Clean Source Data

CLIENT GUIDE 1 OF 4

# Data Licensing FAQ

Straight answers to the questions business owners ask most about licensing their data to AI companies with Clean Source Data.


PREPARED BY
Clean Source Data

DOCUMENT
Data Licensing FAQ

EDITION
November 2026

CLEAN SOURCE DATA | CLIENT GUIDE 1 OF 4

# Data Licensing FAQ

How data licensing works, what qualifies, how you get paid, and how your data and your customers stay protected.

> **A NOTE ON AI BUYERS**
> Every AI company runs its own intake process, and tools, timelines and controls differ from buyer to buyer. This guide describes the standards Clean Source Data holds every buyer to. Before anything moves, your Clean Source project lead confirms the specific buyer's process and walks you through it.

## 01 The Basics

### What is data licensing?

It means turning records your business already creates into an asset that AI companies pay to use. You keep ownership. The buyer gets a defined license to use an approved, de-identified package of your data for purposes such as training, fine-tuning or testing AI models.

### Why do AI companies pay for business data?

AI models have already learned from most of the public internet. What they lack is real-world data showing how businesses actually run: how a call is handled, how a job moves from quote to invoice, how decisions get made. That everyday operational detail is scarce, and buyers pay for it.

### What does Clean Source Data do?

We take you from first call to final payment. We size and value your data, de-identify it on-site so sensitive information never leaves your office, package it for the buyers who value it most, bring you competing offers and manage the deal. You approve every step.

### Who will I work with?

Every client gets a dedicated Clean Source project lead. They are your single point of contact: they build your data inventory with you, review each buyer's requirements, run the on-site de-identification and make sure what is delivered matches what you approved.

## 02 Fit and Eligibility

### What kinds of data qualify?

Depending on scope, rights and sensitivity, a data package can include:

- Call recordings and transcripts, where the right consents are in place
- Email and chat history, and the files attached to them
- CRM, scheduling, dispatch, ticketing and billing records
- SOPs, checklists, reports, spreadsheets and internal documents
- Images, scans and files, with identifying details removed
- Industry-specific operational records unique to your business

### What makes data valuable?

Volume and quality together. Value tends to rise with years of history, consistent workflows, many people contributing, and records that show outcomes, not just activity. Data that is hard to find anywhere else is worth more.

### Is my business a good fit?

Established businesses that run on software are the strongest fit. There is no hard minimum size: smaller businesses can join a pooled dataset with similar companies in their industry, so the combined package earns a strong valuation and each business is paid its share.

### What rights do we need?

You must own the data or have the right to license it, including a plan to remove or transform personal information about individuals. We review rights and any existing contracts with you before anything is offered to a buyer.

## 03 Process and Timeline

### How does the process work?

1. **Discovery call.**
   We discuss fit, your systems, your workflows and potential scope.
2. **Agreement.**
   You sign the Clean Source Data agreement. There is no upfront fee.
3. **Data inventory.**
   With your project lead, you document sources, volume, history and exclusions. No raw data is shared.
4. **Valuation and packaging.**
   We price and package the data for the buyers most likely to value it.
5. **Buyer matching.**
   We present an anonymous profile to vetted buyers and bring you their offers.
6. **Terms.**
   You approve the buyer, price, scope, permitted uses and license terms.
7. **On-site de-identification.**
   We process the approved data on an offline device at your location. You review a sample before release.
8. **Buyer verification and payment.**
   Only the clean package is delivered. The buyer runs its own independent privacy review to confirm nothing sensitive remains, checks it against the agreed criteria, and pays.

### How long does it take?

It depends on the buyer, the size of your data and how quickly the inventory comes together. Smaller packages can close in weeks; larger or more sensitive ones take longer. Having storage figures ready for your discovery call is the biggest thing that speeds things up.

### Does a call commit us to anything?

No. The discovery call and first review only assess fit and scope. Nothing is offered, shared or licensed without your written approval.

## 04 Money

### How is the price determined?

Offers depend on the scope, size, type, history and uniqueness of the data, the permitted uses, and whether the license is exclusive. A buyer makes a final offer only after reviewing your data inventory. Because different buyers value different data, putting your package in front of several of them is how we find the best price.

### When do we get paid?

Most buyers pay after the delivered package is accepted. Review periods vary and are written into each license agreement. Your project lead will confirm the timeline before you sign.

### What does Clean Source Data charge?

Nothing upfront. We earn a success fee only when a deal closes and you are paid. The full fee schedule is set out in your agreement.

## 05 Ownership and Protection

### Do we give up ownership of our data?

No. You license the data; you do not sell it. You keep every right you hold.

### Can we license the same data to more than one buyer?

Often, yes. We aim for non-exclusive licenses so the same data can earn more than once. If a buyer asks for exclusivity, we explain what it costs you and negotiate or price it in.

### How is our data kept safe?

Early reviews use descriptions and counts, never raw records. De-identification happens inside your office on an offline device, to the HIPAA Safe Harbor standard, and only the clean package ever leaves. See guides 3 and 4 for details.

### Does anyone double-check the de-identification?

Yes. After we de-identify the data and you approve it, the AI company runs its own internal analysis to confirm nothing sensitive remains before accepting it. Many brokers let the buyer do the cleaning; with us it is cleaned on-site *and* verified by the buyer: two layers of protection.

**Questions?**

Your Clean Source project lead is your single point of contact.

info@cleansourcedata.com
cleansourcedata.com