



A FREE GUIDE FROM PEWTER

What is your company's data worth?

A plain-English inventory of the records AI labs pay for, and how to size each one yourself in an afternoon.

I. Why AI labs pay for ordinary business records

AI labs are teaching their models to do real work: answer a customer, quote a freight lane, reconcile an account, run a project. To train for that, they build practice environments that copy how real companies operate.

The raw material is the record of a real company doing its job, over years, across the tools it actually uses. That record is expensive for a lab to fake. Made-up data misses the messy parts: the exception, the back-and-forth, the fix. So labs pay companies that have the real thing, and they pay more for records that show complete work from start to finish.

WHAT HAS BEEN PAID

Deal	Amount	Source
Spirit Airlines' de-identified emails and chats, winning bid at a court auction (sale pending court approval)	\$10M	Notice of Auction Results, S.D.N.Y. Case 25-11897, Aug 2026
Michigan HVAC contractor, offer for its records	Up to \$150K	The Information, Jul 2026
cielo24, 13 years of Slack, Jira, email and Drive records	Six figures	Forbes, Apr 2026 ("hundreds of thousands of dollars")

REPORTED BY LICENSING FIRMS, OCTOBER 2026

INDUSTRIAL DISTRIBUTION \$1.2M Communications, product knowledge, order history, supplier records. Paid nine days from bid.	FREIGHT BROKERAGE · 100 PEOPLE \$650K Rate quotes, shipment documents, dispatch records, delivery exceptions and claims.	THIRD-PARTY LOGISTICS \$500K Dispatch, customer messages, SOPs and operating conversations. Paid seven days from bid.
ENGINEERING · 70 PEOPLE \$450K CAD drawings, calculations, technical reports, revisions and specifications.	ROOFING CONTRACTOR \$450K Operating record, anonymized before delivery. Perpetual, exclusive license.	ACCOUNTING · 60 PEOPLE \$350K Five years of workpapers, reconciliations, financial models and tax files.
GENERAL CONTRACTOR \$300K Operating record, anonymized before delivery. Perpetual, exclusive license.	ACCOUNTING GROUP \$280K Workpapers, client workflows, templates and internal communications. Paid five days from bid.	SOLAR INSTALLER \$275K Operating record, anonymized before delivery. Perpetual, exclusive license.

Figures are the publishers' own statements, not independently verified. Pewter did not advise on these deals. Results vary, and smaller firms often receive less. Sources on the last page.

Where would your company land?

The Pewter calculator gives you a range in about a minute. All it needs is your headcount, years in business and region.

[Run the calculator →](#)

II. What buyers actually price on

Two companies with the same amount of data can get very different offers. Five things drive the number.

1. The mix, not just the size

Total storage alone tells a buyer very little. 2 TB that is mostly video backups prices very differently from 2 TB of email, chat and operating records. You need the breakdown by system, not one big number.

2. Internal communications are the anchor

Slack, Microsoft Teams and internal email show how decisions actually get made. They tie everything else together. A deal without internal communications is a weak deal.

3. Depth

Years of history on the same systems. Five years of one CRM is worth more than five CRMs with a year each.

4. Thoroughness

Records that connect. The email about the order, the order in your ERP, the invoice in accounting, the ticket when something went wrong. Complete threads beat scattered fragments.

5. Accuracy

Buyers bid on the numbers you give them. If those numbers fall apart when they look closer, the bid gets pulled. Measure, don't guess. The rest of this guide shows you how.

Know your headcount and years in business?

That's enough for a first range. Your inventory then turns the range into a real number.

[Get your range →](#)

III. Two ways to get your numbers

Option A: admin panels (fastest)

Almost every business system has an admin screen that shows totals: storage used, number of users, number of records. Whoever holds admin access opens it and writes the numbers down. For most companies this takes a few minutes per system. The next section shows where the numbers live in the most common tools.

Option B: your own AI agent (for data spread across many tools)

If your records sit in a dozen places, you can have an AI agent you control walk your systems and total them up. Section V has the setup rules and a prompt you can copy.

THE ONE RULE FOR THIS STAGE

You only need counts. Don't export, copy, download or send any data while you size it. Nothing leaves your systems until there is a signed agreement and the data has been de-identified.

Nobody, Pewter included, needs your passwords or access to your systems to give you an offer.

IV. Where the numbers live, system by system

Six categories, in the order buyers care about them. For each one: what counts, where to find the number, and what to write down on the worksheet in Section VI.

CATEGORY 1 OF 6 • THE ANCHOR

Internal communications

Company email, Slack and Microsoft Teams. This is where decisions, handoffs and fixes happen, so buyers weigh it most. A deal without internal communications is a weak deal.

WHAT COUNTS

Every company mailbox (including shared ones like sales@ and support@), every chat channel, private channels and direct messages, for as many years as you still have.

System	Where the number lives	What to write down
Gmail (Google Workspace) Super admin or storage admin	admin.google.com > Menu > Storage . Hover over the usage bar to see Gmail's share. Under "Users using the most storage," click View all users for per-person totals.	Gmail GB, number of users, the year you moved to Google
Outlook / Exchange (Microsoft 365) Global admin or Reports reader	Microsoft 365 admin center > Reports > Usage > Exchange > Mailbox usage . Shows mailbox count, storage, and per mailbox: item count, storage used and create date . Click Export for a spreadsheet you can total. Switch to "Shared" to include shared mailboxes.	Mailbox count, total GB, total items, oldest create date
Slack Workspace owner or admin	Admin > Workspace analytics (Enterprise: Organization settings > Analytics). Shows members, messages sent, files uploaded, and the split between public channels, private channels and DMs.	Members, messages per month, % in private channels and DMs, your plan, the year you started
Microsoft Teams Global admin, Teams admin or Reports reader	Microsoft 365 admin center > Reports > Usage > Microsoft Teams . The User activity tab shows channel and chat messages per person; Team usage shows activity per team. Pick the 180-day view.	Active users, messages over 180 days, number of teams, the year you started

Watch for: Slack and Teams report recent activity (up to 180 days), not lifetime totals. Write down the monthly rate and the year you started; that is what a buyer uses to estimate volume. On Slack's free plan only the last 90 days are visible and older messages may already be deleted, so check your plan. In Microsoft 365, archive mailboxes aren't in the storage chart: note whether you use them.

File storage

Google Drive, OneDrive, SharePoint, Dropbox and Box. Decks, documents, spreadsheets and PDFs show the finished work behind every conversation.

WHAT COUNTS

Personal drives, shared drives and team folders. Documents, spreadsheets, presentations and PDFs count most. Video, photos and system backups count least.

System	Where the number lives	What to write down
Google Drive Super admin or storage admin	Same Storage page as Gmail. Hover over the bar for Drive's share. Click View all shared drives under "Shared drives using the most storage" for team folders.	Drive GB, shared drive count and GB
OneDrive Global admin or Reports reader	Microsoft 365 admin center > Reports > Usage > OneDrive > Usage tab. Files and storage used per person.	Total files, total GB
SharePoint Global admin or SharePoint admin	Reports > Usage > SharePoint > Site usage tab: files and storage per site. For the company total: SharePoint admin center > Sites > Active sites (storage bar at top right).	Site count, total files, total GB
Dropbox (team plans) Team admin	Admin console > Team usage > Storage . Storage details shows usage over time, the largest team folders and the top members.	Total GB, largest folders and what's in them
Box Admin (reports need Business plan or above)	Admin Console > Account and Billing > Usage for total storage. Reports > User Engagement for files and folders per person.	Total GB, file count

Watch for: Size can mislead. A few camera or backup folders can make up most of your storage while adding little value. Sort by largest folders and estimate how much is video, images and backups versus documents. Box storage also counts old file versions.

Communications and files are the core of most deals

Get a first range now, then come back to finish the inventory.

[Run the calculator →](#)

Knowledge bases

Notion, Confluence, internal wikis and SOPs. They show how your company says the work should be done, which pairs with the record of how it was actually done.

WHAT COUNTS

Pages, SOPs, playbooks, training material and process docs. Current, maintained pages count more than abandoned ones.

System	Where the number lives	What to write down
Confluence (Cloud) Confluence or site admin	Settings > Confluence administration > Data management > Space reports > Create report , then download. Per space: creation date, content items, attachments and storage. Total storage: admin.atlassian.com > your site > Storage .	Space count, total pages, oldest creation date, GB
Notion Workspace owner	Enterprise plan: Sidebar > Settings > Analytics shows page activity and when pages were created. Other plans have no page counter: count teamspaces and top-level pages, or use the agent prompt in Section V.	Approximate page count, teamspaces, the year you started
Other wikis and SOP folders Whoever maintains them	Count the documents in your SOP, training and process folders, and note when each was last updated.	Number of SOPs, how many were updated in the last year

Operating systems

ERP, WMS, MRP, CRM, accounting, ticketing and project management. These hold the structured record of the work: every order, job, ticket and transaction, with dates.

WHAT COUNTS

Records by type (orders, shipments, work orders, deals, tickets, invoices, tasks) and how many years each system has been in use.

System	Where the number lives	What to write down
Salesforce Admin	Setup > Quick Find “Storage Usage” > Storage Usage . Data storage and file storage used, broken down by record type.	Data GB, file GB, top record types, go-live year
HubSpot Admin	Data Management > Data Model > Limits tab > Objects. Record count for contacts, companies, deals and tickets.	Records per object, the year you started
QuickBooks Online Admin	Reports > Standard reports > “Transaction List by Customer” (or by Vendor) > Report period All Dates > Run report.	Transactions, earliest transaction year
NetSuite Administrator	Setup > Company > Service Tier Dashboard . File cabinet size, and your service tier, which is sized on monthly transaction lines.	File cabinet GB, service tier, go-live year
Zendesk Admin	Explore > new report on the “Support - Tickets” dataset > metric Tickets > column Ticket created - Year . Tickets per year since you started.	Tickets per year, first year
Jira / Asana Admin	Jira: Filters > Search work items with an empty search, then click “1000+” for the exact total. Asana: profile photo > Admin console > Insights (paid plans) for task and project counts.	Issue or task count, project count
ERP, WMS, MRP and others Your admin or the vendor	Ask your admin or the vendor's support team for record counts by type (orders, shipments, work orders, purchase orders) and the go-live date.	Counts by record type, go-live year

Watch for: If you switched systems, note whether the old history came across or still sits in the old system. Both count, but a buyer needs to know where each year lives.

Halfway there

Years of operating records connected to your communications are what move an offer up. See your range.

[Get your estimate →](#)

Commerce and transactions

Shopify, order history, invoices, fulfillment records, vendor and supplier data. Transactions tie the conversations to real money and real outcomes.

WHAT COUNTS

Orders, invoices, shipments, returns and purchase orders, plus your supplier and vendor records.

System	Where the number lives	What to write down
Shopify Store owner or staff with reports access	Analytics > Reports > Orders > Orders over time , with the date range set to the life of the store. The Orders page also shows total orders for any date range.	Orders per year, years of history, SKU count
Invoicing and accounting Admin	Use the QuickBooks transaction list above, or your accounting system's invoice report by year.	Invoices per year
Fulfillment and suppliers Operations lead	From your WMS or ERP: shipments per year, active suppliers and purchase orders per year.	Shipments per year, supplier count, POs per year

Work product

Presentations, proposals, reports, client deliverables and design files. This is usually already inside your file storage; the point here is to call it out so a buyer sees it.

WHAT COUNTS

Proposals, decks, reports, deliverables, CAD and design files, templates.

System	Where the number lives	What to write down
Any file storage Whoever knows the folder structure	Search by file type (presentations, PDFs, spreadsheets, CAD) and by folders like Proposals, Clients and Projects. Estimate counts per year.	Rough count by type, years covered

Watch for: Client deliverables can be covered by client contracts. Flag anything that is client-confidential; it may need to be left out or handled separately.

V. Doing it with your own AI agent

An AI agent can save hours when your data is spread across many tools. It can also do damage if you hand it more access than it needs. Run it yourself, inside your own accounts, under these rules.

SECURITY RULES

- **You run it, in your own account.** Don't give an outside person or company access to your systems to do this for you.
- **Read-only access only.** If a tool asks for permission to edit, send or delete, say no.
- **Counts, not contents.** Point it at admin reports and storage totals. It should never open, read or summarize individual emails, messages or files.
- **Turn access off when you're done.** Disconnect the agent and revoke any tokens you created for it.
- **Keep the output to numbers.** If anything comes back with customer names or message text, delete it before you share the results.

PROMPT TO COPY

I want to size my company's business data. Help me take an inventory.

Rules:

- Use read-only access. Never create, edit, send or delete anything.
- Count only. Do not open, read, quote or summarize any email, message, document or record. Use admin reports and storage totals wherever they exist.
- If you can't get a number, say so and tell me which admin screen would show it.

For each system I've connected, report:

1. System name and category (communications, file storage, knowledge base, operating system, commerce, work product)
2. Number of users or accounts
3. Total storage used
4. Item counts (emails, messages, files, pages, records, tickets, orders), whichever apply
5. Earliest and latest date of activity
6. Rough mix by type: documents, spreadsheets, presentations, PDFs, images, video, audio, other

Return one table, one row per system, then a short list of anything you couldn't measure.

Check the agent's numbers against at least one admin panel. If they disagree, trust the admin panel and note the difference.

VI. Your inventory worksheet

Fill in one row per system. Rough is fine for the mix; the totals should come from an admin screen or your agent.

System (tool name)	Users	Earliest year	Volume (GB or count)	Mix (what's in it)	Who has admin access
Email:					
Chat (Slack / Teams):					
File storage:					
Knowledge base / wiki:					
CRM:					
ERP / WMS / MRP:					
Accounting:					
Help desk / ticketing:					
Project management:					
Orders / commerce:					
Other:					

BEFORE YOU SELL, CHECK

- ☐ We've used the same core systems for 3+ years, or we can tell a buyer when we switched.
- ☐ The company owns these accounts. They aren't held by a former vendor, an MSP or a parent company.
- ☐ We've checked client contracts and NDAs for limits on how our records can be used.
- ☐ We've flagged regulated data (patient records, payment card or bank data, student records). It needs extra review and is often left out.
- ☐ One person owns this inventory and can answer a buyer's follow-up questions.

Finished the worksheet?

Run the calculator and request your offer. You can book a call right after. Bring this worksheet and we'll turn it into a real number.

[Get your estimate →](#)

VII. What happens after the inventory

1. **You share the numbers.** The completed worksheet. Counts and totals only, no data.
2. **We confirm them with you.** A short call to check the mix, the years and anything that looks off.
3. **Sample and de-identification (larger deals).** A buyer may ask for a small sample first. It is de-identified before it leaves your hands (names, emails, phone and account numbers removed or replaced), then reviewed by the buyer or an independent third party.
4. **You get a written offer.** Price, license terms, delivery and payment timing, in writing, before you commit to anything.
5. **Delivery and payment.** The full dataset is de-identified and delivered under the agreement. Several published deals were paid within days of the bid.

VIII. How to compare offers

Talking to more than one buyer? Compare the terms, not just the headline number. Ask each one:

Ask	Why it matters
Is the number based on my actual inventory?	A high number from a two-minute form often drops once the buyer sees your real mix or a sample that doesn't match. An offer built on your measured inventory is the one that holds.
Perpetual or term license? Exclusive?	Term licenses (for example 24 months) generally pay less than perpetual ones. Exclusivity should pay more, because it stops you licensing the same data again.
Who de-identifies the data, and who pays for it?	De-identification is real work. Know whether it's on you, the buyer or the broker before you sign.
When am I paid, and how?	At signing, on delivery or in installments. Get dates in writing.
Who does the work on my side?	IT access, paperwork, buyer questions. Some buyers leave all of it to you. Pewter runs the whole sale with you as a single point of contact.

IX. Sharing this with a business owner?

If you advise business owners as an accountant, attorney, broker, consultant or IT provider, this guide is the easiest way to explain the opportunity. Good fits are US companies with 20+ people and 3+ years of records, whether they're operating, downsizing or winding down.

Pewter's partner program pays **10% of what the business is paid, up to \$100,000 per business**, once the deal closes. A \$650,000 deal earns the introducing partner \$65,000.

Introduce a business owner

Apply in two minutes and get your partner code.

pewterdata.com/partners →

Or start with your own company

Headcount, years in business, region. One minute.

Run the calculator →

SOURCES

1. Spirit Airlines auction: [Notice of Auction Results, S.D.N.Y. Case 25-11897](#) (sale pending court approval); also reported by [TIME, Aug 25 2026](#).
2. HVAC contractor: [The Information, Jul 2026](#).
3. cielo24: [Forbes, Apr 16 2026](#).
4. Licensing-firm deals: statements published by licensing firms active in the market, read October 2026. Not independently verified.
5. Admin click-paths, vendor help pages: [Google Workspace storage](#) · [Microsoft 365 usage reports](#), [mailbox usage](#), [Teams activity](#), [OneDrive](#), [SharePoint sites](#), [SharePoint storage](#) · [Slack analytics](#), [analytics data](#) · [Dropbox](#) · [Box usage](#), [user engagement](#) · [Notion analytics](#) · [Confluence space reports](#), [storage](#) · [Salesforce storage](#) · [HubSpot limits](#) · [NetSuite service tier](#) · [QuickBooks Online reports](#) · [Zendesk Explore](#) · [Jira](#) · [Asana](#) · [Shopify order analytics](#), [order reports](#).
6. Admin menus change. Paths checked against each vendor's help pages on October 14, 2026.