A Solopreneur Business Plan for Building Medicaid Fraud Detection Software Aimed at the Early Intensive Developmental and Behavioral Intervention Autism Scam

MAY 31, 2026

🎧 Part I Podcast free on Spotify.

Thoughts on Healthcare Markets & Technology

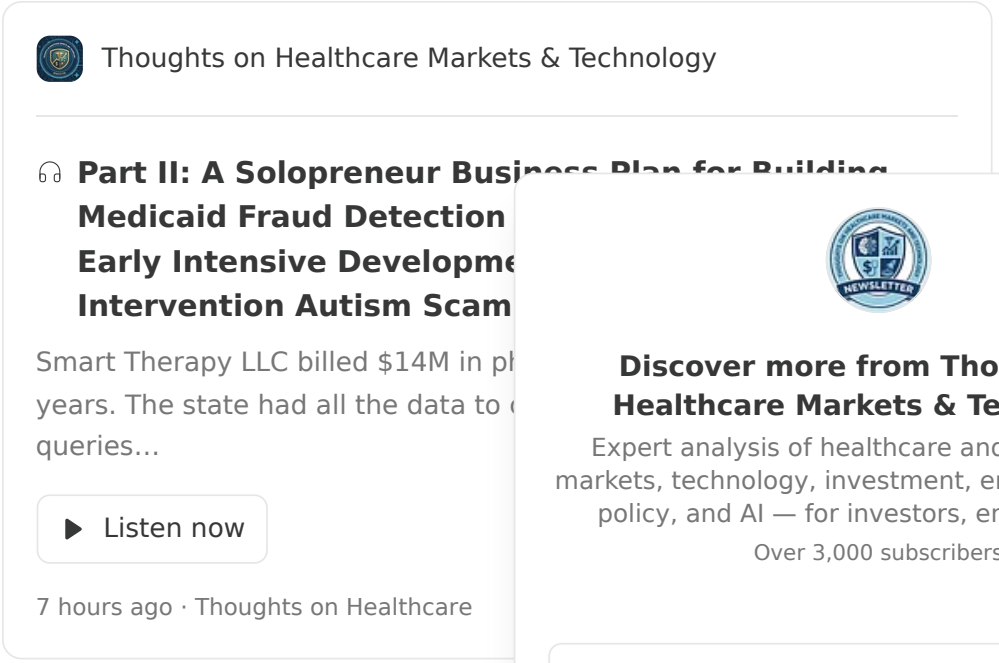
Part I: A Solopreneur Business Plan for Building Medicaid Fraud Detection Software Aimed at the Early Intensive Developmental and Behavioral Intervention Autism Scam

Smart Therapy LLC billed $14M in phantom autism therapy over 5 years. The state had all the data to catch it. Nobody ran the queries...

Listen now

7 hours ago · Thoughts on Healthcare

🎧 Part II Podcast episode for paid subscribers only. Also available on Spotify.



To listen to paid episodes in Apple or Spotify, li
settings on those platforms (instructions inside th
Podcast).

Abstract

A single Minnesota provider, Smart Therapy LLC, allegedly pulled more than $14M out of the state autism benefit by billing one-on-one therapy that mostly did not happen, staffing teenage relatives as clinicians, paying parents $300 to $1,500 a month per kid to keep the census full, and forging supervisor sign-offs while those supervisors were out of the country. The same shell also ran a $465K Feeding Our Future play. None of that is exotic. Every move it made leaves a fingerprint in claims, enrollment, and public registry data. This essay lays out a business plan for one person to build a company that catches exactly this pattern, using Claude Code to compress what used to be a 12-engineer build into a solo effort. Topline numbers worth holding in your head: EIDBI claims went from roughly 600K in 2018 to north of 400M by 2025, Minnesota flagged 14 Medicaid programs as high risk, and a single Optum-run tightening cut spending across those programs by about 29 percent, roughly 165M, in a couple of quarters. The market is not theoretical. The fraud is a data problem wearing a clinical costume.

Table of contents

The scam, reconstructed from the charging document

Strip away the autism framing and the mechanics are almost boring. A 28-year-old registered an LLC with the Minnesota Secretary of State in November 2019, listed herself as sole owner, and enrolled it as a provider in the Early Intensive Developmental and Behavioral Intervention benefit, which is the Medicaid program that pays for ABA therapy for kids under 21 with an autism diagnosis. On paper, Smart Therapy delivered intensive one-on-one behavioral therapy under a Qualified Supervising Professional. In reality, per the information, the techs were often 18 and 19-year-old relatives with a high school education and zero autism credentials, the kids were recruited out of the Somali community with monthly cash kickbacks to parents, and any child who lacked a diagnosis got run through a friendly QSP until they qualified. The line in the charging doc that should make any program-integrity person wince is that there was no child Smart Therapy could not get qualified.

Then comes the billing. Claims went out for the maximum hours Medicaid allowed for a given service, on days when the kid got a fraction of those hours or none at all. Sign-offs from the required providers and supervisors were forged, including from people who did not work there or were physically out of the country on the service date. Drivers hauled kids in and out and billed DHS for transportation, and some of those transport providers were also on the Smart Therapy payroll, which is a tidy little related-party loop. The take was more than 14M from DHS and UCare. Some of it left the country as wire transfers and turned into real estate in Kenya. The same entity, for good measure, enrolled in the federal child nutrition program under Feeding Our Future and claimed it was serving exactly 300 meals a day, seven days a week, scaling up to about 1,200 meals a day, roughly 200,000 meals total, for another 465K. Subtlety was not the brand.

Why this is a software problem the state keeps losing

The reason this case matters for a founder is that it is not a clever fraud. It is a loud one that ran from late 2019 through December 2024 before the first charge landed. The state was not blind because the scheme was sophisticated. It was blind because nobody was running the right queries against data it already held.

The macro picture is brutal and it is the whole pitch. EIDBI claims climbed from a little over 600K in 2018 to more than 400M by 2025 by DOJ's own accounting. Recipients went from about 1,400 in 2020 to more than 5,600 in 2024 while annual cost ran past 300M. Across all autism providers, billings hit roughly 1.6B from 2018 through 2025. Minnesota eventually named 14 Medicaid service categories as high risk for the same reason, and a sibling program, Integrated Community Services, did the identical hockey stick from 4.6M in 2021 to over 170M in 2024. The state legislative auditor later found DHS had the legal authority to investigate kickbacks in the autism program for years and simply did not use it. So the gap is not authority and it is not data. It is the absence of a system that turns the data into ranked, defensible leads before the money goes out the door.

That is the structural failure a software company sells against. The dominant model is pay and chase, where claims get paid fast to keep providers happy and access intact, and recovery happens years later through a Medicaid Fraud Control Unit after the cash is already a duplex in Nairobi. Pay and chase recovers cents. A scoring layer that flags the provider in month three instead of year five is worth orders of magnitude more, and the buyer can do that math without help.

What the product actually is

The product is not an AI that decides who is a fraudster. That framing gets you sued and gets disabled kids cut off from real therapy. The product is a scoring and investigations layer that sits next to the claims pipeline and does two jobs. First, it ranks providers and claims by fraud risk continuously, with every score traceable to specific evidence a human can read. Second, it gives an investigator a workbench where a flagged provider opens into a single view, the entity graph, the billing anomalies, the cross-program links, the documents, all assembled so a case that used to take an analyst three weeks to build takes an afternoon.

Think of it as three layers. An entity-resolution layer that figures out that Smart Therapy, its registered agent, its owners, its transport vendor, and its meal vendor are one cluster of humans even when the paperwork pretends otherwise. A detection layer that runs a library of pattern checks against claims and enrollment. And a case layer that packages the output into something a state attorney or an MCO investigator can act on, with an audit trail, because anything that cannot survive a due-process challenge is worthless.

Scope discipline is the whole game for a solo founder. Do not build a general healthcare fraud platform. Build the thing that catches the EIDBI archetype and its close cousins in the other 13 high-risk programs, which all share the structure of low-credential labor, high-frequency repetitive billing, vulnerable enrollees, and shell entities hopping between programs. One archetype, done so well it is embarrassing for the incumbents, beats a broad platform built by one person and trusted by nobody.

The data you can pull without anyone’s permission

The thing that makes this buildable by one person is that most of the signal lives in public or semi-public data you can ingest before a single customer hands you a claim file. You can build the entity graph cold and walk into a sales meeting already knowing things the buyer does not.

Start with the National Plan and Provider Enumeration System, which gives you every NPI, the rendering and billing names, practice addresses, and taxonomy codes, refreshed monthly and free. Layer the OIG List of Excluded Individuals and Entities, which would have flagged the hidden owner who had been excluded for three years over an adult daycare, and the SAM exclusions list. Pull Secretary of State business registrations to get incorporation dates, registered agents, and officers, which is how you catch the November 2019 LLC that enrolled as a provider weeks later. Add the

Death Master File so a claim signed by a dead clinician lights up. Add USPS address normalization and county property records, because shell entities reuse addresses and mailboxes the way teenagers reuse passwords.

On the claims side, you will not get a state's live MMIS feed on day one, but you do not need it for the demo. CMS publishes de-identified Medicaid research files and T-MSIS analytic extracts, and many states post provider directories and open payment data. That is enough to train and tune detectors and to show a buyer the shape of the answer. The pitch to the first customer is simple. The public data already clusters your high-risk providers. Give us a read-only claims extract under a business associate agreement and we will tell you which ones are billing impossible days. You are asking for the smallest possible data grant in exchange for a ranked list, not a data lake migration. That is a yes a mid-level MCO investigator can actually say.

The detection logic in plain English

Every move in the charging document maps to a detector, and none of them require a neural network to start. The fancy modeling comes later. The first version is mostly arithmetic and graph queries, which is exactly why one person can ship it.

Billing the ceiling is the easiest tell. A provider whose claims cluster at the maximum authorized units, day after day, kid after kid, is statistically screaming. Real clinical delivery is messy and varies. Fraud bills the cap because the cap is the revenue. You compute, per provider, the share of claims sitting at the authorized maximum and the variance of delivered units, and the honest providers form a fat distribution while Smart Therapy sits out on the tail by itself.

Impossible-day capacity is next. Sum the billed direct-service hours per rendering tech per day. When a single 19-year-old is credited with more one-on-one therapy hours than exist between sunrise and pickup, or is billed as present at two sites at once, that is not a productivity miracle. The absent-supervisor version is the same trick applied to the QSP. When a supervising professional is signing off on hundreds of sessions on dates that overlap with, say, a long gap in all their other activity, you flag the credentialing pattern even before you can prove they were in another country.

Related-party loops fall straight out of the entity graph. A transport vendor that shares an owner, an address, a bank routing pattern, or a phone with the therapy provider is the circular billing the prosecutors described. Hidden ownership is the same graph problem. You do not trust the DHS ownership form. You resolve the real cluster from shared identifiers across Secretary of State filings, addresses, agents, and exclusion records, and you surface the excluded party hiding behind a relative's name.

Then the demand-side signals. You cannot see a cash kickback, but you can see its shadow. A provider with abnormally high authorization amounts per child, unusual enrollment volatility as families churn toward whoever pays more, and a QSP whose qualifying rate approaches 100 percent looks like a diagnosis mill, which is what no child we could not qualify actually means in data. The single highest-value detector is cross-program reuse. The same humans showed up in Feeding Our Future, then autism, then housing stabilization, then integrated community services. A graph that spans programs catches the operator on their second scheme, which is where the real money and the real deterrence live.

Building it solo with Claude Code

Here is where the company becomes a one-person job rather than a Series A. The architecture is unglamorous on purpose. A Postgres database with a graph extension or a dedicated graph store for entity resolution, a Python ingestion layer that pulls and normalizes the public datasets on a schedule, a detector library that is mostly SQL and pandas with a scoring service on top, and a thin web app for the investigations workbench. Nothing here is research. It is plumbing, and plumbing is exactly what an AI coding agent is good at.

The realistic loop with Claude Code looks like this. You hand it the NPPES and LEIE file specs and have it write the parsers, the schema, and the loaders, then you point it at the messy reality of the files and have it fix the encoding and the dedupe edge cases, which is where most of the actual hours go. You describe each detector in a sentence, share-of-claims-at-cap by provider, hours-per-tech-per-day exceeding a threshold, and it writes the query, the test fixtures, and the backfill. You give it a sample fraud cluster and ask it to build the entity-resolution scoring, then you spend your judgment correcting the matches, because a false merge that links an innocent provider to a fraud ring is the kind of mistake that ends the company. The workbench, the auth, the audit logging, the PDF case export, all of that is standard app code an agent produces quickly.

What does not get automated is the part that matters, and that is the reason a domain person should build this rather than a generic engineer. Deciding which signals are defensible, where the false-positive cost lands on a real disabled child, how to weight a detector so you flag the operator and not the small honest clinic having a weird month, and how to phrase a finding so it survives a provider appeal. Claude Code collapses the engineering. It does not collapse the judgment about Medicaid, ABA delivery norms, authorization rules, and program integrity law. That asymmetry is the

founder's edge. A solo build that used to need a team is now mostly a question of whether you know which queries to ask.

Who pays, and the contingency versus license fight

There are three buyers and they pay differently. The managed care organizations, the UCare-shaped entities that took capitation and ate the 14M, have Special Investigations Units, real budgets, and a direct financial interest because fraud is their loss now, not just the taxpayer's. They are the fastest yes. State Medicaid program integrity offices and the agencies that run these benefits are the largest spenders but move at the speed of procurement, which is to say geologically. The Medicaid Fraud Control Unit and federal partners are users of the output more than buyers of software, though they make excellent reference logos.

The pricing fork is contingency versus license. Contingency, where you take a cut of recoveries or of prevented spend, is seductive because it aligns you with the customer and lets a cash-poor founder land a deal with no budget line. It is also a trap if you are not careful. Recoveries take years, attribution is a knife fight, and you can do brilliant detection work and watch the case die in a backlog with nothing to invoice. The cleaner model is a software license priced against the spend you are watching, with a measured savings guarantee, because the Optum tightening that cut roughly 165M across the high-risk programs gives you a defensible reference point for what good oversight is worth. A blended deal works best for the first customers. A modest platform fee so you can eat, plus a success component on confirmed prevented payments, with prevented defined tightly enough that you are not arguing about counterfactuals forever. The savings story sells itself. Stopping a 14M scheme in month three rather than year five is not a soft ROI.

Selling it without a sales team

A solo founder cannot run an enterprise sales motion against state government, so do not. The wedge is a report, not a demo. Take the public data you already ingested, build a risk profile of a region's providers in one high-risk program, and turn the most defensible findings into a short, sober briefing that names patterns and not people. This billing-at-cap distribution, these related-party transport loops, these incorporation-to-enrollment timelines under 60 days. Then walk it into an MCO SIU or a sympathetic program-integrity director as evidence that the answer is sitting in data they own. You are not selling software. You are showing them their own house is on fire and you brought a hose.

Design partners come next. One MCO that gives you a read-only claims extract under a business associate agreement, in exchange for a ranked lead list and a fixed low fee, gets you the proprietary signal that makes the product real. Publish the methodology, not the targets. The thought-leadership surface here is enormous because the underlying story, billions across a web of programs, is already in the press, and a credible technical voice explaining how the detection actually works generates inbound from exactly the SIU directors and state CFOs who buy. FOIA the enforcement actions and you get a steady drip of confirmed fraud patterns to validate and market against. The motion is land via a free risk read, prove it on one customer's data, let the savings number recruit the next buyer.

The incumbents, and why a one-person shop still has a shot

This space is not empty. SAS sells fraud analytics to state Medicaid agencies, Optum and Cotiviti run payment integrity at scale, LexisNexis Risk Solutions owns a lot of the identity and provider data graph, Codoxo and Qlarant and Thomson Reuters Pondera all play in government program integrity. A reasonable person asks why a solo founder is not roadkill.

Three reasons. The incumbents sell broad platforms with long implementations and seven-figure price tags, which means they orient toward big national programs and the big claim types, and a narrow, fast, cheap tool aimed precisely at the EIDBI archetype and its 13 cousins is a wedge they are too heavy to bother defending. The incumbents are also generalists about the clinical and policy detail, and the detail is where these schemes hide, so a builder who actually understands CMDE qualification, QSP supervision rules, and authorization ceilings can encode sharper detectors than a horizontal vendor staffed by data scientists who have never read an EIDBI policy manual. And the incumbents are slow, because procurement and enterprise contracting are slow for everyone, which means there is a multi-year window where states are panicking about exactly this fraud and the big players are still scoping statements of work. A solo shop that ships a useful risk report this quarter beats a platform that goes live in 2028. The realistic exit, frankly, is that one of these incumbents buys the company once the data graph and the customer proof exist, which is a fine outcome and shapes what you build.

Moat, unit economics, and a napkin model

The uncomfortable truth is that detection logic is not a moat. Any competent team can write a billing-at-cap query. The moat is the cross-program entity graph, the resolved clusters of humans and shells linked across Feeding Our Future, autism, housing, and the rest, enriched by every confirmed case your customers feed back. That graph gets better with each customer and each enforcement outcome, and a competitor starting fresh has to rebuild years of resolution and labeling. Switching costs help too, because once an SIU runs its caseload through your workbench and trains its process around your evidence packages, ripping you out is painful.

The napkin model is friendly for a solo operator. Costs are basically your time plus cloud and data, call it low five figures a year in infrastructure because public data is free and claims extracts are not enormous. Price a mid-size MCO deal at, say, 150K to 400K a year for the watched spend, which is trivial against a single 14M loss, and gross margin on software like this sits in the 80s once built. Three or four MCO logos is a real business run by one person with maybe a contractor for support. The total addressable market is not small either. Fourteen high-risk program categories in Minnesota alone, multiplied by every other state now waking up to the same exposure after a 259M federal funding pause got everyone's attention, multiplied by the MCOs inside each state, is a long runway for a focused tool. You are not trying to win all healthcare fraud. You are trying to own the vulnerable-population, low-credential, high-frequency-billing corner of it, and that corner is bleeding money in every state that copied Minnesota's benefit design.

The ways this blows up in your face

A sober founder spends real time here because the failure modes are nasty and several of them are ethical, not just commercial. The worst case is not that the product fails. It is that it works badly and a false positive flags an honest clinic into a payment suspension, and a real autistic kid loses real therapy while the appeal grinds. That is not a rounding error. It is the reason the product must be a ranked-lead and evidence tool for human investigators, never an automated denial engine, and why every score has to be explainable down to the specific claim. Build it to make a good investigator faster, not to replace judgment, and write that into the contract.

The rest of the risk list is mundane and survivable. Data access is gated by HIPAA and business associate agreements, so you carry compliance overhead heavier than a normal SaaS and you cannot cut corners on security as a solo operator, which is precisely the kind of unglamorous work to hand Claude Code and then verify carefully. Government procurement is glacial and contingency revenue is lumpy, so you finance the early years on MCO software fees rather than betting the company on

recovery timing. There is defamation and due-process exposure if findings name providers before they are adjudicated, so the language is patterns and probabilities, never accusations. Model drift is real because fraudsters adapt the moment a detector becomes known, so the detector library is a living thing, not a one-time build. And the bus factor is you. One person holding the only knowledge of a system that states depend on is a fragility you have to design around early, with documentation and eventually a second pair of hands, or your acquisition diligence falls apart.

The honest read

The opportunity is real and the timing is unusually good, because the fraud is loud, the dollars are enormous, the states are scared, and the tooling finally lets one capable person build what used to need a funded team. The charging document reads like a product requirements list. Bill the cap, fake the staff, forge the sign-offs, pay the parents, hop the programs, wire it abroad, and every one of those leaves a trace that arithmetic and a good graph will catch.

What it is not is a passive lifestyle business or a clean morality play. It is regulated, the sales cycles fight you, the incumbents will eventually notice, and the same engine that catches a 14M thief can hurt an innocent provider and a vulnerable kid if built lazily. The version of this company worth starting is narrow, explainable, obsessed with false-positive cost, and pointed at the exact archetype Minnesota keeps prosecuting. Done that way, a solopreneur with Claude Code, a domain brain, and the patience to court one MCO at a time has a genuine shot at building something that saves a lot of public money and is worth buying. Done lazily, it is just another black box that pays kickbacks of its own kind in the form of wrongly suspended care. The difference is entirely in the judgment, which is the one part the AI does not write for you.

Thoughts on Healthcare Markets & Technology is a reader-supported publication. To receive new posts and support my work, consider becoming a free or paid subscriber.

1 Like • 2 Restacks

← Previous

Discussion about this post

Comments Restacks

Write a comment...