

TECHNICAL BRIEF

Enterprise Data Foundation Brief

A practical framework for understanding the data route before a law firm evaluates or deploys an AI-enabled workflow.

This brief is an operating checklist, not a security certification or legal opinion. The firm and its advisers remain responsible for regulatory and contractual conclusions.

CLICKOAI LTD | clickoai.com/resources

Version 2 | 15 August 2026

1. The deployment paradox

A capable interface cannot repair an unknown or uncontrolled source system.

An AI product may look convincing in a demonstration because the sample is clean, the retrieval path is known and the expected answer is visible. A live matter environment is different. Documents may sit across inboxes, local folders and a document management system, with inconsistent names and permissions.

Before evaluating the product, map the operating route:

- where the authoritative document lives and who owns it
- how the correct version is identified
- which fields support retrieval and filtering
- which permissions follow the document into another service
- how a missing, duplicate or conflicting record is exposed
- how the reviewer reaches the source behind a material output

If these answers are unavailable, the first project is a data and workflow project. Adding a model may hide the weakness behind fluent output rather than resolve it.

2. The metadata imperative

Use the smallest schema that supports retrieval, review and accountability.

Metadata is useful when it answers a real operating question. More fields are not automatically better. Each field needs a source, owner, permitted values and a correction route.

Field family	Purpose	Control question
Matter and document identifiers	Bind output to the right source	Can two records be confused or overwritten?
Document type and version	Select the right workflow and source	Who decides the label and how is the latest version proved?
Party and role	Support bounded extraction and routing	Are names normalised without changing legal identity?
Dates and status	Control deadlines and state changes	Is the value extracted, calculated or entered by a person?
Source location	Let the reviewer verify output	Can the reviewer open the exact supporting page or clause?
Reviewer and decision	Retain accountability	Who accepted, corrected, escalated or rejected the output?

Create a correction log before expanding the schema. Repeated correction categories show which fields and instructions deserve investment.

3. Build, buy and integration trade-offs

Compare operating obligations, not feature lists.

Question	Existing product	Custom pipeline
Time to first controlled test	May be shorter when connectors and terms fit	Usually longer because the data route and controls must be built
Change and maintenance	Supplier controls release timing	The firm owns monitoring, fixes and model or service changes
Data route	Defined by supplier contracts, configuration and subprocessors	Defined by the firm's architecture and every selected provider
Verification	Depends on available source links, logs and review features	Must be designed and tested explicitly
Exit	Contractual export, deletion and transition terms	Code, credentials, data stores and operational ownership must transfer

For a smaller independent firm, a custom build also creates an ongoing software operation. Treat that proposal as a controlled pilot with a higher evidence bar, not as a shortcut around supplier evaluation.

4. Security, compliance and operating ownership

Ask for evidence that can be retained, tested and revisited.

Supplier labels and certifications may be relevant evidence, but they do not answer the workflow-specific questions alone. Record the exact service scope, date and contractual commitment.

- Data classes: what may enter, what is prohibited and who approves an exception.
- Training use: whether prompts, uploads, outputs, logs or derived data improve any model or shared service.
- Retention and deletion: what is kept, for how long, including backups, logs and derived stores.
- Subprocessors and location: who touches the data, where, and how changes are notified.
- Access and incident handling: who can access records, which logs exist and who owns escalation.
- Human verification: what must be checked, against which source, before an output matters.
- Exit: export format, assistance, deletion evidence, credential closure and the owner of transition.

Name an owner for data, workflow, technical operation and final professional review. One person may hold more than one role, but an unnamed role is an unowned risk.

Next step

Run the organisation-level systems check at clickoai.com/systems-readiness, then use the Strategic Intake Canvas to record the evidence and gaps for one proposed workflow.