GLOSSARY FOR DATA INVESTIGATORS

Term	*Category*	*Definition*
Forensic Image	Collection	A bit-for-bit, read-only copy of a hard drive that captures "deleted" data that a standard "Copy-Paste" would miss.
Write Blocker	Collection	Hardware that prevents any data from being written to a device during investigation, ensuring the evidence remains pristine.
Metadata	eDisclosure	"Data about data." Includes hidden info like who created a file, when it was last modified, and total editing time.
ESI	eDisclosure	The broad term for all digital evidence (emails, Slack, PDFs, etc.).
TAR	eDisclosure	Using AI/Machine Learning to code documents for relevance in massive datasets.
Native Format	eDisclosure	The original file type (e.g., .XLSX). Reviewing in native format is often required to see hidden formulas or metadata.
Load File	eDisclosure	A file used to import data into a review platform, linking the text, images, and metadata together.
Processing	eDisclosure	The stage where raw ESI is converted into a structured format for legal review.
Artefact	Forensics	A digital "breadcrumb" left by the OS or an app (e.g., a shortcut file or a registry key) that proves user activity.
LNK Files	Forensics	Windows "Shortcut" files that prove a file was opened, even if the file itself has since been deleted or moved.
Shellbags	Forensics	Data that tracks which folders a user viewed—essential for proving an employee was "browsing" sensitive directories.
Unallocated Space	Forensics	The "hidden" area of a disk where deleted files livebefore they are overwritten. High-value for recovering "erased" logs.

(Continued)

Term	Category	Definition
Registry Analysis	Forensics	Examining the "brain" of the Windows OS to see what software was installed or what USBs were plugged in.
Ephemeral Data	Forensics	Short-lived data like RAM or self-destructing messages (Signal/WhatsApp) that requires immediate preservation.
Jump Lists	Forensics	Windows features that show "Recently Opened" files per application—proving intent and recent access.
Slack Space	Forensics	The unused space in a file cluster that often contains fragments of old, deleted data.
Hash Value	Integrity	A digital fingerprint (MD5/SHA1). If a single byte in a file changes, the hash changes, proving tampering.
Chain of Custody	Legal	The chronological record of who handled the evidence, ensuring it hasn't been altered from collection to court.
Custodian	Legal	The individual who had physical or logical control over the data (e.g., the departing employee).
Preservation Letter	Legal	A formal notice to an employee or third party to stop all data deletion/rotation because of pending litigation.
De-duplication	Processing	Removing exact copies of the same email or file from a dataset to reduce review time and costs.
Culling	Processing	Using search terms or date ranges to reduce the volume of data before legal review begins.
OCR	Processing	Optical Character Recognition. Turning pictures of text (scans/screenshots) into searchable data.
Exfiltration	Security	The unauthorized transfer of data from a company network to an external location.
Data Mapping	Strategy	Identifying where all an employee's data lives (Laptop, Phone, Cloud, CRM) before starting a collection.