Cortex

The governed AI control plane.

Agents that build, run, and support your software, under your policies, on your infrastructure. The substrate for the systems you simulate before you build them.

ALPHA OMEGA LABS · ENTERPRISE BRIEFING · 2026

• THE BET

Code is a substrate. Like electricity, it changes everything the moment we modulate it.

SOFTWARE THAT SHAPES ITSELF AT RUNTIME, AND SIMULATES REALITY BEFORE WE BUILD IT

WHAT MODULATION IS FOR

Stand up a world. Ask it a question.

Intelligence got cheap. Finding out whether a proposal is true stayed expensive, and it is the only part anyone pays for. So the compute moves: from generating an answer to running the experiment that settles it.

01 · PROPOSE

Nearly free

A design, a molecule, a control policy, an entire service. Generation is commoditizing and the price keeps falling.

02 · TEST

The expensive half

One good hypothesis is ten thousand runs: the sweep, the ensemble, the adversarial suite, the rerun six months later when someone challenges the result.

03 · COMPOUND

It gets cheaper

Each validated result becomes something the organization keeps and reuses, so the next question costs less than the last one.

→ Demand is not saturating. It is **moving**, from generation to verification, and verification is the larger half.

THE CATCH

A substrate you can reshape can burn the building down.

Electricity waited on breakers, meters, and grounding before anyone ran it through a building where people live. Software that rewrites itself needs the same grid, and not having one is why capable AI still stalls in pilots.

GROUNDING

Isolation

A run that can reach anything proves nothing and risks everything. Every experiment needs its own bounded world.

BREAKERS

Authority

Something has to bound what an agent may do, what it may touch, and what it may spend, independently of who is driving it.

METERS

Evidence

A result nobody can reconstruct is not a result. Peer review and the AI Act ask for the same artifact.

→ Enterprises are not blocked on capability. They are blocked on **permission**.

THE PROBLEM

The model was never the problem.

Enterprise AI doesn't fail the technical evaluation. It fails the review that comes after, because a chat window can't answer the three questions that decide whether anything ships.

Q1 · DATA

Where is our data?

Which systems does the AI touch, where does context go, and who else can see it? Shadow AI is already inside the building.

Q2 · AUTHORITY

Who approved that?

When an agent deploys, installs, or deletes: on whose authority? Most platforms cannot produce a single accountable record.

Q3 · EVIDENCE

Why did it act?

Auditors and regulators ask for reconstruction: inputs, policies, approvals, outcomes. Retrofitting evidence is where projects die.

→ Cortex was built so these three answers exist **by construction**.

THE MARKET

Agentic AI is the fastest-scaling enterprise software wave.

Enterprise agentic AI market, USD · Grand View Research 2025 (46.2% CAGR; intermediate years interpolated at stated CAGR). Corroborating: MarketsandMarkets 2025, AI agents $7.8B (2025) → $52.6B (2030).

33%

OF ENTERPRISE SOFTWARE WILL INCLUDE AGENTIC AI BY 2028

Up from under 1% in 2024. Gartner, 2025

>40%

OF AGENTIC AI PROJECTS WILL BE CANCELED BY 2027

Costs, unclear value, **inadequate risk controls**. Gartner, 2025. The survivors will be the governed ones.

95%

OF GENAI PILOTS SHOW NO P&L IMPACT

MIT NANDA, 2025 · purchased platforms succeed ~2× as often as internal builds

→ Near term, the wedge is **governed delivery**. The survivors of that shakeout are the governed ones. Long term, the same substrate is where the **experiments** run.

WHY NOW · REGULATION

The EU AI Act has moved from policy debate to operating requirements.

AUG 2024
In force
The AI Act enters into force

FEB 2025
Prohibitions
Banned practices + AI-literacy duties apply

AUG 2025
GPAI rules
Obligations for general-purpose AI models

AUG 2026
Enforcement starts
Transparency rules and the main enforcement machinery come online

DEC 2027+
High-risk wave
Annex III systems, then regulated products in 2028 under the Omnibus timeline

WHAT DEPLOYERS MUST SHOW

Inventory and classification · data governance (Art. 10) · automatic event logging (Art. 12) · transparency (Art. 13/50) · **human oversight in operation** (Art. 14) · technical documentation.

THE COST OF WAITING

Penalties remain material: up to **€35M or 7%** for prohibited practices and up to **€15M or 3%** for most operator violations. Yet **over half of organizations still lack a basic AI inventory**, the first artifact every program needs.

Sources: European Commission AI Act page + AI Act Service Desk timeline · EUR-Lex Regulation (EU) 2024/1689 Art. 99 · Cloud Security Alliance readiness note, 2026. Full notes: pitch/research/eu-ai-act.md

WHY NOW · SOVEREIGNTY

European enterprises are pulling AI back inside the perimeter.

40%

OF EUROPEAN ORGANIZATIONS ADOPTED SOVEREIGN CLOUD IN 2025

Up from ~30% in 2023–24, with another 31% planning. IDC Digital Sovereignty Survey, 2025

56%

RUN OR PLAN PRODUCTION AI INFERENCE ON PRIVATE CLOUD

vs 41% on public cloud; 83% weigh cloud repatriation. Broadcom Private Cloud Outlook, 2026

54%

NAME DATA SOVEREIGNTY THE TOP INFRASTRUCTURE DRIVER

First time ahead of jurisdiction compliance (51%). Broadcom, 2026

Cortex is one of the few platforms where **cloud and fully on-premise are the same product**, identity, memory, observability, even voice transcription run locally. Model calls go only to the providers you configure, under your policies.

Sources: IDC Worldwide Digital Sovereignty Survey 2025 · Broadcom Private Cloud Outlook 2026 · Grand View Research 2025 (Europe sovereign-cloud market ~$56B in 2025). Full notes: pitch/research/sovereignty.md

THE PATH

Simulating reality is a decade of work. Here is the order it has to happen in.

A substrate nobody is allowed to run is worth nothing, so permission comes first. Everything after this slide is one of these four layers.

01 · PERMIT

Make it allowed

Capability policy, approval gates, audit by construction, sovereign deployment, and the records an AI Act review asks for. Nothing else gets used until someone can say yes.

02 · OPERATE

Make it real

One control plane running agents, deploys, memory, channels, and observability on your infrastructure. A substrate has to be operated, not just described.

03 · PROVE

Make it checkable

An open benchmark, plus compliance evidence produced by normal operation rather than retrofitted for an audit.

04 · COMPOUND

Make it pay off

The research loop: question, governed run, evidence, reusable capability. This is where simulation actually lands.

→ Layers 01 to 03 are the shipped control plane. Layer 04 is where our prototypes and current R&D sit, and it is the reason the first three exist.

WHAT CORTEX IS

One governed control plane between your people, your agents, and your infrastructure.

→ Every action, on every surface, passes the same policies and lands in the same audit trail and memory.

CAPABILITIES

Agents that do real jobs, each under policy.

CODING

Governed Claude / Codex harnesses in the terminal, the IDE, and CI, plan, edit, test, and ship under capability policy.

DEPLOYMENT

Prompt-to-production or dock what you already run. Build → test → release with rollback armed, approval-gated.

SUPPORT

Tickets and events land in one inbox; agents run root-cause analysis, propose remediation, and fix, humans approve.

MONITORING

OpenTelemetry logs, traces, metrics, and session replay from every app and runtime, no vendor wiring.

DEVOPS

Operate VMs, containers, and clusters from one surface, restarts and host actions gated and audited.

RESEARCH

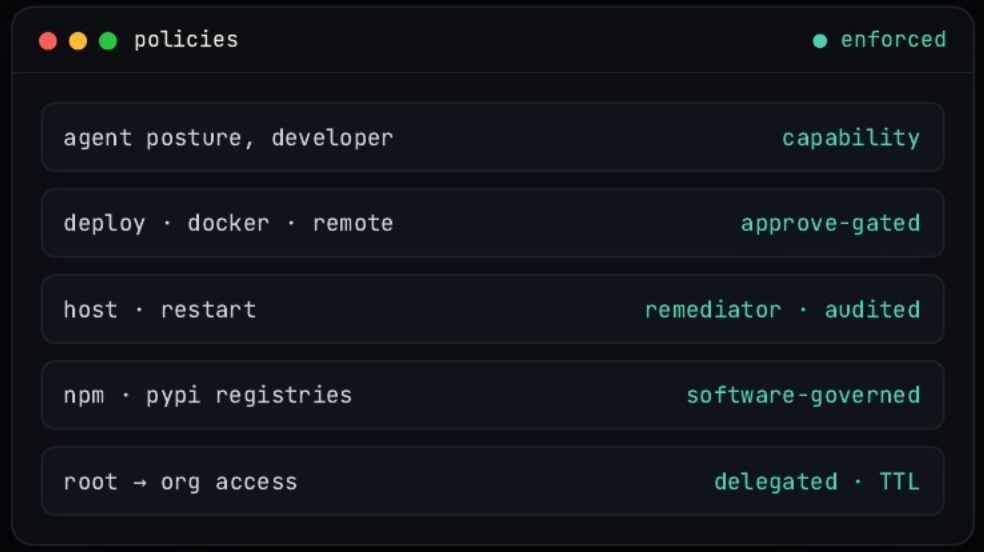
Training runs, GPU provisioning, and experiment tracking, the same platform runs our own R&D stack.

THE CORE IDEA

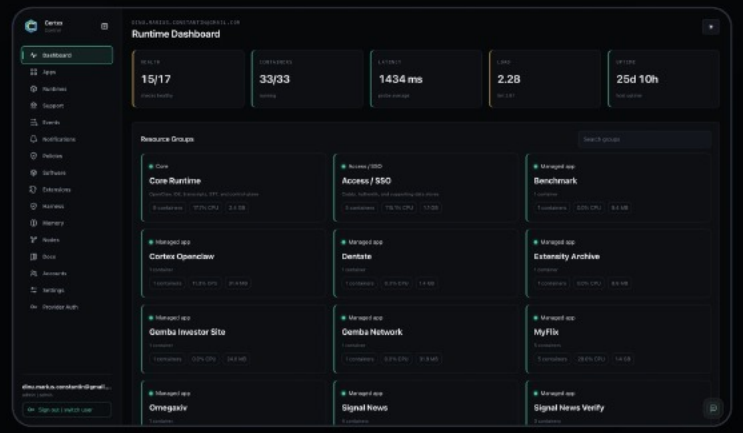
Authority comes from policy, never from whoever is driving.

Electricity only became infrastructure once it shipped with breakers, meters, and grounding. Code that reshapes itself needs the same grid. This is it, not paperwork.

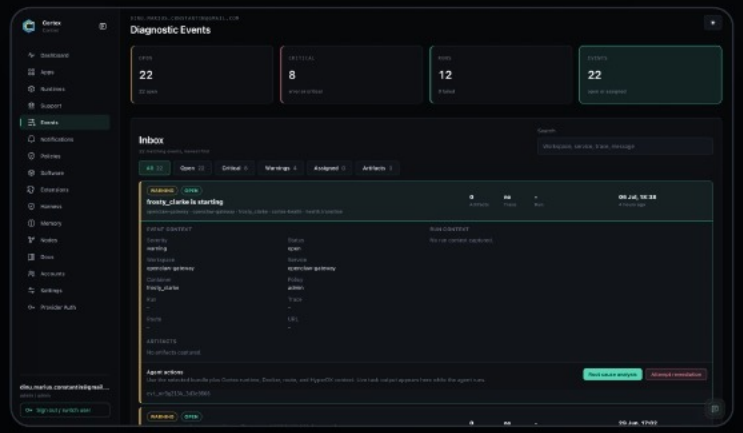
- **Two planes.** Human identity (SSO, roles) is separate from agent capability (policies). No one quietly escalates an agent's authority.
- **Capability ladder.** observer → agent → developer → remediator → admin. Risky actions stop at approval gates on every surface.
- **Software governance.** What agents install is checked against allowed registries, publishers, and licensing policy.
- **Audit by construction.** Every capability check, approval, and denial recorded with actor, target, scope, timestamp.



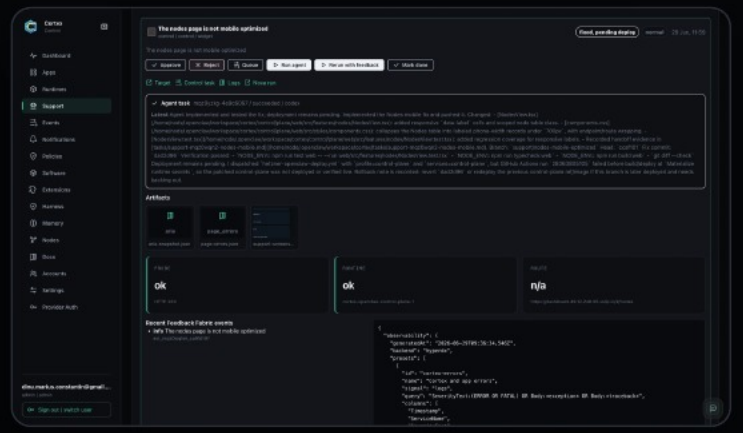
The corporate full stack for R&D and operations.



Runtime dashboard, health, containers, and live metrics for every resource group



Diagnostics, agents run **root-cause analysis** and propose remediation



Support, an agent fixed the ticket, with artifacts, probes, and rollback notes

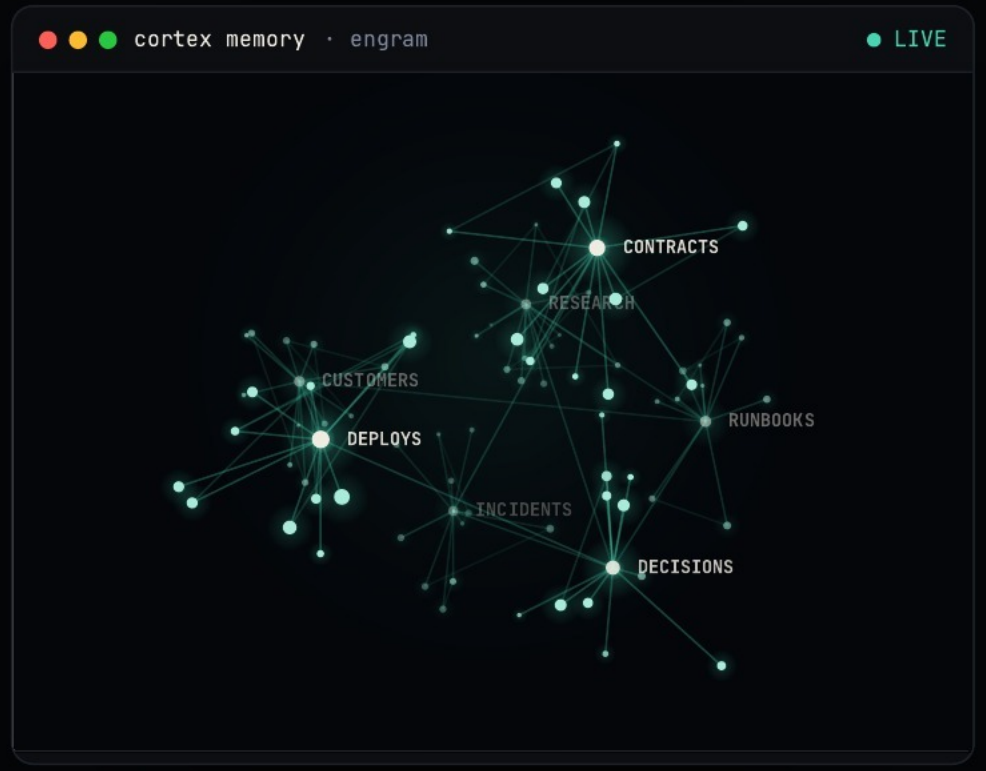
→ Live console of a production stack, not mockups. A full walkthrough is part of the pilot.

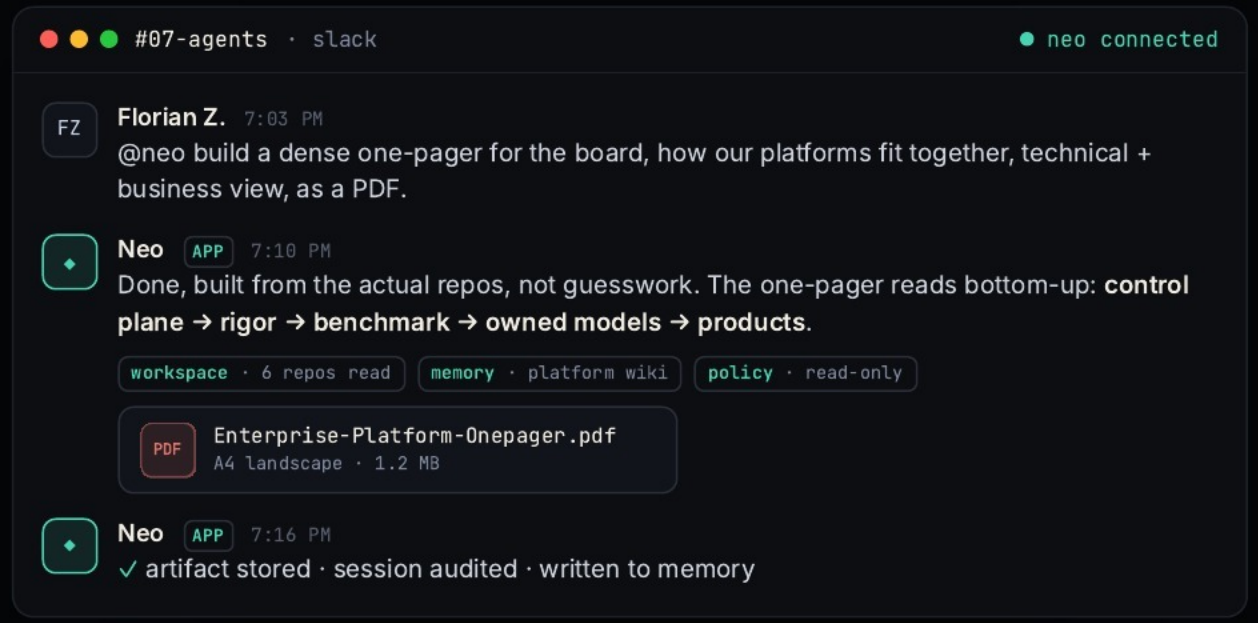
ENGRAM · ORGANIZATIONAL MEMORY

A platform that remembers how your company works.

- **Capture.** Every governed session, threads, coding runs, tickets, distills into a private knowledge graph on your stack.
- **Consolidate.** A gardener agent merges duplicates, links decisions to incidents, and mines repeated work into reusable skills.
- **Recall.** Hybrid lexical + semantic + graph search from Nova, the CLI, or Slack, every answer carries its sources.

Institutional knowledge stops leaving when people do. **It compounds.**





WHERE YOUR TEAMS ALREADY TALK

A thread is a governed session.

- Slack, Telegram, Discord, wired up at signup; name your agent whatever you like
- Grounded in your repos and memory, answers carry sources
- Real deliverables in-thread: documents, diagrams, pull requests
- Approvals and audit inside the conversation, same policy as every surface

MEASURED, NOT CLAIMED

The same model, run inside Cortex, is safer and more capable.

Our open **Gauntlet** benchmark runs an identical frontier model twice, raw, then inside Cortex's governance, across five kinds of real software work. Every run is reproducible.

98/100

GOVERNED SCORE, UP FROM 88 RAW

−47%

JAILBREAKS RESISTED

Malicious-prompt attacks cut by nearly half

+21%

MORE REQUIREMENTS MET

Spec coverage 81% → 98%

+72%

BETTER WHOLE-REPO BUILDS

Entire projects, scored end to end

Paper: **“Cortex: A Fixed-Point Theory of Governed Coding Agents”**. Dinu & Zeba, Alpha Omega Labs, 2026 · results + reruns public

EU AI ACT READINESS

The evidence your auditors ask for, as a by-product of operation.

Art. 10 · Data governance	Dedicated stack keeps data, retrieval, and telemetry inside your boundary, cloud or on-premise.
Art. 12 · Record-keeping	Approvals, denials, and agent actions logged with actor, target, scope, timestamp, automatically.
Art. 14 · Human oversight	Approval gates enforced by the runtime on every surface. Control, CLI, even the Slack thread.
Annex IV · Documentation	Policies, audit records, benchmark results, and reconstructable session trails, no backfilling.

Honest note: no platform makes you compliant by itself. Cortex makes the evidence fall out of normal operation, we work alongside your compliance team.

THE LANDSCAPE

Everyone sells agents. Almost no one owns the operation.

CAPABILITY	CORTEX	ASSISTANT SUITES[1]	FRAMEWORKS & CODING AGENTS[2]	SOVEREIGN MODEL VENDORS[3]
Full on-premise, agents, memory, observability, the whole stack	✓ native	~ cloud-only[1a]	~ self-host varies, DIY ops	~ model & platform only
Capability policies with in-conversation approval gates	✓ every surface	~ admin & DLP controls	× build your own	×
Audit trail per action: actor, target, scope, timestamp	✓ by construction	~ partial logs	× DIY	~
Owns delivery: build → test → release → rollback of your software	✓	×	~ code only, not ops	×
Persistent organizational memory with provenance	✓ Engram	~ RAG add-ons	× DIY	×
Agents native in Slack / Telegram / Discord threads	✓	~ own suite first	×	×
Open benchmark: governed vs raw, reproducible	✓ published	×	×	×

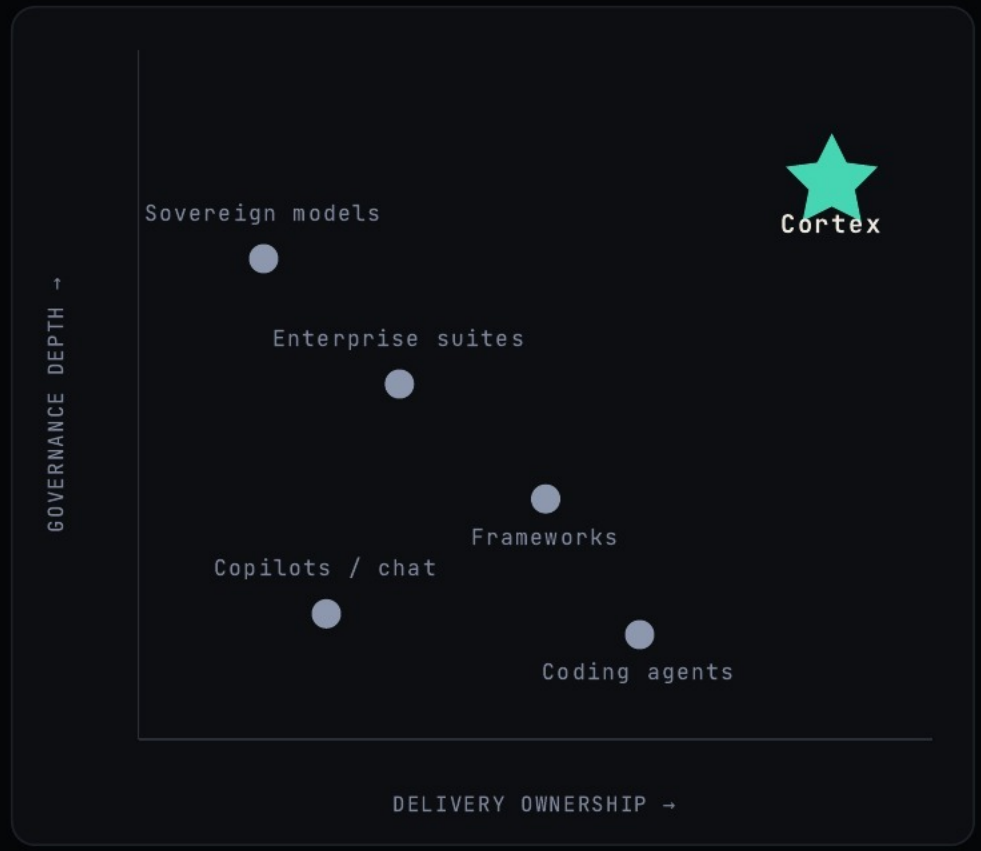
✓ native · ~ partial / add-on / do-it-yourself · × absent [1] ChatGPT Enterprise, Microsoft Copilot, Gemini Enterprise ([1a] Google offers on-prem via Distributed Cloud) [2] Devin, LangGraph, CrewAI [3] Aleph Alpha, Mistral

Assessment from public vendor documentation, July 2026 · full per-vendor notes and URLs: pitch/research/competitors.md

POSITIONING

Deep governance and full delivery ownership, the quadrant is empty up there.

+ **Assistants & copilots** stop at suggestions, you still own integration, deployment, and evidence.
+ **Frameworks** hand you parts, governance, audit, and operations become your engineering project.
+ **Sovereign model vendors** solve residency for the model, not the agents, memory, or delivery around it.
+ **Cortex** ships the whole governed operation: agents, approvals, audit, memory, deploys, observability, cloud or on-premise.



THE LOOP THIS WAS BUILT FOR

Governed delivery is the wedge. The research loop is the reason.

Same control plane, same policies, same audit trail, extended until a question becomes a run, a run becomes evidence, and evidence becomes a capability you keep. omegaXiv and Quarks are current prototypes.

01 · FRAME · PROTOTYPE

omegaXiv

Question → reviewable project

Problem intake, private or public work, papers, sources, code, artifacts, and reruns.

→

02 · RUN · PROTOTYPE

Quarks

Project → evidence

Contract-validated phases, experiments, simulations, traces, and publication packages.

→

03 · ENCODE

Sema

Evidence → executable model

Equations, contracts, effects, assurance, and replay in one AI-native language.

→

04 · DELIVER

Cortex

Candidate → operated system

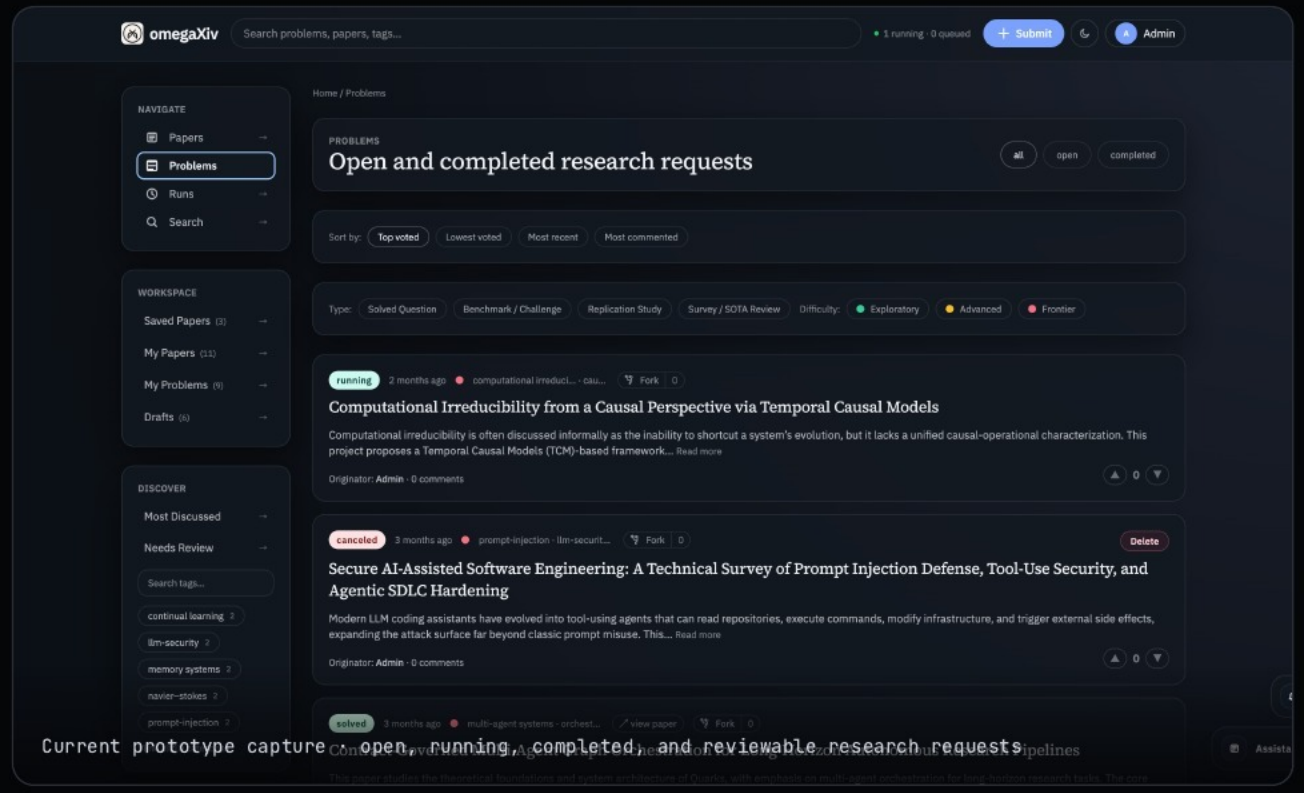
Governed agents, memory, code, deployment, observability, approvals, and audit.

CORTEX CONTROL PLANE · identity · policy · Engram memory · infrastructure · audit

Adopt only the layers your workflow needs.

PROTOTYPE RESEARCH WORKSPACE + EXECUTION ENGINE

omegaXiv makes research legible. Quarks makes it executable.



Current prototype capture: open, running, completed, and reviewable research requests

omegaXiv

PROTOTYPE · FRONT DOOR + DURABLE RECORD

Structured problem intake; autonomous, semi-autonomous, or manual work; private workspaces and public discovery; papers, review, artifacts, and reruns.

Quarks

PROTOTYPE · EXECUTION + ORCHESTRATION

A manifest-driven, contract-validated phase pipeline for experiments and simulations, with provenance, structured traces, and Git-ready result packages.

One lifecycle, distinct responsibilities. Teams inspect and review in omegaXiv; Quarks executes underneath. Contracts validate the run, not scientific truth.

Sources: alphaomega-docs/pitch/Alpha_Omega_Labs_Competitive_Analysis_Deck_Dark.html
alphaomega-docs/quarks-docs/deepwiki.md · alphaomega-docs/quarks-docs/service-api.md

SEMA · RESHAPE THE CODE WITHOUT LOSING THE PROOF

When “it ran” is not enough, make the model and its evidence executable.

models.sema · verified project syntax

```
equation raw_additive_hybrid_value(symbolic_value: f64, learned_correction: f64, gate:
f64) → f64:
    return symbolic_value + gate * learned_correction

def additive_hybrid_value(symbolic_value: f64, learned_correction: f64, gate: f64) !{}:
    require gate ⩾ 0.0 and gate ⩽ 1.0
    return raw_additive_hybrid_value(symbolic_value, learned_correction, gate)
```

equation
Symbolic and numerical model surfaces are language constructs, not prompt conventions.

require · ensure
Bounds, invariants, and evidence gates execute with the program.

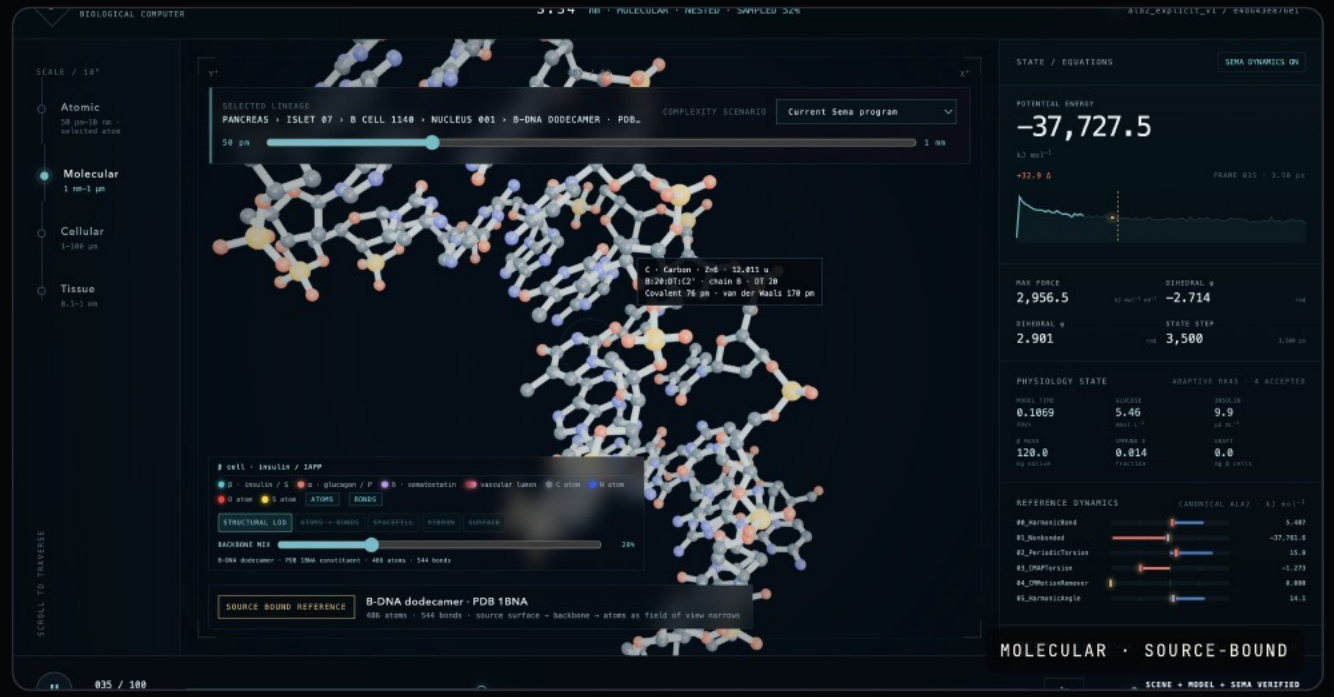
!{effects}
Capabilities are typed, making the Cortex policy boundary explicit when the systems are integrated.

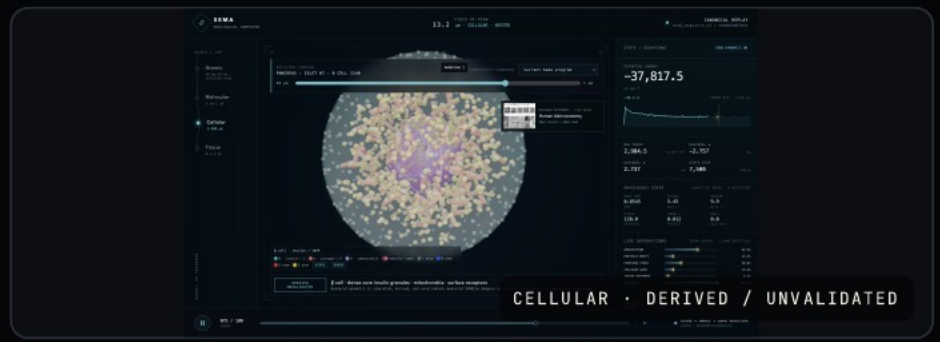
assure · replay
Tests, traces, and deterministic inspection stay attached to the run.

Python-shaped syntax, backed by a Rust reference runtime. Other languages remain supported; Sema is for workflows where the proof boundary matters.

Sources: sema/docs/LANGUAGE.md · sema/examples/biological-computer/src/models.sema

Prototype: one program carrying identity and evidence across scientific scales.





RUNNABLE TECHNICAL VERTICAL

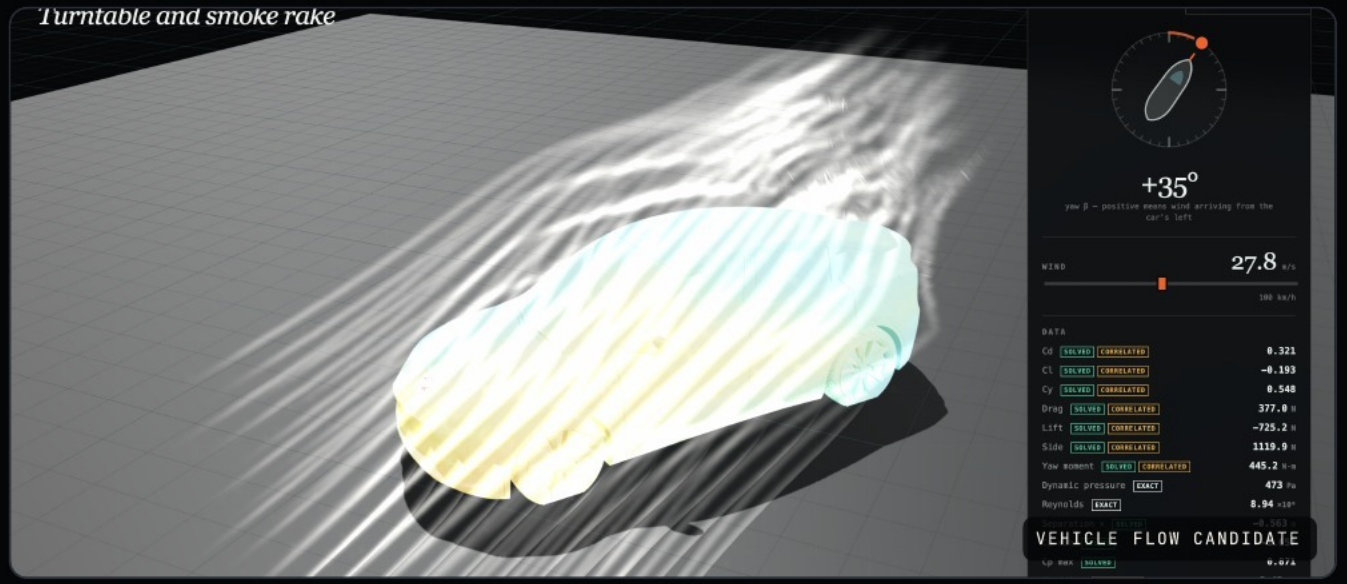
Source-bound structures, simulator adapters, state transitions, and replay stay linked as the view changes scale.

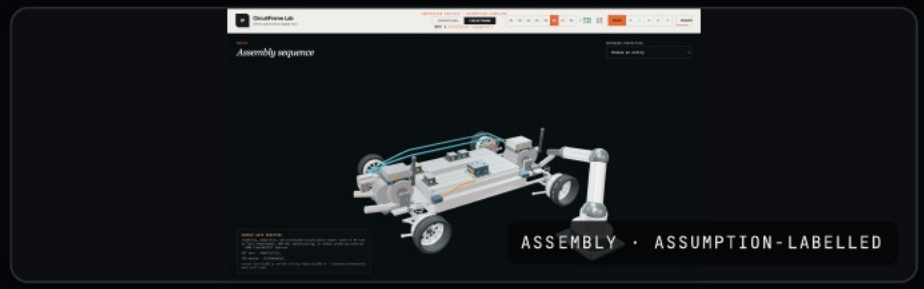
QUALIFICATION STAYS VISIBLE

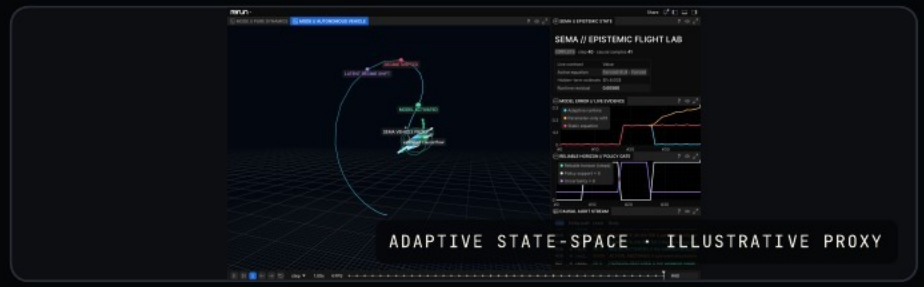
Derived geometry and unvalidated science are labelled explicitly. Technical execution is not clinical or scientific admission.

Source: sema/docs-site/src/content/docs/reference/examples-api/biological-computer.md · prototype captures

The same evidence discipline can move from simulation into development.







MODEL
Equations, bounds, assumptions

SIMULATE
Traceable candidate behavior

REVIEW
Evidence and uncertainty gates

DELIVER
Code, deploy, observe in Cortex

Demonstrated technical prototypes: reduced/surrogate simulation, assumption-labelled reconstruction, and an illustrative state-space proxy, not OEM CAD, governing CFD, robot or plant physics, homologation, clinical, or safety authority.

Sources: sema/docs-site/src/content/docs/reference/examples-api/magna-ev-digital-twin.md · sema/pilot/SHOWCASE.md

INDICATIVE B2B MODEL

One platform license. Clear ownership. Specialist work when needed.

Validate with a fixed outcome, operate Cortex under an annual license, then extend into full R&D without replacing the control plane.

01 · VALIDATE

€45–90k

fixed · 6–12 weeks

- Dedicated governed pilot environment
- One bounded priority use case
- Acceptance, security & KPI baseline

02 · OPERATE

from €25k / month

annual platform license

- Cortex + dedicated deployment
- Updates, operations & standard support
- Cloud, model and compute usage separate

03 · EXTEND R&D

from €55k / month

scoped annual platform + R&D program

- Full Cortex + omegaXiv and Quarks prototypes + Sema toolchain
- Joint roadmap + priority capacity
- Deliverables, SLA and usage set in SOW

BESPOKE ENGINEERING

€175 / hour

Pre-approved cap or fixed-scope work package for integrations, models, simulations and custom SLAs.

CONTRACTUAL IP SPLIT

Platform IP

Alpha Omega retains Cortex, background and reusable platform IP.

Customer property

Customer retains its data; upon payment, it owns SOW-named project deliverables.

Boundaries

Embedded platform and third-party rights remain; exclusivity or assignment is separate.

Indicative · ex VAT · non-binding · annual terms and final IP schedule in contract

DEPLOYMENT & HOSTING

One Cortex product. Three hosting models.

Choose where Cortex runs, inside your perimeter, in a dedicated private cloud, or Alpha Omega-hosted. Independently, use local models, configured public APIs, or both; hosting, support, and usage terms vary.

01 · FULL CONTROL

Fully on-premises

Customer infrastructure

CORTEX → LOCAL

- Existing VMs, Docker hosts or Kubernetes
- Identity, memory, observability and transcription stay local
- Outbound model calls only when configured

Customer hosts · Alpha Omega installs, migrates and supports

02 · DEDICATED

Private cloud

Customer VPC, sovereign cloud or ours

CORTEX → PRIVATE

- Isolated deployment and data boundary
- Customer-owned cloud tenancy or Alpha Omega hosting
- Managed updates and operations available

Customer or Alpha Omega operates

03 · FASTEST START

Hosted / public API services

Managed Cortex · public APIs available

CORTEX → APIs

- We host and operate the Cortex control plane
- Connect configured public APIs, private models or both
- Requests sent to public providers leave the Cortex boundary and follow their residency, processing and usage terms

Alpha Omega hosts · customer approves providers

SAME PRODUCT IN EVERY MODE

Identity · policy · Engram memory · approvals · audit · APIs

SETUP OR MANAGED OPERATION

We architect, install, migrate and operate, on your infrastructure or ours.

Deployment changes responsibility, data path and commercial terms, not Cortex capability

FOR INFRASTRUCTURE & MODEL PARTNERS

You supply the engines. We supply the layer that lets enterprises run them.

+ **We ship no engine and no model.** Cortex is engine-agnostic by construction. Simulators, solvers, provers, and frontier models attach as governed connectors.
+ **The missing piece is the operating layer.** Who may launch a run, against which data, under what budget, and what record survives it. Today that is bespoke glue rebuilt per customer, and inadequate risk controls are among the reasons Gartner expects >40% of agentic AI projects to be cancelled by 2027 (slide 04).
+ **Every environment is demand.** Each governed workspace provisioned is a new, auditable reason to buy compute: capex converted into utilization with a hypothesis attached.
+ **We reach where platforms structurally cannot.** Inside the customer perimeter, under EU AI Act obligations, fully on-premise as the same product, not a downgraded tier.

THE ENGINES

Physics, robotics, molecular, quantum, and frontier-model stacks, from the partners who already own them.

THE GAP

Isolation, authority, budget, provenance, residency. Rebuilt by hand at every enterprise, or the project stalls.

CORTEX

Provision the environment, broker its reach under policy, audit the run, keep the result. One substrate, many engines.

→ The world will be simulated before it is built. We are building where those simulations run.

WHAT YOU GET

Value from week one.

+ **Fast path to a live stack**, provisioned with Alpha Omega support, on our cloud or your perimeter.
+ **No replatforming**, connect the VMs, Docker hosts, and clusters you already run.
+ **One governed surface**, coding, deploys, support, monitoring, DevOps under one policy and one audit trail.
+ **Evidence by construction**. AI-Act-aligned records produced by normal operation.
+ **Compounding memory**, your organization's knowledge stays, links, and grows.

HOW WE START

Validate, then commit.

WEEKS 1–2 · SCOPE

Pick one priority workflow (e.g. support RCA or governed coding). Security & compliance walkthrough on a live stack, bring your compliance team.

WEEKS 3 ONWARD · VALIDATE

Complete the 6–12 week engagement on a dedicated stack. Real workflows, deploys, and audit trails, measured against your KPIs.

THEN · ROLLOUT

Extend policies, channels, and memory across teams. Enterprise plan with multi-node topology, delegation, and SLAs.

a2olabs.com · dashboard.cortex.a2olabs.com/cloud