TRUSTY HOUSE BUYERS

County Records Pipeline User manual

Daily foreclosure, probate, tax-delinquent and surplus-funds leads from official county records in AR, GA, TN, IN, OH, TX, FL and AL: how to browse them, download them, check them and send them to your marketing tools.

Live site: trusty-county-records-pipeline.pages.dev

Prepared by: Poornima Ramakrishnan

Edition: Proof of concept · November 2026

Contents

1. What this is, in one minute
2. A quick tour of the website
3. Finding leads
4. Reading a lead
5. Getting the daily CSV files
6. How the daily run works
7. When something goes wrong
8. Lawful sources only
9. Checking the results yourself
10. BatchLeads, TCPA, Reply Smart, SmartList, GoHighLevel, mail
11. Going live in your accounts
12. Adding a county (for a developer)
13. Security & ownership
14. Frequently asked questions
15. Glossary

HELP & USER MANUAL

Everything you need, in plain English

No technical background needed. Read it start to finish in about ten minutes, or jump to the part you need.

1. What this is, in one minute

Every morning at **7:15 AM Eastern**, this system visits official county sources, collects new **foreclosure, probate, tax-delinquent and surplus-funds** records, cleans them up, removes anything you've already received, and delivers only what is **new or changed**:

- **A daily CSV file**, in your sample format, ready for GoHighLevel or Excel.
- **GoHighLevel contacts**, created or updated with tags your cold callers can filter on.
- **Your marketing tools**: skip trace in BatchLeads, TCPA scrub, Reply Smart texts, SmartList emails and direct mail.
- **This website**, to search, check and download everything without any tools.

> **The golden rule:** we only collect official records that counties publish and allow automated tools to read. No captcha, login or paywall is ever bypassed. When a site says no, we find the county's own file instead, and we tell you.

2. A quick tour of the website

- Overview: today's numbers, the hottest leads, the next run time and a map of coverage.
- Leads: search and filter every lead; click one for the full story.
- Daily files: each day's CSV, ready to download.
- Runs: every daily run, what each source returned, and any alerts.
- Counties & sources: every platform and county reviewed, and why each source is lawful.
- Integrations: the exact requests waiting for your tools.
- Acceptance tests: the clean-run streak, duplicate checks and a random spot check.
- Data model and Roadmap & pricing: for planning and for your developer.

Use the sun / moon button at the top to switch between the light and dark look.

3. Finding leads

1. Open Leads.
2. Type a name, street, case or parcel number in the search box. Partial words work.
3. Narrow down with the filters: lead type, state, signal (e.g. “Absentee owner”), NEW or UPDATED, and the minimum score slider.
4. Sort by *Lead score* to see the most motivated sellers first, or by *Sale date* to see the most urgent.
5. Press **Export results** to download exactly what you see (up to 2,000 rows).

Every filter is saved in the page address, so you can bookmark a search or send it to a team member.

4. Reading a lead

Clicking a lead opens a panel with everything we know:

- **Details**: case or parcel number, owner, decedent, property and mailing address, amounts and dates.
- **People on the case**: owners, heirs, plaintiffs and attorneys.
- **History**: each change the county made, with the old and new value.
- **Source**: a link to the county page and a download of the exact file it came from.
- **Integrations**: what was sent to each of your tools, and when.

The lead score

Each lead starts with a base score by type (Pre-foreclosure 45, Probate 45, Tax-delinquent 40, Surplus 50), then points are added or taken away for each signal. The total is kept between 0 and 100. **70 or more is a hot lead.**

SIGNAL	POINTS	MEANING
absentee owner	+15	Owner mailing address differs from the property (landlord / out-of-area owner)
out of state owner	+8	Owner mailing address is in another state
non homestead	+10	Clerk classifies the foreclosure as non-homestead (rental / second home / vacant)
deceased owner	+15	Owner or record holder is deceased (estate / heirs)
sale soon	+12	Sale date within the next 30 days
sale scheduled	+5	A sale date is scheduled
small balance	+8	Small balance relative to typical values: owner likely holds equity
large surplus	+12	Surplus above $10,000 held for the former owner
hoa lien	+6	HOA / condo association lien foreclosure (often equity-rich owners)
entity owner	-6	Owner is a company (harder to reach, often an investor)
cancelled	-25	Sale cancelled (loan reinstated / bankruptcy / redeemed)
missing address	-15	No property address in the source record
homestead order	+10	Court order determining homestead (heirs inheriting the home)

5. Getting the daily CSV files

1. Open Daily files.
2. Each day shows how many leads were new and updated, by type.
3. Press **Download** next to the file you want: one file with everything, or one per lead type.
4. Press **Preview** to look at the first rows before downloading.

Files are named `trusty_leads_{date}_{scope}.csv` and open cleanly in Excel. Dates are MM/DD/YYYY. The column guide explains every column, and the column names can be changed to match your GoHighLevel import.

Each file has a SHA-256 fingerprint shown next to it. If the fingerprint matches, the file is exactly what the pipeline produced.

6. How the daily run works

1. **7:15 AM Eastern**: GitHub Actions starts the run (adjusted automatically for daylight saving).

2. **Fetch**: each source is downloaded politely, one request at a time, with a clear identity. The file is saved as evidence.
3. **Clean**: names are split, addresses tidied, dates and amounts checked. Rows that fail checks are set aside and counted, never silently dropped.
4. **Compare**: each lead is matched against what we already have. New → NEW. Changed → UPDATED (with what changed). Same → skipped.
5. **Deliver**: the day's CSV files are written, contacts go to GoHighLevel, and the integration queue is filled.
6. **Report**: a run summary is saved and shown on the Runs page, linked to the GitHub log.

A typical run takes under a minute per county.

7. When something goes wrong

You get an email, and a GitHub issue is opened, when:

- a source returns **zero records** on a day it normally has some;
- a parser **fails**, usually because the county changed its page or file layout;
- records still arrive but a key field such as owner, address or case number is **suddenly empty**;
- a source returns far **fewer records** than usual;
- a source has been **stale** for several days;
- a blocked platform **changes its robots.txt**.

Repeat problems update the same issue instead of flooding your inbox. Yesterday's data always stays in place, and the other counties keep running. Under the retainer, broken parsers are fixed within 48 hours, using the saved file to prove the fix.

8. Lawful sources only

Before a county is switched on, we check its robots.txt and terms, and look for the official bulk file behind any auction site. The decision is recorded on the Counties & sources page:

- **Automated**: official bulk files and published lists (e.g. Hillsborough Clerk, Chatham Tax Commissioner).
- **Automated, rate-limited**: public pages and documented JSON (e.g. SRI, Tyler property pages).
- **Licensed or request-based**: captcha, login or paid sites (e.g. Tyler Odyssey court search, GSCCCA). We use a bulk subscription or records request.

- **Not scraped**: sites whose robots.txt forbids bots (RealAuction, GovEase, Linebarger). We use the county's own data and a daily sentinel watches for changes.

9. Checking the results yourself

1. Open Acceptance tests.
2. The ring shows how many clean daily runs in a row (the target is 7). The duplicate checks should all show 0.
3. In **Spot check**, choose a lead type, type any word as the seed and press **Draw 20 records**.
4. Open each *County source* link and compare. **Download CSV** gives you a checklist with a "matches source" column.

The same seed always gives the same 20 records, so you and we can review the same sample together.

10. BatchLeads, TCPA, Reply Smart, SmartList, GoHighLevel and mail

After each run, every new or changed lead is queued for your tools in this order:

1. **BatchLeads / BatchData** finds phones and emails. Names, phones and emails are written back to the CSV and the GoHighLevel contact.
2. **TCPA Litigator List** checks every phone. Litigators and DNC numbers are removed.
3. **Reply Smart** enrolls clean numbers in a text campaign per lead type and state.
4. **SmartList** adds the lead to an email list per lead type and county.
5. **GoHighLevel** gets the contact with tags for your cold callers.
6. **Lob** (or REIPrintMail) sends a postcard or letter, after your approval if you prefer.

A text is never sent unless the TCPA check passed. Each request has a unique key, so nothing is ever sent twice, even if a day is re-run. In this demo every step is in **preview** mode: you can see the exact request on the Integrations page, but nothing is sent.

11. Going live in your accounts

On Day 0 we need four things from you. Each takes a few minutes:

1. **GitHub**: invite us to your repository (Settings → Collaborators → Add people).

2. **GoHighLevel**: create a Private Integration token (Settings → Private Integrations → Create, with Contacts read/write) and copy your Location ID (Settings → Business Profile).
3. **CSV sample**: send the file you use today, so the columns match exactly.
4. **County**: tell us which county you want first. We confirm the sources on Day 1.

Keys are added in your GitHub repository under *Settings → Secrets and variables → Actions*, never in the code or in chat. For the integration add-on, the same is done with your BatchLeads, TCPA Litigator List, Reply Smart, SmartList and mail keys. Each tool is switched on separately.

12. Adding a county (for a developer)

Designed so another developer can add a county in under a day. The README in the repository has the full commands.

1. **Check the source**: read the county's robots.txt and terms; find the bulk file or published list. Record the verdict.
2. **Add the feed**: one entry in `config/sources.yaml` (URL, schedule, method, compliance decision, expected minimum per day).
3. **Reuse or write an adapter**: on a known platform, reuse its adapter. Otherwise add a class in `src/trusty_records/adapters/` with two methods: *fetch* and *parse*.
4. **Save a sample file** in `tests/fixtures/` and add a parser test. Tests replay real county files, so a layout change is caught before it reaches you.
5. **Try it**: `trusty-records replay <feed_id> <file>`, then `trusty-records run --feed <feed_id> --memory`.
6. **Switch it on**: `trusty-records sync-reference`, then merge. The next morning's run includes it.

13. Security & ownership

- All code lives in your GitHub repository from Day 1. You own the code, data and outputs.
- Keys and passwords are stored only as encrypted secrets in GitHub and Cloudflare, never in code.
- The database is locked down: the website can only call a few read-only functions.
- Source files and CSVs are in private storage, served only through the website.
- No third-party licensed code. Open-source libraries only (Python, PostgreSQL).

14. Frequently asked questions

Why do only some counties show leads?

This is a working demo built before the paid test. Two counties (Hillsborough FL and Chatham GA) plus statewide Indiana sales are live. Every other county in the eight states is catalogued and researched, ready to switch on in batches of five.

Why don't foreclosure leads have phone numbers?

Courts don't publish phone numbers. Phones and emails come from your BatchLeads account in the skip-trace step, after which every number is scrubbed against the TCPA Litigator List.

Why is the mailing address sometimes empty?

Court filings usually list the property, not where the owner lives. The mailing address is filled from the county property records or skip trace. Leads with a different mailing address get the “absentee owner” signal.

Will the same lead show up as new twice?

No. A lead is NEW only on the first day it appears. If the county changes it later (for example the sale date moves), it comes back as UPDATED, showing exactly what changed.

Can you collect from RealAuction or GovEase?

Those sites forbid automated access in their robots.txt, so we don't. We take the same cases from the county Clerk's or Tax Commissioner's own files, which is where the auction sites get them anyway.

What if a county uses a captcha?

We never bypass it. We use the county's bulk-data subscription, a state data feed or a standing public-records request instead, and tell you the cost and delay before choosing.

Can it run more than once a day?

Yes. Anyone with access can press “Run workflow” in GitHub Actions. The scheduled run stays at 7:15 AM Eastern so county servers are contacted once a day.

What does it cost to run?

In this setup, nothing: GitHub Actions, Supabase and Cloudflare free tiers cover it. At 100+ counties we recommend Supabase Pro ($25/month) for more storage.

Can it run on our own server or Docker?

Yes. It is a normal Python package with one command, trusty-records run. GitHub Actions is the default because it is free and every run is logged.

15. Glossary

Lead	One record from a county: a foreclosure case, a probate estate, a tax-sale parcel or a surplus-funds claim.
Feed	One source we read from, e.g. “Hillsborough civil filings”.
Adapter	The small piece of code that reads one feed and turns it into clean leads.
Platform	Software many counties share (Tyler, SRI, RealAuction). One adapter can serve all its counties.
NEW / UPDATED	NEW: first time we've seen this lead. UPDATED: the county changed something since yesterday.
Lead score	0–100. How motivated the owner is likely to be, based on the signals below.
Evidence	A saved copy of the county file a lead came from, with a SHA-256 fingerprint proving it hasn't changed.
Sentinel	A daily check that a blocked platform still blocks bots. If its rules change, you get an alert.
robots.txt	A public file where a website says which automated tools may read it.

Skip trace	Looking up phone numbers and emails for a name and address (BatchLeads).
TCPA scrub	Removing phone numbers belonging to known litigators or on Do-Not-Call lists before texting or calling.
Outbox	The queue of requests waiting to go to your tools. Each is sent exactly once.
GitHub Actions	GitHub's free scheduler that runs the pipeline every morning and keeps a log of each run.
Supabase	The hosted PostgreSQL database and file storage holding the leads and evidence.

Questions not covered here? Use the feedback button on any page. It reaches us directly, and we reply within a few hours on US Eastern time.