orders, payments, content posts, permission changes: richer, attributable to accounts, and governed by whatever the developers chose to record: the capability audit of this series applies.
- **Database and transaction records:** the orders, posts and account rows themselves with their timestamps: guide 71's territory, joined to the request trail that created them.
- **Infrastructure layers:** CDNs (edge logs, cache behaviour, the requests the origin never saw), load balancers and WAFs (blocks, bot scores), payment gateways (attempts, 3DS results, gateway-side timestamps): independent corroboration, differently retained.
- **Analytics and tag platforms:** Google Analytics and kin: journey data by design, but sampled, cookie-dependent and privacy-filtered: useful for pattern, treated cautiously for proof, and never the only layer relied on.

§ 0 4 · THE JOURNEY

Reconstructing journeys: sessions, clicks and transactions

- **Sessionisation:** grouping requests by session identifiers, authenticated account, IP-plus-user-agent continuity: turning the log stream into per-visitor narratives with entries, paths and exits.
- **The click-through sequence:** the dispute's minutes reconstructed request by request: product page, basket, terms page (served or not served, status 200 or not), submission, confirmation: the online-contract question answered at the resolution it deserves.
- **Cross-layer verification:** the same journey traced through CDN, server, application and gateway: agreement authenticates; divergence (requests at the edge missing at the origin) is itself a finding.
- **Time discipline:** layers log in different time zones and formats: everything normalised to the guide 70 reference clock before sequences are pleaded, with server NTP status checked where seconds matter.
- **Bots and noise:** crawler and scraper traffic separated from human journeys by agent, pattern and pace: both matter (scraping disputes are bot-journey cases), but conflating them misleads.

§ 0 5 · THE SCREEN

Proving page state: what the site showed and when

- **The gap to bridge:** logs prove the request for /terms; they do not contain the page: what version 200-OK'd at that moment is separate evidence, and online-contract cases live in the gap.
- **Deployment and version history:** release records, source control (guide 76's ladders) and CMS revision histories dating exactly which template, wording and price logic was live: the primary route.
- **Content and pricing records:** CMS audit trails, product and price change histories, A/B test configurations (which variant did this session get: the modern wrinkle, answered from test-platform logs).
- **Archives and captures:** the Wayback Machine and commercial monitors as dated third-party snapshots: coverage is patchy and page-specific, weight is real where they hit.
- **The counterpart browser:** the user's own device: cache, history, screenshots, confirmation emails: the other half of the story per this series' counterpart method, and often the only evidence of a dynamic, personalised render.

§ 0 6 · WHO

Attribution: IP, account, device, human

- **The IP's honest weight:** it names a connection at a moment: shared by households and offices, reassigned by carriers (CGNAT putting thousands behind one address), and cloaked by VPNs and proxies: the starting rung, never the finding.
- **Resolving the IP:** ISP subscriber disclosure via Norwich Pharmacal or statutory routes, with the timestamp-and-port precision carriers need: the request drafted to the second, in UTC, or it fails.
- **Account and session evidence:** the authenticated layer: which account acted, its registration details, credential history, session continuity: attribution's workhorse, tested for compromise before it convicts (the takeover defence is real and examinable: login patterns, resets, impossible travel).
- **Device and fingerprint signals:** user agents, device identifiers, cookies bridging sessions: linking the disputed visit to the defendant's other, admitted visits: quiet and effective.
- **The ladder, stated:** connection, account, device, human: corroborated by the counterpart device (guide 41's artefacts of the visit on their machine) and the correspondence: pleaded to the rung reached, per the standing rule.

§ 0 7 · BEFORE ROTATION

Retention, collection and the preservation race

- **Defaults are short:** server logs rotate in days to weeks; platforms and CDNs retain per plan; analytics per configuration: nothing about litigation is in the defaults: the letter changes that.
- **The layered ask:** preservation named per layer: server access and error logs, application and audit logs, database records, CDN and WAF logs, gateway records, analytics exports: with hosting arrangements identified (who actually operates each layer: the client, an agency, a platform, a host).
- **Third-party operators:** hosts, SaaS platforms and CDNs answer to their customers: the request routes through the site operator, fast, with provider ticket references confirming the freeze.
- **Structured collection:** raw log files with their rotation set, hashed, with the format documented: or platform exports with the reproducibility bundle: guide 71's disciplines, applied to text streams.
- **UK GDPR proportionality:** logs are personal data at scale: preservation is broad, processing and disclosure are minimised: the disputed sessions, not the whole internet's month.

§ 0 8 · THE WIDER MAP

Source architecture: where else the evidence lives

Per this series' source-architecture discipline: the origin server is one witness to a distributed event. The episode also lives in: the CDN and security edge (requests the origin never received, and its own retention); the payment and identity providers (gateway logs, 3DS records, SSO and OAuth events at the identity layer); the analytics and marketing stack (journey echoes, email-click records dating exactly when a user followed a link); backups and archived logs (the rotation-beating layer where shipping to storage was configured); the deployment and content history proving page state (§5's sources); the user's own browser and device (cache, history, autofill, screenshots: the counterpart layer); and the correspondence estate (confirmation emails, complaint threads quoting what the screen said). When the origin's logs have rotated, the edges, the gateways and the counterpart browser rebuild the journey: distributed systems distribute the evidence too.

EVIDENCE	ORIGIN SERVER	CDN / EDGE / WAF	APPLICATION & DATABASE	GATEWAYS & IDENTITY	USER'S BROWSER / DEVICE	DELETED / RECOVERABLE
The requests	Y: rotation-bound	Y: independent copy	Action records	Their legs of the journey	History + cache	Multi-layer survival
The transaction	POST entries	Y	Y: the rows, guide 71	Y: attempts + results	Confirmation artefacts	Database + gateway persist
What was shown	n/a: requests only	Cached copies sometimes	CMS/version history	n/a	Y: the render itself	Archives + counterpart rescue
Who it was	IP + agent	Bot scores, fingerprints	Account + session	Identity provider events	Guide 41's artefacts	Ladder rung-bound
Timing	Server clocks	Edge timestamps	Transaction times	Gateway times	Local artefacts	Cross-layer correction

Fallback strategy in one line: when the origin has rotated, take the journey from the edge and the gateways, the page from the version history and the counterpart browser, and the identity from the account layer: the site was never the only witness to itself.

§ 0 9 · IN THE WILD

Worked examples

EXAMPLE 1 · THE SURCHARGE THE CHECKOUT NEVER SERVED

A haulage customer disputed £84,000 of fuel surcharges, pleading that the booking portal's terms: incorporating the surcharge schedule: were presented at checkout. The layered record said otherwise for the disputed period: the application's deployment history showed the checkout template updated in April, and until the September fix the terms link returned status 404: logged, request by request, for every one of the customer's 211 bookings; the CDN's edge logs agreed; and the customer's own confirmation emails contained no terms attachment in the window. The operator's own logs had disproved its own pleading. The surcharges fell away for the 404 period, and the judgment's observation: that the claimant had evidently not read its logs before drafting: now travels with this laboratory's checkout-audit recommendations.

EXAMPLE 2 · THE REVIEW CAMPAIGN WITH ONE FINGERPRINT

A dental practice was hit by fourteen one-star reviews from "different patients" in nine days, hosted on its own review-enabled site. The application logs unified them: all fourteen accounts registered from three IPs belonging to one residential connection and one VPN exit; twelve shared a single user-agent string down to an unusual browser build; the session cookies showed three accounts operated in one browser within minutes of each other; and the registration emails resolved to sequential aliases at one provider. The Norwich Pharmacal order against the ISP, drafted to the second in UTC, named the subscriber: a former practice manager dismissed that spring. The defamation claim settled with the reviews' removal and costs, on an attribution ladder that never needed to leave the logs until its final rung.

EXAMPLE 3 · THE TAKEOVER DEFENCE, TESTED AND UPHELD

An investor faced a claim over abusive posts made from his account on a trading forum at 03:40. His denial: account takeover: is the standard defence, and this time the logs sustained it: the posting session originated from an IP geolocated abroad, on a user agent never before associated with the account; a password reset had been executed forty minutes earlier from the same foreign IP, using a recovery email changed in the same session; and his own devices' artefacts (guide 41's layer) showed no activity that night, while his home connection's records placed no traffic to the forum. The platform's own security log had flagged the login as anomalous. The claim against him was discontinued: the same examination that convicts a fabricated takeover defence, run honestly, established a real one.

§ 10 · WHERE IT GOES WRONG

Common mistakes and technical limitations

Common mistakes

- **Losing the rotation race:** the layered letter sent after the defaults have eaten the window: the class error again.
- **One layer relied on alone:** a single log file pleaded uncorroborated where the edge, gateway and application layers were available to agree.
- **IP pleaded as person:** §6 skipped: CGNAT, cafes and VPNs turn the finding to sand under cross-examination.
- **Page state assumed from URLs:** "they requested /terms" pleaded as "they saw the terms": Example 1's gap, on the wrong side.
- **Analytics as proof:** sampled, consent-filtered journey data carrying weight it cannot bear.
- **Timestamps unnormalised:** layers in three time zones stitched into one wrong chronology.
- **Takeover defences dismissed unexamined:** Example 3's lesson: the defence is testable, both ways, and testing it is cheaper than losing to it.
- **Whole-log disclosure:** a month of the public's browsing produced unminimised: the UK GDPR problem created in-house.

Technical limitations

- **Logging depth is a development choice:** the application recorded what its builders thought to record: the capability audit precedes every promise.
- **Rotated is usually gone:** absent shipped archives or backups, expired logs do not come back: stated honestly, per the standing rule.
- **Dynamic pages resist reconstruction:** personalised and A/B-tested renders may be provable only from test-platform records and the counterpart browser.
- **Clocks vary by layer:** unsynchronised servers put honest events in the wrong order until corrected.
- **Sophisticated actors launder attribution:** VPN chains and residential proxies cap the ladder's reach: the rung reached is the finding.

§ 11 · INTERROGATORIES & DRAFTING AIDS

Questions to ask · Suggested wording

Ask your client

- Who operates each layer: site, host, CDN, gateway, analytics: and what is each one's retention, today?
- Is log shipping or archiving configured anywhere: and are the disputed period's files already secured and hashed?
- What did the site show in the window: deployment history, CMS revisions, test configurations: can we prove the screen, not just the request?

Ask your opponent

- Please preserve and produce, for [period], the web server access and error logs, application and audit logs, and database transaction records relating to [sessions/accounts/URLs], together with CDN, WAF, load balancer and payment gateway logs, identifying the operator and retention setting of each layer.
- Please produce the deployment history, template and CMS revision records, and any A/B testing configuration showing the content, terms and pricing displayed at [URLs] during [period].
- Please identify the accounts, IP addresses, session identifiers and user agents associated with [the disputed activity], and any security events (resets, anomalous logins, takeover flags) on the relevant accounts.

Ask your eDiscovery / forensic provider

- Do the layers agree: and what does any divergence mean?
- What rung does attribution reach, and what would the next rung need: subscriber disclosure, counterpart device, identity-provider records?
- Can we prove the page as rendered: and does the counterpart browser hold the render itself?

SUGGESTED WORDING · WEB-LOG LIMB FOR THE PRESERVATION LETTER

> "Your client must immediately preserve all website and application records relating to the issues for [period], including: (a) web server access and error logs; (b) application, audit and security logs, including authentication events, password resets and account changes; (c) database records of the relevant transactions, posts or accounts; (d) logs held by content delivery networks, web application firewalls, load balancers, payment gateways and identity providers engaged by your client, whose preservation your client must procure and confirm; (e) deployment, version and content management records sufficient to establish the pages, terms and prices displayed during the period; and (f) analytics data for the relevant sessions. All log rotation, purging and retention-based deletion affecting this material must be suspended today, and where any provider's retention cannot be extended, the data must be exported forthwith and hashed. Please confirm the steps taken, per layer, within [3] days."

§ 12 · QUICK CONTROL

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

The web-log checklist

- ☐ Layer census: operators and retention per layer, day one
- ☐ Layered preservation letter out inside the shortest rotation
- ☐ Raw logs collected with formats documented, hashed, custody logged
- ☐ Timestamps normalised to the reference clock across layers
- ☐ Journeys sessionised and cross-layer verified
- ☐ Page state proved from versions, CMS, archives and the counterpart browser
- ☐ Attribution laddered: IP, account, device, human: rung stated
- ☐ Takeover and compromise hypotheses tested where raised
- ☐ Bot traffic separated from human journeys
- ☐ Disclosure minimised to the disputed sessions per UK GDPR

Red flags

- Retention defaults measured in days while correspondence is measured in weeks.
- Pleadings about what pages showed, unsupported by version history: Example 1's trap.
- Clusters of "independent" accounts sharing IPs, agents or cookies: Example 2's fingerprint.
- Anomalous logins, resets and recovery changes around disputed acts: Example 3's pattern.
- Layers that disagree: edge requests missing at the origin.
- Productions of curated log excerpts resisting the raw files.
- Attribution pleaded from an IP alone.

When to involve a digital forensic expert

At the census, where layers, operators and rotations are mapped against the clock; at collection, where raw logs, formats and hashes are handled per ISO/IEC 27037 and guide 71's bundle; at analysis, where sessionisation, cross-layer verification, page-state proof and the attribution ladder are built (Examples 1-3's machinery); at the procedural steps, where Norwich Pharmacal applications need second-precise, UTC-drafted schedules; and at reporting under CPR Part 35 with every finding on its rung. Through **e-discovery.uk**, log-derived chronologies sit in review beside the correspondence per guide 70.

§ 13 · COMMON QUESTIONS

Frequently asked questions

How long do websites keep their logs?

By default, briefly: server rotations of days to weeks, platform and CDN retention per plan, analytics per configuration: and almost never with litigation in mind. Treat every web matter as a guide 77-style race: census the layers, ask each one's retention, and send the §11 letter inside the shortest answer.

Can logs prove someone read the terms and conditions?

They prove service: the terms URL requested and returned 200 in the session, the checkbox posted, the version live that day (from the §5 history): which is what online-contract law generally needs: reasonable notice and incorporation, not telepathy. What no log proves is reading: plead the served-and-accepted sequence, and let Example 1 remind you to verify the serving before relying on it.

We only have an IP address for the wrongdoer. Is that enough?

It is a start: an IP plus a precise UTC timestamp supports subscriber disclosure from the ISP (Norwich Pharmacal being the usual route), and the account, device and fingerprint layers usually thicken the ladder before and after: Example 2's fourteen accounts fell to exactly that stack. An IP alone, pleaded as a person, is the mistake; an IP used as the first rung is the method.

The defendant says their account was hacked. Now what?

Test it: the takeover defence has a signature either way: login geography and device history, resets and recovery-detail changes, the platform's own anomaly flags, and the counterpart-device question (was the accused's machine active, and did it touch the site?). Example 3 shows the defence surviving honest examination; fabricated versions rarely do, because takeovers leave their own trail and its absence is loud.

Is a Wayback Machine capture admissible to show what a page said?

Routinely received, sensibly weighed: it is a dated third-party snapshot, strong for static public pages at its capture moments, silent between them and for personalised content. Use it as one voice beside deployment history, CMS records and the counterpart browser: the §5 chorus: rather than as the soloist.

Can we get the other side's server logs in disclosure?

Yes, where relevant and proportionate: they are documents within control like any other, and PD 57AD's models reach them: the craft is asking per §11: named layers, defined sessions and periods, raw formats, retention stated: so the request is cheap to comply with and expensive to resist. Expect the proportionality conversation, and win it by narrowness.

§ 14 · REFERENCE

Glossary

Access log
The web server's per-request record: IP, time, URL, status, agent.

CDN
Content delivery network; edge servers with independent logs.

CGNAT
Carrier-grade NAT; many subscribers behind one IP: attribution's diluter.

Norwich Pharmacal order
The disclosure route compelling ISPs to name subscribers.

Rotation
Automatic log replacement on age or size; the deadline.

Session
One visitor's grouped requests; the journey's unit.

Status code
The server's per-request result: 200 served, 404 missing (Example 1's digit).

User agent
The requesting browser/device string; a fingerprint component.

WAF
Web application firewall; blocks, challenges and bot scoring.

3DS
3-D Secure; the payment authentication leg with its own records.

Sources and authoritative references

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

- cflab.uk, digital forensics services (log forensics, journey reconstruction, attribution analysis, CPR Part 35 reporting): cflab.uk/digital-forensics-services · guides library: cflab.uk/guides
- e-discovery.uk, EDRM Phase 02 · Preservation, Phase 03 · Collection and Phase 04 · Processing (rotation-beating preservation; structured log acquisition; log data into review): e-discovery.uk/edrm/preservation · 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, CPR Part 31 and CPR Part 35: justice.gov.uk, PD 57AD · justice.gov.uk, Part 35
- ICO, UK GDPR guidance: ico.org.uk · ISO/IEC 27037: iso.org, 27037
- Internet Archive Wayback Machine: web.archive.org

DISCLAIMER

This publication provides general information about website and application log evidence as at August 2026. Logging depth, retention behaviour and provider capabilities vary by stack, plan and configuration and change on vendors' schedules; verify the current position for the estate 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

Web-log instructions at the laboratory open with the layer census and close rotations before they close the case: raw collection with formats and hashes, timestamps normalised, journeys sessionised and verified across the independent layers. Page-state proof is built from deployment and CMS history, archives and the counterpart browser; attribution climbs the ladder with subscriber-disclosure schedules drafted to the second where the next rung needs the court; and takeover hypotheses are tested honestly in both directions, per Example 3. Findings report under CPR Part 35 to their rungs, minimised to UK GDPR standards, and through **e-discovery.uk** the reconstructed journeys sit in review beside the correspondence on one chronology. Conflicts checks, confidential scoping discussions and fixed-fee estimates are routine; and if the dispute happened on a website, the retention question goes to its operators today, before the logs answer it for you.

INSTRUCT THE LAB

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