A ONE-PAGE BRIEF FOR REVENUE OPERATIONS

Clean the CRM Before You Turn On the AI

Every AI feature in the revenue stack scores, personalizes, and forecasts from CRM records exactly as it finds them.

THE ASK

Two focused weeks of CRM cleanup before the next AI feature goes live, plus a monthly re-check. No new headcount, no platform swap. The alternative is paying for AI tooling that confidently reads records we already know are duplicated, blank, and out of date.

WHY THIS CAN'T WAIT

- AI tools read records **as given**, with no dedupe step of their own. Two records for one company means two scores and two forecast entries.
- Missing data is not flagged, it is **silently guessed**. A blank employee count reads as "small," so enterprise accounts land in the wrong sequence.
- The errors arrive **polished**: full sentences, confidence scores, first names. A bad report looks bad. Bad AI output looks great, so it goes unquestioned for weeks.

FOUR WAYS DIRTY DATA BREAKS AI

Duplicate accounts	Engagement and open deals double-count. One $60K deal on three records reads as $180K.
Missing firmographics	Models guess. Sparse enterprise records score like SMBs and land in low-touch nurture.
Stale contacts	Outreach personalizes to titles people left, citing the old role as its evidence.
Inconsistent picklists	"SaaS," "Saas," and "software" split one segment into five; the model finds no pattern.

"AI-READY DATA," DEFINED

- ✓ **Uniqueness.** One record per real company and person, surviving rebrands and domain changes.
- ✓ **Consistency.** Controlled values in every field a model reads: industry, geography, lifecycle stage, lead source.
- ✓ **Completeness where it counts.** The 5 to 8 fields that drive scoring are filled. If the vendor can't name them, that's an answer too.
- ✓ **Recency.** Records carry a last-verified date. A title from 2021 is a hypothesis, not evidence.
- ✓ **Detectable anomalies.** Impossible values are surfaced, not averaged in, before they bend the model.

WHAT ONE MISSED MATCH COSTS

One contact, **two records**: a pricing-guide download in March, then a webinar signup in August under a rebranded company name and new domain. Nothing connected them. Downstream:

- ✗ **Wrong score, twice.** 58 and 87 for the same person; combined, she scores in the 90s.
- ✗ **Wrong title, wrong segment.** One sequence used her pre-rebrand role; the other guessed her industry from a webinar topic.
- ✗ **Wrong forecast.** A phantom deal sat at 60% for two months, then was marked lost.

Four wrong outputs from one missed match. No error logged anywhere.

THE TWO-WEEK PLAN, IN ORDER

1. **Deduplicate first.** Merging later means enriching, scoring, and paying for the same company twice.
2. **Standardize the fields scoring reads.** The 5 to 8 decision-driving fields, on controlled lists.
3. **Fill the gaps that change decisions.** Employee count on ICP accounts, yes. Fax number, no.
4. **Flag impossible values.** Fix or quarantine before a model averages them in.
5. **Then connect the other systems.** Clean first, then sync, so every tool reads the same customer.
6. **Set a threshold, re-check monthly.** Data quality decays; define "good enough to score on" as a number.

PROVE IT IN 20 MINUTES

THE BASELINE EXERCISE

Export 1,000 records. For **account name, industry, and employee count**, count duplicates, blanks, and value variants. That number makes the case by itself.