



EXPERT INTERVIEW

UPDATED
10/02/2025

Levi Lian, CEO of Raycaster, on why vertical AI is workflows first & chat last

TEAM

Jan-Erik Asplund
Co-Founder
jan@sacra.com

DISCLAIMERS

This report is for information purposes only and is not to be used or considered as an offer or the solicitation of an offer to sell or to buy or subscribe for securities or other financial instruments. Nothing in this report constitutes investment, legal, accounting or tax advice or a representation that any investment or strategy is suitable or appropriate to your individual circumstances or otherwise constitutes a personal trade recommendation to you.

This research report has been prepared solely by Sacra and should not be considered a product of any person or entity that makes such report available, if any.

Information and opinions presented in the sections of the report were obtained or derived from sources Sacra believes are reliable, but Sacra makes no representation as to their accuracy or completeness. Past performance should not be taken as an indication or guarantee of future performance, and no representation or warranty, express or implied, is made regarding future performance. Information, opinions and estimates contained in this report reflect a determination at its original date of publication by Sacra and are subject to change without notice.

Sacra accepts no liability for loss arising from the use of the material presented in this report, except that this exclusion of liability does not apply to the extent that liability arises under specific statutes or regulations applicable to Sacra. Sacra may have issued, and may in the future issue, other reports that are inconsistent with, and reach different conclusions from, the information presented in this report. Those reports reflect different assumptions, views and analytical methods of the analysts who prepared them and Sacra is under no obligation to ensure that such other reports are brought to the attention of any recipient of this report.

All rights reserved. All material presented in this report, unless specifically indicated otherwise is under copyright to Sacra. Sacra reserves any and all intellectual property rights in the report. All trademarks, service marks and logos used in this report are trademarks or service marks or registered trademarks or service marks of Sacra. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any report is strictly prohibited. None of the material, nor its content, nor any copy of it, may be altered in any way, transmitted to, copied or distributed to any other party, without the prior express written permission of Sacra. Any unauthorized duplication, redistribution or disclosure of this report will result in prosecution.

Published on Oct 02nd, 2025

Levi Lian, CEO of Raycaster, on why vertical AI is workflows first & chat last

By Jan-Erik Asplund



Background

This year, we've tracked the evolution of vertical AI from legal (Harvey) to finance (Hebbia).

To learn more about how vertical AI as a category is developing, we reached out to Levi Lian, co-founder & CEO of enterprise AI for life sciences startup Raycaster (YC F24).

Key points via Sacra AI:

- **Foundation model companies OpenAI & Anthropic commoditized the first wave of vertical AI companies that reasoned over public data—like web PDFs, patents and earnings calls—with agentic products that browse the web, do deep research and build datasets, spurring the next wave of vertical AI companies that focus on integrating frontier models with vertical-specific workflows, embedded with proprietary telemetry that captures document diffs, user corrections & tool calls.**

“The “chat + wide research + longer compute” pattern is everywhere. It’s useful but quickly becomes a feature, not a company. . . If your edge is earnings calls, patents, or web PDFs, horizontals will catch you. OpenAI, Anthropic, Manus, and others will fill spreadsheets on demand. We tried that early for life sciences legal and R&D and concluded differentiation lasts about a quarter.”

- **In highly-regulated, risk-averse industries, less than 1% of manual workflows have been replaced with AI as companies fear IP exposure through third-party models, unauditible changes to high-stakes documents, and AI-generated content they can't verify or trace back to source material, creating the opportunity for companies like Raycaster for biotech (Seed, Y Combinator) alongside bigger players like Harvey for law (\$21M Series A, Sequoia) and Hebbia for finance (\$30M Series A, Index Ventures).** "In legal and finance, real workflow adoption and penetration is still early. In biotech it's even earlier. Not because people are anti-AI, but because the documents are the company—IP, process, and know-how. Leakage, unverifiable edits, and missing context are non-starters. This is why life sciences professionals aren't even using ChatGPT. "
- **Today's emerging vertical AI playbook: 1) embed engineers a la Palantir's forward-deployed model to build AI workflows and automate away manual processes, 2) hire domain experts as internal evaluators through a Mercor-like contractor network to tune & improve outputs, and 3) capture customers' usage data to continually improve the system.** “Early on it’s hands-on by design. Map the workflows, and the inputs and outputs. Codify the company-specific context. Encode the plan and run pilots against said workflows, then iterate. Eventually this becomes a productized module. We’ve also experimented with the Mercor style contractor recruiting, where we bring on domain SMEs as validators. For some high profile customers, that’s how we deliver immediate utility without making them feel like guinea pigs.”

Interview

What is Raycaster in short and what inspired you to start the company?



Raycaster is AI for life sciences—think Harvey or Hebbia, but aimed at development, where documents are the product. We handle drafting, processing, and editing across the drug development lifecycle. You could also think of us as Devin for documents: companies set up their document context (how they draft, edit, review, and validate), and then agents can take up requests to complete difficult, long-running tasks.

Most people hear “AI for life sciences” and jump straight to discovery—new molecular structures, novel targets. While important, the real bottleneck is the 8-10 years after discovery: clinical trials, manufacturing, quality, regulatory. Even when things go well, odds of success are roughly 10 percent and the spend is in the billions.

The spark came from watching smart friends—PhDs who expected to do frontier science—spend their weeks reconciling Word files and spreadsheets in the industry. My parents are physicians; my dad later moved into pharma, so dinner table stories were about how drugs actually get to patients. At Stanford I saw how development drags and how much of it is “document maneuvering.” They were essentially just maneuvering documents. We talked to a lot of teams and realized there’s a huge gap—and yet the problem’s fundamentally tractable with AI.

So walk me through some of the challenges and problems in that workflow in the eight to ten years after the discovery gets made. What did you see there?

Three big challenges come up.

First you have the massive coordination problem. You’re moving work across sponsors, CDMOs/CROs, sites, central labs, and regulators. Choosing and onboarding a CRO/CDMO can take a quarter or two. A small change in a clinical protocol cascades across site packets, the schedule of assessments, ICFs, the SAP, CRFs, and the eTMF. Every handoff adds places to misalign and redo work.

Then comes the quality aspect of drug development. Quality/CMC is the heaviest lift in the dossier. Day to day that means keeping specs and methods aligned across sites, writing deviation and CAPA narratives, running change-control impact analysis, comparability when something changes, and



making sure RIM/publishing metadata matches the documents. At manufacturing sites, QA/QC is often one of the largest teams, and biotechs feel it because a lot of scientist time turns into document prep and review.

Lastly, you not only have multiple stakeholders, but artifacts depend on each other. Tech transfer packs feed master batch records. Executed batch records generate deviations and campaign summaries. Those roll into PPQ/CPV packages and stability reports. All of that ladders into Module 3 sections that get published. If one link is missing or inconsistent, you trigger rework and regulatory friction. On the CMC side that shows up as information requests, clock stops, or a second review cycle; on the clinical side, similar drift shows up as protocol amendments.

Can you lay out all the steps at a high level—the entire eight to ten years process?

Sure. Preclinical → IND → Phase 1 → Phase 2 → Phase 3 → submission → approval → commercial. IND opens human testing; Phase 1 tests safety; Phase 2 looks at dose and signal; Phase 3 confirms efficacy and safety at scale; then you file an NDA or BLA and, if approved, move into manufacturing at commercial scale with post-market commitments.

As you move through phases, the cast and the workload expand. You bring on CROs for trial operations, central labs and bioanalytical partners for samples and assays, and CDMOs for process scale-up and tech transfer. Regulators across regions weigh in along the way—FDA, EMA, MHRA, PMDA—not just at the end. You're also dealing with the physical world: site start-up and enrollment, site comparability, and multi-site method alignment. The cost per approved drug lands in the low billions over 8–10 years, and a surprising share of the drag isn't bench science—it's all the coordination, trial operations, supply readiness, and document work that ties everything together.

What makes this compelling to apply AI to? Is AI really what makes this addressable as a market, given that it was under-addressed for so long?

Traditionally you had Veeva and IQVIA as system of record—like Salesforce or ServiceNow in their worlds. They're essential, they keep a history of trial, quality and regulatory



docs. But the “what’s missing, what’s inconsistent, what needs to be written next” work still happens ad hoc. That’s where timelines slip. Roughly 3 in 4 protocols see at least one amendment, averages are around 3 per study, and half are avoidable. Each one typically costs in the hundreds of thousands and adds weeks to months. About one in four drugs doesn’t pass on the first review cycle. If you run on milestones, a quarter slip can mean missing a financing window. AI lets us add two layers on top of the system of record:

System of intelligence—continuously reads the current documents and data, flags gaps, checks conformance to templates/SOPs/acceptance criteria—think “continuous checks” for documents instead of end-of-cycle fire drills.

System of action—drafts or edits the next artifact, generates tables/diagrams, runs small code transforms where needed, and files back with page-linked sources.

And there’s a deeper win for AI: living context engineering. Communication cost is the biggest human cost in doc-heavy orgs—every handoff and “who owns this?” sync is a tax. With AI, design docs, SOPs, and acceptance tests stay in sync with the actual artifacts; decisions and rationales are captured once; the agent proposes the next step with evidence. You get a leaner but sharper company where organizational wisdom acts as a standing reality check—and everyone’s average performance goes up.

What made Veeva originally successful?

Veeva went deep on the pharma vertical. They started with a life-sciences CRM built on Salesforce, then built Vault to manage the documents that regulators actually inspect—eTMF, QualityDocs, RIM, Submissions. When you own the system of record for regulated content, you’re mission-critical, not just a sales tool. Today they’re moving CRM onto Vault to keep data and workflows under one roof.

IQVIA came from the other direction—clinical data and services. IMS Health’s data plus Quintiles’ CRO gives them end-to-end coverage, from trial design to commercialization. Like Veeva, they package software and services into integrated offerings, which makes them sticky and hard to displace.



Can you speak to regulatory concerns and how AI adoption might be slower in this industry? Is that a challenge or opportunity for Raycaster?

Both. Biopharma runs on cGMP, 21 CFR Part 11, audit trails, and validated systems. Veeva and IQVIA are the systems of record for a reason—they're the compliant memory of the organization.

We operate as a complementary layer. The drafting and pre-submission phases are typically non-GMP and non-Part 11—no signatures yet—which creates a safe lane to move fast.

How we chose to operate is similar to the Devin analogy. Before work even starts, we help customers define the workspace, parameters, schemas, SOPs, and who approves what. Our job is context engineering—wire the tools, templates, acceptance criteria, roles, and evaluators so users don't have to prompt. The UI mirrors how teams already work: click to draft, click to edit, click to review.

With this in mind, we were able to launch with two primary use cases:

On the biotech side—tech transfer authoring and QA, CTD-aligned specs, and RFPs for CDMOs/CROs. Results: fewer redlines and re-runs, faster supply readiness, fewer last-minute surprises.

For providers like CROs and CDMOs—ingest messy sponsor packets and return consistent responses and questions, site plans, and procedures with sources. Results: shorter vendor selection and onboarding, which often eats a quarter or two.

How do you think about context engineering and making this simple for users? With long documents, it's challenging to manage context in ChatGPT or Claude without losing the needle in the haystack.

We tried chat with agentic flows. Helpful, but brittle. The real win is context engineering.

For us, context engineering means wiring the actual organization into the agent: your repositories and permissions, your schemas and templates, your tool specs and plans, and



your review roles and acceptance criteria. The agent works in your objects—specs, methods, batch records, tech transfer checklists—flags what’s missing or inconsistent, cites the exact page, proposes a fix, and routes it to the right owner. Experts stay in the loop as evaluators. Early on, they approve actions, tune the checks, and teach the system what “good” looks like. Their feedback updates templates and rules, so drafts get more trustworthy, handoffs shrink, and each decision improves the next one.

Why leadership cares is simple: context cuts communication cost. Most organizational delay is handoffs, clarifications, and “who owns this” loops. A good context layer propagates leadership intent and turns the org’s accumulated wisdom into a standing reality check. People make faster, sharper decisions because the system keeps the plan and the work in sync.

A coding analogy: on a new codebase, if you keep design docs in lockstep with implementation, the agent can move from product vision down to table design without losing the bigger picture. Companies can do the same with SOPs, design docs, and acceptance tests tied to real artifacts. That becomes a durable memory system—and why investing in AI plus context pays for itself in attention saved and rework avoided.

Where do you fall on fine-tuning your own models versus using off-the-shelf foundation models?

We’re model-agnostic. Today we use frontier models off the shelf and win on workflows, tool design, context engineering and evaluators, not on training a new base model.

The durable piece is a portable context layer—repositories and permissions, domain schemas and templates, tool specs and plans, and acceptance tests. That keeps us independent of any one vendor, and lets customers swap models without ripping out workflows.

We chose to build a context layer deeply embedded within organizations. As horizontal players become more vertical and vertical AI becomes more horizontal, having a portable context layer independent of foundation models provides strategic and defensible value. This approach avoids vendor dependence



while enabling organizations to choose underlying models while maintaining our context layer.

Is there a Palantir-esque forward deployed engineer aspect when onboarding companies?

Yes, early on it's hands-on by design. Map the workflows, and the inputs and outputs. Codify the company-specific context. Encode the plan and run pilots against said workflows, then iterate. Eventually this becomes a productized module.

We've also experimented with the Mercor style contractor recruiting, where we bring on domain SMEs—former CMC, QA, regulatory writers—as validators. For some high profile customers, that's how we deliver immediate utility without making them feel like guinea pigs.

Are there any vertical AI companies that have jumped out at you for inspiration?

A few, but mostly for what they didn't do at first and what they've shifted toward.

Lesson 1, chat commoditizes fast. In legal and finance, the “chat + wide research + longer compute” pattern is everywhere.

It's useful but quickly becomes a feature, not a company. The UI often collapses into spreadsheet views running research agents. That's not durable.

Lesson 2, public-data plays get raced to the bottom. If your edge is earnings calls, patents, or web PDFs, horizontals will catch you. OpenAI, Anthropic, Manus, and others will fill spreadsheets on demand. We tried that early for life sciences legal and R&D and concluded differentiation lasts about a quarter.

Lesson 3, adoption is still single-digit where work is regulated. In legal and finance, real workflow adoption and penetration is still early. In biotech it's even earlier. Not because people are anti-AI, but because the documents are the company—IP, process, and know-how. Leakage, unverifiable edits, and missing context are non-starters. This is why life sciences professionals aren't even using ChatGPT.



With Harvey, we saw the shift from chat to packaged workflows with firm-grade guardrails and content partnerships. With Hebbia, we saw a strong control surface for multi-agent retrieval that non-prompt-engineers can actually drive. The common thread: workflows first, evaluators second, chat “anything” last.

This helped us with product direction:

1. We focus on internal documentation—tech transfer, specs and methods, batch records, change control, Module 3 slices—not just web scraping.
2. We ship workflow modules, not prompts: click to draft (RFP, Module 3, campaign summary), click to review, click to tag SMEs.
3. We win on context engineering.

How do you think about the potential of MCP for what you're doing at Raycaster?

Thinking about MCP through the lens of context is exactly right. As an application, the goal is win-win—every time foundation models improve, our product should get better overnight.

MCP helps that happen. It standardizes how models discover, call, and audit tools, so model providers have an incentive to promote it. That buys portability, cleaner eval, and consistent traces. If Veeva or IQVIA expose MCP endpoints, great—we spend less time on bespoke glue and more on customer value.

But MCP is about designing tools, not the entire context. The hard part in life sciences is semantics and governance. We still have to supply the ontology—specs, methods, batches, stability studies, Module 3 sections, the enterprise-specific processes and change control, along with indexing the documents that are in different formats. That’s where our focus lies.

What is the flywheel around being very deep in customers' documents and processes such that it becomes hard to take Raycaster out?



The flywheel encompasses several elements.

First is context. Every draft, edit, comment, and approval teaches the system how this org actually works across Veeva, IQVIA, SharePoint, LIMS. Templates, reviewers, edge cases, naming quirks—it all becomes living context. That's the product stickiness.

Second is evaluation. Trust mounts when you start using AI for first draft, but then for review, then final check and even submission. You already see this play out fast in the coding world with copilots to autonomous agents in a three year span. Pass/fail history becomes a quality signature for the org and a training signal for the agent. Over time you see hard numbers move—fewer draft cycles, fewer avoidable amendments, faster QA turnarounds. Essentially, people will then trust Raycaster for any document work, not just the drafting.

Lastly, it's this concept called data currency. Travis May, CEO of Liveramp, once wrote about six moats of data, with the most important being a data currency like FICO or Nielson. I see a potential for Raycaster to define a new standard in pharma manufacturing. Sponsors and providers can reference the score in SLAs—"submit at ≥ 85 with no red flags." Once it becomes an acceptance primitive between two parties, it's sticky.

Under the hood we log tool-use trajectories—plans, tool calls, diffs, and fixes—and SMEs label tough cases. These de-identified traces and labels become high-signal eval sets for our own post-training and, where appropriate, even for foundation model labs.

So the flywheel is simple: context makes better drafts, evaluators create trust, the score becomes currency. The longer we run, the smarter and safer the workflows get, and the harder it is to go back to manual.

In the long term, could developing this AI across all their documents make it possible to enter other areas like R&D and drug discovery?

I definitely see huge potential here, but only if the documents become a living system, not just better search. Today, we technically already assist R&D teams with paper/patent mining,



but it's increasingly commoditized. The real lever is linking internal R&D artifacts (protocols, assay reports, negative results) to downstream manufacturing and clinical context. Our job is to build that organizational graph from day one—protocols, specs/methods, batch records, stability/PPQ, Module 3—with impact analysis and evaluation so tech transfer is always up to date with the science.

This means near term we're still Devin for documents for drug development, but then medium term can check upstream so the entire R&D pipeline gets design-for-manufacturing and regulatory signals in-line.

Long term, the tool-use traces from the AI that tie early R&D choices to approval outcomes become high-signal training data for the frontier labs. I do see a symbiotic relationship with the labs, where the real inputs and outputs are the training gym.

This lets us, a vertical player, “meet in the middle” with horizontals: they advance general reasoning; we supply enterprise context, tests, and the outcomes that matter.

Disclaimers

This transcript is for information purposes only and does not constitute advice of any type or trade recommendation and should not form the basis of any investment decision. Sacra accepts no liability for the transcript or for any errors, omissions or inaccuracies in respect of it. The views of the experts expressed in the transcript are those of the experts and they are not endorsed by, nor do they represent the opinion of Sacra. Sacra reserves all copyright, intellectual property rights in the transcript. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any transcript is strictly prohibited.