Why Origami

Partner program for integrators and consulting firms. Origami AI, 2026.

When a practice picks a platform, it's really picking whose infrastructure the next five years ride on, and whose engineers it gets to call when a client asks something hard.

Who built it

Cody Launius and Jared Dyson invented SDO and hold the patent themselves, personally. It matters more than a legal footnote usually does: the inventors are still reachable, still the ones who answer when the core needs a decision.

Jared Dyson, co-inventor of SDO. Jared built the core technology that now powers Walmart Connect and Sam's Club Connect. That's production infrastructure at retail scale, where margin errors are measured in billions. Before Origami, Walmart Global Tech. He's the primary name on the patent paperwork, so when one of your clients asks who stands behind this thing, you get to give them a person.

Cody Launius, co-inventor of SDO. Cody built the platform from 2020 and authored the IETF vCon use-cases draft. The model lineage grew out of collaboration with researchers from Google DeepMind, Stability AI, Runway ML, and Black Forest Labs, plus a partnership with IBM: fourteen models trained, three serving production today. Those weren't advisory seats on a slide. It was research work, done together, and you can see it in how the models were built.

Every deployment needs an implementation team. That's by design.

Most vendors paper over implementation with a promise of easy setup. Nobody in procurement believes it. The dashboards get built, the data keeps piling up, and nothing about the job has changed.

Origami's endpoints are APIs and SDKs. Writing code against them is the product, not a setup chore. So instead of pretending otherwise, we split the toolbox from the hands: we build and run the platform, certified firms do the implementation, and both jobs get done properly that way.

For an integrator, that's the whole pitch. The platform requires firms like yours, and services aren't an add-on bolted onto the sale. They're how the product ships.

The infrastructure

An AI infrastructure company, with our own models, compute, and products. We run our own models, our own inference, our own weights. The company started in 2020 as a researcher collective; product work incorporated in 2023 around the IETF vCon standard.

We've trained fourteen models and three run in production today (Tani, Yama, Seishin). A good part of that training happened on consumer hardware, back when the industry was renting datacenters.

The compute is ours: bare metal, single-tenant, AMD SEV, on dedicated racks in colocation facilities we own and run. Automated provisioning, zero-human operations.

Nobody holds the keys. The architecture is public-key-only, so no human at Origami can open customer data, and our own database contains none of it.

The failures get documented too. We ship the write-ups next to the successes instead of burying them, which is why the evidence in this brief is stronger than the marketing.

The Karpathy convergence. The SDO data model is a latent-space vector graph. Karpathy published the same structural principle realized in interlinked markdown. Two independent discoveries, different mediums, same structure. That's how you tell a fundamental shape from proprietary luck. We hold the patent; he published the structure independently. If you're going to build a practice on a data model, this is the one the industry has now converged on twice.

The stack: six products, one deployment

Enterprise customers don't buy products. They buy a deployment: SDO goes onto the customer's data residency host, and everything else connects back to it.

1. SDO, the foundation. Ingests from any source. EL without the T: it learns structure and builds its own tooling, then runs an always-on data scientist agent that surfaces patterns before anyone knows to look for them. What you get is a command and control dashboard and the API everything else connects through.

2. Personal Assistant, included at no licensing fee. Email, calendar, tasks, documents: a 24/7 assistant, one user per instance. Everything it sees flows into the customer's SDO instance. We leave it free on purpose; every user makes the intelligence layer richer.

3. Virtual Executive, the AI manager. One per department, handling goal setting, work planning, assignment, and tracking, coordinated across departments on shared SDO data. It's forked from our internal agent orchestration system, with one addition: it hires and trains agents for the roles itself.

4. Vision Builder, the development platform. Five specialized agents: Architect, Engineer, QA, Reviewer, DevOps. The Reviewer actually uses the feature like a person would, clicking through the application instead of checking assertions. Code review is enforced, and builds deploy to the customer's host or VPC.

5. Research Platform, lab management for research-intensive organizations: academic, enterprise R&D;, government, biotech. The PI just asks their assistant for status; the team uses the web app for structured data, provenance, and compliance. Licensed per lab.

6. Ambient Conversational Intelligence: phone, messaging, and ambient voice capture. Each conversation is sealed into a signed, tamper-evident vCon container with a consent record, then structured and fed into SDO. It's a capture layer rather than an intelligence system: it records and structures, SDO does the thinking. Installed hardware, enterprise only.

The ecosystem loop: SDO creates intelligence, the products act on it, and every product feeds SDO back. Every product makes the others smarter. For your practice, that changes how an account grows. The second product doesn't need a new sale; it's already sitting in the deployment you built.

Deployment: four modes, every one an engagement

Sized to the customer's rated usage profile, from full cloud to fully air-gapped. Each mode is a different engagement for whoever builds it.

Cloud: the full stack lives in Origami colocation. The data residency host is run by us, under standard licensing. It's the fastest way to start.

Hybrid: a full copy of the platform minus GPU sits at the customer site; inference stays with Origami. On-prem hardware is leased from Origami. We say so up front so it doesn't surface late in

procurement.

Tethered: customer site with a controlled connection back. The middle path for regulated data that still wants platform updates.

Air-gapped: no connection at all. The full stack sits inside the customer's perimeter. It's the hardest deployment to build and the easiest one to defend in a security review.

Compliance here is architecture, not badges. PHI is physically isolated on dedicated erasure sets; that's how the system is built, not a policy someone agreed to. The architecture is public-key-only and no human holds data keys. Our facility has SOC 2 and we'll show it, as supporting evidence rather than the headline. We don't claim certifications we don't hold, and to a security-literate buyer that reads better than a badge would.

Proof: three engagements, three kinds of impossible

The client names are under embargo, so there are no logos here. The work does the talking, and every one of these started as an implementation project rather than a signup.

Revenue intelligence, CPG suppliers: the backend everyone said was unreadable. Corruption at scale left a revenue recovery firm's backend unreadable. We recovered the data using forensic accounting techniques, and now we're replacing the whole backend and rebuilding the workflow around it from the ground up. Partner read: when a client's data looks unfixable, that's an implementation project, not a subscription.

Medical equipment manufacturing, Germany to the US: thirty years of German, made queryable. A German medical equipment manufacturer changed hands across the Atlantic. Thirty years of German-language business and technical documents, basically the company's institutional memory, went into SDO and came out as a queryable corpus for the American successor. Partner read: cross-lingual legacy ingestion, scoped as a deployment instead of a consulting estimate with open hours.

Critical infrastructure, operational technology: a predictive failure model for physical assets. A contractor that secures OT environments, defense, nuclear, utilities, needed to know when physical assets were about to fail. We rebuilt their digital twin system: BMS telemetry, ambient conversation, and other inputs, fused together in SDO. Partner read: multi-source fusion across OT data. The kind of engagement that runs for quarters.

Three sectors, same shape each time. What looked like unusable data turned into a scoped deployment, and now it runs. That's most of what an Origami practice will look like.

The commercial model: recurring on both sides

Two-part enterprise pricing: licensing covers the right to use the software, resources covers the customer's dedicated intelligence compute and storage. Inference never shows up in licensing; all AI usage draws from the customer's resources pool.

Licensing is a right-to-use fee, billed along each product's actual value axis: per data domain, per manager, per lab, per device. It isn't a flat enterprise tax.

Resources is the compute pool: intelligence plus data residency, floored at one person-month of compute and calibrated through a Proof of Value engagement.

Both parts recur, so your revenue does too: implementation, custom connectors, ACI hardware installation, Proof of Value work, and ongoing services against a deployment that keeps growing.

Partner economics are confirmed at certification. The algorithm that sets enterprise pricing is still being tuned, and we'd rather confirm your margin structure once, when it's real, than publish a number we might have to walk back later.

The program

How the program works. It follows the AWS model: toolbox, certification, directory. We're building it deliberately, and the first names in the directory are the ones buyers find first.

1. Certify. Certification on the Origami product stack: deployment modes, APIs and SDKs, the data residency host, the ecosystem loop. The first cohort of outside firms is in training now.

2. Get listed. Certified firms show up in the public integration directory, which is where customers come looking for implementation capacity. It's how a firm gets found by buyers who need hands.

3. Co-market. Co-marketing through the directory and joint materials. We tell the deployment story together, and the proof points in this brief are available to certified partners building their own case studies.

4. Sell and implement. Certified partners sell and implement Origami products without needing an ownership stake. Deal structure and partner economics get confirmed in the certification conversation itself.

Who this is for

We're keeping the program small on purpose.

Firms with genuine integration practice: data engineering, systems integration, OT and building systems, regulated environments. Teams that have shipped code against someone else's APIs and lived to tell it.

Firms that want a practice rather than a price list. The value compounds with every deployment you learn, and recurring revenue rewards the firms that stay.

Resellers should skip it. There's no margin-only motion here. If your firm does implementation work, this program is for you; if you're after volume licensing, it isn't.

Talk to our engineers first. Tell us what you integrate and who you've shipped for. Certification starts when your technical leads talk to ours: partners@origamiai.cloud. The full program page lives at https://origamiai.cloud/partner.