DATA CONFIDENTIALITY

Confidentiality is the first question. Answer it before anyone asks

The risk is rarely the model. It is staff using unapproved tools with deal data, and a fund that cannot show what was used, where, and on which terms.

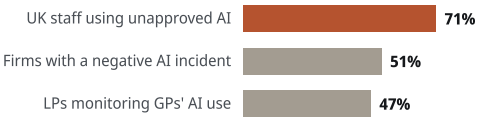
KEY MESSAGE

Remove the reason for shadow AI with better safe tools, clear rules and logs, and turn confidentiality into a trust asset.

3 things to know

1. **The leak is behavioural, not technical:** 71% of UK staff have used unapproved AI tools at work [3].
2. **Better safe tools remove the reason for shadow AI**; bans without full-power alternatives make it worse.
3. **LPs are watching:** 47% monitor GPs' AI use [4]. A clear policy is a trust asset.

EXHIBIT 1 **Signals of AI data risk, % of respondents (three surveys)**



Microsoft 2025 via Aurascape [3]; McKinsey 2025 [6]; PEI LP Perspectives 2026 via Blott [4].

3 in 20 days
Samsung leaks via ChatGPT, 2023 [1]

71%
UK staff used unapproved AI at work [3]

47%
of LPs monitor GPs' AI use [4]

4 vs 2
AI risks actively managed, 2025 vs 2022 [6]

Where data actually leaks

Not through model hacks. Through people pasting a CIM, a management presentation or a model into a personal account on a tool whose terms allow training or retention.

Samsung learned this in 2023: three incidents in twenty days involving source code, meeting notes and test data, followed by a ban, a prompt-length limit and an internal model [1]. For a fund, the same act can breach an NDA.

What enterprise tiers give you

Enterprise and API tiers of the main providers do not train on business data by default; zero-retention options exist; private cloud routes (AWS Bedrock, Azure, Google Vertex) keep data in your tenancy [2].

They add single sign-on, role-based access and audit logs. They remain third-party services, with sub-processors and retention settings that counsel should read [2].

The six controls a fund needs

1. An approved-tools list and no personal accounts for deal data. **2.** Enterprise contracts reviewed by counsel. **3.** Data classes (public, confidential, inside information) with a rule for each.

4. Pseudonymisation of company and people names before upload where possible. **5.** Logs of who used what, on which deal. **6.** A one-page policy you can send to LPs and sell-side advisers.

Sources on page 3.

CarveLines © 2026

Start outside-in

A large share of early work needs no confidential input: sector theses, competitor maps, pricing, customer sentiment, hiring signals. Running these on public data first lets a team learn the tools with zero exposure.

It also brings the team to the data room with sharper questions. Confidential material enters only once the six controls are in place.

EXHIBIT 2 Data classes and the rule for each (CarveLines view; dots = level of control)

DATA	RULE	CONTROL
Public	Any approved tool	●○○
Firm internal	Enterprise tier, logged	●●○
Deal confidential	Enterprise, no retention, named users	●●●
Personal data	Pseudonymise before upload	●●●
Inside information	Approved tools only, need-to-know	●●●

CarveLines framework, consistent with [25].

Regulation and LPs

Regulators focus on what firms say and do with AI: the SEC has pursued false AI claims by advisers (see No. 12), and the EU AI Act classifies uses by risk. Nearly half of LPs already monitor how their GPs use AI [4].

McKinsey finds organisations now actively manage four AI risks on average, against two in 2022 [6]. A fund with a clear answer turns a compliance question into a reason to trust it.

> "generative AI, deployed without privacy-preserving architecture, is a data exfiltration channel that operates through the users themselves."
>
> **Donovan Vanderbilt**, founder, Stealth Cloud Intelligence · 2026 [1]

5 hints: the average fund vs full power

	THE AVERAGE FUND	FULL POWER
1	Bans consumer tools	**Provides better enterprise tools first, then enforces**
2	Writes a long policy	**Publishes one page with data classes and rules**
3	Trusts vendor marketing	**Reads contracts: training, retention, sub-processors**
4	Uploads names and numbers	**Pseudonymises by default**
5	Cannot answer an LP	**Has logs and a one-page answer ready**

Watch out

A ban without an alternative. Bans push usage to phones and personal accounts, where nothing is logged. Provide a safe tool first, then enforce.

Questions for the partners' meeting

1. Which AI tools are approved, and is the list known to every team?
2. Which contracts exclude training and retention, in writing?
3. Who holds the logs, and could we answer an LP question tomorrow?
4. Which work could run on public data only?

Sources on page 3.

CarveLines © 2026

7 recipes: who, how, what you get

01 One-page policy

COMPLIANCE

- Define data classes: public, internal, confidential, personal, inside information
- Map approved tools to each class
- Name contacts and the incident route
- Send to all staff; keep ready for LPs

Output one-page policy
Time 2 days

02 Contract review

COUNSEL

- Check training, retention, sub-processors, location, logs, breach notice
- Request changes where needed
- Record terms in the register
- Review yearly

Output contract register
Time 1 week

03 Shadow AI check

COMPLIANCE + IT

- Run an anonymous survey on tools used
- Review network data for consumer AI domains
- Ask why people use them
- Close each gap with a better approved tool

Output gaps closed
Time 1 week

04 Pseudonymisation step

AI LEAD

- Build a skill or local tool that swaps names for codes
- Make it the default in deal workflows
- Keep the key locally
- Test on three documents

Output default pseudonymisation
Time 2 days

05 Public-data-first plan

DEAL TEAM

- Before the NDA, list what can be done on public data
- Run the outside-in workstream
- Arrive at the data room with sharper questions
- Add confidential data only in approved tools

Output outside-in workstream
Time 1 hour to plan

06 Log review

COMPLIANCE

- Look at logs monthly for unusual uploads
- Check access by deal
- Record incidents and fixes
- Report to partners quarterly

Output incident register
Time 1 hour a month

07 LP answer pack

COMPLIANCE

- Two pages: tools, contracts, controls, incidents
- Align with the policy
- Approve with partners
- Send with DDQ answers

Output ready for DDQs
Time 1 day

Monday morning

Ask compliance for three answers by Friday: which AI tools are approved, which contracts exclude training and retention, and who holds the logs. Any answer of “we don't know” is the first project.

Sources: [1] Stealth Cloud Intelligence, The Samsung-ChatGPT incident, 2026; [2] Layer3Labs, AI and confidential information compliance guide, 2026; [3] Microsoft study 2025, cited by Aurascape, Jun 2026; [4] PEI LP Perspectives 2026, cited by Blott, 2026; [5] Aurascape, Private equity AI security: regulatory and cyber risk, Jun 2026; [6] McKinsey, The state of AI in 2025, Nov 2025. Quotes verbatim from the sources cited. Figures as published; survey scopes and dates vary. Vendor-reported figures are labelled; [E] = CarveLines estimate; ‘CarveLines view’ = our interpretation.

CarveLines © 2026