

THE FULL OWNER'S GUIDE

Before You Sell Your Data

What every owner and CEO should know about licensing company records to AI labs.

Company records may have value to AI buyers, but private pricing and unfamiliar terms can make an offer hard to judge. This guide helps owners, CEOs and counsel understand the opportunity, the conditions and the questions to resolve.

What buyers may want

Records of skilled people doing real work: emails, drafts, revisions, tickets, quotes, drawings and decisions. Examples include insurance verification, claims handling, project coordination, dispatch, estimating and client communication. Review ordinary workflows as well as any commercially sensitive material.

Five factors to discuss with a buyer

- **Exclusivity:** the right to be the only buyer using the agreed records for specified purposes. Price that restriction explicitly.
- **History:** years of records, including work created before widespread AI use. More history is useful only if it fits the buyer's needs.
- **Visible decisions:** the steps and choices behind a result.
- **Volume and consistency:** usable records in well-kept systems.
- **Clean rights:** permission to license what is delivered.

What is still uncertain

There is no public price list. Offers depend on the buyer's current needs and may change quickly. Some datasets attract little interest; others attract much more.

What to do

- Get a price band from a real buyer before committing. Distinguish a supported offer from a broker's estimate.
- Treat exclusivity as a major negotiating point. Accept it only at a price that reflects what you give up.

What you own versus what you only hold

Having records on your servers does not automatically give you the right to license them.

Material	Who has a say	What to check
Your operational records	Your company	Ownership, privacy obligations and other restrictions.
Client files and communications	Clients, through contracts and NDAs	Limits on use, disclosure and reuse.
Employee messages and personnel records	Employees, through law and policy	What staff were told in handbooks and privacy notices.
Third-party documents and vendor content	Authors, vendors and licensors	Whether the original license permits reuse.

Check the software terms too

Export access does not settle whether you may use the export for AI training. Slack API restrictions illustrate why this matters. Review the current terms and permitted export method for every system involved.

What is still uncertain

Older client contracts may not mention AI training. Broad confidentiality terms can still restrict it. Do not assume removing identities also removes a client’s confidentiality rights.

What to do

- Have counsel review your main client agreements, NDAs and standard terms before sharing a sample.
- Check whether each system permits export for the intended purpose.
- Exclude material you cannot clear. A smaller, clearly permitted dataset can be easier to transact than a larger disputed one.

Protect the people in the records.

Applicable rules depend on the records, the parties, the use and the jurisdictions involved.

Removing names and emails is not enough by itself. Job titles, dates, locations and writing style can still identify someone. Use separate firms to de-identify the records and independently assess the result.

Records to flag	Rule to check	Practical implication
Patient or health-plan records	HIPAA	Where HIPAA applies, Safe Harbor and Expert Determination are de-identification routes. Counsel must confirm the permitted disclosure.
Substance-use treatment records	42 CFR Part 2	Additional confidentiality rules may apply. Consider excluding these from a first deal.
Customer financial information	GLBA	Restrictions may apply to sharing nonpublic personal information.
Student records	FERPA	Check consent requirements, valid exceptions or compliant de-identification.
People in the EU or UK	GDPR / UK GDPR	Pseudonymized information remains personal data.
California and other US states	State privacy laws	Check notices, opt-outs and sensitive-data requirements.

What is still uncertain

There is no single universal test that clears every email or chat dataset for every use. A new AI-training use also needs review against the original collection purpose and notices.

What to do

- Map the applicable rules before scoping the sale. Leave out higher-risk categories for a first deal.
- Use a de-identification provider and a separate independent reviewer.
- Honor existing privacy promises to employees and customers.

Sell rights a buyer can rely on.

A useful dataset needs a defensible license, not just a folder of files.

Why copyright matters

Whether AI training without permission is fair use remains disputed. Courts consider the facts of each case. A rights-cleared license can reduce uncertainty, but litigation does not establish the price of your records.

Two copyright cases to understand

Bartz v. Anthropic: the court distinguished training and lawfully purchased books from retaining pirated books in a central library. A \$1.5 billion settlement received final approval in July 2026, with approximately \$3,000 per covered book.

Thomson Reuters v. ROSS: in September 2026, the Third Circuit found that using Westlaw headnotes for a competing legal-search product was not fair use. It considered harm to existing services and a developing AI-training licensing market.

What is still uncertain

The ROSS ruling did not resolve the wider questions about generative AI. These cases do not mean every training use requires a license, or that every company archive is valuable.

What to do

- Identify your original work separately from material owned by clients or third parties.
- Check purchased reports, stock images and client deliverables before including them.
- Expect the buyer to ask you to promise that you can grant the license. Promise only what you have checked.

The owner's takeaway

Your offer should describe both the records and the rights being granted. Make the permissions as clear as the deliverables.

Agree the protections before delivery.

Deleting source files does not necessarily remove their influence from a trained model. Do not rely on being able to reverse training later.

Term	What it should establish
Exclusivity	Which records and uses you cannot license elsewhere, for how long, and the price paid for that restriction.
Use restrictions	The permitted training and evaluation uses, plus explicitly prohibited uses.
No resale or sublicensing	No onward transfer of the raw data unless expressly authorized. Distinguish this from an agreed sale of a training environment.
Derivative works	Who owns and may sell what is built from the records, including training environments.
Delivery and deletion	How data is transferred and stored, who can access it, and when raw copies must be destroyed.
Audit rights	How compliance with use restrictions and downstream reporting can be checked.
Representations and indemnities	Each party’s promises and who pays for specified claims or breaches.
Liability caps	The agreed limits on what each party could owe.

What is still uncertain

There is no single standard AI-data license. Definitions of derivative works and model outputs vary. Monitoring and audit rights matter because use restrictions can be difficult to verify.

What to do

- Use counsel experienced in data licensing.
- Do not deliver actual records before the license is signed and the agreed delivery conditions are met.
- Keep promises narrow, cap exposure and seek reciprocal buyer protections.

Check coverage before extraction.

Do not assume your existing policies cover a deliberate license of data to an AI buyer.

Policy or issue	Question for your broker
Cyber	Does it cover this agreed transfer, or only breaches and attacks? How do intentional-act provisions apply?
Technology errors and omissions	Does contractual-liability wording exclude promises or indemnities in the data license?
General liability	Are there AI exclusions or endorsements that narrow cover? Check changing 2026 forms and endorsements.
Directors and officers	Could the policy respond to relevant claims against management?
Specialist AI insurance	Is the policy designed for AI users or for a seller licensing data? Does it address privacy and rights claims?

Match the coverage to this deal

Specialist policies, including some placed through Lloyd’s, are not interchangeable. Coverage for a data seller’s privacy, copyright or contractual exposure may need to be arranged for the specific transaction.

What is still uncertain

Wording, exclusions and pricing are evolving. Different brokers may assess the same risk differently. A policy name is not confirmation that a claim is covered.

What to do

- Ask your broker to assess cyber, E&O and D&O cover against the actual license. Get the assessment in writing.
- Put the required coverage in place before extraction starts.
- Align the policy with the contract’s promises and indemnities. Identify exclusions, limits, deductibles and any uninsured exposure.

Insurance can support an indemnity; it does not replace the contract or automatically cover every obligation.

Negotiate both the first fee and the future upside.

If a buyer builds a product from your records, the commercial discussion should cover that downstream sale too.

1. The data-license fee

The buyer pays for permission to use the agreed records. Establish the price, payment milestones and what you receive after fees and costs.

2. A share of what is built

A builder may turn the records into a training environment and sell it to a frontier lab. Under Pewter's model, the seller keeps a written share of that downstream sale. It could exceed the first payment, but no amount or future sale is guaranteed.

What is still uncertain

Private pricing makes comparisons difficult. Bundled products can make it hard to attribute revenue to your records. Tax treatment depends on how the arrangement is structured.

What to do

- Ask whether the buyer will resell what it builds. Put your revenue share in writing.
- Require downstream reporting and audit rights covering bundles, renewals and products derived from your data. Specify that the buyer pays audit costs if an underpayment is found.
- Agree a downstream-sale deadline or minimum payment, and what happens if the buyer misses it.
- Ask your CPA to review the structure before you sign.

Make the calculation visible

Define the percentage, the revenue it applies to, any permitted deductions, payment dates and how bundled sales are allocated. A percentage without these terms is incomplete.

Get the right people involved early.

A deal can stall inside your company when a risk appears without a clear answer or owner.

Who	Their concern	What to prepare
General counsel	Privacy, client contracts and copyright	Rights review; independent de-identification report or certificate; contract and liability caps.
Board and owners	Price, reputation and future company sale	A buyer-backed price band, clear economics and change-of-control terms.
Security lead	Records leaving company systems	Controlled extraction, secure processing and monitored access.
Finance	Payment and tax	Payment terms, net proceeds and a tax review.
Employees and clients	How their information will be used	De-identification, honest communication and necessary exclusions or permissions.

What is still uncertain

There is no universal communications playbook. A lawful deal can still concern employees, clients or the public.

What to do

- Involve general counsel while the deal is being shaped, rather than only at final approval.
- Use the headline test: could you explain the deal publicly in two clear sentences and stand behind it?
- Name one internal owner from the first call to final payment.

A specialist report supports a decision.

It does not replace counsel’s review, the company’s approvals or the agreed allocation of responsibility.

Give each step a named owner.

The role, deliverable and responsibility should be explicit in the contracts. Hiring a specialist does not automatically transfer all legal liability.

Role	Purpose and responsibility
Licensing counsel	Review rights, regulation and contract terms; define the permitted license.
Insurance	Arrange coverage for agreed third-party risks and contractual exposure, within policy terms.
Extraction and packaging	Export only the agreed records; control security and scope.
Storage and processing	Maintain custody and securely prepare large or video datasets.
De-identification	Remove or transform identifiers and sensitive details.
Independent audit	Separately test de-identification and document remaining identification risk.
Permitted-use monitoring	Check use against the license’s restrictions.
Training-environment builder	Build the agreed environment and manage the downstream sale.
Revenue-share auditor	Check reported downstream revenue and amounts owed.
Final buyer	Set requirements, agree price and fund its purchase.

Specialist offerings and prices are still developing. Agree scope and fees early; without coordination, one provider can end up waiting on another.

What to do

- Name the responsible party for every step before records move.
- Agree the scope, responsibilities and delivery conditions before execution begins.
- Require full transparency: every party sees every other party’s share.

09 / RED FLAGS

Know when to slow down.

These warning signs mean you need clearer answers before proceeding.

Records requested before the agreement

Descriptions and volume estimates are fine for the first discussion. Under this model, actual records wait until the license is signed and the agreed delivery conditions are met.

A price with no buyer behind it

Ask whether the figure is a committed offer or an estimate. A quoted number alone is not secured demand.

Unclear buyer or intended use

Know who the counterparties are and what uses are permitted before signing. Identify the end buyer or define the approved route to one.

A hidden split

You should be able to see what each party earns and judge your own share.

No independent privacy check

The firm de-identifying the records must be separate from the firm auditing that work.

No downstream-sale plan

If a product built from your records will be resold, require your revenue share and audit rights in writing.

No owner for insurance or claims

Ask who responds to a privacy or rights claim, what the contract says and which coverage applies. Do not leave the answer implicit.

18 questions before you sign.

If an answer is missing, make that the next item to resolve.

Before you start

- ☐ 01. Which systems hold the records, how much is there and how many years are covered?
- ☐ 02. Which records are competitively important and should stay out?
- ☐ 03. Is a company sale likely in the next two years?
- ☐ 04. Who owns this deal internally from first call to final payment?

Rights and privacy

- ☐ 05. Do client contracts, NDAs or software terms limit this use?
- ☐ 06. Which privacy rules apply?
- ☐ 07. Who de-identifies the data, and who independently checks the work?
- ☐ 08. What have we promised employees and customers in privacy notices?

The deal

- ☐ 09. Is there a committed buyer with a written price band?
- ☐ 10. Is the license exclusive, and does the price reflect that restriction?
- ☐ 11. Which uses are allowed and which are prohibited?
- ☐ 12. What do we promise, what is our liability cap, and what does the buyer promise?
- ☐ 13. Does insurance respond, or is new coverage in place before extraction?

Getting paid

- ☐ 14. Are the payment amounts, dates and acceptance conditions clear?
- ☐ 15. Will the buyer resell what it builds, and is our share in writing?
- ☐ 16. Do we have downstream audit rights and an agreed deadline or minimum?

After delivery

- ☐ 17. How will we monitor compliance with permitted uses?
- ☐ 18. Can the license and revenue share transfer to a future company owner?

One partner to coordinate the whole deal.

For enterprise organizations with 1,000+ employees, Pewter designed a coordinated system to speed up enterprise data deals.

Pewter brings legal, insurance, security, privacy and engineering specialists into one coordinated process. Clear responsibilities and aligned handoffs are designed to reduce delays and move enterprise deals forward faster.

The conditions of a Pewter deal

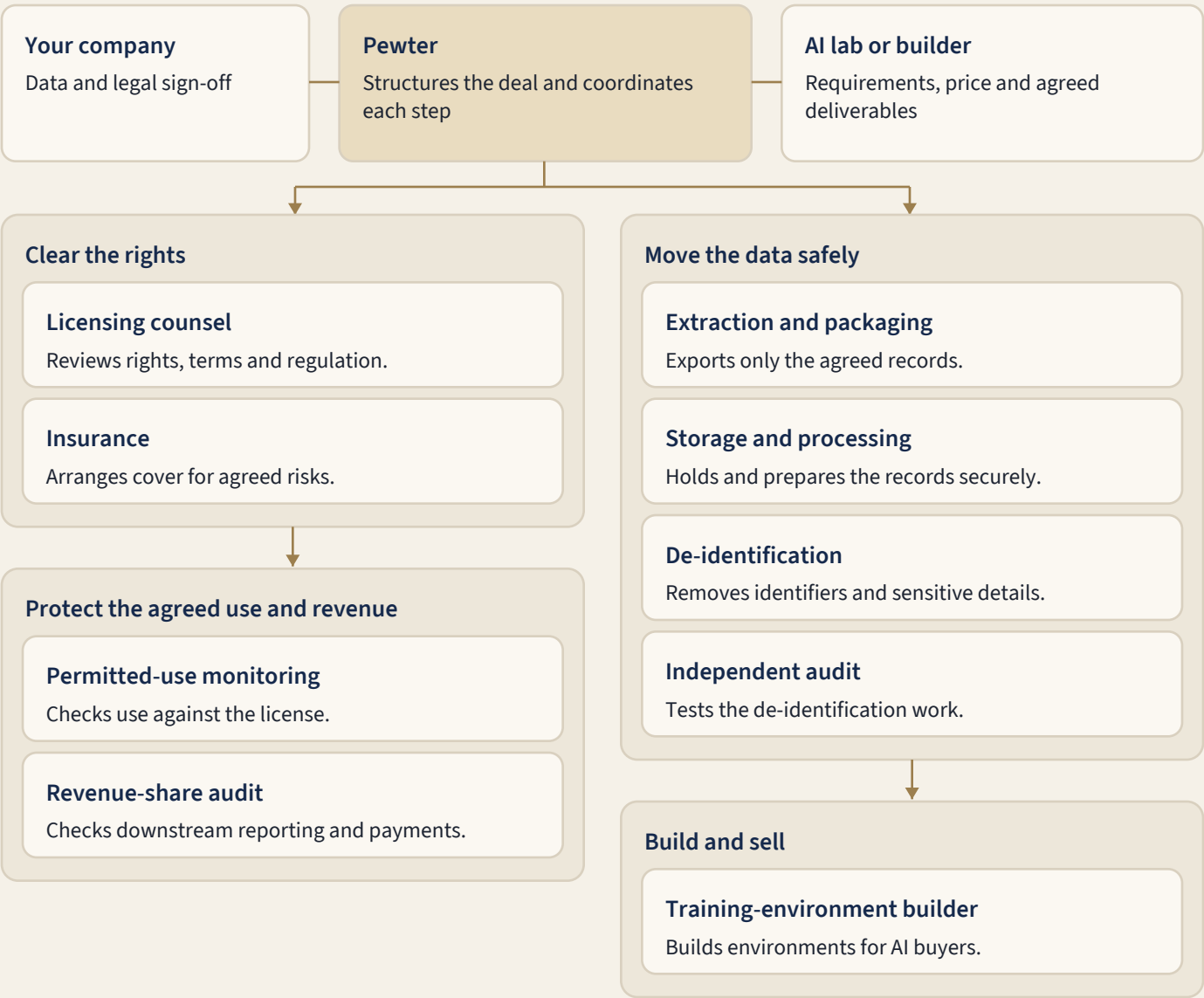
- **A committed buyer:** execution starts only after the buyer has committed and signed.
- **Exclusive, defined rights:** every license under this model is exclusive and limited to agreed uses. The contract defines its scope and duration.
- **Specialist accountability:** every specialist stands behind its own step, with responsibilities documented.
- **Visible economics:** every party sees every other party's share of the deal.
- **Downstream participation:** you retain a share of what is built from your data, supported by audit rights.

The timing target

The model targets closing the agreed data transaction in about five weeks after legal sign-off, with the agreed approvals and access ready. This is a planning target, not a guarantee. Environment construction and a later sale require their own schedule. For planning, Pewter compares this with six to twelve months for an uncoordinated deal; this is not a verified market benchmark.

Faster enterprise deals.

For organizations with 1,000+ employees. Pewter designed this system to coordinate specialists, reduce delays and speed up enterprise data deals.



One coordinating partner. Defined responsibilities. A system designed for faster enterprise execution.

Next step: Book a confidential offer review or email lawrence@pewterdata.com. Discuss likely value, obstacles and whether a deal makes sense. A review is not a guaranteed valuation or offer.

General information, not legal, tax or insurance advice.