

From Public Data to Pipeline: Automated FCC Lead Generation with Clay

Built by FastForward · Case Study · GTM Automation · Regulatory Tech

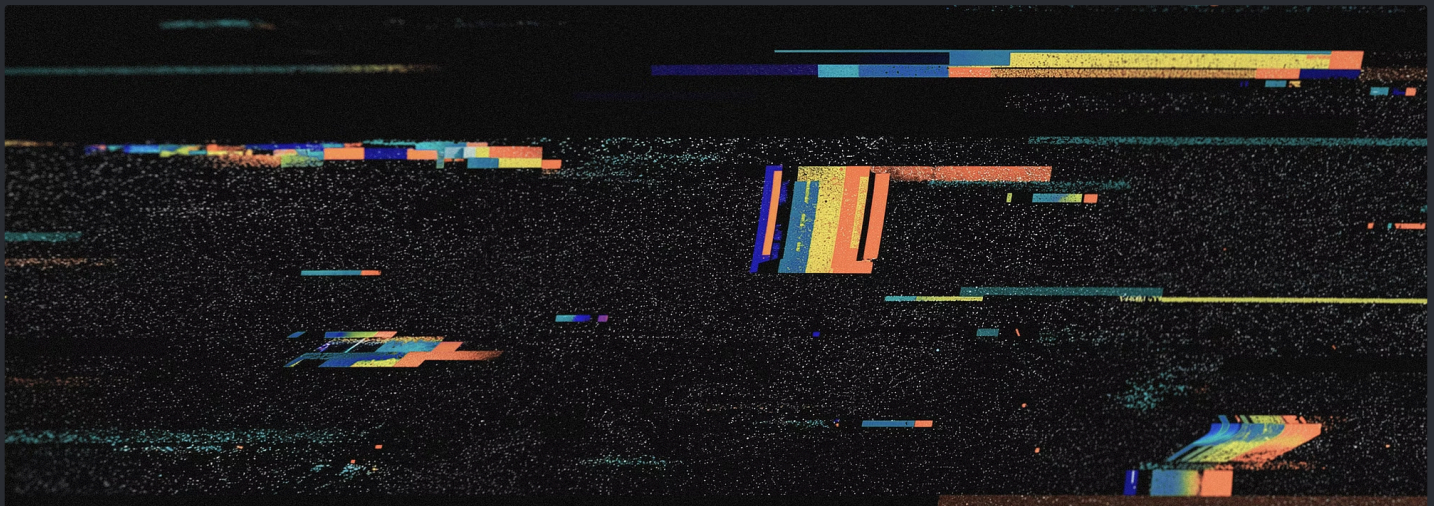
The Business Problem

Our client, Approve-IT helps manufacturers and importers navigate regulatory approvals across 250+ countries. Their BD team knew that companies actively filing FCC equipment authorization applications were among their most qualified prospects. These filings happen before products reach market, making them an early-stage, high-intent buying signal.

The challenge: the FCC's public database contains this data openly, but monitoring it manually isn't viable at scale. It was taking the Approve-IT 100's of hours per year to pull this data. Hundreds of new filings appear daily. Without a systematic and automated approach, the team had no way to consistently identify net-new prospects at the exact moment they most need compliance services.

The Signal

- High-intent: companies actively preparing for launch.
- Verified data: The applicants contact information is included in the publicly-available website.
- Early-stage: maximum lead time for outreach.
- Public: zero data acquisition cost
- Volume: hundreds of filings per day, globally.
- ICP Fit: These contacts represent the highest likelihood to convert for Approve-IT as they are actively filing for approvals. Perfect fit for their services.



The Solution at a Glance

FastForward built a fully automated daily pipeline that monitors new FCC filings, extracts contact data, pulls it into Clay, enriches records with company and people intelligence, and routes qualified leads to outreach with no manual intervention.

10–50

New leads per day

~\$0.50

Daily enrichment

~3 min

Daily runtime

\$0

Hosting cost

How It Works

Two layers - a GitHub Actions scraping layer & a Clay enrichment layer - connected by webhook, firing daily.

Stage	Component	Function	Technology
1	Cron trigger	Fires daily at 7:00 PM EST	GitHub Actions
2	FCC scrape	Paginates results, extracts 16 fields per filing	Python + Playwright
3	Filter & deduplicate	Excludes China filings, skips processed IDs	Python (sent_ids.json)
4	Webhook ingest	Routes clean records into Clay	Clay Webhook
5	731 report scrape	Retrieves TCB application page per filing	ZenRows
6	AI contact extraction	Parses name, title, email, phone	Clay AI
7	Company lookup	Finds domain via AI search; falls back to email	Clay AI
8	DNC check	Blocks enrichment if existing client	Clay table lookup
9	Output routing	Routes to three output tables	Clay + Surfe

Enrichment Pipeline: Four Steps to a Qualified Lead

1

Scrape the 731 TCB Report

ZenRows retrieves the Form 731 TCB report page per filing, bypassing bot-detection barriers that block standard HTTP requests.

2

AI Contact Extraction

Clay's AI extracts structured fields: first name, last name, job title, email, and phone from raw HTML. Handles variation across TCB lab form formats.

3

Company Website Resolution

Clay AI searches for the company's primary domain. Falls back to extracting domain from the contact's email address if lookup fails.

4

Do Not Contact (DNC) Check

Company domain is checked against Approve-IT's current client list. If matched, all downstream enrichments halt immediately.

Output Routing: Three Paths to Outreach

Filing Contacts

Extracted directly from Form 731 TCB application. Email verification runs before routing to outreach.

People at Company

Additional contacts at the same company via Clay people search, filtered by relevant job titles. Monthly cadence.

Surfe Contacts

Broader search layer for job titles not surfaced by Clay's standard people finder. AI qualification step recommended before activation.

Technology Stack & Challenges

Technology Stack



GitHub Actions — Free-tier cron scheduler, zero infrastructure overhead



Playwright + Python — Headless browser scraping of FCC JS-rendered results



Clay — Central enrichment platform; webhook ingest, enrichment, output routing



ZenRows — Managed proxy for 731 report retrieval, bypasses bot-detection



Surfe — Extended contact search for broader job title coverage

Known Limitations & Roadmap

1. Contact Deduplication Across Runs — Same individual may re-enter on future filings. Options: rolling email blocklist in Clay or CRM-side suppression.

2. Surfe Table Activation — Built and populated but not yet activated. Add AI job title qualification step before enabling.

3. DNC Automation — Currently maintained manually. Should connect directly to CRM for automatic updates.

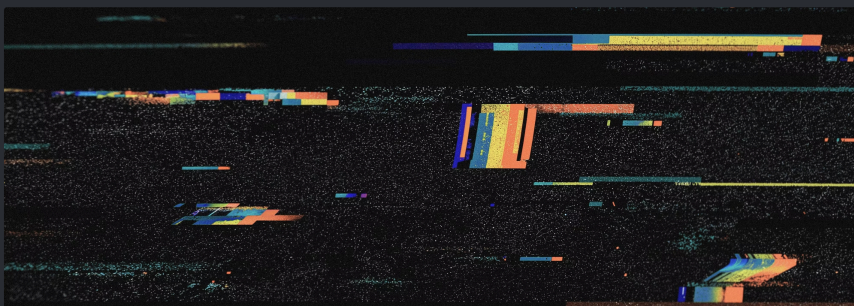
4. DNC Coverage Gap — DNC check not yet applied to the people search output table (Output 2).

5. Historical Data — ~20 years of FCC filings in a separate table. One-time enrichment pass not yet determined.

Built by FastForward

This pipeline was designed, built, and documented by FastForward for Approve-IT as part of an ongoing GTM engineering engagement. FastForward builds GTM automation systems for B2B teams by identifying high-intent data signals, designing enrichment pipelines, and connecting them to sales workflows.

If your team has an untapped data signal, a manual prospecting process that doesn't scale, or a CRM full of unenriched records, we'd like to talk. Engagements like this typically reach full production in weeks, not months.



What FastForward Delivers

- End-to-end GTM automation architecture
- Signal identification and pipeline design
- Clay, Python, and webhook engineering
- Documentation and handoff for in-house teams
- Ongoing iteration and optimization