COMPUTER FORENSICS LAB

London digital forensics & e-discovery · Established 2007

Forensic Collection of Exchange Online Mailboxes

Active, Archived, Deleted and Recoverable Mail: Reaching Every Layer of the Mailbox, with Metadata Intact

The mailbox remains litigation's central exhibit, and Exchange Online has quietly made it deeper than most lawyers realise: beneath the folders a user sees sit online archives, recoverable-items stores holding weeks of deletions, hold-preserved layers keeping what users think they destroyed, and message properties recording far more than sender-recipient-date. This guide covers the mailbox's full anatomy, how each layer is preserved and collected defensibly, what deleted really means at each depth, the permissions and delegation records that decide who-read-what disputes, and the metadata disciplines that keep an email exhibit standing when its authenticity is challenged.

⏱ Estimated reading time: 26 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. Mailbox work has run through its casework since the on-premises era: today's practice spans Exchange Online preservation and collection, deleted-mail recovery through the layered stores, header and authenticity analysis, and delegation and access disputes. Its eDiscovery arm, **e-discovery.uk**, threads and deduplicates mailbox collections at disclosure scale.

ESTABLISHED 2007 · LONDON | ISO 17025-ALIGNED PROCEDURES | EMAIL & HEADER ANALYSIS | DELETED MAIL RECOVERY | CPR PART 35 EXPERT REPORTS | FULL CHAIN-OF-CUSTODY DOCUMENTATION

§ CONTENTS

In this guide

§ 0 1 · ORIENTATION

Executive summary

THE HEADLINE POINT: A MAILBOX IS A LAYERED STORE, AND MOST "DELETED" MAIL IS A LAYER QUESTION

An Exchange Online mailbox is not a folder tree: it is a stack. The **visible mailbox** (what the user sees) sits above an **online archive** (older mail moved by policy, searched separately and forgotten by users routinely); beneath both runs the **recoverable-items structure**: the staged afterlife of deleted mail, where items removed from Deleted Items persist for a retention window (14 days by default, extensible to 30), and where **holds convert the window into permanence**: a mailbox on litigation hold keeps everything, including what the user purges, including prior versions of items they edit, in layers invisible to them. The consequences run both ways: for recovery, "he deleted the emails" is the start of a layer walk, not a conclusion: the window, the hold status and the hold's placement date determine what survives, and guide 51's week-one hold decides most of it; for authenticity, the mailbox's depth is where fabrication dies: a genuinely sent message exists in transport-stamped, hold-preserved, counterpart-corroborated layers a pasted-up exhibit cannot fake. Collection must reach the whole stack deliberately: visible plus archive plus recoverable stores, with message metadata (headers, MAPI properties, timestamps at their true precision) preserved through export mechanics chosen for the purpose (guide 52's discipline), and the mailbox's **permission and delegation record**: who else could read, send-as or manage this mailbox, and when: collected alongside, because who-saw-what disputes are decided there and in the audit layer (guide 57). The standing instruction this guide justifies: never accept, or produce, "the mailbox" without stating which layers were included and what the hold history was: the two facts that govern everything else.

- **Hold status and hold date are the first questions** of any mailbox: they determine whether deletions were preserved or merely delayed: ask them before promising or conceding anything.
- **The archive is part of the mailbox:** policy-moved mail that custodians forget and rushed collections miss: named explicitly in every scope.
- **Recoverable-items is a store, not a rumour:** its contents are enumerable and collectable, with deletion dates: the purge itself becomes chronology.
- **Shared, delegated and departed mailboxes carry extra records:** permission history and delegation logs frame every "his assistant sent that" and "I never saw it" dispute.
- **Counterpart copies remain the great corroborator:** every external message lives in another tenant; guide 51's architecture rule applies to mail with special force.

§ 0 2 · FIRST PRINCIPLES

The problem in plain English: the mailbox beneath the mailbox

Users experience a mailbox as what they can see, and lawyers inherit the assumption: the Inbox, some folders, Deleted Items, done. Exchange Online's reality is closer to geology. Mail the user "deleted" last month may sit two strata down, dated and intact; mail from four years ago lives in an archive the custodian has genuinely never opened; the message a witness swears they never received may carry a delivery record and a read event; and the thread someone edited before forwarding exists in both versions because a hold was quietly keeping originals. None of this surfaces unless collection is instructed to reach it: a scope that says "the mailbox" gets the surface, and the surface is where weak disclosure lives.

The professional habit this guide installs is stratigraphic: every mailbox request, instruction and disclosure statement names its layers: visible, archive, recoverable, hold-preserved: and states the hold history. Two sentences of precision, and the difference between arguing about what a mailbox contains and knowing it.

§ 0 3 · ANATOMY

Mailbox anatomy: the layers and what each holds

LAYER	WHAT IT HOLDS · HABITS THAT MATTER
Visible mailbox	The user's folder tree, calendar, contacts, drafts, rules. Rules and auto-forwards are evidence in their own right (the quiet exfiltration mechanism guide 57's audit events corroborate). Drafts record composition without sending: intent evidence with its own timestamp semantics.
Online archive	Policy-moved older mail in a separate store with its own folder tree. Custodians forget it exists; searches and exports must include it by name; its move policy's dates explain why "the 2021 mail is missing" usually is not.
Deleted Items	Ordinary first-stage deletion; user-visible, user-emptied, unremarkable: except that its emptying pattern around key dates is chronology.
Recoverable items (Deletions)	Second stage: items removed from Deleted Items or shift-deleted, held for the retention window (14 days default, 30 configurable), user-purgeable but enumerable with deletion timestamps.
Recoverable items (Purges / Versions), under hold	The hold layers: purged items retained indefinitely while hold persists; edited items' prior versions preserved. Invisible to users; the layer where guide 51's Example 1 lived.
Mailbox configuration	Rules, forwards, delegates, permissions, auto-reply history: exported as exhibits, because each is a finding waiting to be read.

§ 0 4 · DELETED, REALLY?

What deleted means, layer by layer

- **Deleted to Deleted Items:** not deleted: moved. Fully collectable, with the move itself often datable from item properties.
- **Removed from Deleted Items / shift-deleted:** in Recoverable Items (Deletions) for the window: collectable with deletion dates; after the window, unheld: gone from the mailbox, and the architecture paths take over.
- **Purged by the user from Recoverable Items:** unheld: the strongest user-side deletion, gone at window's end or immediately on purge processing; held: preserved in the Purges layer regardless, with the purge attempt itself evidencing intent: the deletion that convicts.
- **Edited under hold:** the original preserved as a version: the before-and-after pair that decides tampering allegations without leaving the mailbox.
- **Admin-deleted mailboxes:** guide 51's leaver trap: the whole stack survives the grace period, then goes unless retention policies or holds independently preserved content: the question is factual, urgent and answered from tenant records, not memory.
- **The honest sequencing rule:** every recovery answer is "hold status, hold date, window length, deletion date": four facts, obtainable in an hour, that replace all speculation.

§ 0 5 · GETTING IT OUT

Collection: reaching every layer with metadata intact

- **Scope the stack explicitly:** the collection specification names visible mailbox, online archive and recoverable-items structures per custodian: Purview reaches them when told to; "when told to" is the discipline.
- **Choose export mechanics for mail:** per guide 52: format decisions govern how message properties, attachment relationships and folder provenance survive; the specification and metadata reports are exhibits; and threading downstream (the processing guides ahead) depends on identifiers exported now.
- **API and independent routes for the contested and the structural:** item-level properties beyond compliance export, and operator independence where guide 52's framework points there: mailboxes of accused custodians are the classic independent-collection population.
- **Collect the configuration alongside:** rules, forwards, delegation and permission state exported with the content: the one-page exhibit that repeatedly reframes cases (the forward-to-personal rule created the week before resignation).
- **Departed custodians first:** their grace clocks and licence dependencies make them the priority population in every mailbox plan: guide 51's leaver playbook, applied.

§ 0 6 · OTHER HANDS

Permissions, delegation and who-read-what

- **The permission map:** full-access grants, send-as and send-on-behalf rights, folder-level shares: exported as the mailbox's map of possibility, because "only he could have sent that" is checkable and often false.
- **Delegation in the message record:** sent items carry sender/from distinctions where delegates acted; the properties distinguish the principal from the hand: read them before pleading authorship.
- **Access events in the audit layer:** non-owner mailbox access, item reads and admin access are audit-log events (licence- and configuration-dependent: guide 57's territory): the who-read-what record for leak and privacy disputes, and the layer that caught guide 51's Example 3.
- **Shared mailboxes get shared-asset framing:** team inboxes with many hands are guide 43's shared folder translated to mail: the report maps the pool by permission and session before any singular attribution.

§ 0 7 · IS IT REAL?

Message metadata and authenticity

- **Headers are the message's passport:** transport hops, authentication results (SPF, DKIM, DMARC), message identifiers and per-hop timestamps: written by servers, not users, and the first read in any authenticity dispute.
- **Item properties extend the record:** creation, submission and delivery times at true precision, client information, conversation identifiers: the internal consistency a fabricated item struggles to fake and a genuine one exhibits effortlessly.
- **The layered-store corroboration:** a real message aligns across visible/hold layers, both parties' mailboxes, and transport logs; a fabricated exhibit exists in one place, incoherently: the mailbox's depth doing authenticity's work.
- **Counterpart comparison closes it:** the same message collected from both ends, hash- and header-compared: the two-tenant checkmate this series' email-forgery examples run on, and the standing reason to collect from every side you can reach.

§ 0 8 · THE WIDER MAP

Source architecture: where else the evidence lives

Per this series' source-architecture discipline: never examine the mailbox in isolation. Mail evidence coexists in: local caches on every device the account synced (OST files and mobile app stores: the offline copy that survives tenant deletion, per guides 41-42); counterpart mailboxes in other tenants (every external message, with their hold histories); journaling and archive platforms receiving mail in parallel (guide 49's better-first-stop); third-party tenant backups (monthly mailbox states across years); transport logs (message-trace records within their window: delivery facts without content); the audit layer (access, rules changes, deletions as events); personal accounts where forwarding rules or manual practice moved mail (guide 69 ahead); and printed or exported copies in the paper and file record. When the mailbox layer has genuinely forgotten: window expired, unheld, account gone: these paths, led by counterparts and backups, reconstruct the correspondence.

EVIDENCE	MAILBOX STACK	DEVICE CACHES (OST/ MOBILE)	COUNTERPART TENANTS	BACKUP / JOURNAL / ARCHIVE	TRANSPORT + AUDIT LOGS	DELETED / RECOVERABLE
Message content	Y, per layer + hold	Y: as last synced	Y for external mail	Y where deployed	Trace: facts, not content	Window + hold governed
Deletion events	Recoverable stores: dated	Cache divergence: inferential	n/a	Bracketed by sets	Audit: dated events	The purge often outlives the purger's intent
Rules / forwards	Configuration export	Client-side rules: sometimes	Received-side traces	Point-in-time states	Audit: creation events	Varies
Who read / accessed	n/a (content only)	Device artefacts	Their audit layer	n/a	Y: the record	Licence-bound window
Sent-item authenticity	Properties + hold versions	Cached copies	Y: the checkmate	Historical states	Headers + trace	Fabrication rarely survives convergence

Fallback strategy in one line: when the mailbox has forgotten, ask the OST, the counterpart, the backup and the journal: and remember the deletion's own record (recoverable stores, audit events) frequently matters more than the deleted content.

§ 0 9 · IN THE WILD

Worked examples

EXAMPLE 1 · THE PURGE PRESERVED BY THE HOLD

A sales director, litigation hold on his mailbox since the claim landed, spent a weekend "tidying": Deleted Items emptied, then Recoverable Items purged through the recovery interface: thoroughness that would have worked in 2015. The hold's Purges layer kept all 412 items with purge timestamps; the audit log recorded the sessions; and the purged set, reviewed, contained the pricing correspondence his witness statement denied existed. The tribunal heard the deletion chronology before the content: 22:40 to 01:15, the night after disclosure requests arrived: and the content after. Costs followed the conduct. The layer he could not see had been the only one that mattered.

EXAMPLE 2 · THE ARCHIVE THAT HELD THE MISSING YEARS

A partnership dispute's disclosure was attacked for a suspicious gap: almost nothing from 2020-2022, the formative years. The producing firm's collection had been honest and shallow: Purview search of visible mailboxes, archives never named in scope. The re-run, specified stratigraphically, found the years intact in every custodian's online archive, moved there by a two-year policy nobody remembered setting. Eleven thousand documents entered disclosure late, with a costs conversation about the first exercise: and the opposing solicitor's gap-analysis letter, which had read the absence as concealment, was answered by a move policy's date arithmetic. Nothing had been hidden except by scope drafting.

EXAMPLE 3 · THE EMAIL THAT EXISTED IN ONLY ONE MAILBOX

A claimant produced a decisive email: the defendant director agreeing terms: from her own mailbox; the director denied sending it. The stack analysis ran both ways. Her copy: present only in the visible layer, no counterpart in her hold-preserved structures consistent with normal receipt, headers showing an authentication chain that did not match the defendant tenant's outbound configuration for the period. His tenant: no trace in any layer, no transport record in the trace window, no hold-version residue: and his mailbox had been under hold since before the alleged send date, which made absence evidential rather than convenient. The laboratory's report stopped short of the word: the layered silence, the header incoherence and the one-sided existence did not need it. The claim was discontinued a fortnight later.

§ 10 · WHERE IT GOES WRONG

Common mistakes and technical limitations

Common mistakes

- **Surface-only collection:** Example 2's gap: archives and recoverable stores outside the scope nobody wrote down.
- **Hold questions unasked:** recovery promised, or loss conceded, without the four facts (§4) in hand.
- **Departed mailboxes deprioritised:** the grace clock running while active custodians are collected first.
- **Configuration ignored:** the forwarding rule that was the whole case, unexported.
- **Authorship pleaded past delegation:** "he sent it" where the properties say the assistant did.
- **Single-mailbox authenticity findings:** exhibits pronounced genuine or forged without counterpart and layer corroboration: Example 3's method skipped.
- **PST habits from guide 51's FAQ:** the unmanaged export as strategy, again.
- **Trace-window delay:** transport records aged out while correspondence debated whether to ask.

Technical limitations

- **Windows are windows:** unheld deletions past retention are gone from the stack: architecture paths, not indignation.
- **Audit and trace reach are licence- and time-bound:** guide 57's caveats govern who-read-what claims.
- **Header analysis has floors:** internal mail hops less; forwarded and copied items carry their transformation's marks; the examiner states what the headers can and cannot show per message.
- **Hold versions are not a full edit history:** preservation captures states the platform versioned, not every keystroke: claims sized accordingly.
- **Encrypted and protected messages** add rights-management layers to collection and review: identified early, handled deliberately, occasionally a stated limitation.

§ 11 · INTERROGATORIES & DRAFTING AIDS

Questions to ask · Suggested wording

Ask your client

- For each custodian: hold status and date, archive policy, recoverable-items window: the four facts, in writing.
- Which mailboxes are shared or delegated, and who has held what access since [date]?
- Which custodians have left, and where is each mailbox in the deletion pipeline?

Ask your opponent

- Please state, per custodian: litigation-hold status and placement date; whether collection included the online archive and recoverable-items structures; and the recoverable-items window in force.
- Please disclose mailbox configuration for [custodians]: rules, forwarding, delegates and permission grants, with change history from the audit log.
- For the message at [exhibit], please produce the complete item with full headers and properties from every layer in which it exists, and confirm the transport-trace position.

Ask your eDiscovery / forensic provider

- Does the collection specification name every layer, per custodian: and what will the export mechanics preserve for threading and authenticity work?
- What does the recoverable-items enumeration show around the key dates?
- For the disputed messages: what does the layer-and-counterpart analysis support, in report-ready language?

SUGGESTED WORDING · MAILBOX LIMB FOR THE COLLECTION SPECIFICATION

> "For each custodian at Schedule A, collection shall include: (a) the primary mailbox including all folders, calendar, contacts, drafts and Deleted Items; (b) the online archive in full; (c) the recoverable-items structures, including Deletions and, where holds apply, Purges and Versions, with item-level deletion timestamps preserved; and (d) mailbox configuration: rules, forwarding, delegation and permissions: as at collection and, from audit records, across [period]. Export shall preserve full message headers, item properties and conversation identifiers per the format specification at Schedule B, with metadata reports and hashed manifests exchanged. Hold status and placement dates per mailbox shall be stated in the collection report."

§ 12 · QUICK CONTROL

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

The Exchange Online checklist

- ☐ Four facts per mailbox: hold status, hold date, window, archive policy: documented day one
- ☐ Departed mailboxes triaged first; grace clocks stopped
- ☐ Collection scope names every layer, per custodian, in writing
- ☐ Configuration exported: rules, forwards, delegates, permissions, with change history
- ☐ Export mechanics chosen for headers, properties and threading; specifications exhibited
- ☐ Recoverable-items enumerated around key dates; deletion chronology built where it matters
- ☐ Delegation read before authorship pleaded
- ☐ Authenticity work run on layers + counterparts + headers, never one copy alone
- ☐ Transport trace requested within its window
- ☐ Opponent's layer-and-hold position obtained in correspondence and diarised

Red flags

- Disclosure described as "the mailbox" with no layer statement.
- A hold placed suspiciously late, or released mid-dispute.
- Gaps that align with an archive policy nobody has disclosed.
- A decisive email existing in exactly one place.
- Forwarding rules to personal accounts in the configuration export.
- "We can't recover deleted mail" from a tenant under hold.
- Delegated mailboxes pleaded as single-author.

When to involve a digital forensic expert

At scoping, to write the stratigraphic specification and secure the four facts; at recovery, where the layer walk and its enumeration are laboratory work; at authenticity, always: Example 3's method is the standard this series holds email exhibits to, and single-copy findings do not meet it; and at the who-read-what and delegation disputes where mailbox records meet the audit layer. Reported under CPR Part 35, with the layers, windows and licence bounds stated: and through **e-discovery.uk**, threaded and deduplicated into review so the correspondence reads as conversations, not fragments.

§ 1 3 · COMMON QUESTIONS

Frequently asked questions

The custodian deleted emails months ago. Can they be recovered?

Run the four facts: if a hold predated the deletion, yes: the Purges/Versions layers kept them, dated; if no hold and the recoverable window has passed, not from the mailbox: and the architecture paths (device caches, counterpart tenants, backups, journals) take over, frequently successfully. The answer is never "probably": it is an hour's tenant lookup, and the honest report states which case applies and why.

Can a user or admin permanently destroy mail on a held mailbox?

Users, no: every user-level deletion route ends in a preserved layer while the hold stands, and the attempts are enumerable (Example 1). Admins can release holds and alter retention: actions that are themselves audit-logged and, mid-dispute, approximately confessional. The practical assurance: a hold placed early and left alone converts the mailbox into an append-only record: which is why the week-one hold is this block's recurring hero.

How do we prove the other side received and read an email?

In layers of strength: transport trace (delivery to their tenant: strong, window-bound); their mailbox's collected copy (existence at collection); audit read-events where licensing and configuration recorded them (guide 57: genuine read evidence, not universal); and conduct corroboration (replies, forwards, actions consistent with knowledge). Read receipts are weak and refusable. The report stacks what exists and names each layer's weight: "delivered, present, and acted upon" usually suffices where "read at 14:32" is unavailable.

What about the OST file on the laptop: is it worth collecting?

Frequently decisive: the OST is the mailbox as last synced: which means it can preserve what tenant-side deletion later removed, capture the pre-purge state of a mailbox whose hold came late, and survive the account's deletion entirely. Guides 41 and 48's imaging disciplines apply; the comparison of OST contra tenant is itself a deletion chronology. In leaver and spoliation matters, the device cache is standard scope, not an afterthought.

Are auto-forwarding rules really that significant?

Repeatedly: a rule forwarding to a personal account is a continuous exfiltration channel with a creation date, an author context and (in audit) a paper trail: and its discovery reframes matters instantly ("what left, since when, and where did it go" becomes computable). The configuration export costs minutes; check it in every departure, leak and fraud matter before anything else is theorised.

Our exhibit email is challenged as fabricated. What should we commission?

The Example 3 workup, from your side: full item with headers and properties from every layer of your mailbox; counterpart collection or requests at the other end; transport and authentication analysis; and layer-coherence review (does the item sit in the stores a genuine message of its date would occupy, including hold structures?). Genuine messages pass this convergently and fast; the report then carries CPR 35 weight the bare exhibit never could. Commission it before the hearing, not after the insinuation lands.

§ 14 · REFERENCE

Glossary

Delegation
Send-as / send-on-behalf and access rights; the other hands in a mailbox.

Four facts
Hold status, hold date, window length, archive policy: the mailbox's governing constants.

Headers
Server-written transport and authentication record; the message's passport.

Hold (litigation hold)
Preservation converting deletion routes into hidden retention layers.

Message trace
Transport-log delivery record; window-bound facts without content.

Online archive
The policy-moved second mailbox custodians forget.

OST cache
Device-side synced copy; survives tenant-side deletion.

Purges / Versions
Hold layers preserving purged items and pre-edit states.

Recoverable items
The staged post-deletion stores beneath Deleted Items.

Stratigraphic scope
Collection specified layer by layer; this guide's habit.

Sources and authoritative references

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

- cflab.uk, digital forensics services (Exchange Online collection, deleted-mail recovery, header and authenticity 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 (stratigraphic mailbox collection; threading and deduplication 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
- Microsoft documentation: recoverable items and holds: learn.microsoft.com, recoverable items · learn.microsoft.com, Purview eDiscovery
- PD 57AD and CPR Part 35: justice.gov.uk, PD 57AD · justice.gov.uk, Part 35
- ICO, UK GDPR guidance: ico.org.uk

§ HOW A SPECIALIST LABORATORY CAN ASSIST

Working with Computer Forensics Lab

Mailbox matters run through the laboratory end to end: the four facts secured in day-one tenant lookups; stratigraphic collection specifications and their execution by the route guide 52 selects; recoverable-items enumeration and deletion chronologies (Example 1's engine); archive-gap arithmetic before anyone alleges concealment (Example 2's rescue); configuration exhibits that surface the forwarding rules and delegation realities; and the full authenticity workup: layers, counterparts, headers, transport: that Example 3 models, reported under CPR Part 35. Device-side OST work joins tenant-side collection under one custody record, and through **e-discovery.uk** the corpus is threaded and deduplicated so review reads conversations. Conflicts checks, confidential scoping discussions and fixed-fee estimates are routine; if a departed custodian's grace clock is running, that lookup comes first.

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