

PRIVATE AI FOR THE ENTERPRISE

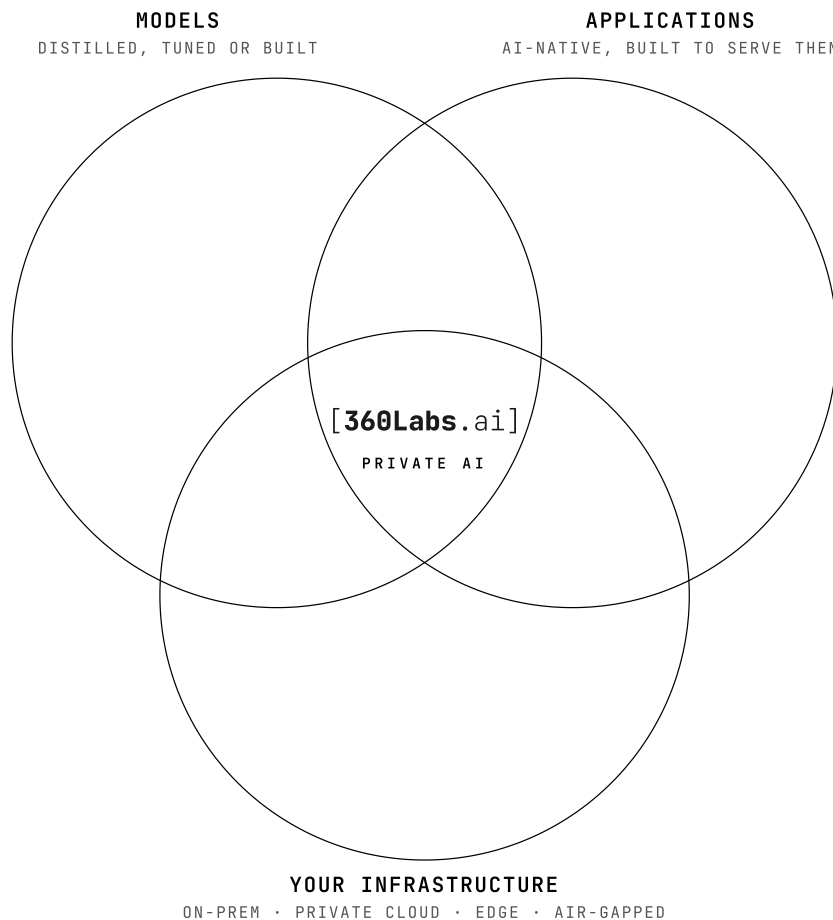
AI ENGINEERING & RESEARCH LAB
MODELS · APPLICATIONS · SYSTEMS
DEPLOYED ON YOUR INFRASTRUCTURE

COMPANY BRIEF & CASE STUDIES
COPYRIGHT ©2026
LIGHTYEAR 360 LABS PVT LTD

360Labs

What is 360 Labs?

One of the world's first AI Engineering & Research Labs, focused on deploying private AI systems for large enterprises.



WHAT WE DO

01

Distill & fine-tune open models

Turn strong open models into specialists for one domain, so a smaller model you own does the job of a large one you rent.

02

Build small & medium language models

Trained from scratch when the language, domain or security bar calls for a model that is entirely yours.

03

Build AI-native applications & platforms

The software that serves those models to your people, inside the workflows they already use.

We envision a world where AI models & applications are as democratised as operating systems.

What Is Private AI, and Why Are Large Enterprises Adopting It?

Most AI today runs on public cloud, with intelligence routed through APIs from providers like OpenAI and Anthropic.

For regulated industries such as banking, finance, law, government, healthcare and real estate, that route rules out much of what AI can do. **Private AI** runs the intelligence on infrastructure the enterprise controls.

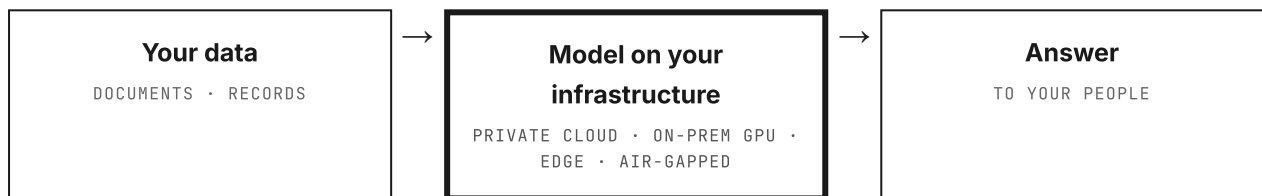
PUBLIC CLOUD AI

DATA LEAVES YOUR PERIMETER



PRIVATE AI

DATA NEVER LEAVES



WHY ENTERPRISES ARE ADOPTING IT

01

Data security & regulatory compliance

Sensitive and regulated data stays inside your perimeter, so it can finally be put to work with AI.

02

A hedge against frontier-lab token cost inflation

Costs track the compute you own or lease, not per-token prices a vendor can raise.

03

Ownership of AI-related tech IP

The models, fine-tunes and systems built on your data belong to you, not to a provider.

Same intelligence. Your infrastructure, your data, your IP.

Scale & reach

A global engineering footprint across 4 continents, building for Nifty-500 and Fortune 500 clients, with private-AI deployments live across India, the UK, the GCC and the USA.

DELIVERY

50+

CUSTOMERS

30+

DEPLOYMENTS

RESEARCH & TEAM

100+

AI PRODUCTS

10+

PAPERS

20+

AI ENGINEERS



GLOBAL FOOTPRINT

12+

CITIES

7+

COUNTRIES

4+

CONTINENTS

24/7

COVERAGE

VANCOUVER • SAN FRANCISCO • HOUSTON • TORONTO • LONDON • MANCHESTER • DOHA
DUBAI • ABUJA • DAR ES SALAAM • DELHI • MUMBAI • HYDERABAD • BENGALURU

TEAM FROM & AFFILIATED WITH



WORK DELIVERED FOR



Who are we?



Prithvi Raj Agrawal CO-FOUNDER & CEO

BSc Mathematics from the University of Toronto, followed by **BSc IT and Digital Innovation** at City University of London, specialising in AI, machine learning and digital product design.

Previously a core member of the product and growth team at **DeepReel.com** (\$10M+), leading enterprise sales and delivery for **Accenture** and the **Aston Martin Aramco F1 Team**. Co-founded and raised funding for 2 AI SaaS startups before founding 360 Labs.



Saurabh Kumar CO-FOUNDER & CTO

Hardware engineer and systems programmer with 12 years across government-scale and consumer-scale systems. Built real-time instrumentation for the **Indian Meteorological Department**. Engineered core systems behind **IDFC Bank's** customer banking application. Architected **SLM360**, India's first **sub-50MB** language model, running on-device at **39ms latency**.

Indian patent holder in sensor-based measurement and a published author, with **Med360** and **AgentGuard** released publicly. **Quantum in Machine Learning and AI Systems**, IIT Delhi.

Team Info

NAME AND ROLE

BACKGROUND

Akshat Bhatia

Regional Partner & Chief Legal Officer

Lawyer and legal technologist. Qualified as **Solicitor (England and Wales)** from **Herbert Smith Freehills Kramer** (London), advising **UBS, Credit Suisse and Morgan Stanley** on roughly **\$10Bn of transactions and disputes**. Founded the firm's **Digital Committee**, shipping AI tooling into live legal matters including various **\$1Bn+ legal matters**. Involved in procurement and pilots on **Harvey, Legora, Orbital** and other AI vendors. **LLB, LLM, MSc AI** from various UK universities.

Suryansh Srivastava

Lead Engineer & AI Researcher

Previously **Senior Software Associate** at **Capital One**, where he led backend engineering for a platform launched across **North America** at **NADA '26**. Received multiple **Director awards** and **Team awards**. B.Tech and M.S. by **Research in Quantum and AI** from **IIIT Hyderabad**, with a first-authored publication in the **Journal of Physics A**, a **patent** in progress, and **AWS Certified Solutions Architect**.

Rahul Gupta

Chief of Staff & Lead Project Manager

Technical Product & Project Leader with experience building **0→1 SaaS platforms** across fintech, CRM, ERP and cybersecurity. Previously Sole Product Lead at **MortgagX**, an early **AI-native UK mortgage platform** that engaged with major institutions including **Lloyds** and **Nationwide Building Society**. Delivered **200+ production releases** with **zero critical defects** in a year and **300+ third-party integrations** across diverse business domains.

Led a critical cybersecurity programme for a **Belgian government-services cybersecurity firm Sertalink**. Previously managed operations across the **UK and Europe** at **Van Technologies**, where he designed and delivered one of the **UK's first online load-planning systems** for the removals industry.

Samyak Jain

Senior Full-Stack AI Engineer

Previously **Senior Engineering Leader** with **5+ years at Red Hat**, where he led **Fedora Linux Release Engineering** for **1.2M+ users**, overseeing infrastructure processing **50,000+ packages daily** across **8 major Fedora releases**. **Founding member of Konflux**, a Red Hat platform powering **OpenShift**, one of the world's leading enterprise cloud platforms.

Recipient of **2 Red Hat Impact Awards**, with **10+ years contributing to Debian** and mentoring **GSoC & Outreachy** projects. A **published IEEE researcher in computer vision**, and speaker at leading global technology conferences including **FOSDEM, Flock, DevConf, GNOME Foundation and FOSSASIA**. **CNCF-certified, Certified Product Owner from Scrum Alliance, Magna Cum Laude graduate**, with prior experience at **Nokia** and an **MBA from NMIMS**.

Aditya Malik

Full-Stack AI Engineer

Full-Stack AI Engineer with **5+ years** deploying machine-learning systems into production. Experience spans AI applications, **retrieval-backed systems, AWS infrastructure** and secure enterprise platforms, with contributions to **Neovim, Palantir's Blueprint, Hermes Agent, Alacritty, OpenAI, Unsloth, ngrok** and **Open Jarvis**.

Nityam Tulani

Full-Stack AI Engineer

Full-Stack AI Engineer specialising in **data ingestion, retrieval and multimodal AI**. Selected for **Amazon ML Summer School** and **McKinsey.org Forward**, placed **2nd nationally at the Samsung PRISM Hackathon**, and contributes to **Gemini open source** and **Google Summer of Code**.

Case Studies

Technical deep-dives into private AI: models, applications and systems deployed on our clients' own infrastructure, across industry, government, finance and research.

[CLICK TO SEE SECTION]

SECTION 01 Client Builds

01	Government	02	Logistics
03	Manufacturing	04	Defense
05	Real Estate & Construction	06	Energy & Infrastructure
07	Healthcare	08	Banking & Finance
09	VC & PE Funds	10	Aviation & Aerospace
11	Legal		

SECTION 02 AI Research

12	AI Research
----	-------------

SECTION 03 Hardware, IoT & Robotics

13	Hardware & IoT	14	Robotics
15	Meteorology		

SECTION 04 Cybersecurity

16	Cybersecurity
----	---------------

FULL PORTFOLIO

View full at 360labs.ai/portfolio



SECTION 01

Client Builds

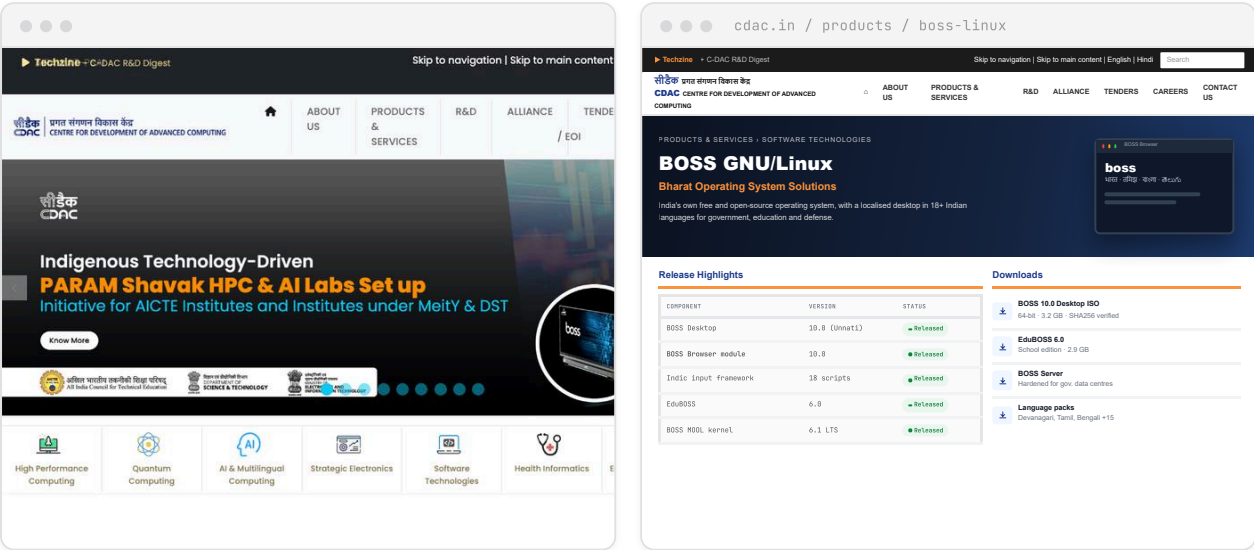
Client production AI shipped to clients across the world, deployed privately on the clients' own infrastructure.

Sovereign Browser for India's National Operating System (BOSS Linux)

Client: Centre for Development of Advanced Computing (CDAC), Govt. of India

We built the browser that ships inside India's own national operating system, so millions of government machines can get online without depending on any foreign software vendor.

- Government
- Sovereign Software
- Open Source
- Linux
- National OS



CLIENT PROBLEM

India's government machines needed a browser that was part of a nationally owned OS, not a foreign product that could change terms, harvest data, or be switched off from abroad.

OUR SOLUTION & ARCHITECTURE

We developed and contributed the **browser module** for BOSS Linux (Bharat OS), C-DAC's Debian-based national OS. It integrates with the BOSS desktop, meets government IT-security standards, and renders 18+ Indian scripts.



WHY IT MATTERS

Sovereign by design	Core software owned and controlled inside India
No foreign dependency	Runs without any overseas vendor in the path
Language inclusion	18+ Indian scripts rendered natively

BENEFITS REALISED

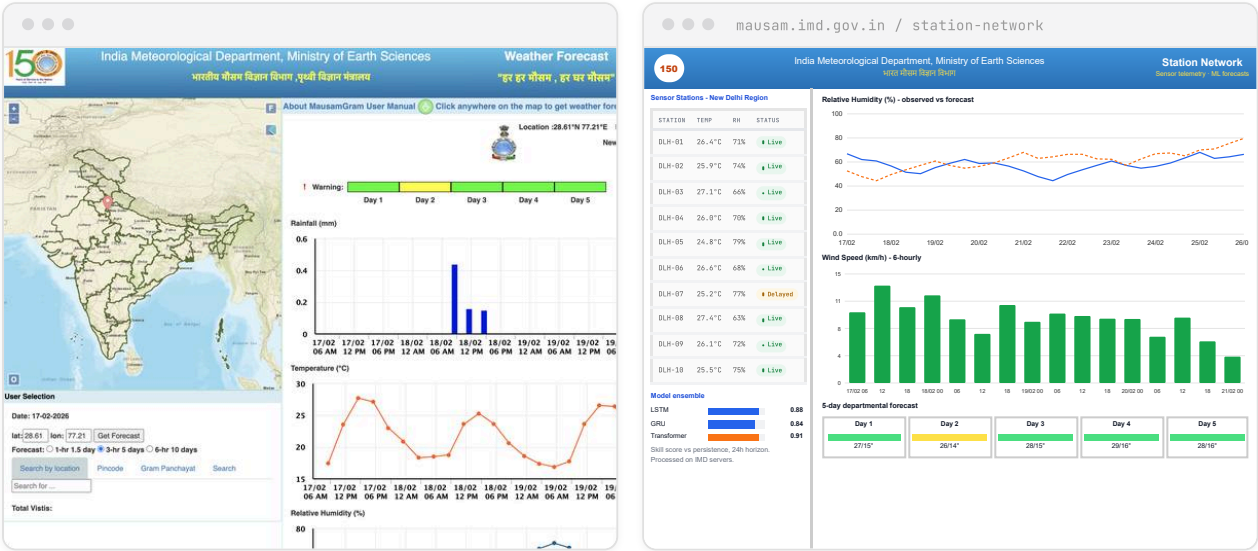
6M+	Government, education & defense devices
18+	Indian languages supported
100%	Open-source, auditable codebase

On-Premise Weather Sensing & Forecasting for a National Weather Office

Client: Indian Meteorological Department (IMD), Govt. of India

We built the weather sensors and the forecasting AI for India's national weather office, and all of it runs on the department's own machines rather than a foreign cloud.

- Government
- ML / Deep Learning
- Time Series
- IoT Sensors
- On-Prem

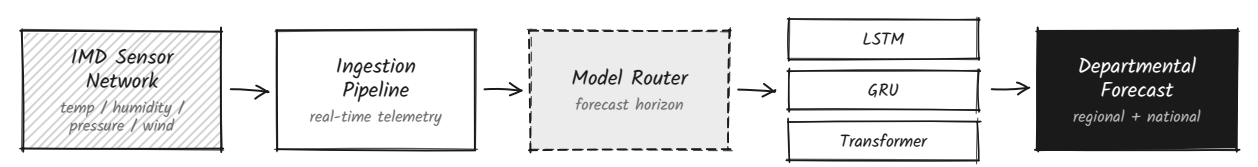


CLIENT PROBLEM

National weather forecasting cannot depend on sending raw sensor telemetry to third-party clouds, and off-the-shelf global models do not resolve India-specific conditions well.

OUR SOLUTION & ARCHITECTURE

We delivered custom **sensor hardware** for environmental data capture and an **ML forecasting stack** (LSTM, GRU, Transformer ensembles) running on IMD infrastructure, with a real-time ingestion pipeline and multi-horizon forecasts at the departmental level.



WHY ON-PREM

Data stays departmental	Sensor telemetry never leaves IMD systems
India-tuned models	Trained on local conditions, not generic globals
Government-grade uptime	Reliability and accuracy to national standards

BENEFITS REALISED

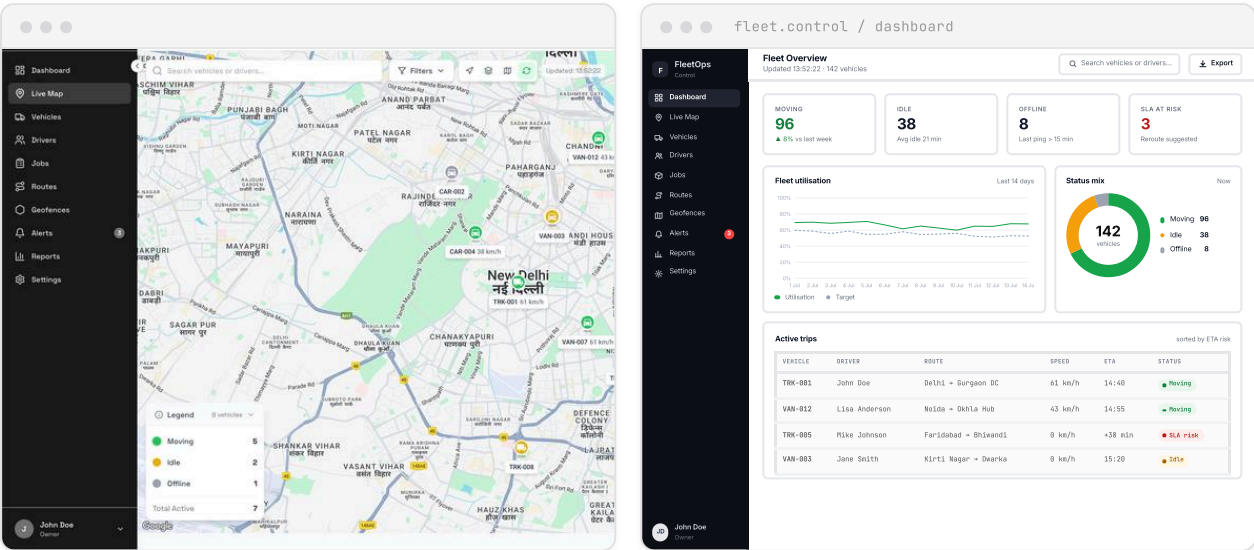
Multi-horizon	24-72h departmental forecasts
Real-time	Live sensor-to-forecast pipeline
100%	Processed on IMD infrastructure

Fleet Management Platform

Client: One of the World's Largest Logistics Companies (Undisclosed Under NDA)

We gave a global logistics operator one live map of every truck, driver and route, so dispatchers stop guessing where their fleet is and start running it from a single screen.

- Logistics
- GPS Tracking
- Route Optimization
- Real-Time
- Private Deployment



CLIENT PROBLEM

There was no centralized system for tracking vehicles or dispatching jobs. Drivers were unreachable, routes were planned by hand, fuel was wasted, and SLAs slipped as the fleet scaled.

OUR SOLUTION & ARCHITECTURE

A 3-layer platform: a **web dashboard** with a live map, an **optimization engine** for route planning, and a **mobile driver app**. Real-time GPS, geofencing, intelligent dispatching, and fleet analytics, hosted inside the client's own environment.



WHY PRIVATE AI

Fleet data stays in-house	Vehicle and route data on the client's infrastructure
No vendor lock-in	Owned stack, not a rented tracking SaaS
Works at scale	Built for thousands of daily dispatches

BENEFITS REALISED

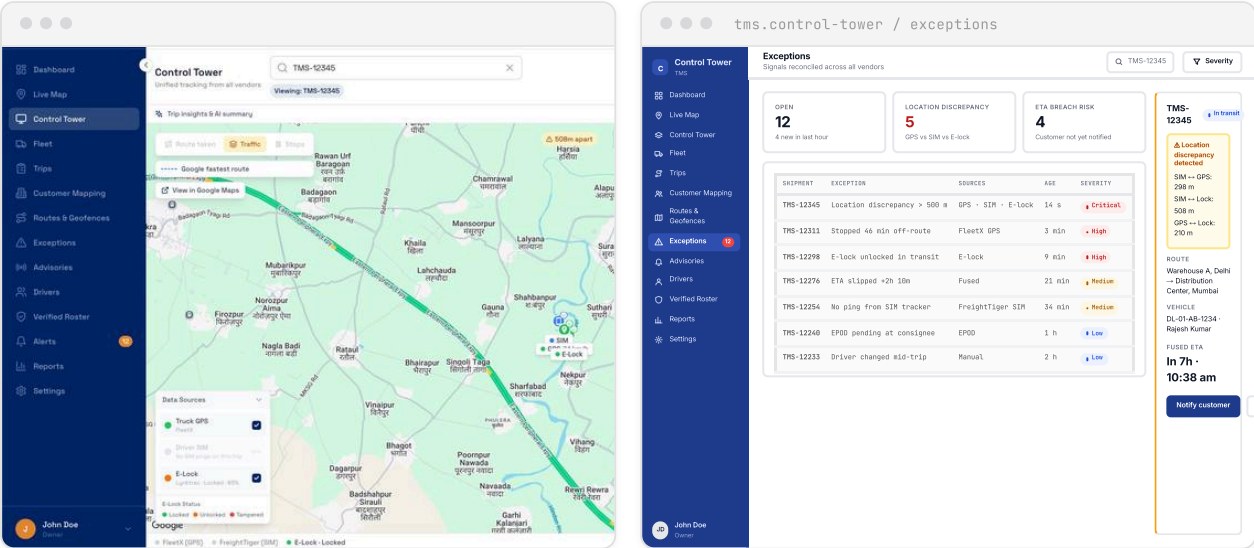
15s	GPS refresh interval
100%	Fleet visibility on one screen
Fewer	Missed SLAs as fleet scaled

Transport Management Platform

Client: One of the World's Largest Logistics Companies (Undisclosed Under NDA)

We fused messy tracking signals from many different carriers into one honest view of where every shipment actually is, so operators stop chasing status across a dozen vendor portals.

- Logistics
- TMS
- Control Tower
- Signal Fusion
- Real-Time



CLIENT PROBLEM

Shipments move across many carriers, brokers and telematics vendors, each publishing signals in different formats and cadences. Operators chased status in spreadsheets and portals with no single reliable view.

OUR SOLUTION & ARCHITECTURE

A **unified control tower** that fuses every vendor signal into one real-time shipment view. It normalizes formats, reconciles conflicts, drives a shipment state machine, and surfaces exceptions before they become customer escalations. One map, one ETA, one truth per shipment.



WHY PRIVATE AI

Carrier data consolidated privately	No shipment data shared across vendors
Single source of truth	One canonical timeline per shipment
Exceptions caught early	Before they reach the customer

BENEFITS REALISED

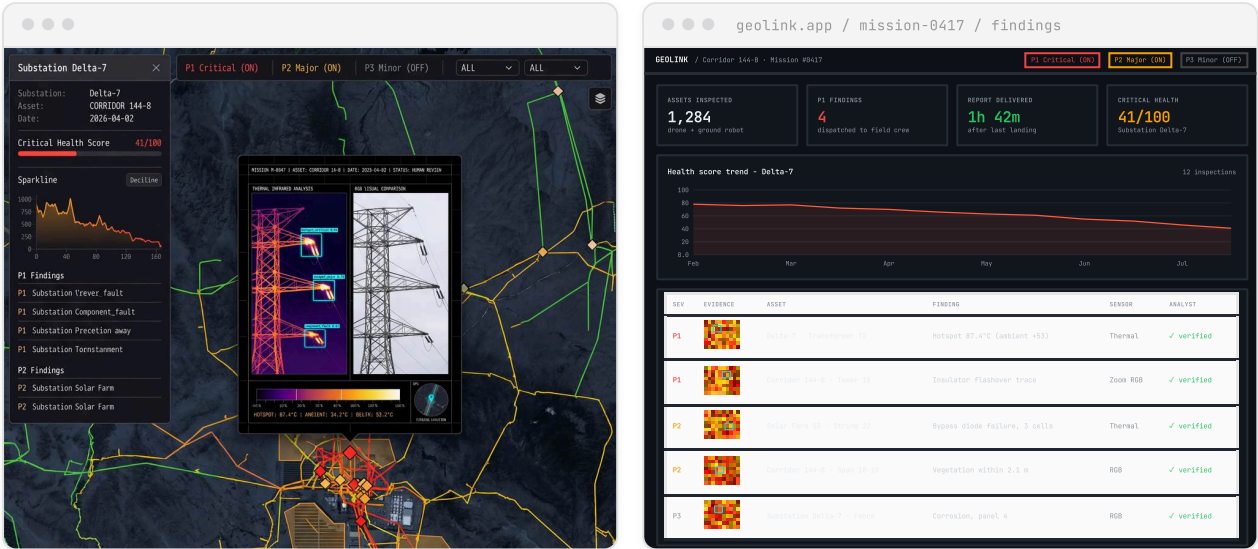
Multi-vendor	GPS, EDI, telematics, EPD fused
Real-time	Live state machine per shipment
1 view	Replaces a dozen vendor portals

Computer Vision Defect Detection System for Power Grids & Pipelines

Client: Connect Aerospace LLP

We turn drone and robot inspection footage of power grids and pipelines into a clear, GPS-mapped defect report within 2 hours, instead of a messy PDF that lands weeks later.

- Infrastructure Intelligence
- Drone + Robot
- Computer Vision
- Thermal Imaging
- SaaS



CLIENT PROBLEM

Critical infrastructure is inspected on fixed cycles using helicopters and ground crews; the output is slow, costs \$3,000 to \$8,000 a day, and produces unstructured PDF reports days or weeks after the flight, with no GPS context or defect tracking.

OUR SOLUTION & ARCHITECTURE

A 3-layer inspection-intelligence platform: field collection (drones + ground robots with thermal/RGB/zoom), an **AI detection layer** (YOLOv8 with human-in-the-loop review), and a client portal with a GPS asset-health map, severity-ranked findings (P1-P4) and historical comparison. Reports in 2 hours, not 2 weeks.



WHY IT MATTERS

Human-verified AI	Every detection reviewed before it reaches the client
Defects tracked over time	Not a one-off PDF with no history
Infrastructure-grade	Grids, solar farms, pipelines, ports

BENEFITS REALISED

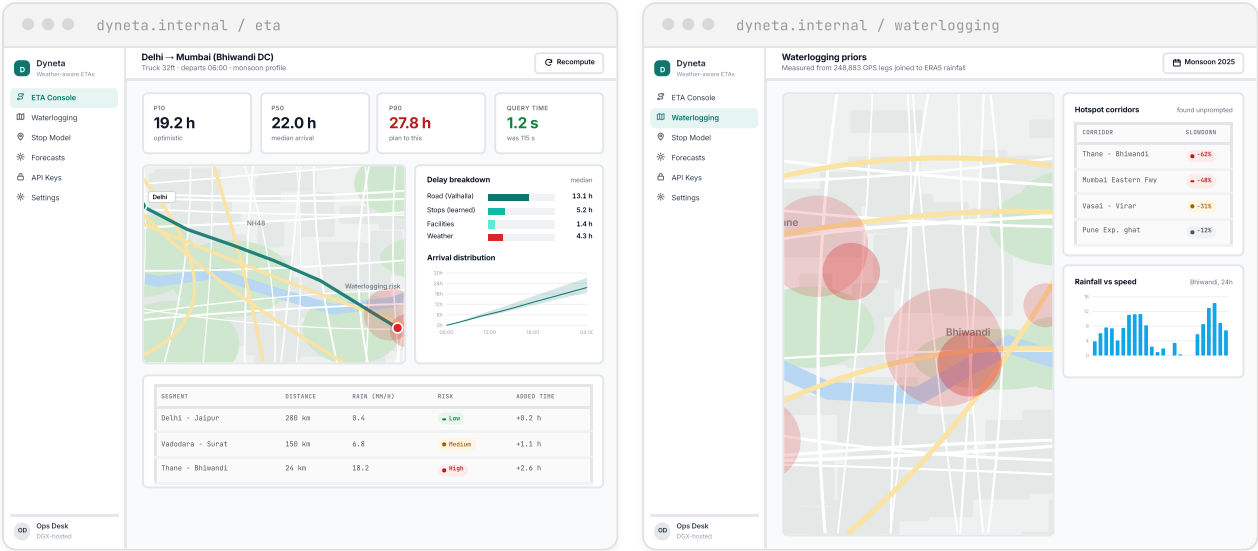
< 2 hr	Report turnaround vs 5-10 days
\$3-8k/day	Cost of the methods it replaces
P1-P4	Severity-ranked, GPS-mapped findings

Weather-Aware ETA Prediction Engine for Indian Logistics

Client: 360 Labs Product, Built on a Global Freight Forwarder's Fleet Data

We built an ETA engine that knows Indian roads flood in the monsoon, so instead of one optimistic arrival time it gives a realistic range and tells you why a truck will be late.

- Logistics
- ETA Prediction
- Machine Learning
- On-Prem DGX
- Published Research



CLIENT PROBLEM

Truck ETAs in India collapse in the monsoon: worst-case journey times differ by 457% between dry and rainy days, and the incumbent telematics estimate assumed a flat ~17 km/h regardless of weather.

OUR SOLUTION & ARCHITECTURE

An ETA engine that predicts a **range** of arrival times (P10-P90) broken down by road, stops, facilities and weather. It learns real stop behaviour from 8,846 fleet stops and measures waterlogging risk from 248,883 GPS legs joined to rainfall, finding the Mumbai-Thane-Bhiwandi flood corridor on its own. Trained and served on 360 Labs' own NVIDIA DGX.



WHY ON-PREM

Runs on owned DGX	No per-query cloud cost; Rs. 0/month to run
Fleet data stays private	Client GPS never leaves the environment
Research-grade honesty	Error bars, not a single false-precision number

BENEFITS REALISED

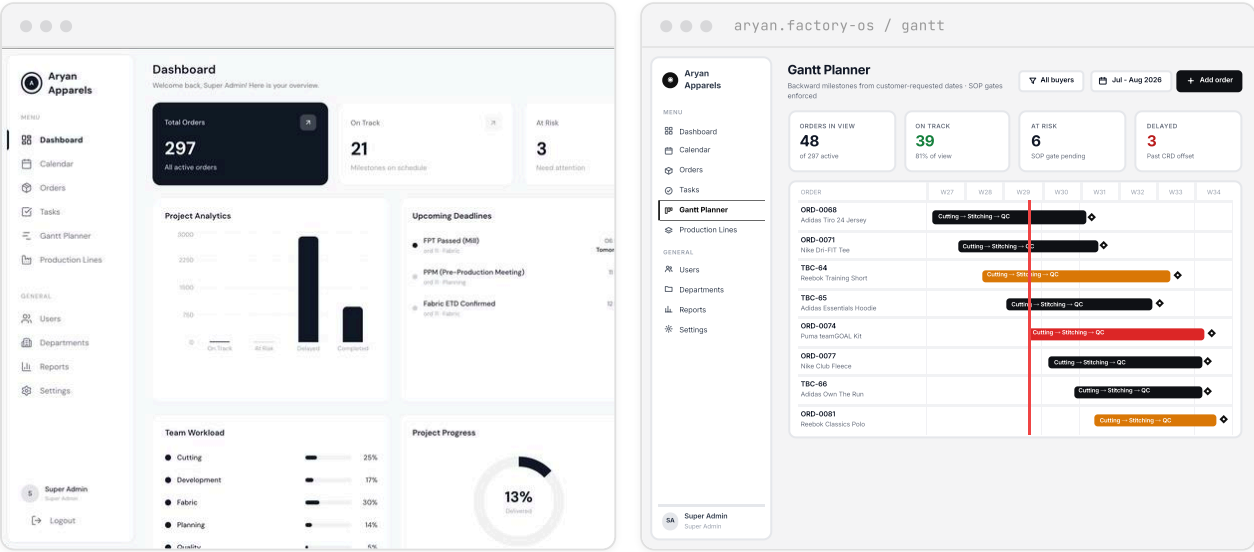
63%	Lower median ETA error (2.85h vs 7.79h)
1.2s	Route query, down from 115s
IJFMR 2026	Peer-reviewed & published

AI-Native Production Planning System for Garment Factories

Client: Aryan Apparels Pvt Ltd

We replaced the spreadsheets a garment factory used to track hundreds of orders for Adidas and Nike with one live board that plans backwards from the deadline and blocks work that skips a quality gate.

- Manufacturing
- Production Planning
- SOP Gates
- Capacity Simulation
- Next.js



CLIENT PROBLEM

Garment manufacturers producing for Adidas, Nike and Reebok tracked hundreds of orders across 9 departments in Excel: no automated milestone planning, no quality-gate enforcement, no capacity simulation, and no real-time visibility into delays.

OUR SOLUTION & ARCHITECTURE

A dual-track milestone engine with **SOP quality gates** and capacity simulation. Track A computes backward milestones from the customer-requested date; Track B triggers from real events like GRN receipt. 9 mandatory pre-cutting gates with evidence uploads, plus what-if capacity simulation with overtime.



WHY IT MATTERS

One source of truth	Replaces Excel across 9 departments
Quality enforced	No order proceeds past a failed gate
Plan before you commit	What-if simulation with overtime

BENEFITS REALISED

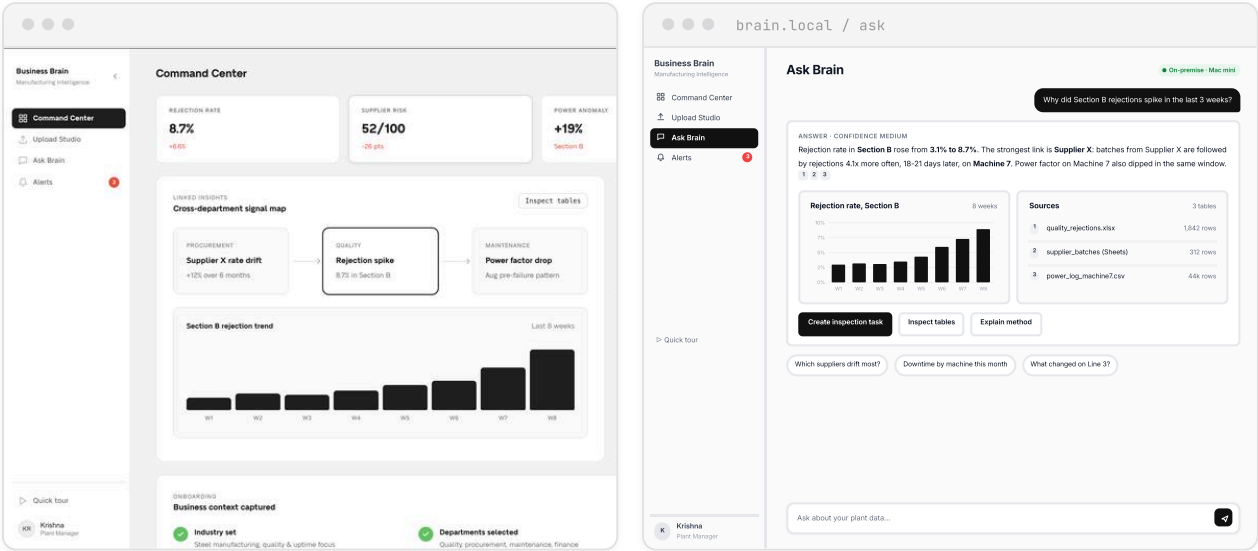
100s	Orders tracked live
9	Mandatory SOP quality gates
Real-time	Line utilisation & delay visibility

On-Premise Data Intelligence Platform for Manufacturers

Client: Undisclosed (Under NDA)

We stitched a factory's scattered spreadsheets into one brain that runs entirely on a Mac Mini on-site, and it surfaces the surprising connections people would never spot by eye.

- Manufacturing
- Knowledge Graph
- On-Prem
- 7-Layer Pipeline
- Private AI



CLIENT PROBLEM

Manufacturing teams run on a sprawl of spreadsheets: production logs, supplier records, quality reports, downtime, each in its own format. The real signal lives in multi-hop relationships across these silos, which generic BI tools either drown in trivia or miss entirely.

OUR SOLUTION & ARCHITECTURE

An **on-premise** platform that stitches every spreadsheet into one knowledge graph and surfaces only insights that matter. Column fingerprinting auto-classifies every column; a 7-layer pipeline enumerates thousands of candidate analyses and filters to ~18 high-quality findings scored on surprise, magnitude and variance. The whole stack runs on a Mac Mini on-site.



WHY PRIVATE AI

Runs on a Mac Mini	A full intelligence stack entirely on-site
No spreadsheets leave	Proprietary operational data stays in the plant
Finds hidden signal	Multi-hop correlations across silos

BENEFITS REALISED

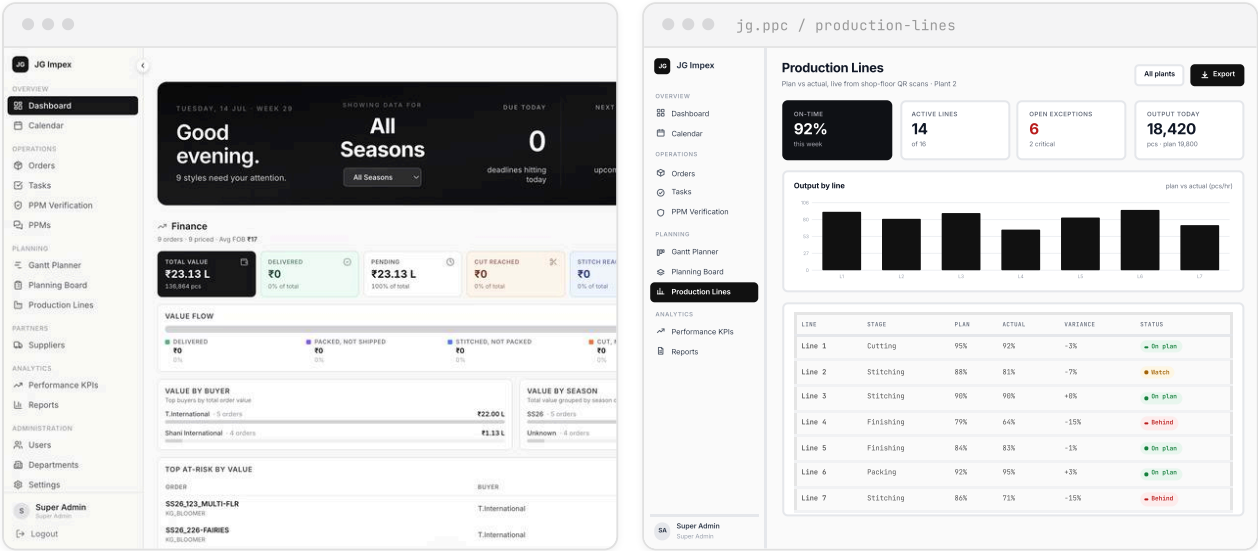
~18	High-signal findings from thousands
5-dim	Interestingness scoring
100%	On-premise, zero data egress

Real-Time Production Planning & Control Platform for Garment Factories

Client: JG Group of Companies

We built the control room for a garment group's factories, so leaders can see every cutting, stitching and finishing line live against the plan instead of reading yesterday's history.

- Manufacturing
- PPC
- Capacity Planning
- Shop-Floor
- Role Dashboards



CLIENT PROBLEM

Production planning ran on spreadsheets and disconnected processes with no single source of truth. Capacity was set by hand, progress across cutting, stitching and finishing was invisible until a delay had already happened, and leadership read history rather than acting on live status.

OUR SOLUTION & ARCHITECTURE

The **operational control layer** for JG Group manufacturing: it plans the work, tests whether capacity can carry it, governs execution with SOP enforcement, and shows the floor live against the plan. Role-based dashboards span multiple plants, business units and production lines.



WHY IT MATTERS

One source of truth	Replaces spreadsheets across plants
Live, not historical	Act on status, don't read the past
Configured to the business	Role dashboards per plant and line

BENEFITS REALISED

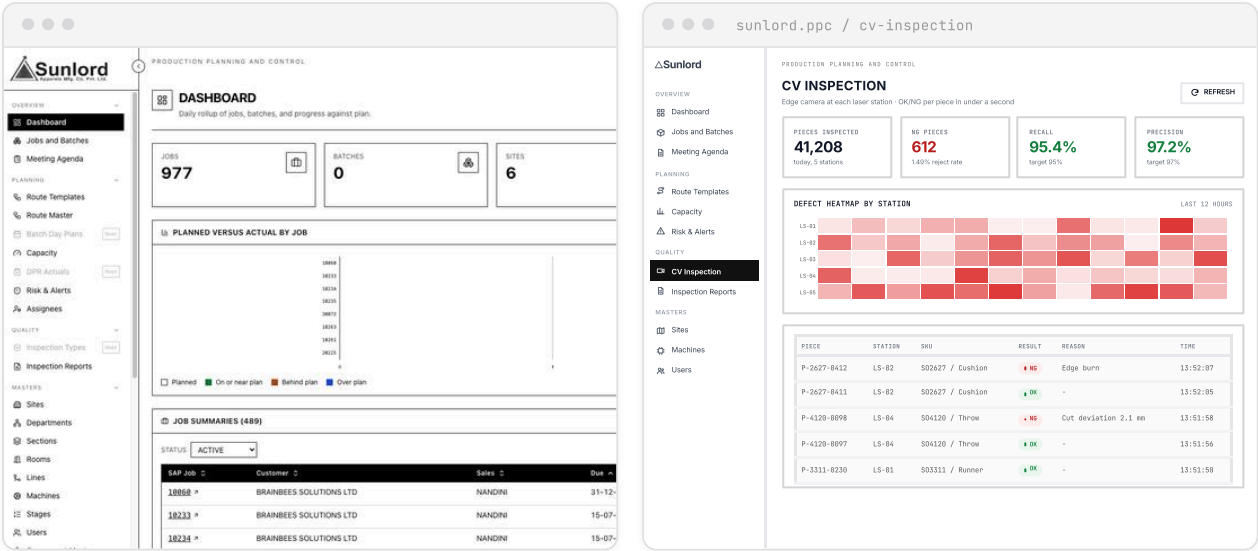
Multi-plant	Business units and lines in one view
Real-time	Shop-floor QR/barcode tracking
Plan vs actual	Exceptions ranked by severity

Edge-AI Manufacturing Platform for Garment Production Lines

Client: Sunlord Apparel's Mfg Co Pvt Ltd

We wired 4 systems into one factory brain: planning, a camera that checks every cut piece, a sensor that counts stitches, and an AI co-pilot, with the cameras and sensors running right on the floor.

- Manufacturing
- SAP B1
- Edge Computer Vision
- IoT Sensors
- On-Prem

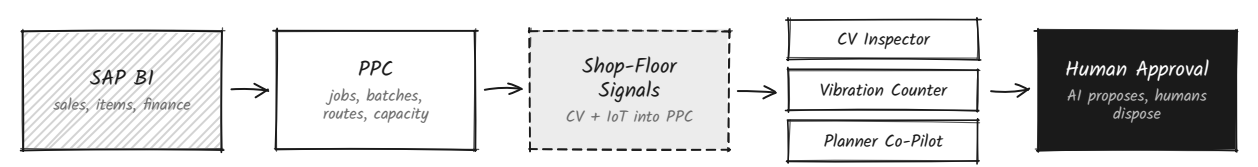


CLIENT PROBLEM

Sunlord ran planning on spreadsheets that broke at scale, had no live view of jobs on the floor, reconciled costs by hand at month-end, and had no clean way to feed shop-floor signals (defects, piece counts, machine issues) into the plan.

OUR SOLUTION & ARCHITECTURE

4 connected systems with **PPC at the centre**: SAP B1 keeps sales and finance; computer vision and IoT push live data into PPC continuously; and an AI intelligence layer reads from PPC and the live feeds, makes suggestions, and routes them through a human-approval workflow. The CV inspectors and vibration counters run on **edge compute on the floor**.



WHY PRIVATE AI

Inference on the floor	CV and IoT run on edge compute, not the cloud
Factory data stays local	No production imagery leaves the plant
AI proposes, humans dispose	Every change passes an approval workflow

BENEFITS REALISED

97% / 95%	CV precision / recall per piece
Per-machine	IoT stitch counting, calibrated
On-prem	4 systems, one factory brain

SAP-Integrated AI Demand Forecasting Platform

Client: Undisclosed (Under NDA)

We gave a manufacturer one trustworthy data estate across 3 companies and SKU-level forecasts, where the AI only picks which reviewed number to run and never writes back to SAP.

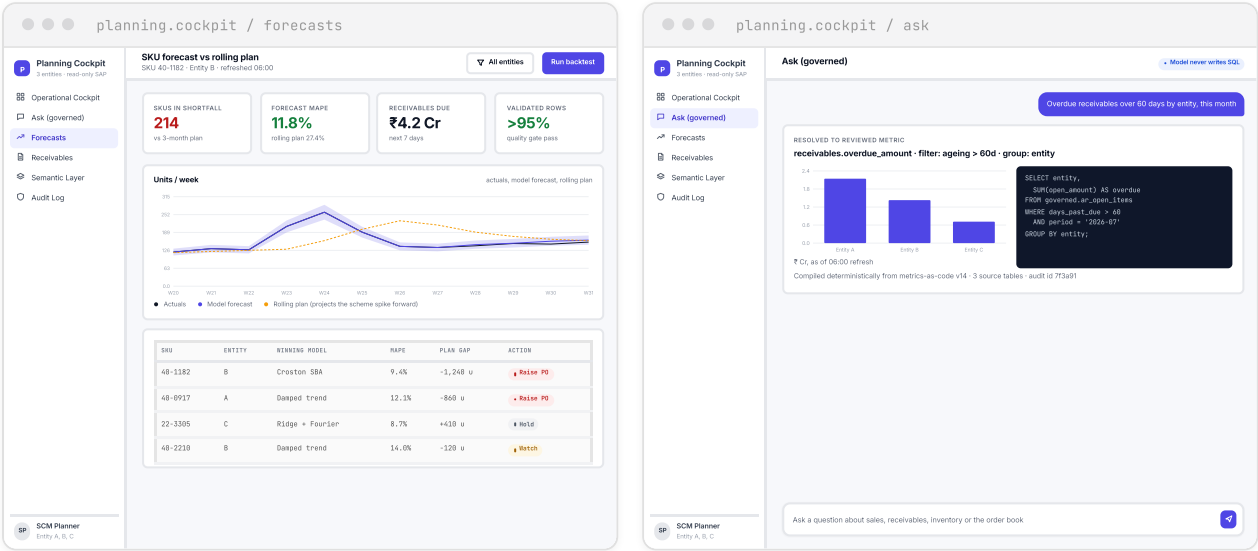
Manufacturing

SAP S/4HANA

Demand Forecasting

Semantic Layer

Read-Only



CLIENT PROBLEM

Commercial operations ran on SAP S/4HANA across 3 entities with a layer of spreadsheets refreshed a day behind, so a receivable was often already paid before anyone acted. A single month of one register ran past 100,000 invoice lines, and a trailing 3-month rolling forecast left hundreds of SKUs in shortfall.

OUR SOLUTION & ARCHITECTURE

One **governed data estate**, mapped once and reviewed, with accuracy coming from a deterministic compiler rather than model behaviour. A semantic layer maps cryptic SAP fields to reviewed business meaning, so at question time the model only chooses which reviewed metric to run. Forecasts are per-SKU and measured against the client's own plan. Nothing is ever written back to SAP.



WHY PRIVATE AI

Read-only to SAP	The model never writes, never sees raw rows
Deterministic accuracy	Metrics compile to exact SQL, not guesses
Scoped by role	Data up to one picture, access down by role

BENEFITS REALISED

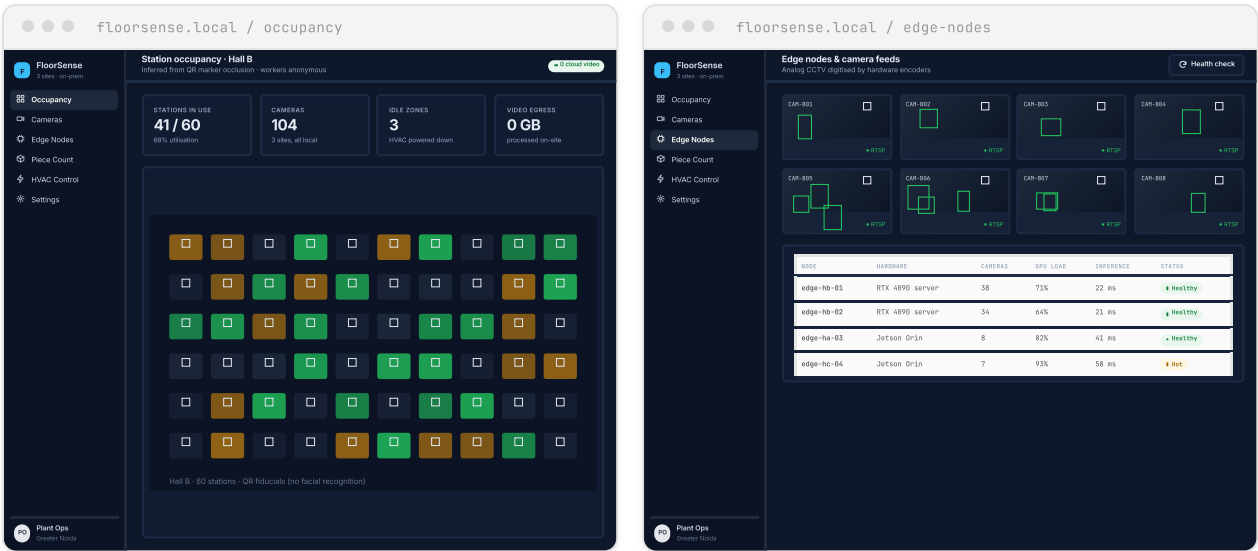
>95%	Records passing validation
Per-SKU	Forecasts vs the client's own plan
Terabytes	Governed across 3 entities

Privacy-Preserving Camera Analytics & Workforce Tracking for Factories

Client: One of India's Largest Toy Manufacturers (Undisclosed Under NDA)

We made 100+ old factory cameras smart without sending a single video frame to the cloud, and tracked who is working where using QR tags instead of face recognition.

- Computer Vision
- Edge AI
- Zero Cloud Video
- Privacy-Preserving
- On-Prem

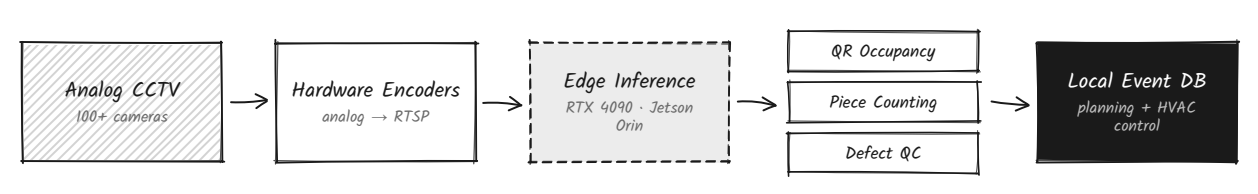


CLIENT PROBLEM

Legacy analog cameras streamed signals incompatible with modern CV, and a full replacement was not viable. Streaming 100+ high-res feeds to the cloud would have meant severe lag and bandwidth cost, and facial recognition for worker tracking was rejected outright on privacy grounds.

OUR SOLUTION & ARCHITECTURE

An **edge-first** spatial-intelligence layer with zero cloud video. Hardware encoders digitised the analog estate into RTSP streams without replacing a camera. Edge nodes process every feed on-site, and QR codes on machines act as fiducial markers: when a worker occludes a marker, the desk is confirmed in use. Workers stay anonymous and no raw video ever leaves the facility.



WHY PRIVATE AI

Zero cloud video	Every frame processed on-site
Privacy by design	QR fiducials, never facial recognition
No camera rip-&-replace	Encoders reuse the analog estate

BENEFITS REALISED

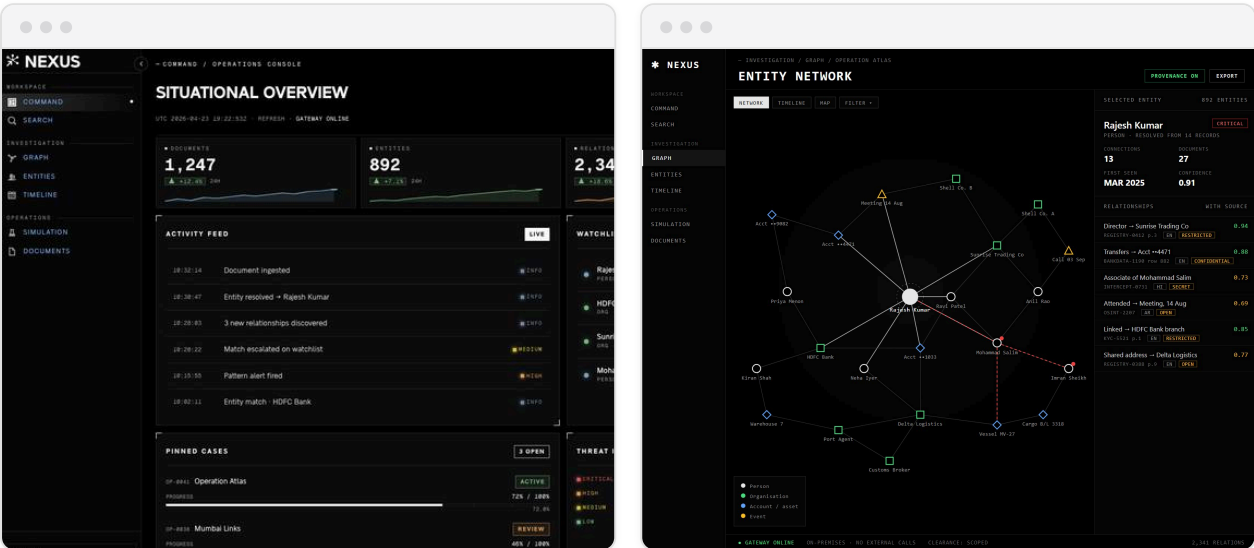
100+	Cameras, zero cloud streaming
3	Factory sites on edge nodes
0	Raw video leaving the facility

AI Data Intelligence for Government

Client: Undisclosed (Under NDA)

We built a system that reads intelligence from any format, language or classification level and draws the hidden network of who is connected to whom, all inside the agency's own walls.

- Defense
- Data Intelligence
- Entity Networks
- Knowledge Graph
- On-Prem



CLIENT PROBLEM

Investigative teams work across fragmented sources (documents, databases, intercepts, OSINT) in many formats, languages and classification levels. Building a connected picture of entities, relationships and events takes weeks of manual work.

OUR SOLUTION & ARCHITECTURE

A modular intelligence platform that turns messy data into actionable **entity networks**. It ingests raw data of any format, links it across sources, and surfaces entity, relationship and event networks **with provenance**, at classification levels appropriate to the operator. Deployed inside the agency perimeter.



WHY PRIVATE AI

Inside the perimeter	Classified material never leaves the agency
Provenance on every link	Answers trace to their source
Classification-aware	Operators see only their level

BENEFITS REALISED

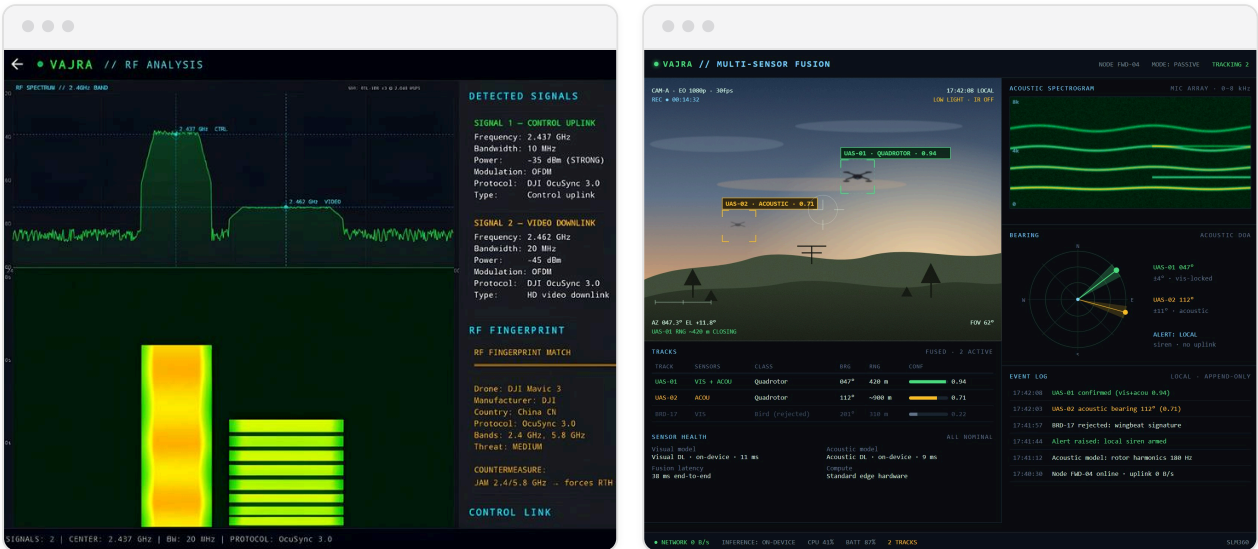
Any format	Documents to intercepts to OSINT
Weeks → live	Manual network-building automated
Air-gap ready	No external dependency

AI Counter-UAS System

Client: Undisclosed (Under NDA)

We built a drone-detector that sees and hears incoming drones using AI running on the device itself, so it works at a forward post with no network and no cloud.

- Defense
- Counter-UAS
- On-Device ML
- Edge AI
- Zero Network

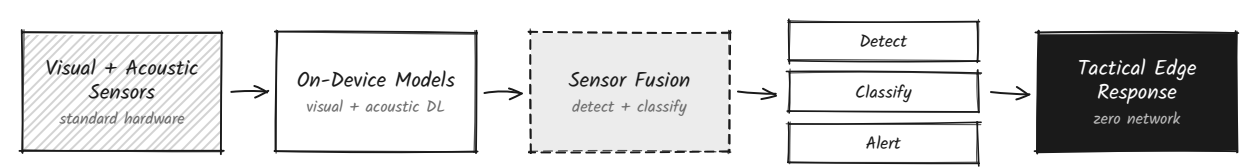


CLIENT PROBLEM

Small unmanned aerial systems are a growing threat across critical installations. Existing counter-UAS solutions are expensive, dependent on external infrastructure, and slow to deploy at forward positions.

OUR SOLUTION & ARCHITECTURE

A lightweight **on-device** counter-UAS stack combining visual and acoustic deep-learning models. Multi-sensor detection runs entirely on standard hardware with **no network dependency**, deployable at the tactical edge in denied or degraded environments.



WHY PRIVATE AI

Runs on-device	No cloud, no external infrastructure
Works denied/degraded	Zero network dependency at the edge
Standard hardware	Deployable at forward positions

BENEFITS REALISED

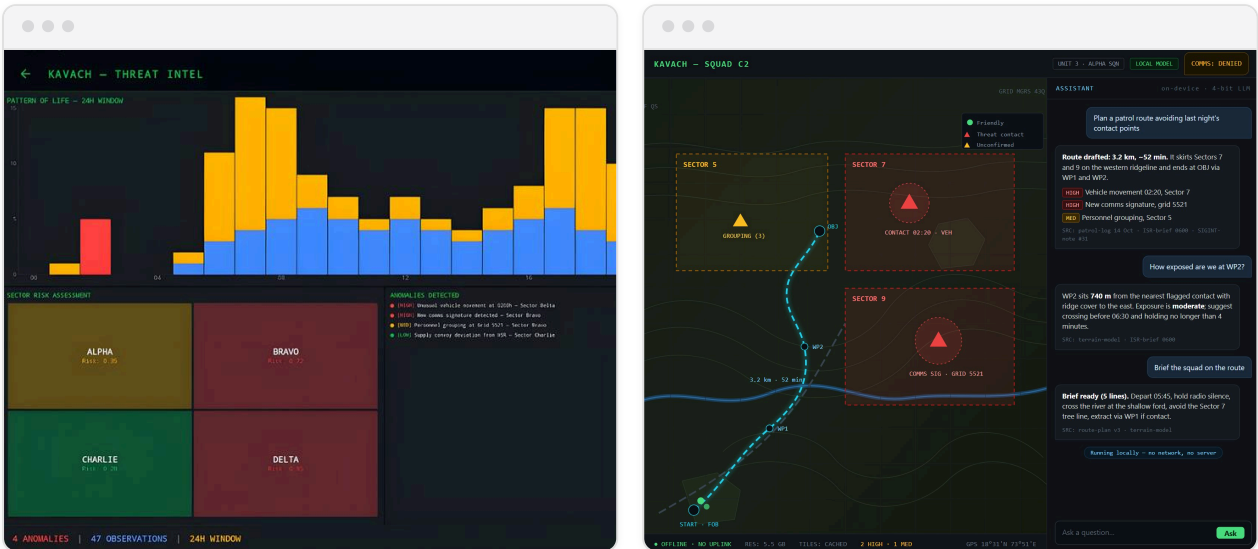
Dual-sensor	Visual + acoustic deep learning
0	Network dependency
Edge-ready	Denied and degraded environments

AI Tactical Intelligence Platform

Client: Undisclosed (Under NDA)

We put a command-and-control assistant on a squad's own device that takes plain-language orders, writes the reports and plans threat-aware patrols, with no server and no network behind it.

- Defense
- Tactical AI
- Command & Control
- Edge AI
- Zero Network



CLIENT PROBLEM

Existing C2 systems require persistent network connectivity, dedicated server infrastructure, and cost hundreds of thousands of dollars per node, inoperable at the tactical edge where squad-level units need real-time situational awareness.

OUR SOLUTION & ARCHITECTURE

An **on-device** tactical-intelligence platform for the operator's own device: natural-language command and control, ISR processing, automated reporting and threat-aware patrol planning, all running locally with no dependency on external networks.



WHY PRIVATE AI

On the operator's device	No server, no persistent network
Edge-survivable	Works where connectivity does not
Low cost per node	Not six figures per installation

BENEFITS REALISED

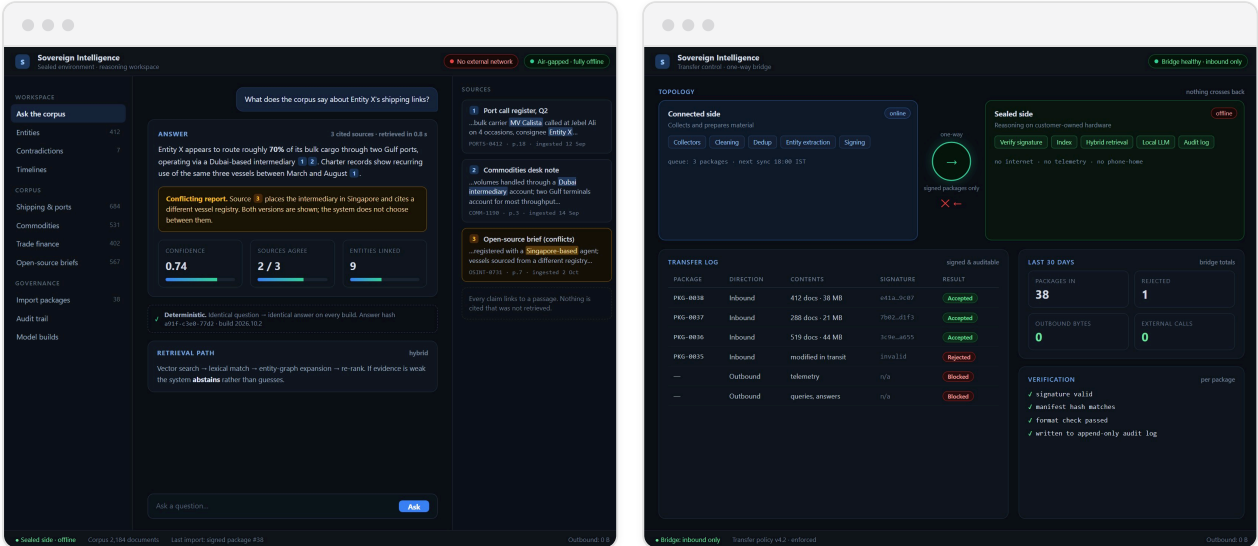
Natural-language	C2 at the squad level
0	Network dependency
Local	ISR, reporting & patrol planning

Sovereign Intelligence Platform

Client: Undisclosed (Under NDA)

We split the system in 2: an online side that collects and a sealed, offline side that thinks, joined by a one-way bridge, so sensitive analysis happens on the customer's own air-gapped hardware and nothing ever phones home.

- Defense
- Sovereign AI
- Air-Gapped
- Hybrid Retrieval
- Determinism



CLIENT PROBLEM

Family offices, commodities businesses and defense-adjacent analysts cannot put confidential material in the cloud. Hosted tools send documents off-site, produce fluent answers grounded only by accident, cite nothing, hide conflicting reports, and answer the same question differently each time.

OUR SOLUTION & ARCHITECTURE

2 environments joined by a **one-directional bridge**: a connected side that collects and a sealed side that thinks. Nothing crosses except signed, auditable packages, and nothing crosses back. Every claim traces to a source, contradictions are shown rather than resolved, and the system abstains rather than fabricates. Identical questions on an identical corpus produce identical answers on every build.



WHY PRIVATE AI

Air-gapped by design	Sealed side has no internet, ever
Nothing phones home	Outbound is nothing - no telemetry
Deterministic	Same corpus, same answer, every build

BENEFITS REALISED

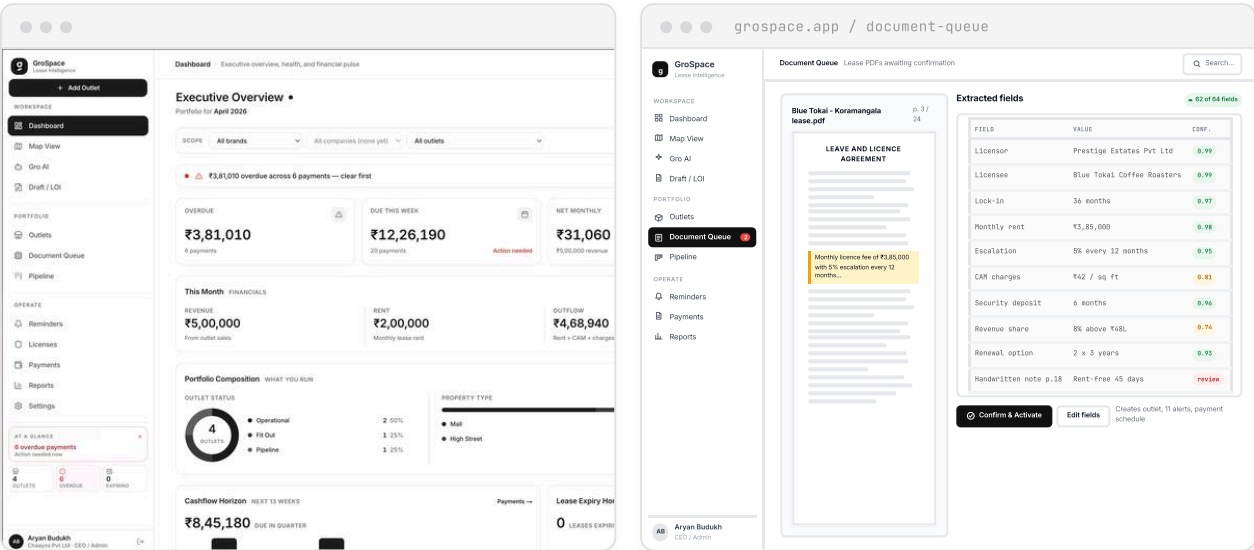
One-way	Signed packages in, nothing out
Cited	Every claim traced; abstains when thin
~33k → 1M+	Retrieval quality held as corpus grows

AI-Native Lease Management

Client: GroSpace Global Pvt Ltd

We trained an AI model to read a retail lease the way a property manager would, pulling 60+ details out of scanned PDFs and then watching every rent date, obligation and risk across the whole portfolio.

- Real Estate AI
- Document Extraction
- Obligation Engine
- Portfolio Risk
- Private Deployment



CLIENT PROBLEM

Multi-brand retail operators managing 50-500+ outlets (Blue Tokai, Domino's, McDonald's) tracked leases by hand. Extracting structured data from scanned or handwritten PDFs, tracking obligations and managing payments across rent models meant manual review of every document.

OUR SOLUTION & ARCHITECTURE

AI extracts **60+ structured fields** from lease PDFs, including scanned and handwritten documents. A "Confirm" and "Activate" workflow auto-creates outlets, obligations, alerts, and payment schedules. A 6-stage deal pipeline, natural-language portfolio chat, and multi-model payment generation sit on top, with per-tenant, row-level security.



WHY PRIVATE DEPLOYMENT

Lease data stays tenant-scoped	Row-level security isolates each operator
No manual review bottleneck	AI reads scanned & handwritten PDFs
Obligations never missed	Rules fire before a date slips

BENEFITS REALISED

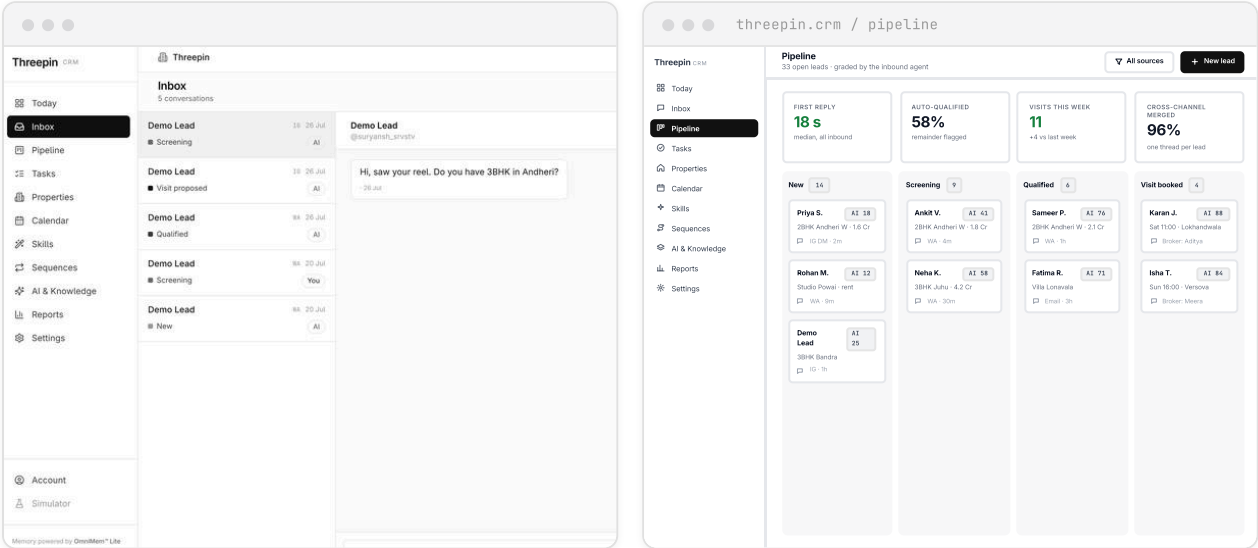
60+	Fields extracted per lease
50-500+	Outlets per operator
11	Alert types auto-generated

AI-Native Lead Management

Client: 3Pin Realty (Three Pin)

We put an AI agent on the brokerage's WhatsApp and Instagram that answers every enquiry in seconds, qualifies it, matches a listing and books the site visit, so hot leads stop going cold in the gap between channels.

- Real Estate
- AI-Native CRM
- WhatsApp
- Instagram
- Lead Scoring



CLIENT PROBLEM

A high-volume, Instagram-led brokerage: content drives millions of monthly views, but inbound lands as an Instagram DM, moves to WhatsApp, and leads go cold in the gap. No screening layer, and the system of record was spreadsheets, so consultants spoke to prospects with half the story.

OUR SOLUTION & ARCHITECTURE

An AI-native CRM with an **inbound lead agent** on WhatsApp and Instagram. Every enquiry is captured within a second, de-duplicated across channels, answered in under 30 seconds, graded against the requirement, matched to the catalogue and booked into a site-visit slot. Brokers operate the whole system by chatting on WhatsApp.



WHY IT MATTERS

Property DB stays client-owned	Read through a connector, not copied out
No lead lost in the gap	One thread across Instagram and WhatsApp
Brokers keep the close	AI does capture, humans do deals

BENEFITS REALISED

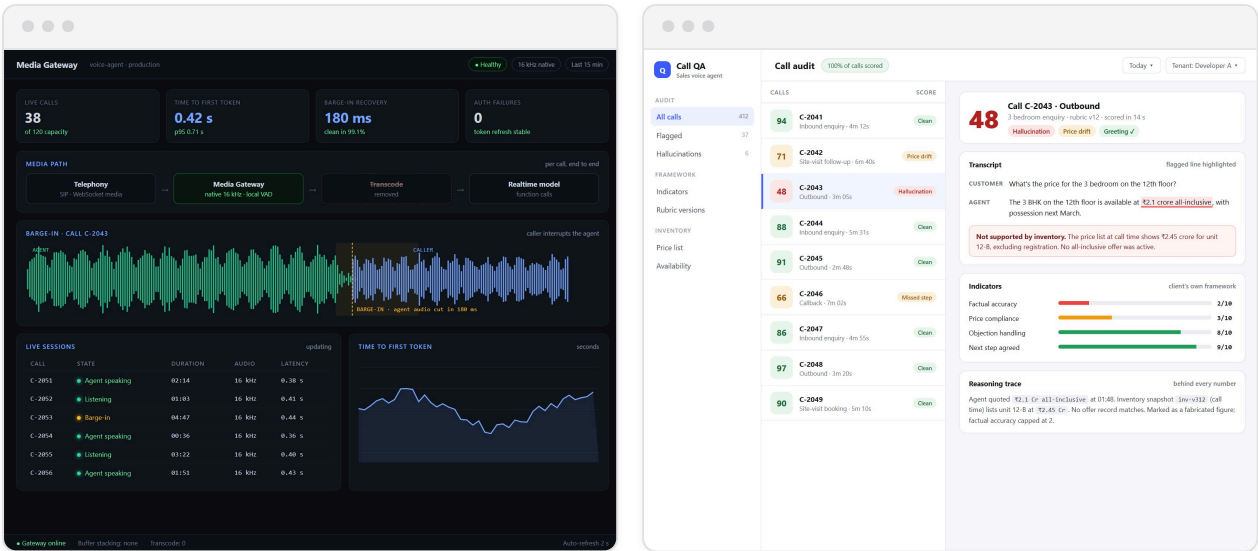
< 30s	First reply on every inbound
>95%	Cross-channel leads held in one thread
up to 60%	Auto-qualified before a broker touches it

Voice AI Infrastructure & Automated QA

Client: Undisclosed (Under NDA)

We fixed a property developer's sales-call voice bot so it interacts naturally, then built an auditor that scores every single call against the client's own quality rules instead of a lucky sample.

- Real Estate
- Voice AI
- Realtime Streaming
- LLM-as-Judge
- Multi-Tenant RLS



CLIENT PROBLEM

Pre-sales ran on a hosted voice agent carrying hundreds of calls a day, and the media path worked against it: a transcode broke barge-in, pre-warming stacked audio buffers into spikes, and token refresh caused auth failures. QA was manual: 2 auditors grading a fraction of volume, so an agent quoting a price that didn't exist was caught only by chance.

OUR SOLUTION & ARCHITECTURE

Fix the media path, then judge every call. A gateway carries native 16 kHz audio end-to-end with the transcode deleted and a local voice-activity detector cutting interruptions at the gateway. An **audit engine** scores every call against the client's indicator framework, with a reasoning trace behind each number, reading the same versioned inventory the agent reads, so a spoken price is checked against what was true at call time.



WHY PRIVATE AI

Per-tenant isolation	Force row-level security on rubric and inventory
Every call judged	Not a 2% manual sample
Grounded in truth	Price checked against call-time inventory

BENEFITS REALISED

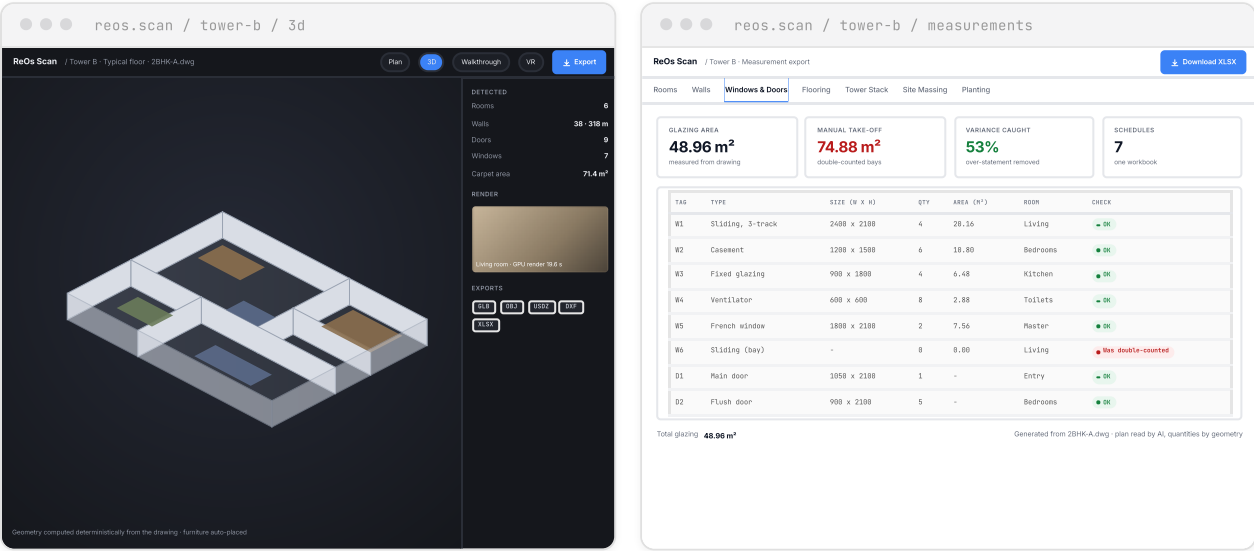
100%	Of calls scored, not sampled
Per-call	Reasoning trace behind every score
Hallucination	Alerts when the agent invents a fact

AI Drawing-to-3D Model & Bill of Quantities Platform

Client: Listed Mumbai Real Estate Developer

We built a tool that reads an architect's drawing and, in minutes, produces a furnished 3D model and a measured bill of quantities: work 4 departments used to redo by hand.

- Construction
- Drawing-to-3D
- Quantity Take-off
- Deterministic Geometry
- GPU Render

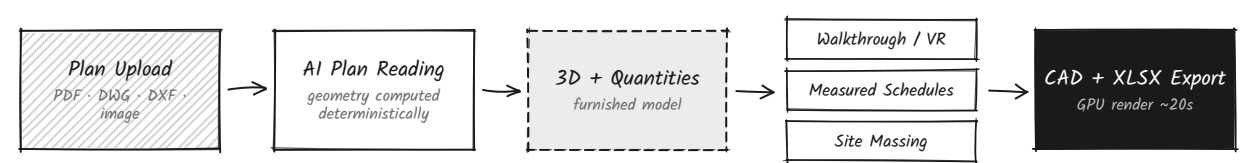


CLIENT PROBLEM

Architecture, BIM, interiors and landscape teams each rebuilt the same drawings by hand to get 3D views, quantity take-offs and site massing: the same work, 4 times, for every project.

OUR SOLUTION & ARCHITECTURE

ReOs Scan reads PDF, DWG, DXF or image plans and produces a furnished **3D model** with walkthrough, VR and CAD exports, plus measured schedules in Excel, window/door modules, tower stacking and site massing with FSI. Plans are read by AI; the geometry is computed **deterministically**, and photorealistic GPU renders return in about 20 seconds.



WHY IT MATTERS

Deterministic geometry	AI reads; maths measures - no hallucinated quantities
One upload, 4 teams	Architecture, BIM, interiors, landscape
Catches costly errors	Found a 53% glazing over-count

BENEFITS REALISED

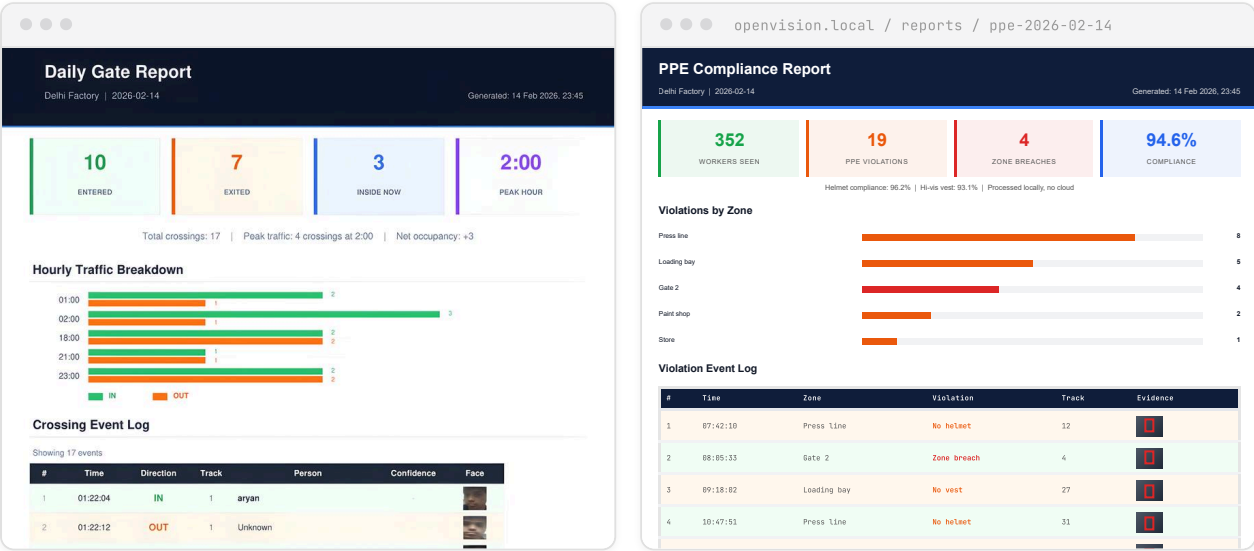
~20s	Plan to GPU-rendered 3D
7	Measured schedules per drawing
53%	Glazing over-count caught vs manual

Computer Vision PPE & Safety Monitoring for Construction Sites

Client: The Bansal Group

We put a desktop app on-site that watches the existing CCTV for missing helmets and vests, people in restricted zones and SOP breaches, with no cloud, no server and no internet.

- Computer Vision
- PPE Detection
- On-Premise
- YOLOv8
- Zero Cloud



CLIENT PROBLEM

Industrial facilities with 350+ workers ran manual gate registers, no PPE compliance tracking, no zone-violation detection, and hours of weekly safety reconciliation, with zero real-time occupancy or safety visibility.

OUR SOLUTION & ARCHITECTURE

A desktop application powered by **open vision models** processing live camera feeds locally. YOLOv8 handles PPE detection (helmet, vest), restricted-zone monitoring and person tracking with IN/OUT counting. All events land in SQLite; daily PDF and Excel reports are auto-emailed. No cloud, no servers, no internet required.



WHY PRIVATE AI

Fully local desktop app	No cloud, no servers, no internet
Footage never leaves site	All events written to local SQLite
Open models	No per-frame API cost

BENEFITS REALISED

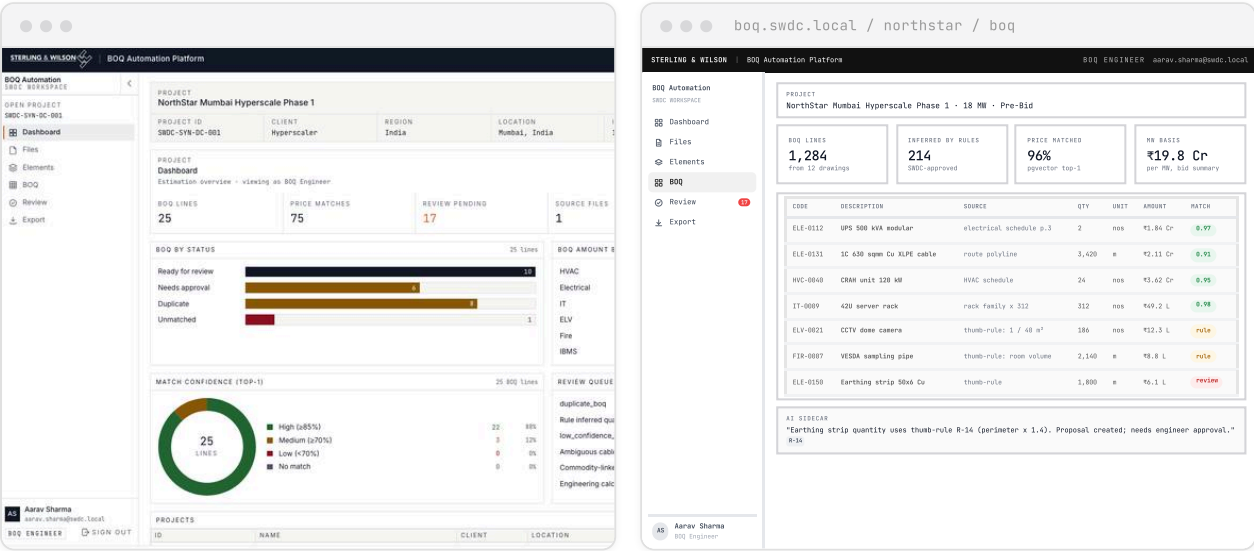
350+	Workers monitored live
0	Cloud or internet dependency
Daily	Auto-emailed compliance reports

AI Bill of Quantities Automation for Data-Centre Tenders

Client: Sterling and Wilson Renewable Energy Ltd

We turned data-centre tender drawings and 3D models into normalized bills of quantities automatically, including the items experts used to add from memory, with every AI line citing its source or waiting for sign-off.

- Data Center
- CAD / BIM Extraction
- Rule Engine
- Semantic Matching
- FastAPI



CLIENT PROBLEM

Data-centre tender estimation at SWDC was manual and knowledge-dependent. Drawings took days to convert into usable BOQs, CAD tools missed inferred items, and many line items depended on discipline-specific thumb rules held in expert heads.

OUR SOLUTION & ARCHITECTURE

An AI-assisted estimation engine for SWDC's workflow. It extracts structured quantities from 3D sources, applies **SWDC-approved rules** for items not explicitly drawn, normalizes multi-format BOQs, and semantically matches lines against historical rates and vendor quotes. Every AI action either answers with citations or creates a proposal requiring human approval.



WHY IT MATTERS

Expert rules encoded	Inferred items no longer live only in heads
Cited or approval-gated	Every AI action is traceable
Deterministic calculators	Weights, amperage, MW-basis exposed

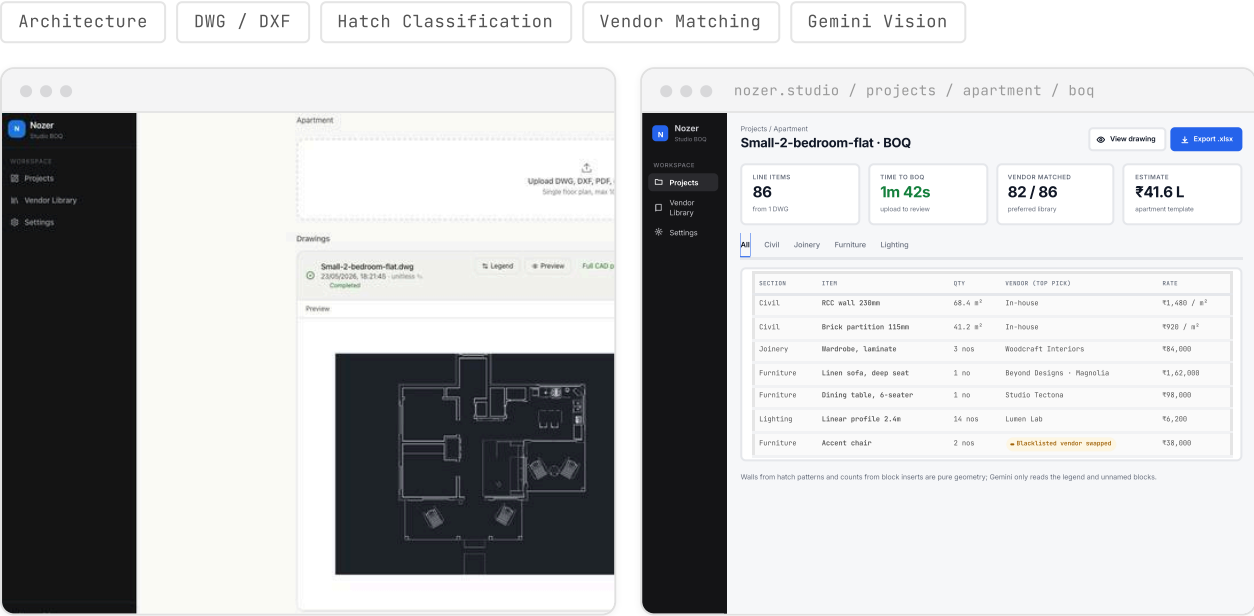
BENEFITS REALISED

3D → BOQ	Structured quantities from models
Days → fast	Manual conversion automated
Human-approved	Every write becomes a proposal

CAD-to-Bill-of-Quantities Automation for Architecture Studios

Client: Nozer Wadia Associates

We built a tool for a boutique architecture studio that reads a CAD drawing, counts the walls and furniture, matches the studio's preferred vendors, and spits out a filled bill of quantities in under 2 minutes.



CLIENT PROBLEM

Nozer Wadia Associates went from AutoCAD drawing to BOQ by hand. Walls were differentiated only by hatch pattern, furniture placed as blocks, and the studio's preferred and blacklisted vendors lived in seniors' heads. Every project started BOQ generation from scratch, one line at a time.

OUR SOLUTION & ARCHITECTURE

Upload a DWG, get a reviewable BOQ in under 2 minutes. The system reads walls from hatch patterns and counts furniture from block inserts **deterministically**, uses Gemini vision only where it earns its keep (reading the legend, unnamed blocks), matches every item against the studio's vendor library, and emits a filled BOQ from the studio template. Every quantity and vendor pick is human-editable.



WHY IT MATTERS

Geometry is deterministic	Wall areas and counts are maths, not AI
AI only where it earns it	Vision reads legends and odd blocks
Studio knowledge captured	Preferred/blacklisted vendors encoded

BENEFITS REALISED

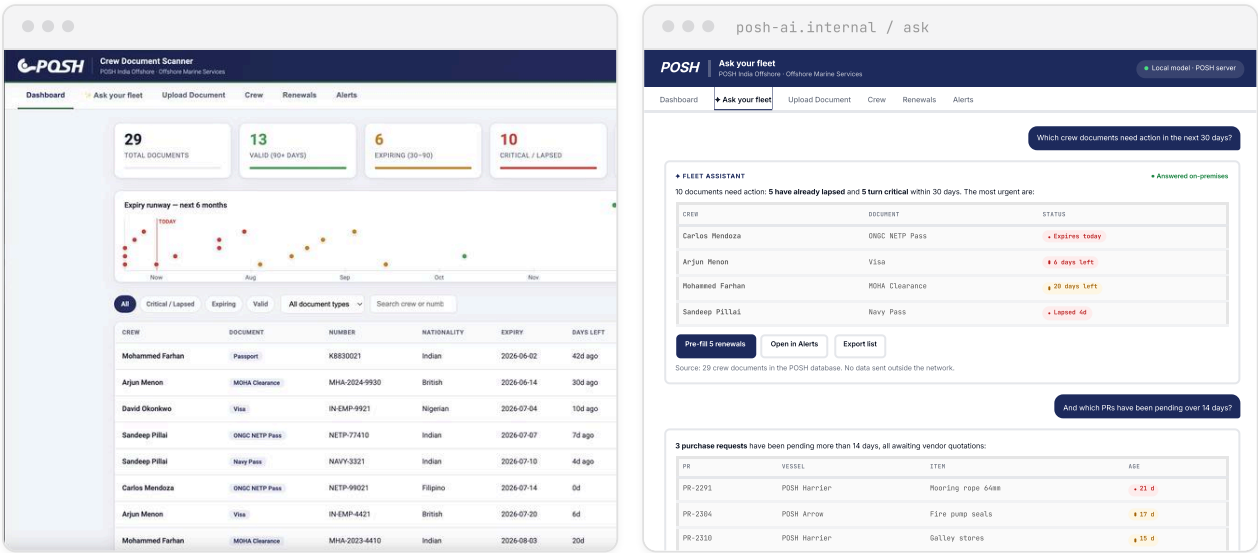
<2 min	Drawing to reviewable BOQ
Top-3	Vendor picks per line
100%	Human-editable before export

Private AI Document Automation for Crew & Supplier Paperwork

Client: POSH India Offshore Pvt Ltd

We put an AI on POSH's own servers that reads crew passports, visas and supplier paperwork the way a person would, tells the team what is about to expire, and none of that paperwork ever leaves POSH.

- Maritime
- Private AI
- On-Prem
- Open-Source LLM
- OCR
- PostgreSQL



CLIENT PROBLEM

POSH runs flagship vessels at day rates of \$600K-\$1M, where a lapsed crew pass, a stalled PR or a misplaced vendor quote costs more in downtime than preventing it. Critical office workflows ran on email chains, Excel and the memory of a few people.

OUR SOLUTION & ARCHITECTURE

6 lightweight AI tools deployed **entirely on POSH infrastructure**: no data leaves the network. A reading AI scans crew documents, quotations, PRs, GRNs and vendor paperwork, checks them against simple rules, and drives an expiry dashboard and a single morning exception report of only what needs action today. Open-source Llama 3 / Mistral run locally on the POSH server.



WHY PRIVATE AI

Crew IDs stay in-house	Never sent to any outside AI
Vendor prices stay private	Seen only by POSH
No cloud dependence	Runs on the POSH server, even offline

BENEFITS REALISED

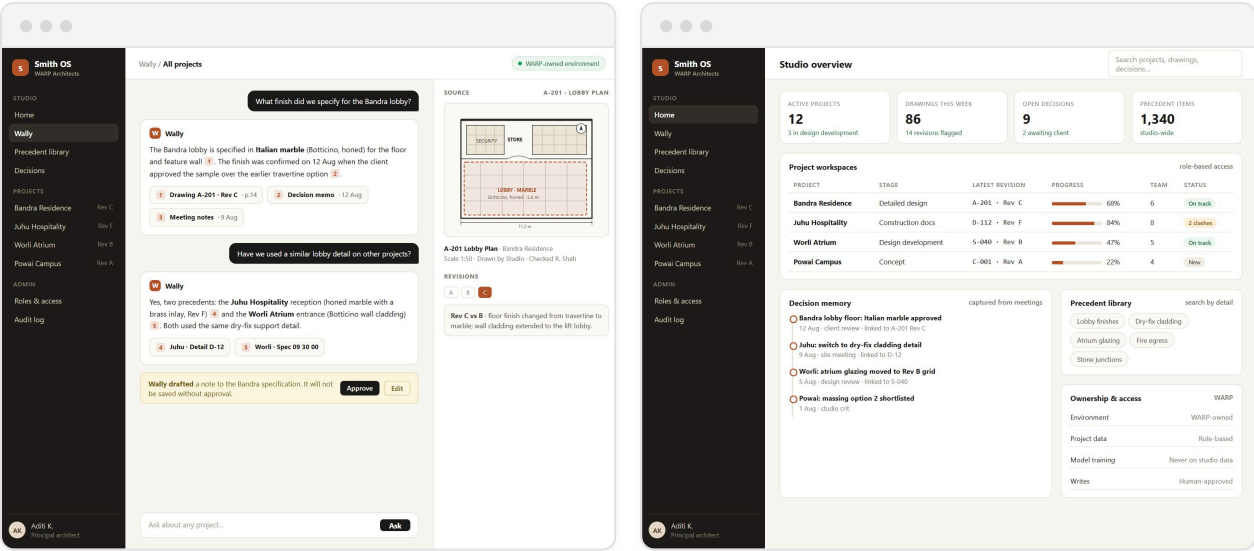
70%	Less manual document processing
90%	Faster document handling
100%	Of data processed on POSH infrastructure

AI-Native Operating System for Architecture Studios

Client: WARP Architects

We built an architecture studio its own private operating system with an assistant that answers questions about every project with citations, keeps the studio's IP in-house, and can later be sold on to other firms.

- Architecture
- AI-Native OS
- Cited Answers
- IP Ownership
- Multi-Tenant



CLIENT PROBLEM

A 45-person architecture and design studio wanted one place for projects, drawings, decisions and precedent, with full ownership of its IP and data, and a product it could later sell to other AEC firms as SaaS.

OUR SOLUTION & ARCHITECTURE

Smith OS with Wally, an assistant that answers with citations across project workspaces, plus drawing and revision intelligence, meeting and decision memory, and a studio-wide precedent library. Human approval on every write, role-based access, and a multi-tenant base ready for SaaS, with full IP transfer and data that never trains shared models.



WHY PRIVATE AI

Full IP transfer	WARP owns the environment and the data
Never trains shared models	The studio's work stays the studio's
Human-approved writes	AI drafts; people decide

BENEFITS REALISED

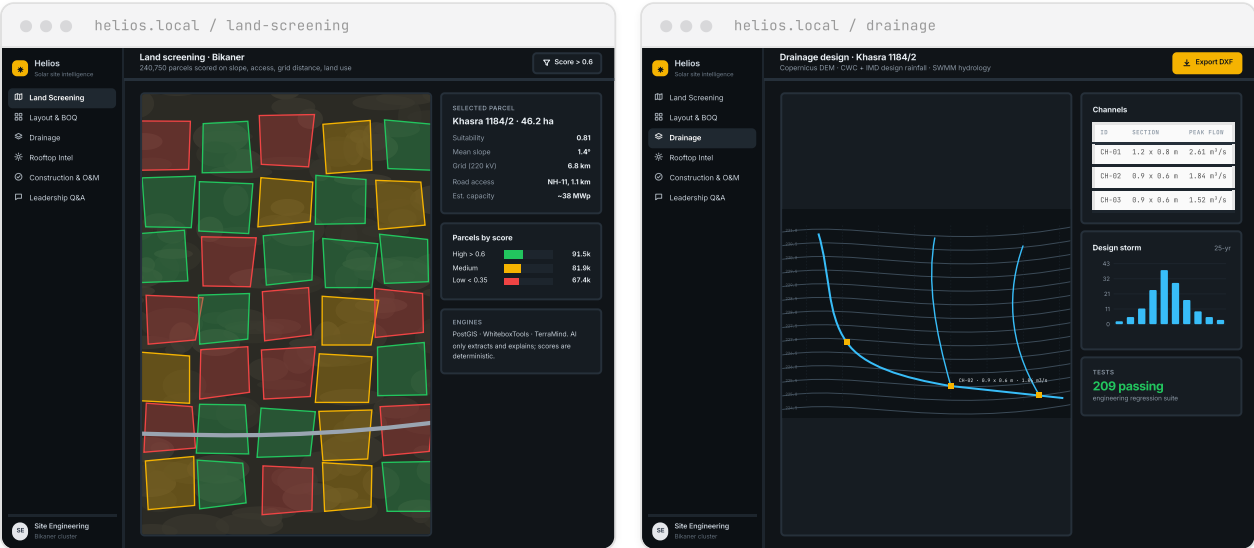
Cited	Every answer traces to a drawing or memo
Multi-tenant	Base ready to resell to other firms
Chosen	Over competing builders

AI Solar Site Screening & Layout Design Platform

Client: Leading Indian Solar Manufacturer

We built a platform that screens land, lays out panels, estimates the bill of quantities and even designs the site drainage for utility-scale solar: work that used to take weeks of manual GIS per site.

- Energy
- Solar PV
- GIS
- Deterministic Engineering
- On-Prem DGX

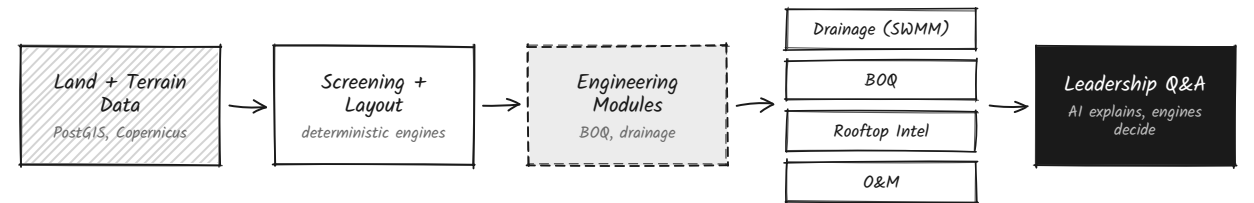


CLIENT PROBLEM

Screening land, designing layouts, estimating BOQ and planning drainage for utility-scale solar takes weeks of manual GIS and engineering work per site.

OUR SOLUTION & ARCHITECTURE

A 6-module platform: land screening, layout and BOQ, drainage design, rooftop market intelligence, construction/O&M tracking, and a leadership Q&A layer. Engineering values come from **deterministic engines**; AI only extracts and explains. Automated drainage design with CAD export is a capability no mainstream solar tool offers. Runs on 360 Labs' NVIDIA DGX.



WHY ON-PREM

Runs on owned DGX	No per-site cloud GIS bill
Deterministic engineering	AI explains; engines compute the numbers
Data layers stay in-house	Land and terrain data kept private

BENEFITS REALISED

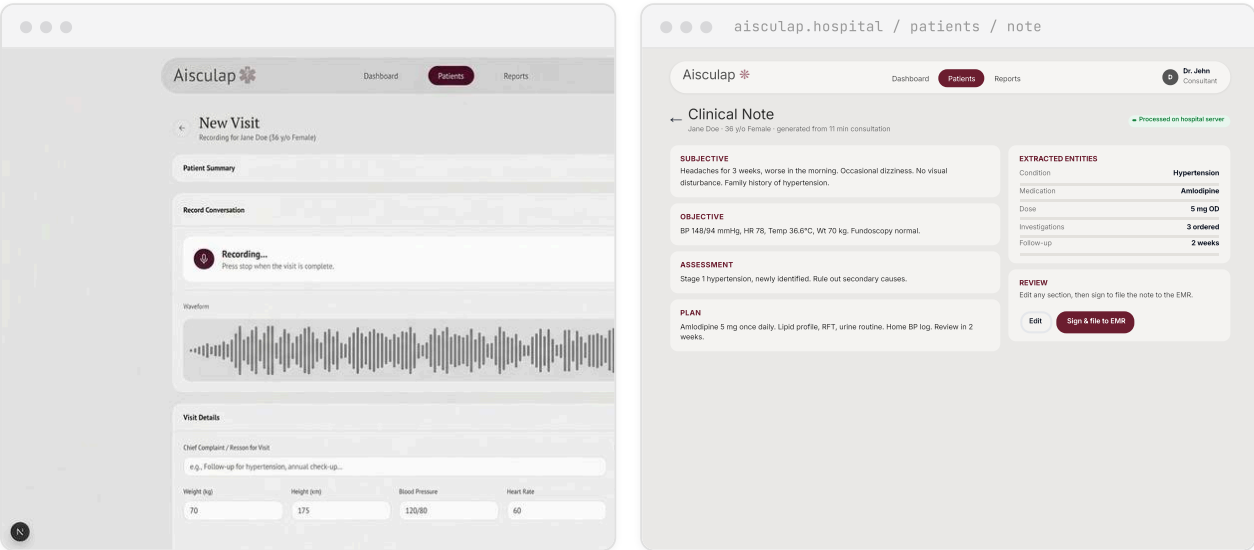
2.4 lakh	Bikaner parcels screened live
209	Automated engineering tests passing
Drainage	A capability no rival solar tool offers

AI-Native Clinical Notes

Client: Private Hospital (Undisclosed Under NDA)

We built a medical scribe that quietly listens to the doctor-patient conversation and writes the clinical note for the doctor to sign, giving them back the time they used to spend typing.

- Healthcare AI
- Medical Scribe
- Speech-to-Text
- Clinical NLP
- Private Deployment



CLIENT PROBLEM

Clinicians spend a disproportionate part of their day writing notes after each consultation. Manual documentation eats into patient face-time, drives burnout, and introduces inconsistency. Existing dictation tools need structured input and miss clinical context.

OUR SOLUTION & ARCHITECTURE

An AI medical scribe that runs in the background of every consultation. It listens, transcribes, extracts clinically relevant information, and produces structured notes (SOAP, problem lists, prescriptions, follow-up plans) automatically, ready for the doctor to review and sign, deployed privately for the hospital with integration-ready EMR output.



WHY PRIVATE DEPLOYMENT

Patient audio stays in-hospital	Sensitive consultations kept private
Doctor signs off	AI drafts; the clinician owns the record
EMR-ready	Integrates with existing systems

BENEFITS REALISED

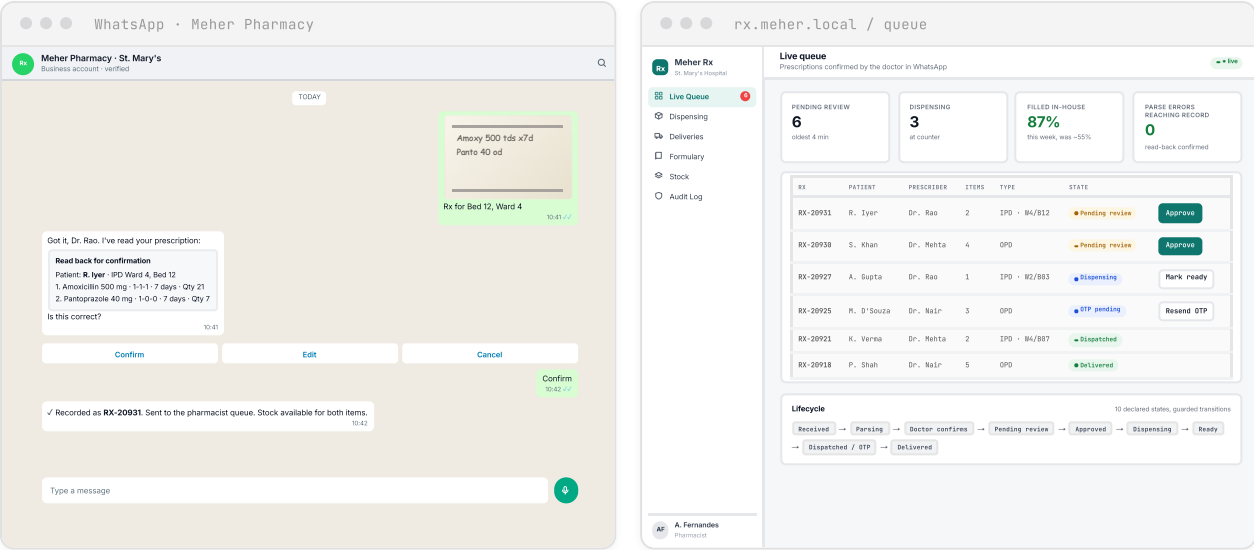
80%	Documentation time saved
Real-time	Scribe runs during the consult
SOAP	Structured notes, problem lists, Rx

WhatsApp-Native Hospital Pharmacy

Client: Meher Pharma (India) Limited

We moved prescriptions into WhatsApp with a strict safety machine behind them, so a doctor just sends a script, confirms what the AI read back, and the pharmacy fills it, with nothing to install.

- Healthcare
- WhatsApp Cloud API
- State Machine
- AI Parsing
- PostgreSQL RLS



CLIENT PROBLEM

The patient left with a paper slip and filled it anywhere, so nothing recorded how many in-house prescriptions were filled in-house. The pharmacy learned of a prescription only when paper reached the counter, and dosage was re-keyed off handwriting. A hospital cannot absorb a parsing error.

OUR SOLUTION & ARCHITECTURE

Capture inside WhatsApp, with a formal **state machine** behind it. The doctor sends a prescription as a form, voice note, photo or text: no install, no login. It is parsed, **read back and confirmed by the doctor** before recording. The lifecycle is 10 declared states accepting only valid transitions, so an invalid step is impossible to express. PostgreSQL row-level security isolates each hospital.



WHY IT MATTERS

Safety by construction	Invalid transitions cannot be expressed
Confirmed before recorded	No parsing error reaches the record
Per-hospital isolation	Row-level security at the database

BENEFITS REALISED

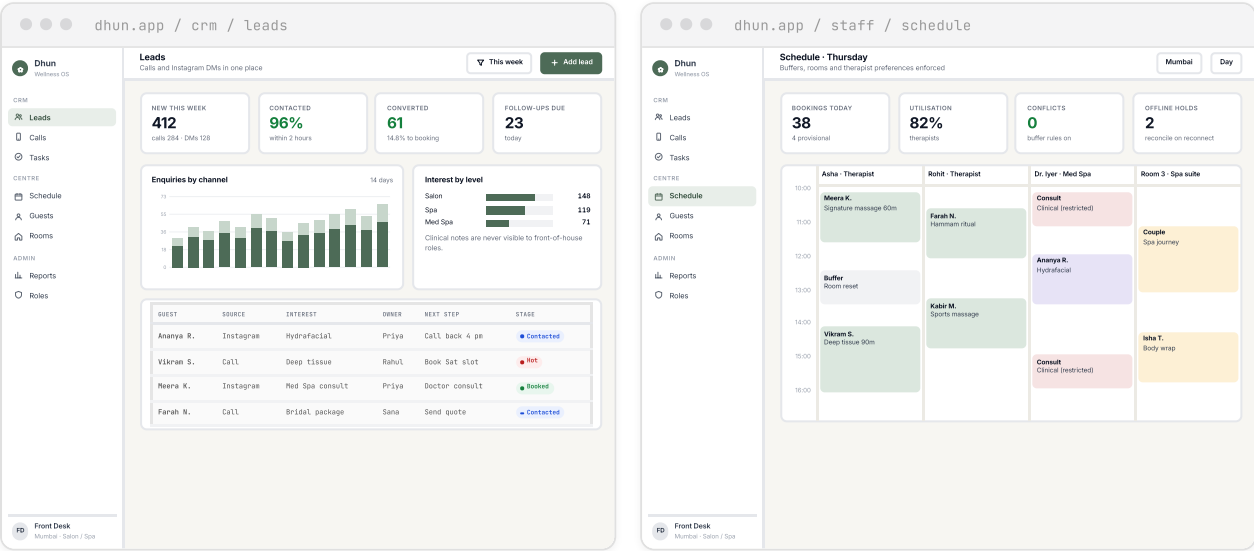
10-state	Guarded prescription lifecycle
Zero-install	Nothing for the doctor to set up
Audit	Every pharmacist action logged

AI-Native Management System for Wellness Centres

Client: Dhun Wellness

We gave a multi-level wellness chain one system of record with 3 role-locked apps on top, so the front desk never sees a doctor's clinical notes and a guest's history follows them across every centre.

- Healthcare
- CRM
- Role-Based Access
- Offline-First
- System of Record



CLIENT PROBLEM

The doctor's clinical notes were visible to the front desk, with no role boundary. Online bookings didn't flag back-to-back appointments or room clashes. 100 calls a day and 200 Instagram DMs lived in Excel and memory, and the multi-level Salon/Spa/Med Spa model had no clean way to carry a guest across levels while keeping clinical data private.

OUR SOLUTION & ARCHITECTURE

One system of record, **3 surfaces**. The centre's database stays the source of truth; an operating layer sits on its API. The CRM, staff app and consumer app are all surfaces on that one layer, so a change anywhere reconciles to a single source. Clinical data is walled off by role across 5 staff roles, and guest profiles carry across centres and levels.



WHY IT MATTERS

Clinical data walled off	Front desk never sees doctor's notes
One source of truth	3 surfaces reconcile to one record
Guests carry across levels	Salon to Spa to Med Spa, cleanly

BENEFITS REALISED

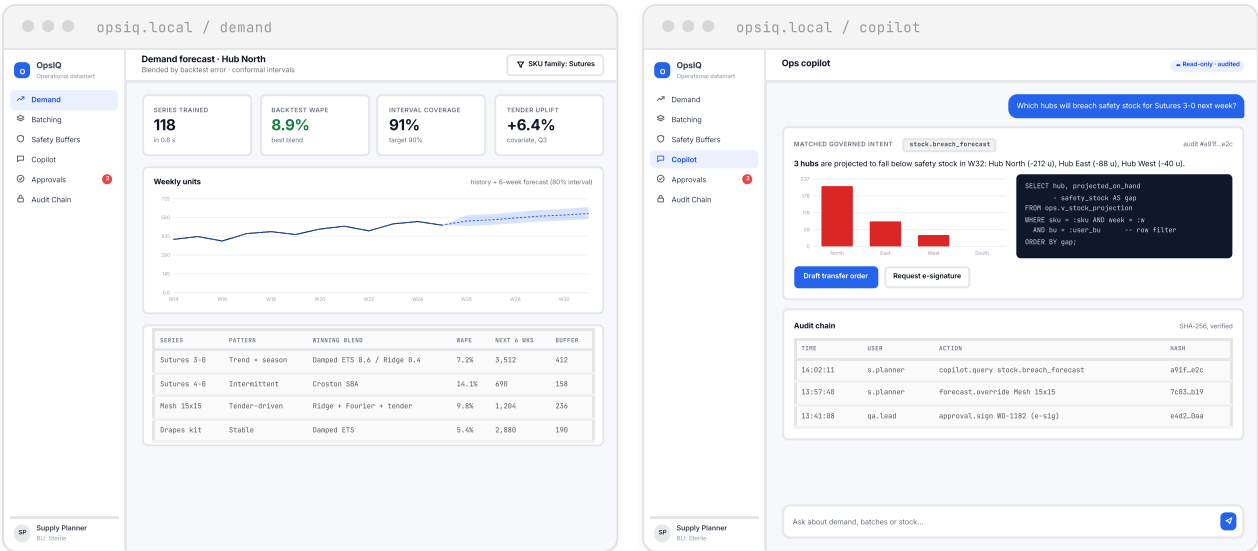
5	Role-scoped staff views
Offline-first	Keeps operating through downtime
Multi-centre	Mumbai and Delhi, no duplicates

Pharma Operations Intelligence Suite

Client: Undisclosed (Under NDA)

We built one governed layer that forecasts demand, schedules sterile production, and answers plain-English questions over operations data, with every answer audited and every write e-signed.

- Healthcare
- Demand Forecasting
- Sterile Batching
- Governed NL-to-SQL
- Audit Chain

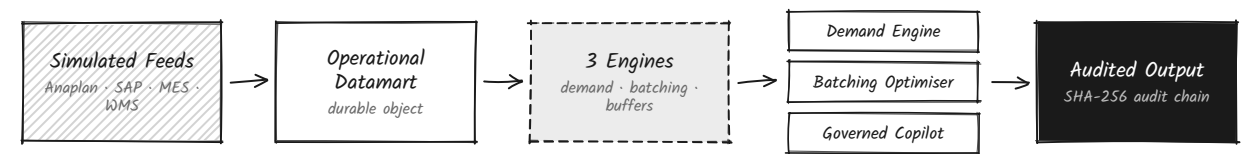


CLIENT PROBLEM

Forecasting demand across every SKU and region, scheduling sterile production across cleanrooms with sequence-dependent changeovers, and answering questions over operations data all sat in separate systems and spreadsheets, with no governed path from question to audited answer.

OUR SOLUTION & ARCHITECTURE

One operational datamart, 3 engines, one **governed query layer**. A demand engine forecasts every series; a batching optimiser sequences production and packs sterile loads; and an ops copilot answers in natural language through 8 governed intents, with free-form questions passing a read-only guard. A hash-chained audit log and e-signature approvals wrap the whole suite.



WHY PRIVATE AI

Governed query path	Free-form questions pass a read-only guard
Everything audited	Hash-chained log, e-signature approvals
Row-level filters	Business-unit scoping on every path

BENEFITS REALISED

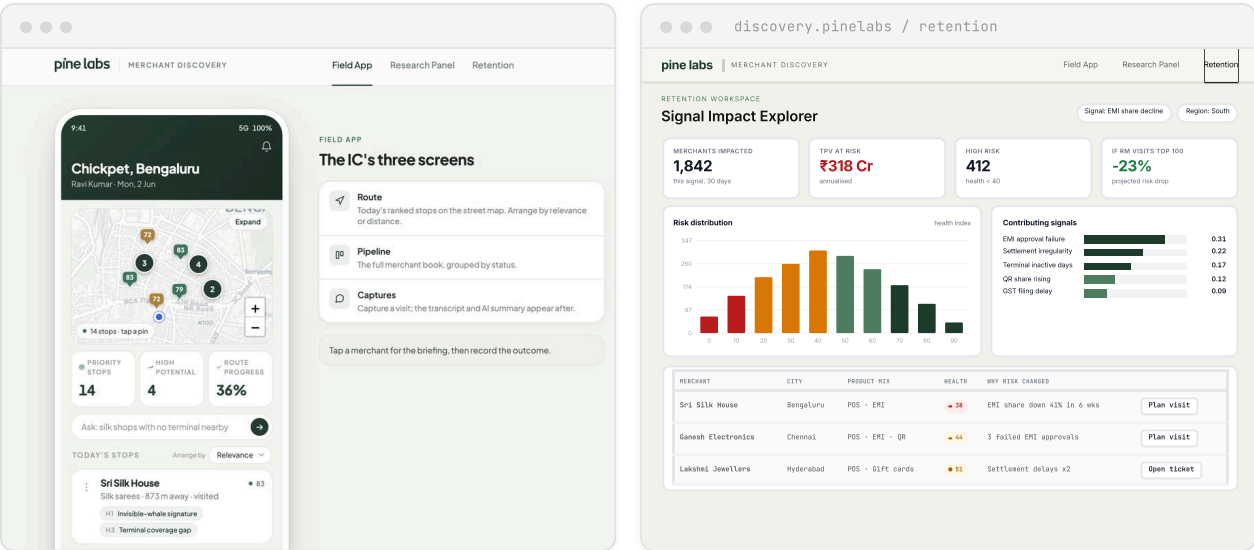
<1s	~120 series trained
8	Governed intents to parameterised SQL
SHA-256	Hash-chained, verifiable audit

Pine Labs Merchant Discovery + Retention

Client: Pine Labs Ltd

We built a platform that continuously measures how healthy each Pine Labs merchant relationship is and flags the ones about to churn early, using Pine-specific signals a generic churn model would miss.

- FinTech
- Merchant Retention
- Churn Prediction
- Field App
- Signal Explorer



CLIENT PROBLEM

Pine Labs merchant churn was managed through manual downloads and spreadsheet analysis, so at-risk merchants were identified late. Generic churn models miss Pine-specific signals across POS, QR, EMI, BNPL, gift cards and settlements, with no unified view of relationship health.

OUR SOLUTION & ARCHITECTURE

A **Merchant Health & Retention Intelligence Platform**, not a generic churn model. It continuously measures relationship health, detects product-level disengagement early, and surfaces targeted interventions. 3 surfaces: a field discovery app for IC visits, a research panel for portfolio analysis, and a retention workspace with a Signal Impact Explorer and What-if Simulator.



WHY IT MATTERS

Pine-specific signals	EMI, settlement, attachment - not generic churn
Early, not late	At-risk merchants flagged before they leave
Actionable	What-if simulator projects the fix

BENEFITS REALISED

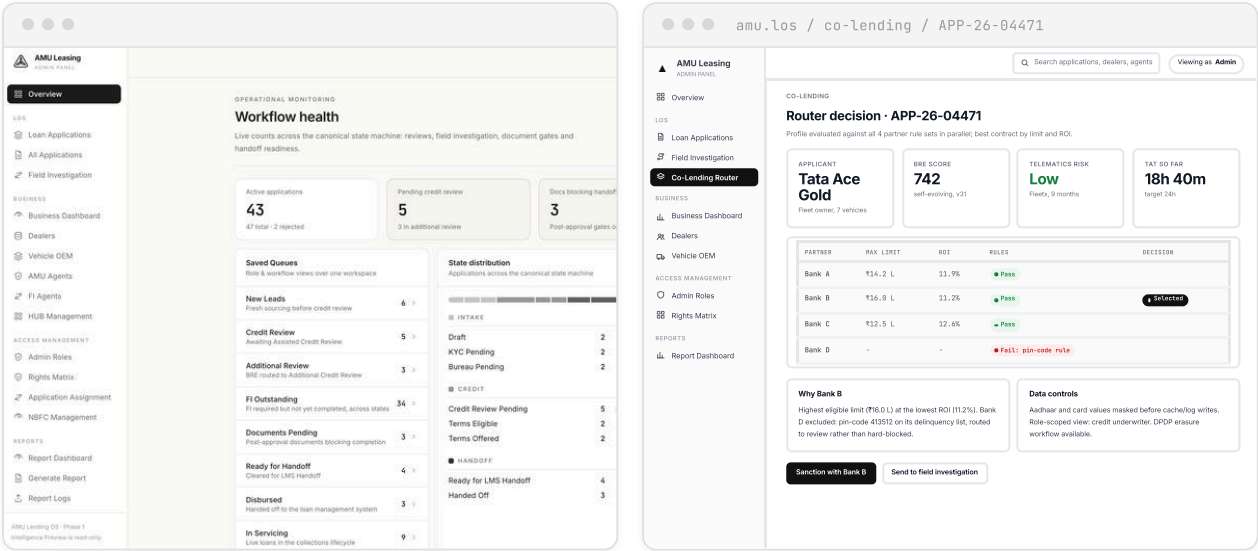
2M+	Pine Labs touchpoints analysed
Per-merchant	360 view of contributing factors
TPV-at-risk	Quantified by signal, geography, product

Autonomous Lending Platform for a Leasing Company

Client: AMU Leasing Pvt Ltd

We replaced a 5-10 day manual loan process with an autonomous lending platform that decides in 24 hours, learns from every loan it has already made, and routes each deal to the cheapest of 4 banks.

- FinTech
- Digital Lending
- Self-Evolving BRE
- Co-Lending
- RBI DLD · DPDP



CLIENT PROBLEM

AMU Leasing ran a manual, fragmented underwriting process taking 5-10 days, driven by a stagnant 120+ variable checklist with no live feedback loop between disbursed loans and incoming decisions. Multi-bank routing, regulatory controls and telematics all sat outside the core workflow.

OUR SOLUTION & ARCHITECTURE

An AI-native autonomous lending architecture with a **24-hour TAT**. An event-driven platform with a **self-evolving BRE** trained on post-disbursement feedback at 4, 6 and 12 months, a real-time co-lending router across 4 partner banks, and telematics-linked risk. Strict RBAC, data masking and DPDP-compliant erasure are baked in.



WHY IT MATTERS

Regulatory controls built-in	RBI DLD, DPDP erasure, data masking
Learns from itself	BRE retrains on post-disbursement outcomes
Masking before writes	Aadhaar/card values obfuscated in logs

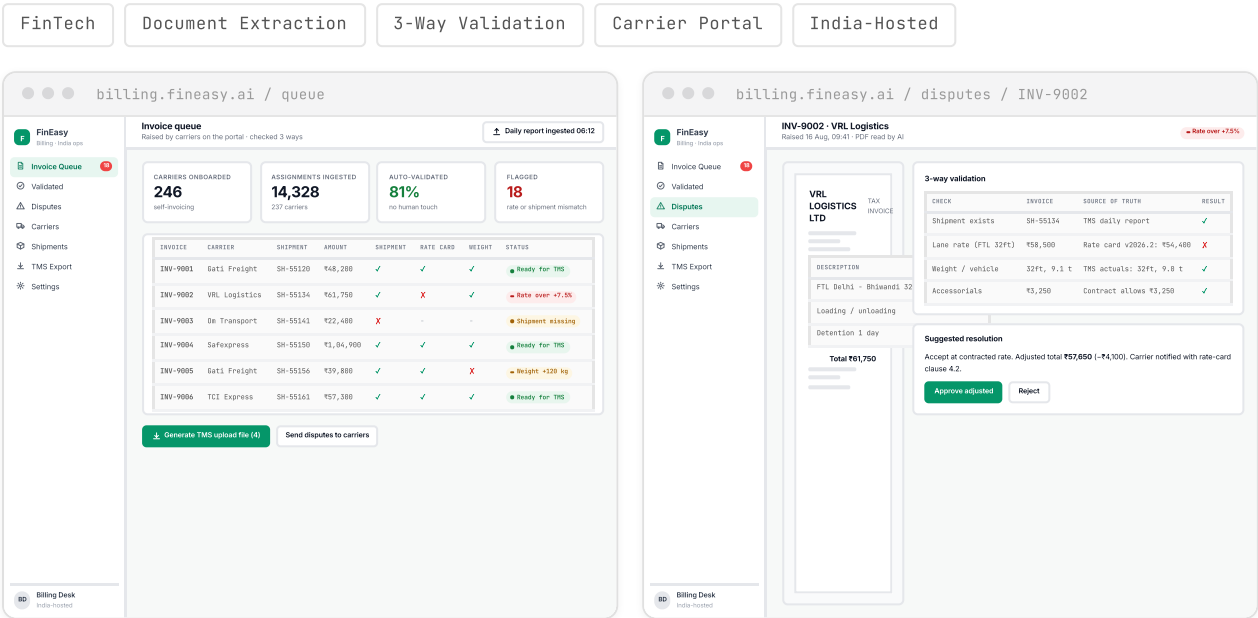
BENEFITS REALISED

24h	TAT, down from 5-10 days
4	Banks in the real-time co-lending router
Self-evolving	Scoring adapts every 4/6/12 months

AI Freight Invoice Validation Platform

Client: Global Freight Forwarder, India Operations

We let carriers raise their own freight invoices on a portal, then had AI read each invoice and check it 3 ways against the shipment and the rate card, replacing a desk that checked every bill by hand.



CLIENT PROBLEM

The client's global transport management system had no API and no carrier self-invoicing, so every freight invoice was checked by hand against shipments and rate cards.

OUR SOLUTION & ARCHITECTURE

A **carrier portal** where vendors raise their own invoices. AI reads the invoice PDFs, validates each one 3 ways against shipment actuals and rate cards, and produces a file ready to upload to the TMS. Daily shipment reports are ingested automatically by email. Now run as a managed platform, India-hosted.



WHY IT MATTERS

Bridges a closed TMS	Self-invoicing where the TMS had no API
3-way checked	Every invoice vs shipment vs rate card
India-hosted	Financial data kept in-region

BENEFITS REALISED

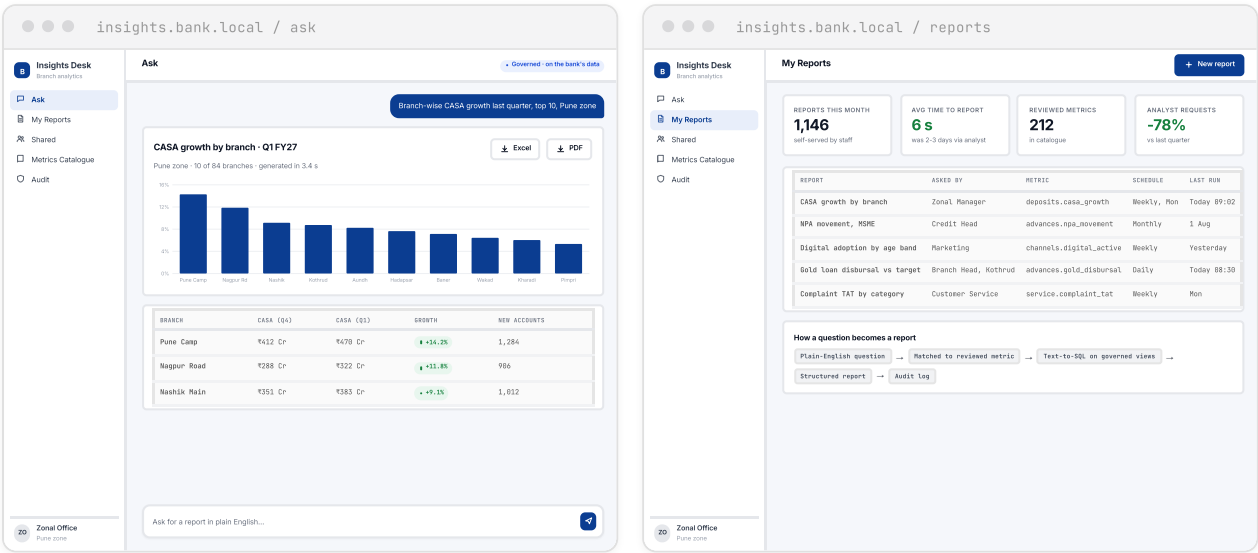
246	Carriers onboarded
14,328	Assignments across 237 carriers
Managed	Now a running platform

Natural-Language Reporting for a Bank

Client: \$7B Public Sector Bank

We let bank staff ask for a report in plain English and get a structured answer back from the bank's own data, so decisions stop waiting in a queue behind a human analyst.

- Banking
- Text-to-SQL
- Report Generation
- Governed Access
- Reusable Pattern



CLIENT PROBLEM

Teams depended on analysts to turn data requests into reports, which slowed every decision down to the speed of a queue.

OUR SOLUTION & ARCHITECTURE

A natural-language **query-to-report layer**: staff ask a question in plain English and get a structured report back from the bank's data. The same pattern now powers reporting for our manufacturing clients: a proven, reusable governed text-to-SQL approach.



WHY IT MATTERS

On the bank's data	Governed access, no raw export
No analyst bottleneck	Decisions stop waiting in a queue
Proven pattern	Reused across manufacturing clients

BENEFITS REALISED

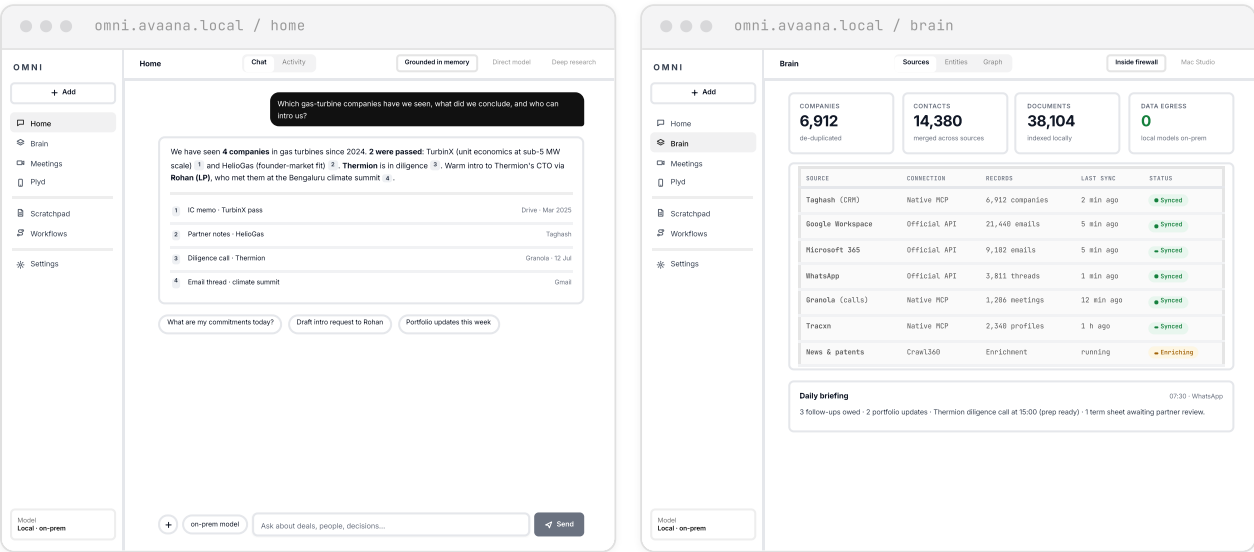
Plain-English	Question to structured report
Seconds	Instead of an analyst queue
Reusable	One pattern across deployments

On-Premise Enterprise Brain for a Venture Capital Fund

Client: Avaana Capital: \$150M Venture Capital Fund, Led by the Co-Founder of Nexus Ventures

We put one brain over all of a climate fund's scattered systems, running on a Mac Studio inside their firewall, so a partner can ask one question and get a cited answer across every email, call and database they already pay for.

- VCPE
- Enterprise Brain
- On-Prem Mac Studio
- Inside the Firewall
- Human-Approved



CLIENT PROBLEM

Avaana's operating knowledge sat across Taghash CRM, Google Workspace, M365 email, WhatsApp, Zoom calls in Granola, Drive, Tracxn and LinkedIn Sales Navigator. A question like 'which companies have we seen in gas turbines, what did we conclude, and who can I speak to' could not be answered reliably from those systems; it was answered from memory or not at all.

OUR SOLUTION & ARCHITECTURE

We deployed the Omni agent platform and **Enterprise Brain** as an AI-native back-end on Avaana's own **Mac Studio, inside the firewall**. It reaches each system through the right door (native MCP, official API, or our Crawl360 browser only where no channel exists), builds a de-duplicated entity graph, and answers in natural language citing the source of every fact. Proprietary data and local models never leave the machine; human approval gates every write.



WHY PRIVATE AI

Inside the firewall	Enterprise Brain runs on Avaana's Mac Studio
Proprietary data never leaves	Local models on-prem; cloud only via Avaana's keys
Human-approved writes	Agents read freely; every write is approved

BENEFITS REALISED

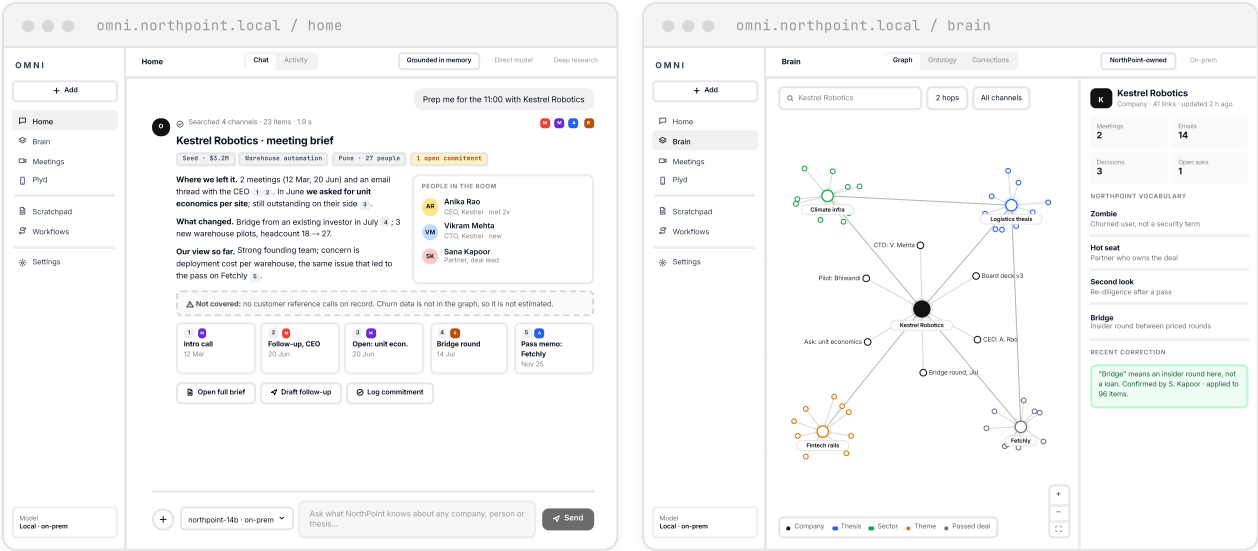
1	Queryable layer over every source
Cited	Source named on every fact
Daily	Briefings running from day one

Private Intelligence Layer for an Investment Fund

Client: NorthPoint Capital: \$155M SEBI-Registered Investment Fund

We built an investment fund its own private intelligence layer, trained on the fund's real memos and vocabulary, so meeting prep becomes one cited question instead of a hand-assembled search across 5 systems.

- VCPE
- Investment Ontology
- On-Premise
- NorthPoint-Owned
- Cited Answers



CLIENT PROBLEM

NorthPoint's hardest material (diligence notes, legal documents, sector research and the reasoning behind a pass) lives as text, in meetings and in people. A schema handles the metrics but not the part that decides investments. Partners meet 5-10 companies daily, and preparation is assembled by hand across several systems every day.

OUR SOLUTION & ARCHITECTURE

A **NorthPoint-owned** deployment: Omni runs on infrastructure NorthPoint owns, with the knowledge store, document store and connector credentials inside the fund's perimeter. We built a **NorthPoint investment ontology** from the fund's own memos, internal terms and real partner questions. Collection runs on the research seats the fund already pays for, and every answer names its source, stating what it does not cover rather than fabricating.



WHY PRIVATE AI

NorthPoint-owned	Runs on the fund's own account and keys
No licence revocation	No server that can switch it off
Refuses over fabricates	States gaps rather than guessing

BENEFITS REALISED

Fund-specific	Ontology from NorthPoint's own memos
Cited	Every fact names its source
Owned	Standing export and source escrow

AI Visa & Travel Rules Intelligence Engine

Client: Global Visa and Travel Services Provider

We built an engine that gathers scattered official visa rules and answers 'do I need a visa, which type, how do I apply', checking every answer against the exact source text it came from.

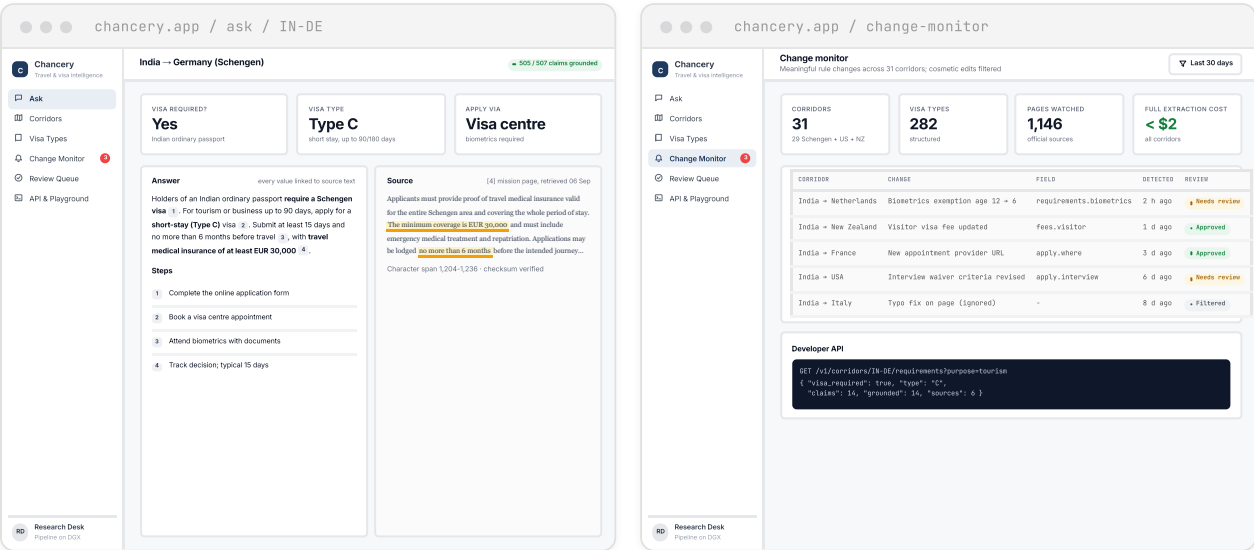
Travel

Visa Intelligence

Grounded Extraction

Knowledge Graph

On-Prem DGX



CLIENT PROBLEM

Visa rules are scattered and inconsistent across embassies and service pages, so travellers and agents can't get reliable answers to the 3 questions that matter: do I need a visa, which type, and how do I apply?

OUR SOLUTION & ARCHITECTURE

An engine that harvests official sources, extracts **structured visa rules**, and checks every answer against the exact text it came from. It includes a knowledge graph of travel corridors, change monitoring that flags meaningful rule changes, a review queue, and a developer API with a live playground. The extraction pipeline runs on 360 Labs' NVIDIA DGX.



WHY ON-PREM

Grounded, not guessed	Every claim traced to a source character span
Runs on owned DGX	Full extraction for under \$2
Human review queue	Meaningful changes flagged for a person

BENEFITS REALISED

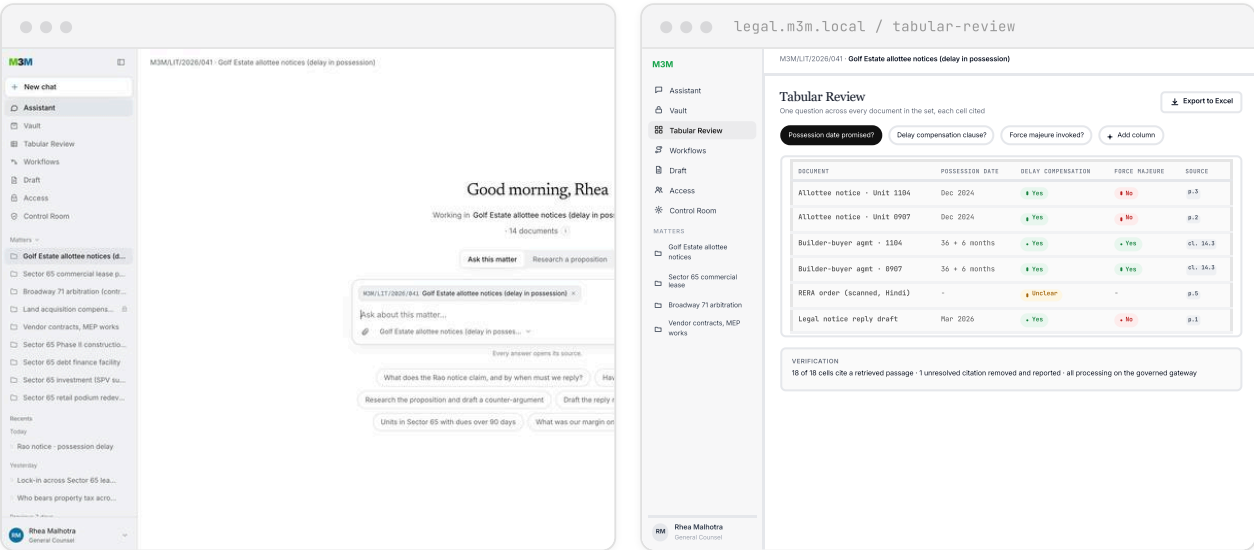
31	Travel corridors mapped
282	Visa types structured
505/507	Claims traced to source text

Legal Intelligence Platform

Client: Undisclosed (Under NDA)

We built an in-house legal team a private AI workspace that reviews documents at portfolio scale and does precedent research, and it refuses to show a citation it cannot resolve to a real source.

- Legal AI
- Document Review
- Citation Verification
- Clause Library
- On-Prem



CLIENT PROBLEM

Legal teams need fast, confident answers across their own matters and the case-law corpus, but hosted tools produce fluent answers grounded only by accident, cite nothing reliably, and cannot be trusted with privileged documents on someone else's cloud.

OUR SOLUTION & ARCHITECTURE

Verifiability before capability. Every assertion traces to retrieved sources; citations that do not resolve are removed and reported. Hybrid lexical and dense retrieval with reranking runs over matter documents and the Indian court corpus, behind one governed gateway with an audit trail for the General Counsel.



WHY PRIVATE AI

Privileged docs stay on-prem	Never sent to a hosted model
Citations verified	Unresolved ones removed and reported
Governed gateway	Audit trail for the General Counsel

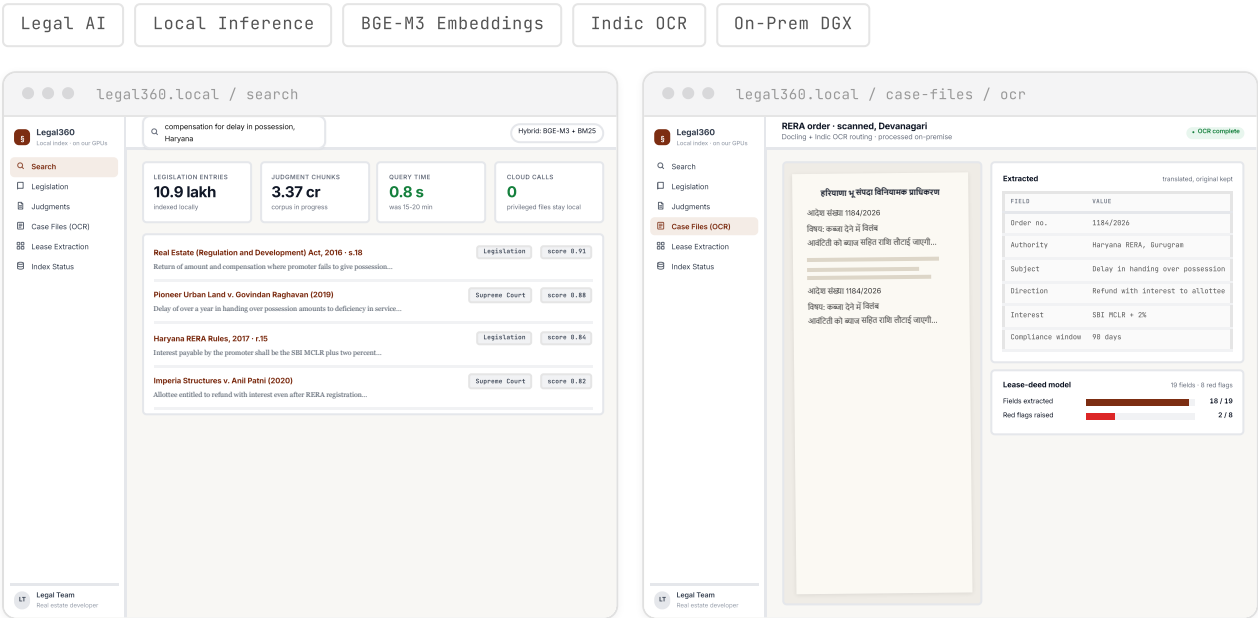
BENEFITS REALISED

Portfolio-scale	One question across a whole document set
Verified	Every citation resolves or is removed
On-prem	Hybrid retrieval over the client corpus

On-Premise Legal Intelligence

Client: Real Estate Developer's Legal Team, Plus 360 Labs' Own Indian Law Index

We built a fully local legal search index: over a million pieces of Indian legislation plus the client's own case files, running on our own GPUs, so privileged documents never touch the cloud.

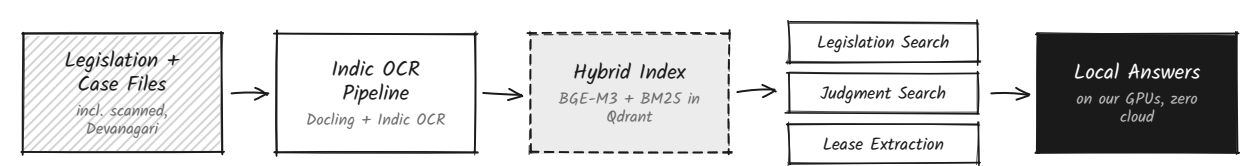


CLIENT PROBLEM

Legal teams need fast search across Indian law and their own case files without sending privileged documents to the cloud. The client's existing setup took 15 to 20 minutes per query.

OUR SOLUTION & ARCHITECTURE

A fully **local legal search index**: BGE-M3 embeddings combined with keyword search over Indian legislation and judgments, running on our own GPU hardware. An OCR extraction pipeline handles scanned case files including Devanagari documents, and a lease-deed extraction model reads 19 fields and 8 red flags, all on-premise, with vLLM serving for faster embedding throughput.



WHY PRIVATE AI

Privileged files never leave	Local inference on our own GPUs
No per-query cloud cost	Embeddings served locally with vLLM
Indic-capable	OCR for Devanagari and mixed-script

BENEFITS REALISED

10.9 lakh	Legislation entries indexed
3.37 cr	Judgment-chunk corpus in progress
30%	Faster embedding throughput with vLLM

SECTION 02

AI Research

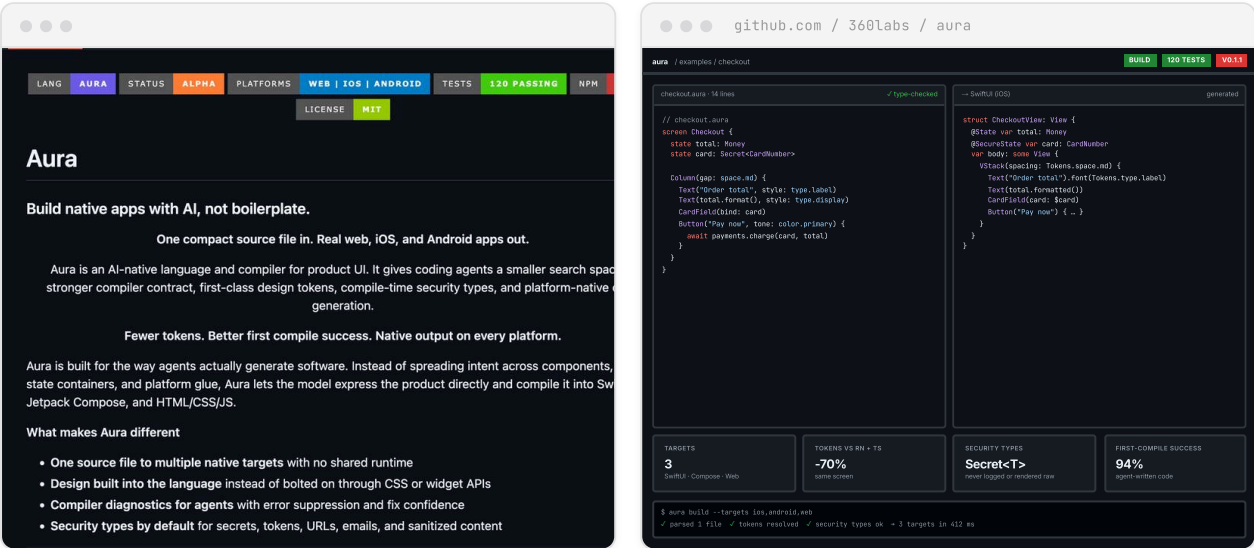
Original research and internal builds: foundation models, compilers, personal AI, on-device language models, published science and private-AI infrastructure.

Cross-Platform AI-Native UI Language & Compiler

Client: 360 Labs Research

We built a single UI language that compiles one source into native iPhone, Android and web apps at once, so a product built for AI to write needs one codebase instead of 3.

- Programming Language
- Compiler
- SwiftUI
- Jetpack Compose
- Rust



CLIENT PROBLEM

Building the same product across iOS, Android and web means writing 3 separate UI codebases in 3 paradigms. Cross-platform frameworks compromise on native feel and bloat token counts when used as LLM context, while design tokens and security constraints live outside the code and drift.

OUR SOLUTION & ARCHITECTURE

One compact source language that compiles to all 3 native targets. Aura's grammar is designed for both human and AI authoring: typically **70% fewer tokens** than equivalent TypeScript + React Native. Design tokens and security types are first-class language constructs, enforced by the compiler before any platform code is generated.



RESEARCH HIGHLIGHTS

Built for AI authoring	70% fewer tokens than React Native
Security as a type	Enforced by the compiler, not convention
True native output	SwiftUI and Compose, not a webview

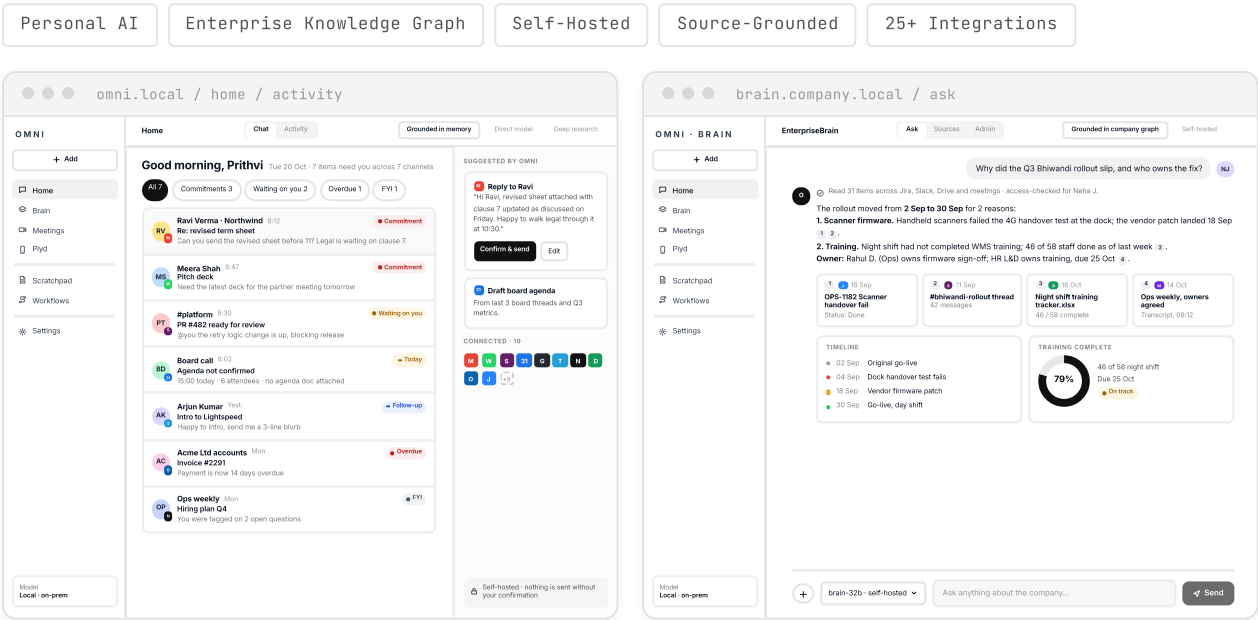
BENEFITS REALISED

3	Native targets from one source
70%	Fewer tokens than TS + React Native
120+	Grammar, type and codegen parity tests

Self-Hosted Personal Assistant & Enterprise Knowledge Graph

Client: 360 Labs Product

We built 2 products on one memory graph: a personal assistant across 25+ channels and a company brain for teams, both self-hosted, so every byte of your memory stays on hardware you own.



CLIENT PROBLEM

Personal AI assistants either live in the cloud and own your data or ignore the channels where work actually happens. Enterprise knowledge sits scattered across Slack, email, docs, tickets and repos with no single, source-grounded layer that answers with citations.

OUR SOLUTION & ARCHITECTURE

One memory graph, 2 products. **OmniAgent** centralizes communication across WhatsApp, Gmail, Slack, Telegram and 15+ more into a private, confirmation-first assistant. **EnterpriseBrain** scales the same graph company-wide with source-grounded answers across 25+ integrations. Both run **self-hosted**, spin up in 60 seconds, and never call external APIs for search.



WHY PRIVATE AI

Runs on your hardware	"Spin it up on a Pi in your closet"
No external search calls	Nothing mined, nothing leaves
Source-grounded	Every answer traces to original data

BENEFITS REALISED

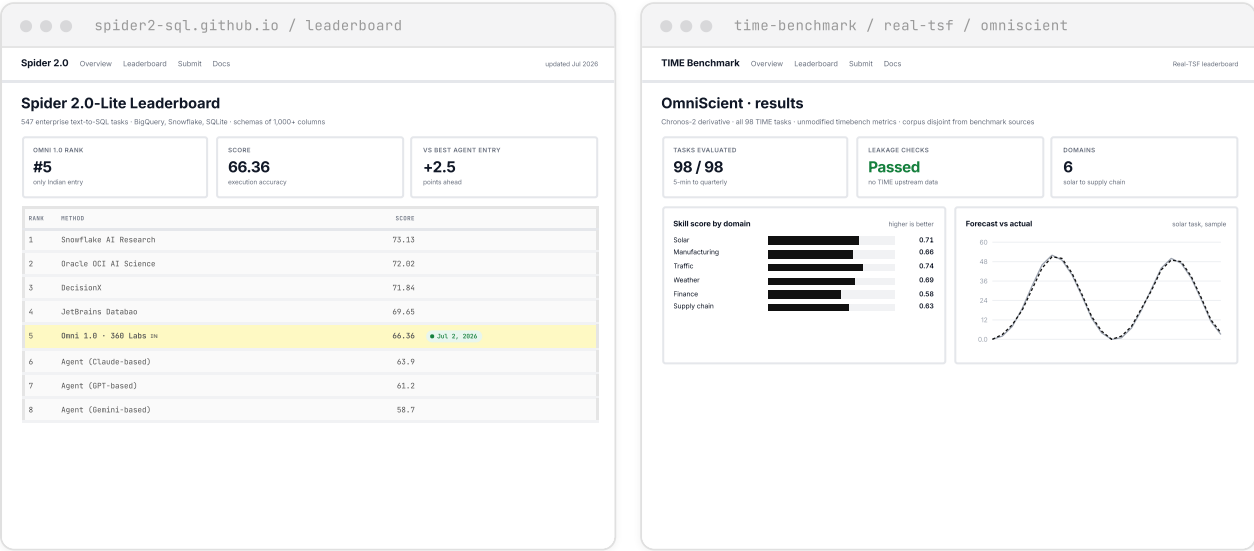
25+	Integrations across the stack
60s	Cold-start: upload a sheet, get a graph
Self-hosted	Both products, zero data egress

Foundation Models for Enterprise SQL & Time-Series Forecasting

Client: 360 Labs Research

We trained 2 of our own foundation models and put them on public leaderboards: one that writes enterprise SQL and one that forecasts business time-series, proving an Indian lab can compete at the frontier.

- Research
- Foundation Models
- Text-to-SQL
- Time Series
- Leaderboards



CLIENT PROBLEM

Most models break on enterprise data: huge real databases with 1,000+ columns, multi-step queries over 100+ lines of SQL across different dialects, and time-series forecasting across unlike domains, and reported scores are often inflated by benchmark leakage.

OUR SOLUTION & ARCHITECTURE

2 360 Labs foundation models on public leaderboards. **Omni 1.0** (enterprise text-to-SQL) ranked #5 on Spider 2.0-Lite at 66.36, the only Indian entry, beating every Claude/GPT/Gemini agent submission. **OmniScient** (time-series forecasting) was evaluated across all 98 TIME benchmark tasks, with fine-tuning corpora fully disjoint from benchmarks and their upstream sources: leakage discipline by design.



RESEARCH HIGHLIGHTS

Leakage discipline	Fine-tuning corpora disjoint from benchmarks
Enterprise-shaped	1,000+ column schemas, 100+ line SQL
Indian frontier model	Only Indian entry on Spider 2.0-Lite

BENEFITS REALISED

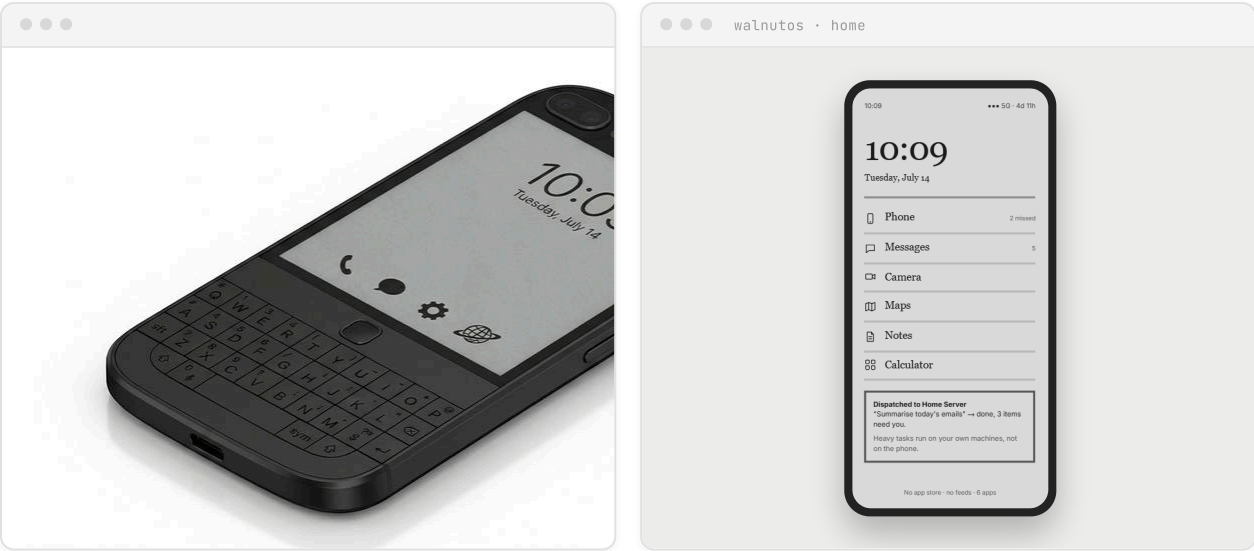
#5	Spider 2.0-Lite at 66.36
98	TIME benchmark tasks evaluated
6+	Domains: solar to supply chain

AI-Native Phone & Operating System Concept

Client: 360 Labs Product

We designed a phone that helps you fight brain rot: 6 apps, no app store, 5-day battery, and an AI-native OS that offloads heavy work to your other computers instead of doing everything on the phone.

- Hardware
- Focused Phone
- AI-Native OS
- Minimalism
- Consumer



CLIENT PROBLEM

Over 20 years, the smartphone made us the most connected and the most distracted we've ever been. Today's phones are weaponised for attention extraction, serving AI-generated slop optimised for engagement. The device meant to be a tool became a competitor for the user's focus.

OUR SOLUTION & ARCHITECTURE

A phone with only 6 apps (Phone, Messages, Camera, Maps, Notes, Calculator), no app store, 5-day battery, and **AI-native by default**. Walnut dispatches tasks to the user's other computers (home, cloud, desktop) rather than doing everything on-device, keeping what a phone existed for and removing the surrounding economy of distraction.



WHY IT MATTERS

A clean client, not every machine	Offloads to the user's own computers
No attention economy	No app store, no engagement slop
Simpler is more advanced	AI-native gateway, not a do-everything brick

BENEFITS REALISED

6	Built-in apps, nothing else
5-day	Battery life
On-device	AI-native OS by default

SLM360: On-Device Language Model Under 50MB

Client: 360 Labs (Internal)

We built India's first sub-50MB language model that runs entirely on the device, answering in 39 milliseconds with no server and no connection, private by default because nothing ever leaves the phone.

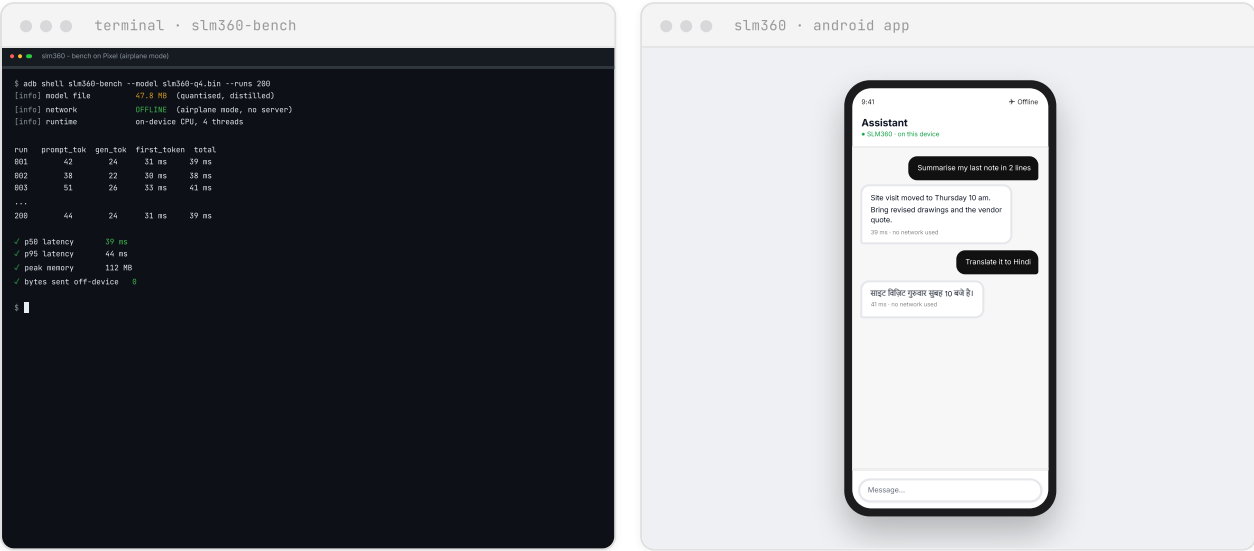
On-Device

Sub-50MB

39ms Latency

Distillation

Fully Offline



CLIENT PROBLEM

Language models are too large for deployment on constrained devices, and server round-trips break the offline, private-by-default posture that on-device AI is supposed to provide.

OUR SOLUTION & ARCHITECTURE

A **sub-50MB** model that runs entirely on-device, answering at **39ms** with no connectivity and no server. Quantised and distilled for device-class hardware, SLM360 delivers interactive responses fully offline: no data ever leaves the device.



WHY PRIVATE AI

Nothing leaves the device	No server, no connectivity in the path
Private by default	On-device inference is inherently private
India's first	Sub-50MB on-device language model

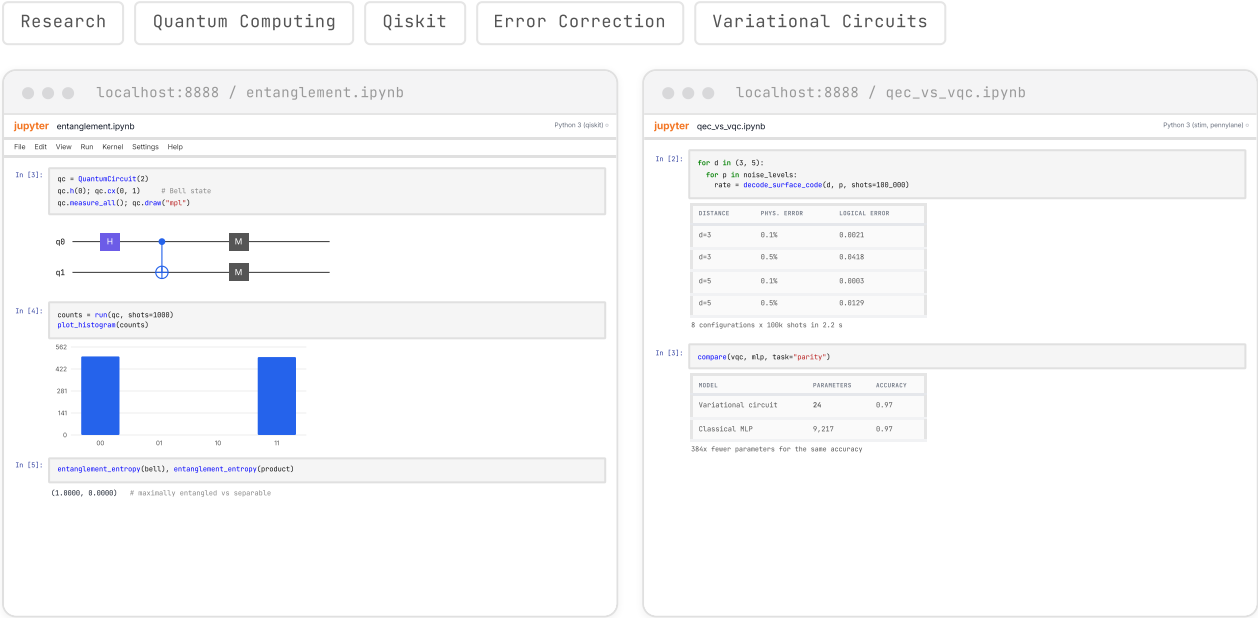
BENEFITS REALISED

<50MB	Model footprint
39ms	On-device latency
100%	Offline, zero data egress

Simulating Quantum Entanglement

Client: 360 Labs Research

We simulated entangled quantum states and measured what makes them special, built the standard quantum gates from scratch, and showed a quantum circuit matching a task with 384 times fewer parameters than a classical network.



CLIENT PROBLEM

Understanding where quantum methods beat classical ones means building and measuring quantum systems, not just reading about them, from distinguishing entangled from unentangled states to benchmarking error-correction decoders.

OUR SOLUTION & ARCHITECTURE

Simulated entangled states (Bell states), measuring **entanglement entropy** to tell them apart from unentangled ones, and built standard gates from first principles. Benchmarked a surface-code error-correction decoder across 2 code distances and 4 noise levels, and compared a variational quantum circuit against a classical neural network.



RESEARCH HIGHLIGHTS

Built, not just read	Gates and states from first principles
Measured advantage	Entanglement entropy distinguishes states
Quantum vs classical	VQC matched the task, far fewer params

BENEFITS REALISED

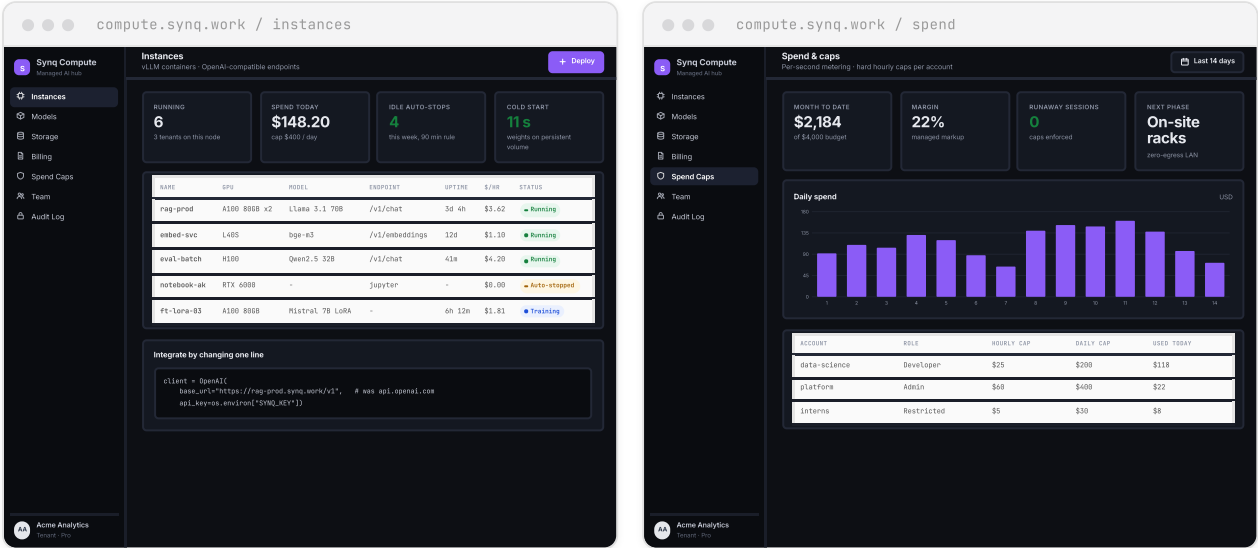
2.2s	100k-shot QEC baseline, 8 runs
384x	Fewer parameters than the classical MLP
d=3,5	Surface-code distances benchmarked

Managed AI Compute Platform

Client: Synq Managed Spaces Pvt Ltd (Synq.Work)

We turned managed office property into an AI hub: rented GPU compute with automatic billing now, graduating to a bare-metal sovereign cloud on-site later, so tenants can run AI without their data ever leaving the building.

- AI Infrastructure
- GPU Compute
- Multi-Tenant
- vLLM
- Sovereign Cloud



CLIENT PROBLEM

Standing up an AI hub the conventional way means property conversion, a data-centre buildout and hardware procurement before a single rupee of revenue. Meanwhile compliance teams reject any service whose proprietary inputs travel over public channels, and idle cloud GPUs can burn thousands over a weekend.

OUR SOLUTION & ARCHITECTURE

Launch in the cloud, then graduate to metal. A **white-labelled provisioning portal** over leased GPU capacity with automated billing gives market presence at zero hardware capex. Once demand is proven, physical GPU racks deploy inside the facility as a **zero-egress local cloud**, so the same tenants move from rented compute to hardware that never lets their data leave the building.



WHY PRIVATE AI

Graduates to zero-egress	On-prem racks keep data in the building
Single-tenant isolation	Hardware-isolated, not shared clusters
Run-away cost control	Idle instances auto-stop at 90 minutes

BENEFITS REALISED

<60s	Self-serve instance launch
OpenAI-compatible	Integrate by changing the base URL
Zero-egress	LAN once on bare metal

SECTION 03

Hardware, IoT & Robotics

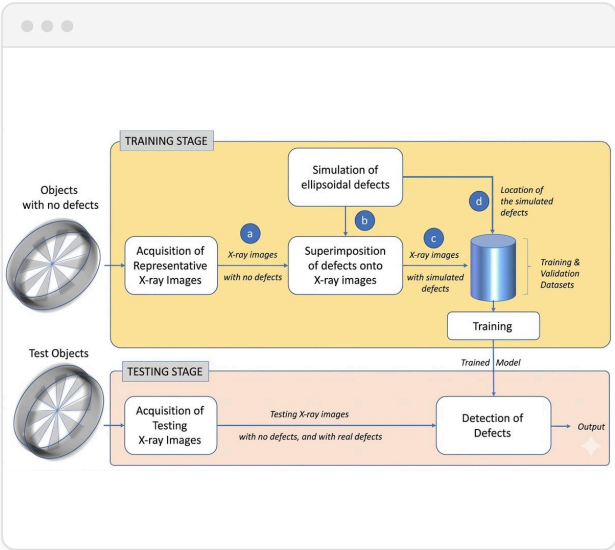
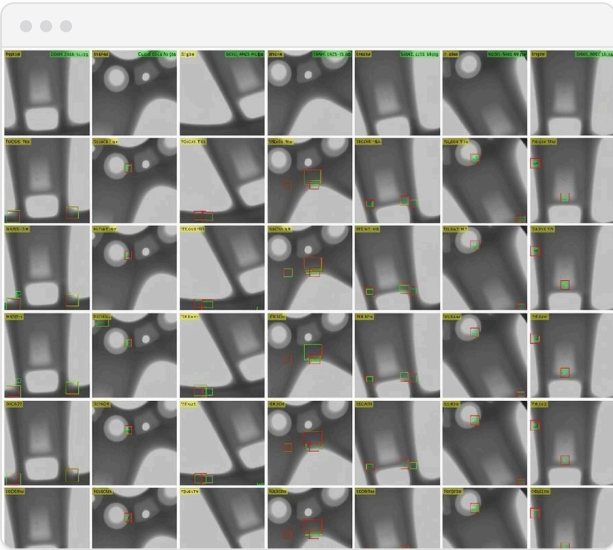
Edge devices, embedded prototypes, on-device vision and autonomy:
intelligence that runs where the data is created.

AI X-Ray Defect Inspection for Metal Castings

Client: 360 Labs (Internal)

We built an X-ray inspector for metal castings that spots defects right on the production line, running on edge hardware so the X-ray images never have to leave the plant.

- Computer Vision
- X-Ray
- Defect Detection
- Edge Inference
- On-Prem



CLIENT PROBLEM

Defect inspection of castings depends on operator judgement, and image data cannot leave the plant for cloud processing.

OUR SOLUTION & ARCHITECTURE

An **on-device vision model** inspects X-ray captures at the line, flagging defects in real time with no cloud round-trip. The model is trained on plant data and deployed on edge hardware, so defect flags route into line QC with no image egress from the facility.



WHY PRIVATE AI

Inference on the line	Edge hardware, no cloud round-trip
Images never leave	No X-ray egress from the plant
Trained on plant data	Tuned to the real production line

BENEFITS REALISED

Real-time	Per-casting defect flagging
Edge	Runs on the production line
0	Image egress from the facility

AI Camera Defect Inspection for Manufactured Panels

Client: 360 Labs (Internal)

We built a camera inspector that checks manufactured panels for defects at full line speed, scoring every panel on the device itself so production imagery stays on-site.

- Computer Vision
- Defect Detection
- Edge QA
- Line Speed
- On-Prem



CLIENT PROBLEM

Panel QA at line speed needs inspection faster than manual review, and production imagery stays on-premise.

OUR SOLUTION & ARCHITECTURE

A vision model runs QC on the line itself, scoring every panel **on-device** without moving data off-site. On-device defect detection covers the full panel surface and keeps pace with production throughput, with all imagery processed locally.



WHY PRIVATE AI

On-device scoring	Every panel judged locally
Keeps line speed	Matches production throughput
Imagery stays on-site	Processed locally, no egress

BENEFITS REALISED

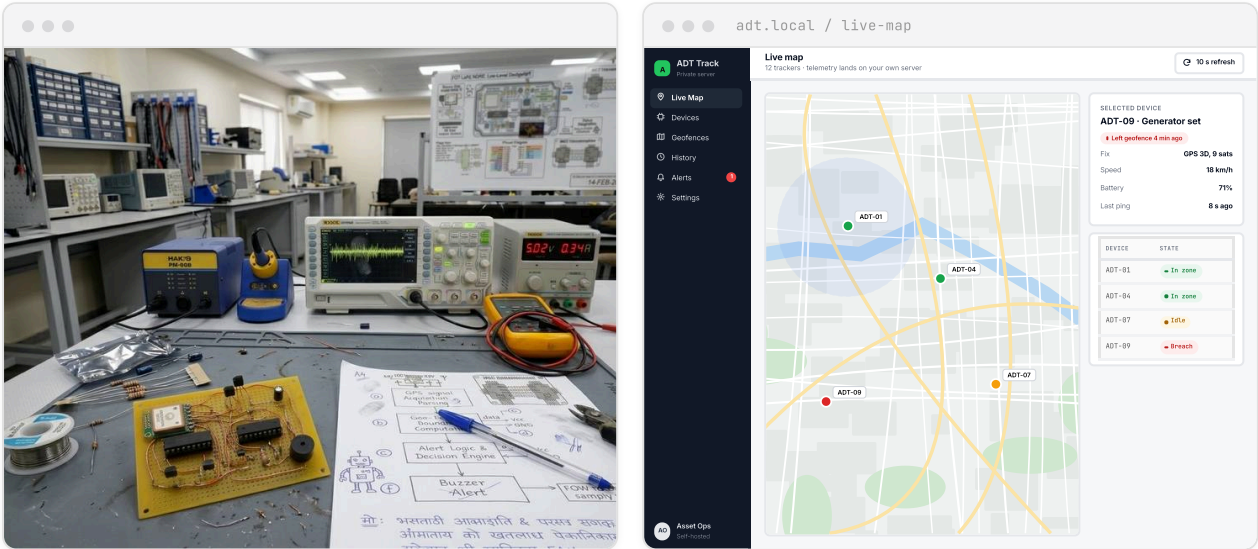
Line-speed	Inspection without the bottleneck
Full-surface	Defect detection per panel
On-prem	All imagery processed locally

Custom GPS Asset Tracker, from Circuit Board to Dashboard

Client: 360 Labs (Internal)

We designed a GPS asset-tracker from the circuit board up, so defense, logistics and industrial users can track their assets on their own dashboard at a fraction of the cost, without routing position data through a third-party cloud.

- Embedded
- GPS
- Asset Tracking
- Hardware
- Private Data Path



CLIENT PROBLEM

Commercial asset trackers are costly and route position data through third-party clouds, so you pay more and give up control of where your assets are.

OUR SOLUTION & ARCHITECTURE

An **embedded GPS tracker** (built from scratch at the hardware level on a hand-soldered perfboard with an embedded microcontroller and GPS receiver) streams position to a private dashboard at a fraction of the cost, with telemetry landing on the client's own infrastructure and geo-fencing built in.



WHY PRIVATE AI

Private data path	Telemetry lands on the client's own infra
Built from the board up	Hardware-level control, not a resold tracker
Fraction of the cost	Versus commercial alternatives

BENEFITS REALISED

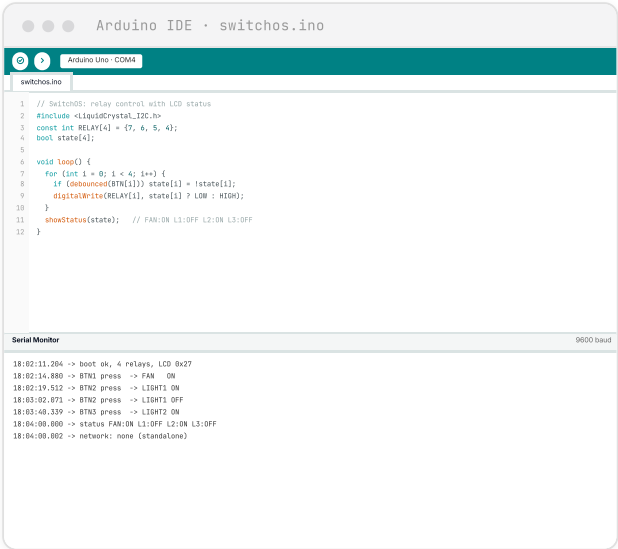
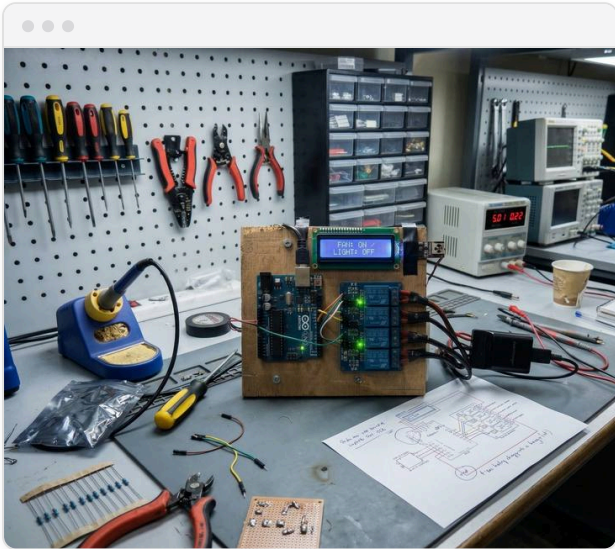
Embedded	GPS + microcontroller, hand-built
Geo-fencing	Real-time position + zones
Client-owned	Telemetry, not a vendor cloud

Low-Cost Standalone Smart Switchboard Controller

Client: 360 Labs (Internal)

We built a smart home-automation controller from off-the-shelf parts for under \$10 that retrofits onto existing Indian switchboards and works standalone: no Wi-Fi hub, no cloud account, no app subscription.

- Embedded
- Arduino
- Home Automation
- IoT
- Standalone

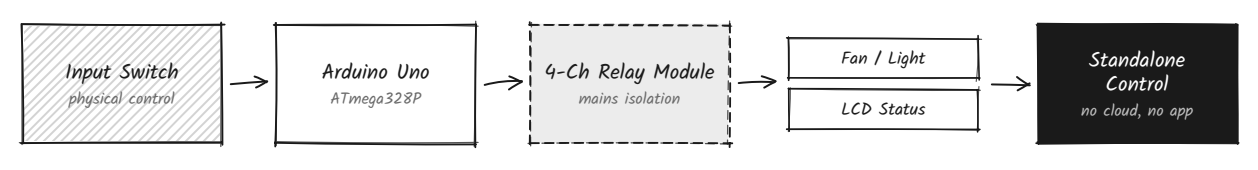


CLIENT PROBLEM

Commercial smart-home systems are expensive, require proprietary ecosystems, and don't integrate with older Indian household wiring. Most need Wi-Fi hubs, cloud accounts and app subscriptions.

OUR SOLUTION & ARCHITECTURE

A fully functional smart-home controller built from off-the-shelf components for **under \$10**. No cloud dependency, no app; it operates standalone with physical and programmable control, retrofitting into existing Indian switchboards without rewiring, deployable in any home without an electrician.



WHY IT MATTERS

Fully standalone	No cloud, no app, no subscription
Retrofit-friendly	Mounts alongside existing switches
Under \$10	Off-the-shelf components

BENEFITS REALISED

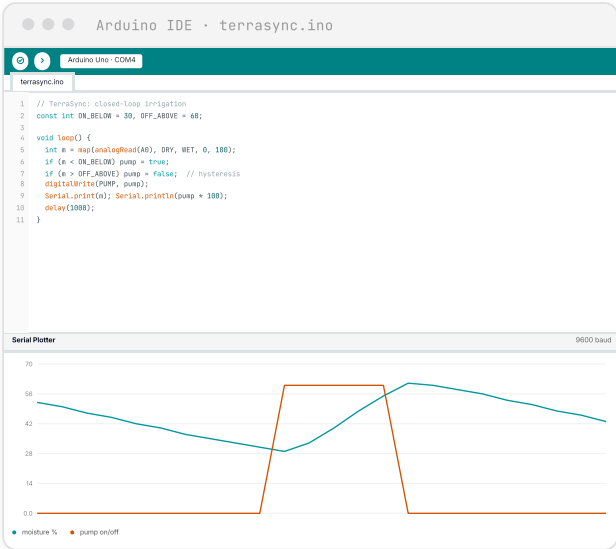
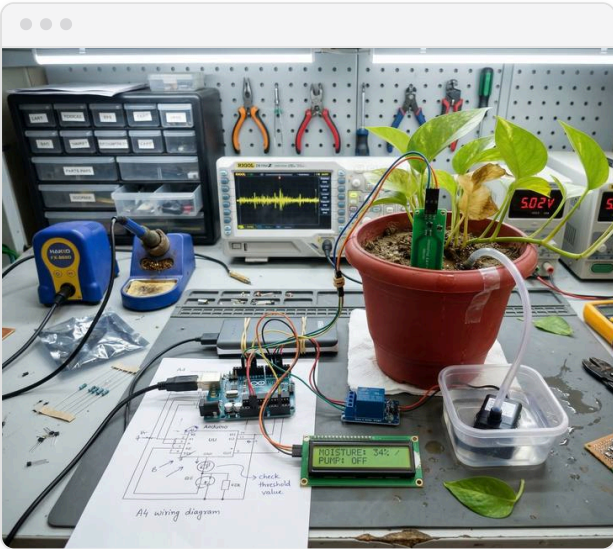
<\$10	Hardware cost
Standalone	Physical + programmable control
Retrofit	No rewiring, no electrician

Autonomous Soil-Moisture Irrigation Controller

Client: 360 Labs (Internal)

We built a self-contained irrigation system that reads soil moisture and waters the plant automatically when it dries out: battery-powered, fully on-device, with no Wi-Fi, app or cloud.

- Embedded
- Soil Moisture
- Smart Irrigation
- IoT
- Battery-Powered



CLIENT PROBLEM

Plants die from inconsistent watering, and manual irrigation is unreliable for people who travel or keep many plants. Commercial smart-irrigation systems are expensive and over-engineered for small-scale use.

OUR SOLUTION & ARCHITECTURE

A complete **closed-loop irrigation system** built for under \$8. No Wi-Fi, no app, no cloud; it reads soil moisture, decides when to water, and acts entirely on-device. Portable power-bank operation means it deploys anywhere (balconies, rooftops, remote farms) without mains electricity, with hysteresis to prevent rapid pump cycling.



WHY IT MATTERS

Reads, decides, acts on-device	No Wi-Fi, app or cloud
Deploy anywhere	Battery-powered, no mains needed
Closed-loop	Hysteresis prevents pump chatter

BENEFITS REALISED

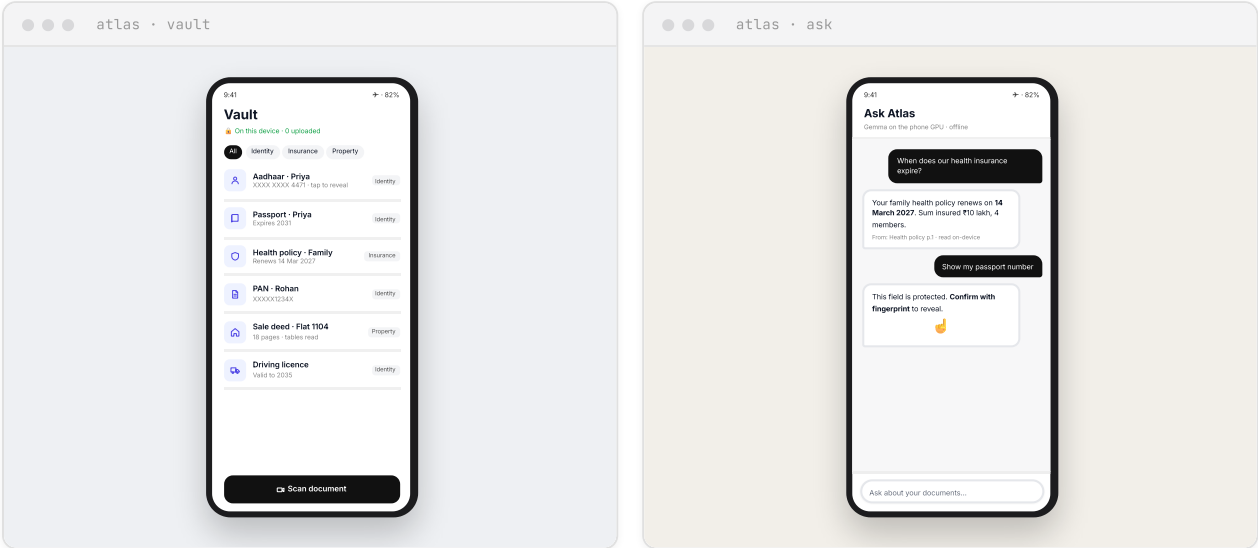
<\$8	Hardware cost
Battery	Field-deployable anywhere
Autonomous	Reads and waters with no human

On-Device AI Document Vault

Client: Consumer AI Start-Up

We built a document vault that reads your Aadhaar, passports and insurance papers entirely on the phone's own chip, so your most sensitive documents are understood by AI without ever being uploaded anywhere.

- Hardware & IoT
- On-Device AI
- Vision-Language Model
- Privacy Gate
- Zero Egress



CLIENT PROBLEM

Families keep identity, insurance and property documents scattered across phones and drawers, and cloud apps ask them to upload their most sensitive papers, the exact documents you least want on someone else's server.

OUR SOLUTION & ARCHITECTURE

A fully **on-device document vault**: OCR and a vision-language model run on the phone's GPU to read Aadhaar, PAN, passports, licences, insurance and property papers. Extraction stays consistent across re-scans, a privacy gate and biometric reveal protect sensitive fields, and an on-device assistant answers questions about the vault, with zero documents leaving the device.



WHY PRIVATE AI

Runs on the phone's GPU	OCR and VLM entirely on-device
Zero documents leave	Nothing uploaded, ever
Biometric-gated	Sensitive fields revealed only on auth

BENEFITS REALISED

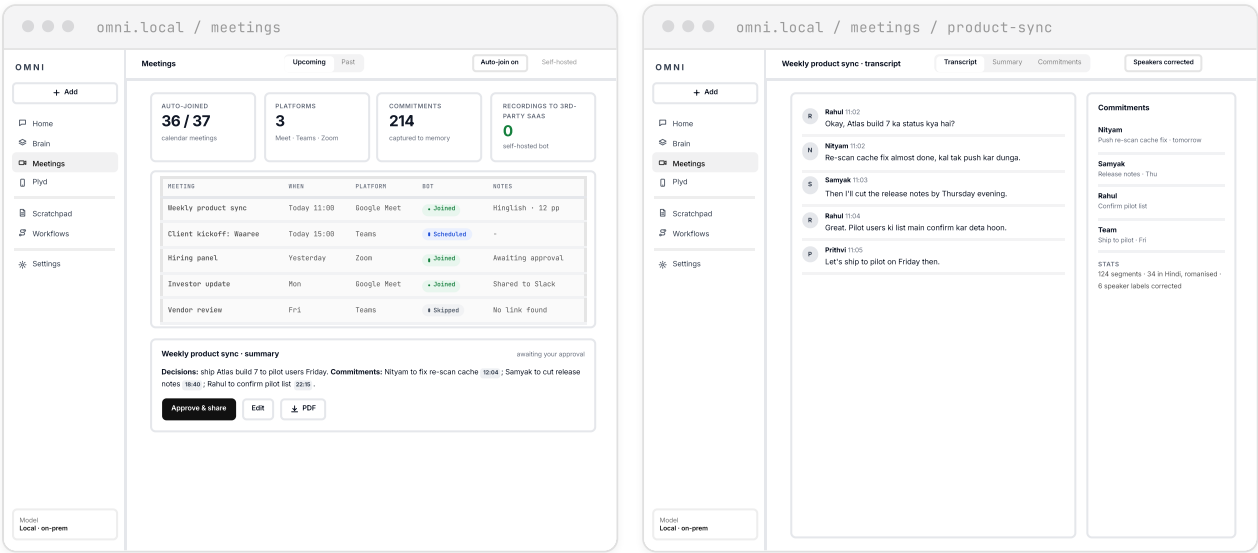
0	Documents leaving the device
38	Document entity types recognised
On-device	OCR + vision-language model

Self-Hosted AI Meeting Notetaker for Hinglish

Client: 360 Labs Internal Product

We built a self-hosted meeting bot that joins your calls, writes clean Hinglish notes, and pulls out who-owes-what into a company memory, without sending a single recording to a third-party SaaS.

- Hardware & IoT
- Self-Hosted
- Meeting Notes
- Hinglish
- Memory Graph



CLIENT PROBLEM

SaaS notetakers send every meeting recording to third parties and handle mixed Hindi and English poorly: a privacy problem and a quality problem at once.

OUR SOLUTION & ARCHITECTURE

A **self-hosted** bot that joins Meet, Teams and Zoom calls from the calendar, then produces Hinglish transcripts, corrects speaker labels, extracts commitments into a company memory graph, and shares summaries to Telegram, PDF and Slack once the owner approves. It ran on a private network, auto-joining 36 of 37 calendar meetings.



WHY PRIVATE AI

Self-hosted	Recordings never go to a third-party SaaS
Private network	Runs on infrastructure you control
Owner-approved sharing	Nothing leaves without a person's nod

BENEFITS REALISED

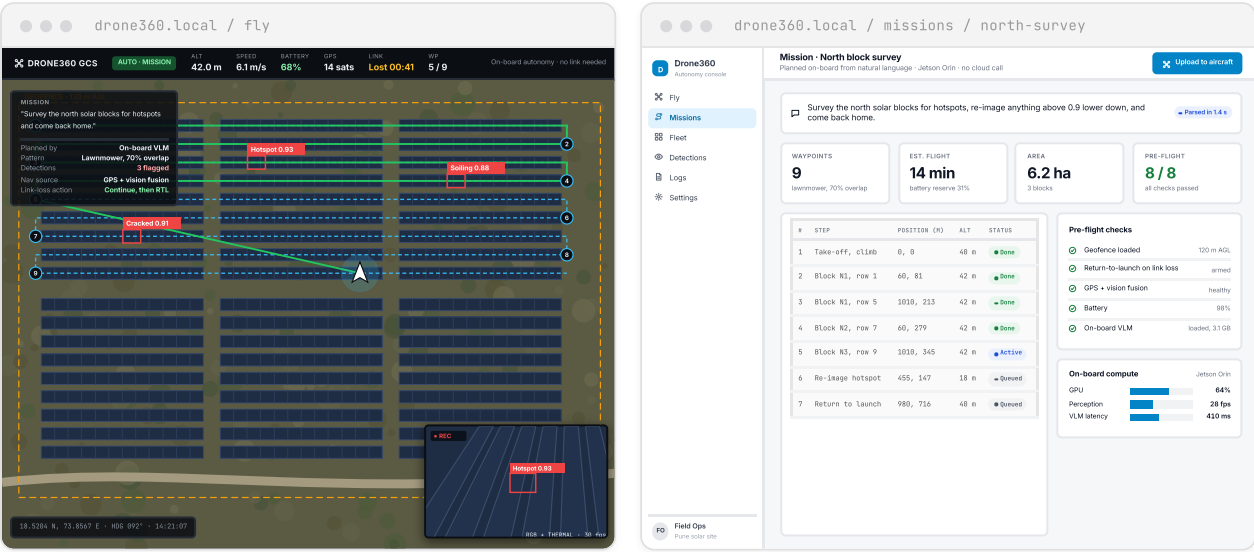
36/37	Calendar meetings auto-joined
Hinglish	Mixed-language transcripts, clean
Memory graph	Commitments captured per meeting

Autonomous AI Drone with Natural-Language Missions

Client: 360 Labs R&D

We built the full brain-and-body stack for a drone that takes missions in plain language and flies them on its own, keeping the intelligence on-board so it works where there's no pilot and no cloud link.

- Robotics
- Autonomy
- On-Board VLM
- Sensor Fusion
- Edge Compute

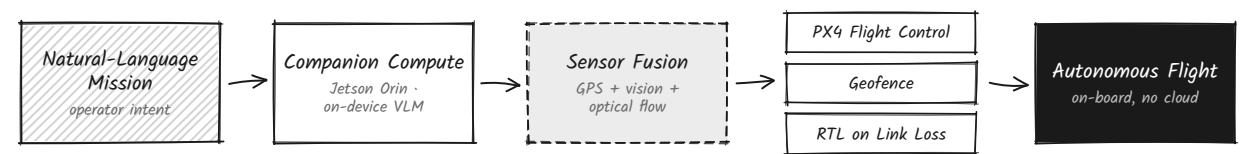


CLIENT PROBLEM

Drones that need a pilot or a cloud link can't operate in GPS-poor or disconnected environments, exactly the conditions where autonomous inspection is most valuable.

OUR SOLUTION & ARCHITECTURE

A full **autonomy stack**: flight control, a companion computer running perception and an **on-device vision-language model** for natural-language missions, sensor fusion that moves from GPS to GPS-plus-vision and optical flow, and an architecture built for on-board visual SLAM. Geofencing and return-to-launch on link loss are built in; it was delivered as a complete 16-week architecture and working monorepo.



WHY PRIVATE AI

Intelligence on-board	Perception and VLM run on the drone
Works disconnected	No pilot, no cloud link needed
Fail-safe built-in	Geofence and return-to-launch on link loss

BENEFITS REALISED

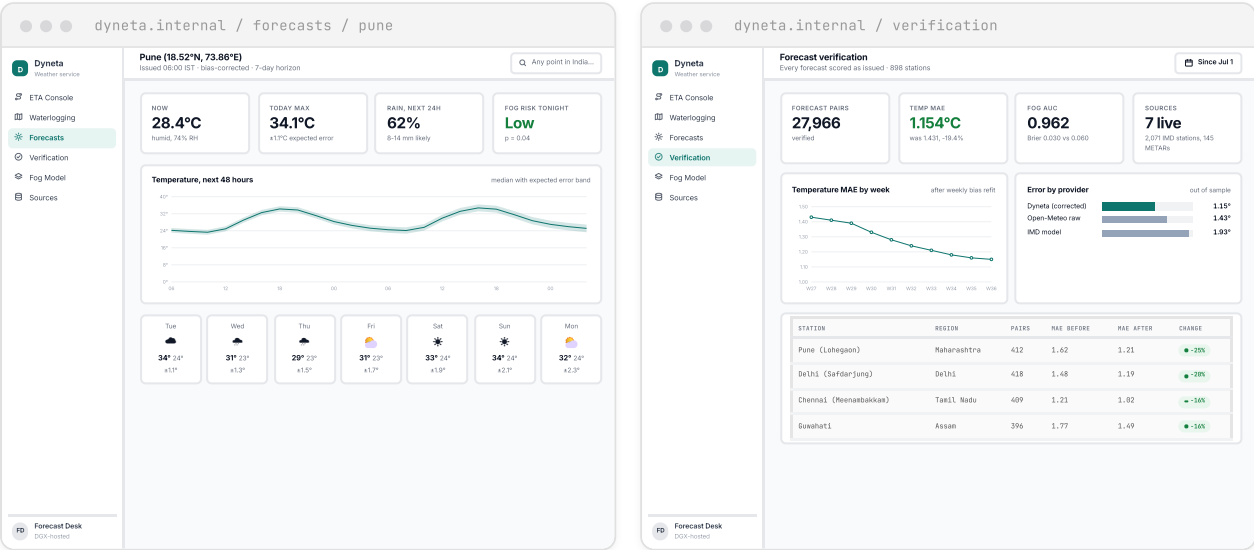
On-board	Natural-language mission planning
16-week	Architecture + working monorepo
GPS-poor	Fusion from GPS to vision + optical flow

Probabilistic Weather Forecasting Service for India

Client: 360 Labs Product

We built a weather service that gives an honest forecast for any point in India out to 7 days, and, unusually, tells you how wrong that forecast is likely to be, all verified on our own hardware.

- Meteorology
- Forecasting
- Bias Correction
- Fog Model
- On-Prem DGX



CLIENT PROBLEM

Global models don't resolve Indian convection, and official forecasts are often less accurate than free alternatives. Worse, nobody tells you how wrong a forecast is likely to be.

OUR SOLUTION & ARCHITECTURE

A weather service for **any location in India**, from now to 7 days ahead, with **honest error bars**. It ingests 7 live sources including 2,071 IMD stations and 145 airport reports, verifies every forecast as issued, corrects temperature bias site-by-site every week, and predicts fog with its own model, all running on 360 Labs' NVIDIA DGX.



WHY ON-PREM

Runs on owned DGX	No per-forecast cloud cost
Honest about error	Error bars, not false precision
India-tuned	Resolves local convection and fog

BENEFITS REALISED

19.4%	Temperature error cut out-of-sample
0.962	Fog prediction AUC
27,966	Forecasts verified at 898 stations

SECTION 04

Cybersecurity

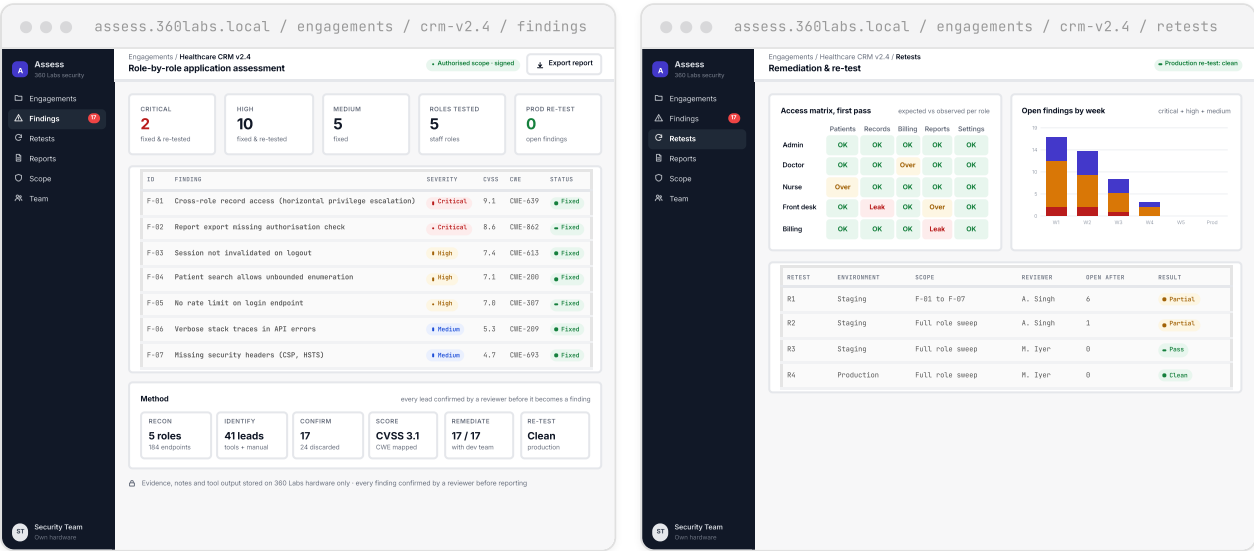
Offensive-security research and air-gapped model infrastructure, run entirely on 360 Labs' own hardware with nothing leaving the premises.

Offensive Security & Red Teaming

Client: 360 Labs Security Research (Authorised Assets Only)

We run authorised penetration tests and red-team assessments on applications before they ship, finding and fixing the serious holes, and we do it on our own hardware so the attack data never goes to anyone else's cloud.

- Cybersecurity
- VAPT
- Red Teaming
- Air-Gapped
- Responsible Disclosure



CLIENT PROBLEM

Applications handling sensitive data need rigorous security testing before launch, but commercial AI tooling refuses offensive-security work, and sending attack data and findings to a third-party cloud is itself a risk.

OUR SOLUTION & ARCHITECTURE

360 Labs runs **authorised** penetration tests and red-team assessments entirely on its own hardware. A healthcare CRM was tested role-by-role across 5 staff roles, every finding confirmed against real tool output and remediated. Methodology is deliberately human-reviewed: recon, identify, score, remediate, with automated output treated as a **lead**, **not a finding**, and only non-destructive proof-of-concept permitted on in-scope assets.



WHY PRIVATE & RESPONSIBLE

Own-hardware testing	Attack data never leaves the premises
Authorised assets only	Scope limited to what we own or are approved to test
Leads, not findings	Every claim confirmed before it is reported

BENEFITS REALISED

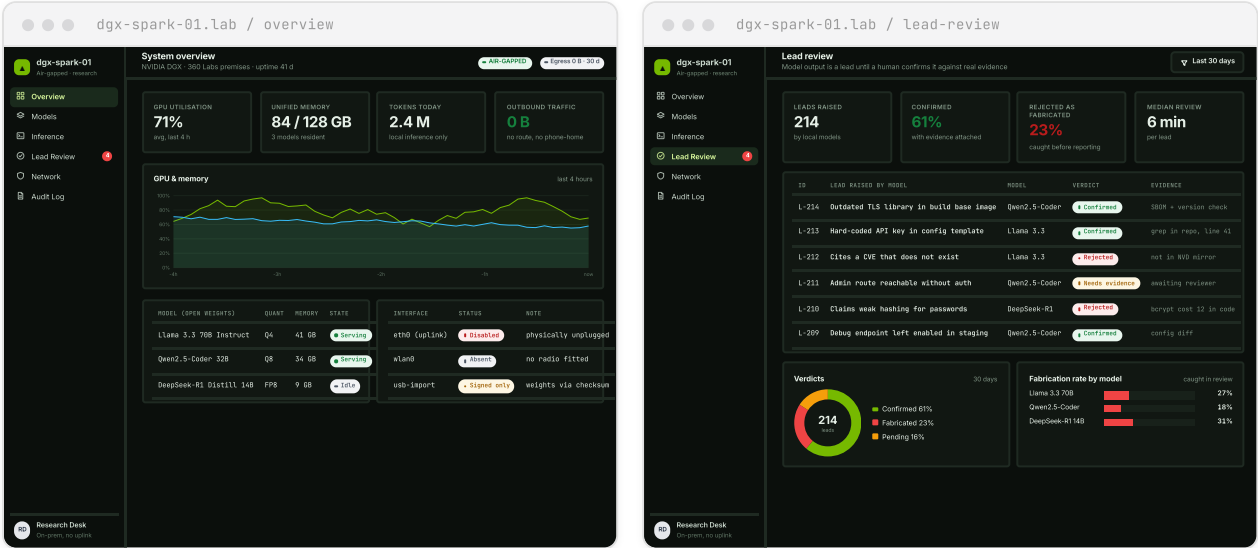
2 / 10 / 5	Critical / high / medium found & fixed
0	Findings on the production re-test
Non-destructive	Proof-of-concept only

Air-Gapped Security Research Models

Client: 360 Labs Security Research

We run open-weight models for security research entirely air-gapped on our own NVIDIA DGX, so the sensitive work of adversarial testing happens on hardware we control with nothing ever leaving the building.

- Cybersecurity
- Air-Gapped
- Open-Weight Models
- NVIDIA DGX
- On-Prem



CLIENT PROBLEM

Security research needs models that will engage with adversarial and offensive-testing questions, which hosted commercial models refuse, and sending that work to the cloud is a confidentiality risk in itself.

OUR SOLUTION & ARCHITECTURE

An open-weight model stack run **entirely air-gapped** on 360 Labs' own NVIDIA DGX for red-teaming and adversarial testing, with no data leaving the premises. The same discipline that governs our application penetration testing applies here: model output is treated as a **lead**, **not a finding**, because ungrounded models fabricate, and every claim is confirmed against real evidence before it is acted on.



WHY PRIVATE AI

Fully air-gapped	Runs on our own DGX, no network path
Nothing leaves	No data egress, no phone-home
Grounded by discipline	Output is a lead until confirmed

BENEFITS REALISED

Air-gapped	Open-weight models on-prem
0	Data leaving the premises
Lead-not-finding	Fabrication guarded by human review

2 Ways to Run Fully Private

Every engagement in this deck runs the same zero-egress architecture. You pick the ownership model that fits your security posture: the intelligence comes to your data either way.

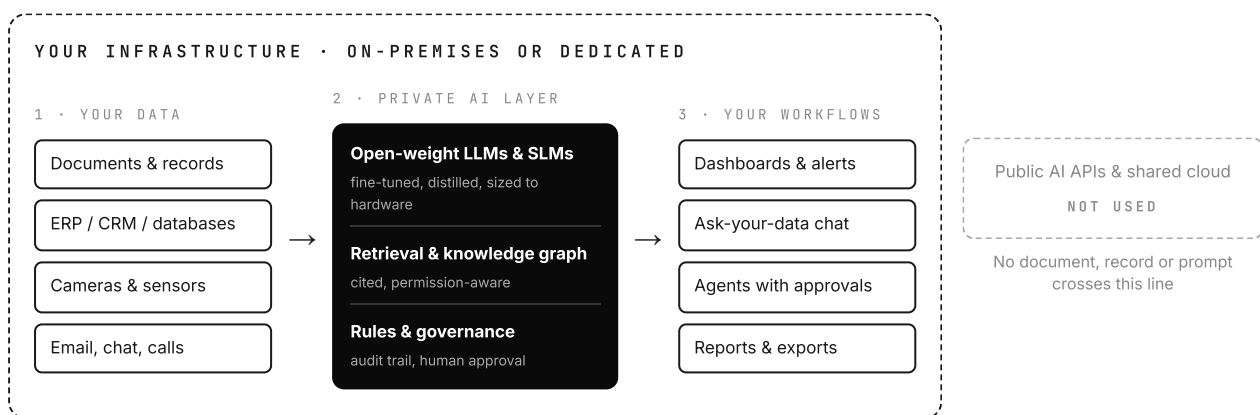
SERVERS YOU OWN

We architect, deploy and operate the full stack on GPUs you purchase and host in your own data centre. Models, vector stores, adapters and audit logs stay inside your perimeter; updates arrive as signed offline packages.

HARDWARE YOU RENT FROM US

Dedicated, single-tenant hardware operated by 360 Labs. No shared cloud tenancy, no multi-customer clusters; your workloads alone run on the machines, with the same no-egress, audit-trailed posture.

THE PRIVATE AI PERIMETER



TECHNICAL DETAILS

Model Serving

Open-weight models served locally with no hosted model interface; on-premise LLMs and custom SLMs across client estates.

Data Isolation

Row-level security at the database, on-premise retrieval, and a single governed gateway with audit trails for every query.

Air-Gapped Updates

Signed offline packages and hash-chained audit, so capabilities improve without opening a network path.

Hardware Scale

From a Mac Mini or Mac Studio running a full intelligence stack, to RTX 4090 and Jetson Orin edge nodes, to dedicated GPU racks and NVIDIA DGX for enterprise serving.

GET IN TOUCH

If you're a large enterprise looking to adopt private AI, deployed on your infrastructure, reach out.

Reach out to our founder: **pra@360labs.ai**

360LABS.AI · LIGHTYEAR 360 LABS PVT LTD

[360Labs.ai]