
LITIGATION SUPPORT AUDIT GUIDE

The 5-Question Pre-Litigation Records Audit

What Teams Need to Know Before Starting Historical Document Collections

Laura Kirbyson

Historical Research & Litigation Support Consultant | April 2026

Historical litigation files have complex records management challenges. Document collections for these files serve many people for many purposes: litigation counsel, expert witnesses, researchers building chronologies, and staff preparing productions for example. Every collection must be responsive to the issues identified in the Statements of Claim or Defence — and it must remain so as those issues evolve over what is often a years-long file.

The consequences of getting this wrong are not abstract. Blown budgets, missed records, inadmissible documents, privilege breaches, and embarrassing gaps at discovery are all real risks — and most are mitigated with the right structure in place before work begins.

After more than twenty years working on historical and fiduciary litigation files, I have regularly used the same core pillars to successfully manage the life cycle of these document collections. What follows are the questions that lead to an organized five-question audit framework that can be applied at the outset of any historical document collection project.

If you cannot answer these questions, you likely need a better structure and process for managing your data.

Q1

Have you identified every repository where potentially relevant records could exist?

Why This Question Matters

A comprehensive collection is the foundation of everything else. If you have not mapped potential sources before you begin collecting, you cannot build a defensible budget, a realistic timeline, or a credible research plan and, at worst, your collection could be hit or miss. More

importantly, if opposing counsel or an expert witness produces a document or set of documents that you do not have, the reliability of your entire collection comes into question.

What to Look For

- A documented research plan identifying repositories by name, review priority, and rationale
- Coverage of records in the power and possession of your own client (classification and retention systems are frequently overlooked)
- Assessment of third-party and restricted-access collections — archives with ATIP/FOIP requirements, records requiring band or community consent, language-specific collections, or materials held in remote locations
- Identification of access conditions: location, hours, privacy restrictions, language (French, English, or other), and legibility constraints
- Gaps explicitly acknowledged — if areas of uncertainty exist, they should be named, not assumed away

RED FLAG

No written record of where you planned to look, who would review it, or the priorities assigned to different repositories or collections based on the likelihood of relevant content.

Q2

Can you demonstrate where you looked, what you found, and why some records were not collected?

Why This Question Matters

The litigation team must be able to demonstrate at any point in the file which records were identified and reviewed, which were not collected and why, and what the expected yield from high-priority repositories was, as well as the reasons lower priority files were not reviewed. Historical files can stay open for years. Issues change. New records surface. The team needs a living record of its search methodology — not just a snapshot.

What to Look For

- Records in the power and possession of your client, including active and archived holdings
- The organization's classification and retention systems — both current and as they existed during the relevant period — to establish what should exist, what may have been lawfully destroyed, and whether any gaps require explanation
- A files database built from the research plan, with each source traceable back to its repository, collection, fonds, or series, as well as who reviewed it and when
- Confirmation that reviewers could read historical handwriting well enough to recognize potentially relevant documents

- Confirmation that reviewers were given sufficient context to recognize relevance — issue summaries, key names or places, timelines
- A process for revisiting previously reviewed files if issues change over the life of the file
- Documentation of review decisions: what was collected, what was excluded, and the rationale
- Notation of whether each source will be shared with other parties to the litigation
- A mechanism for tracking ongoing collection management as new records surface

RED FLAG

High-priority repository reviews that did not yield anticipated results, or collection numbers that are unexpectedly low — both invite questions about what is missing and why.

Q3

Can every document in your collection be traced back to its original source?

Why This Question Matters

Defensible provenance is the evidentiary bedrock of a historical collection. Counsel may ask: can you prove where this document came from and that it has not been altered? Source tracing also becomes critical when the same document appears in multiple repositories — sometimes with annotations, different dates, or missing pages. The team needs a framework for identifying the authoritative version and managing variants in production, including whether duplicates or variants will be produced.

What to Look For

- A structured files database that links each document to its source file, repository, collection, and retrieval record
- A defined approach to duplicate and variant documents: which version is authoritative, how variants are flagged, and whether the team needs to confirm that all recipients of a document actually received it
- Confirmation that data exchange formats are compatible with opposing counsel's platform and any court-mandated production formats — format incompatibility is a common and preventable problem
- Chain of custody documentation sufficient to support admissibility arguments if challenged

RED FLAG

Documents in the collection with no clear source notation, or multiple versions of the same record with no indication of which is authoritative.

Q4

Is your coding accurate, consistent, and searchable enough to find what you need — and avoid producing what you should not?

Why This Question Matters

Coding is the infrastructure of your collection. If coding is inconsistent — names entered in different formats, dates handled without a standard, source fields incomplete — your search results will be unreliable. Unreliable searches create two problems: you miss what you are looking for, and you risk accidentally producing privileged or protected materials.

What to Look For

- A coding manual that defines standards for all the fields in use — name format, date format (including how to handle unknown dates and partial dates), titles, document types and source format, and special characters that could hamper data exchange
- A clear protocol for OCR handling — particularly how your platform treats handwritten documents, which are common in historical collections and which may not work well
- A privilege review protocol: historical collections, especially government records, regularly surface documents that may be solicitor-client privileged, cabinet-confidential, or subject to Crown privilege. Someone must be responsible for this, and team decisions must be documented. Note that a legal document from the 1870s available in a public repository is unlikely to retain privilege — but that determination must still be made consciously, not by default
- Regular tracking and coding audits: Are all field contents entered consistently across the team and adhering to the coding manual? Are there patterns of missing data? Did privilege/produciability review take place?

RED FLAG

Name or date inconsistencies discovered during audit, or a production that inadvertently included documents that should have been withheld.

Q5

Are your images legible, your data exportable, and your collection structured to remain reliable as the file evolves?

Why This Question Matters

A collection that cannot be navigated — by counsel, by researchers, or by expert witnesses — is a liability. Image quality, data portability, and long-term collection management are not

afterthoughts. They determine whether the work you have done holds up under scrutiny, including on the stand.

What to Look For

- A quality review process for images: Are all documents legible? Where they are not, have better copies been sought, or has an unofficial or official transcription been arranged?
- Consideration of whether expert witnesses — historians, anthropologists, others — can navigate the collection independently and defend their report or analysis methodology if challenged
- Data export readiness: Can the coded data be exported cleanly into other platforms? Are special characters, field formats, or encoding choices that could create problems on exchange identified in advance?
- A process for keeping the collection current: as new records surface, as issues change, or as file priorities shift, what is the mechanism for updating and re-reviewing? Is there a mechanism, preferably in the database, to record which documents have been processed and which remain outstanding?

RED FLAG

Images that are illegible or unreadable in the database, or a collection structure that an expert witness cannot independently navigate and defend. Not knowing from within the database which records have been produced.

The Reliable Structure Behind Every Answer

Each of the five questions above corresponds to a structural element of a well-managed historical document collection. Below is a summary of those elements and what they must contain.

Research Plan

Provides the documented foundation for Questions 1 and 2. Must include:

- Description of the litigation file: parties, timelines, key issues
- Key people, places, topics, and events relevant to the claims
- Chronology of identified events
- Repository list with descriptions of relevant fonds, collections, and series
- Identify the organizations' classification and retention systems and use it to understand file blocks, which records are unlikely to have been retained, and which may have been moved to off-site storage/archives
- Review priority assigned to each repository or target collection
- Access requirements: location, hours, language, privacy or consent conditions
- Staffing plan: who will review, how many reviewers, timeline

- Known gaps or areas of uncertainty are explicitly named

Files Database

Supports Questions 2 and 3. Built from the research plan. Must include fields that allow tracing each document to its original source, and notation of whether each source will be produced to other parties (consider FOIP and Privilege). Records of results from file review will be recorded in here including when the review took place, who performed it and whether or not anything potentially relevant was identified and collected.

File Review Protocol

Supports Questions 2 and 4. Defines the processes by which reviewers assess and collect records and includes a mechanism to flag whether previously reviewed files need to be revisited if issues change. Details are recorded in the Files database.

Document Review, Coding, and Privilege Protocol

Supports Question 4. A coding manual and privilege review workflow are both required. The coding manual must address the formats for all fields in the database that are used in the Documents database. This will include at least: dates (including unknown and partial dates), names, document types, document titles, source format, and data exchange constraints (special characters). The privilege protocol must assign clear responsibility and document decisions.

Quality and Longevity Review

Supports Question 5. Covers image legibility standards, transcription processes where needed, expert witness navigation requirements, data export readiness, and a process for ongoing collection management over the life of the file.

A Note on Working at Scale

The framework above applies regardless of collection size — but the stakes of getting it wrong scale sharply with volume. In my experience with large historical files, the disciplines described here are not optional. They are what allow a team of reviewers to produce consistent, defensible, and usable results under litigation conditions.

No collection is perfect at the outset. The goal of this audit is not to achieve perfection — it is to surface problems early, when they are still manageable, rather than at production or at trial.

Laura Kirbyson