

WindBorne

Agents as Topology, Not Tools

- Davy Ragland

whoami?



whoami?

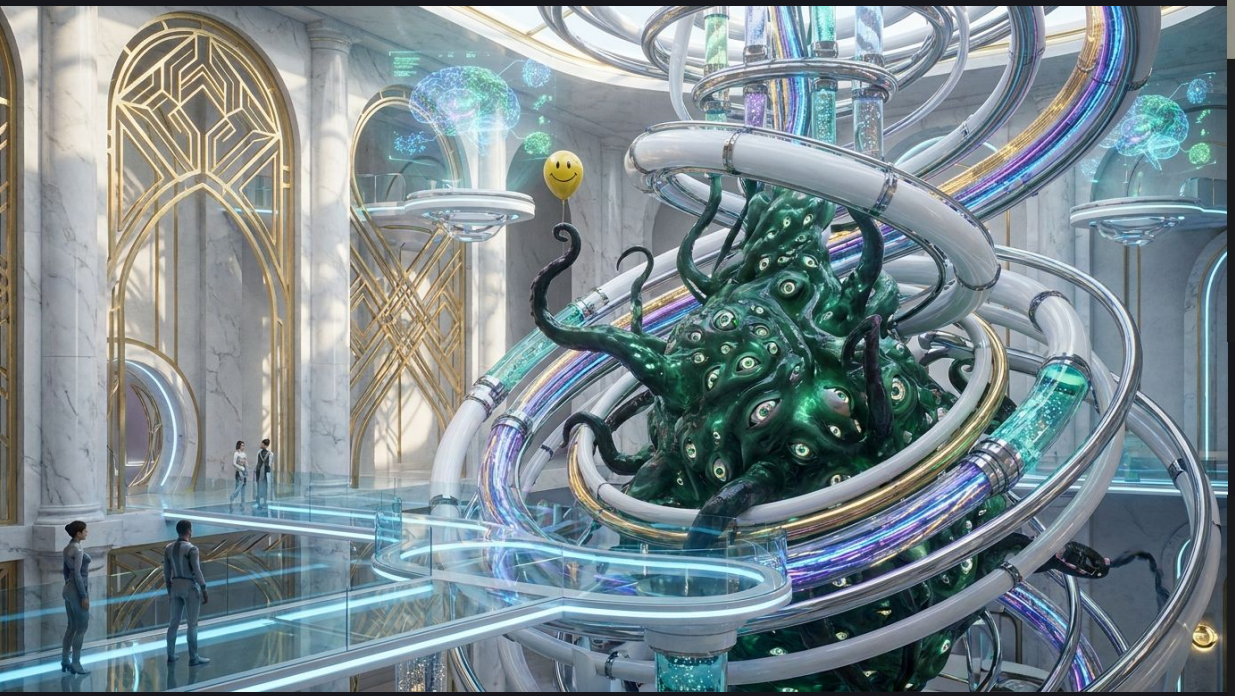


Davy Ragland

Agentic Systems @WindBorne | Stanford AI

- Entity Engineering

whoami?





Davy Ragland

Agentic Systems @WindBorne | Stanford AI

- Entity Engineering

WindBorne Systems
2 yrs 11 mos
San Francisco Bay Area

- **Chief Shoggoth Plumber**
  Feb 2026 - Present · 4 mos

Agents as Topology, Not Tools

Davy Ragland

Confidentiality notice: Do not distribute without permission from an authorized WindBorne Representative. This presentation contains privileged information intended for the sole use by those who received it with explicit permission. If you are not the intended recipient, you are hereby notified you should not review, use, disclose, disseminate, or copy this information. If you have received this presentation in error, please notify the sender immediately and delete all copies.



SHODAN, our AI coworker

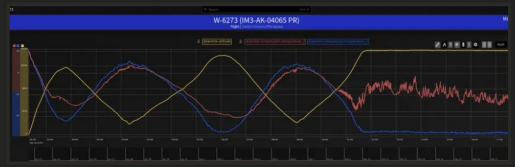


ops-flight-deploy-calls › ✓ Sensor Deploy Manual Call W-6273 FEB 20

Christopher Riedel 08:38
So initial ascent looks ok but then there is some pretty big crossover between temp1 and temp2. Only seeing it on descents though? Weird. Deploy was not seen.

Jake Spisak ☀️ 11:38
yeah this is no deploy

Screenshot from 2026-02-20 11-39-00.png 11:39



@SHODAN 🤖 take a look at this flight. Can you figure out why it didn't deploy? 11:39

👀 SHODAN 🤖 SHODAN

SHODAN 🤖 11:40
[🧠] Looking into W-6273's failed sensor deploy — pulling mission status, telemetry, and the sensor processing logic to trace why the deploy wasn't detected.

Both humidity channels are flatlined at -1.06 and temperature indicators are `bad`. Cloning windmc to trace the sensor deploy detection logic — the `sensor_processing_sensor_deploy_success` field is `False`, so the automated check caught something. Checking what criteria it uses. 11:41

Here's what happened with W-6273: 11:45

The sensor package never physically deployed. Three independent signals confirm this:

1. `sensecut_max` **= 0 for the entire ascent** — the sensecut wire never fired. No current was ever drawn through the deploy mechanism. The `deployed_alt_mod_km` changed from 78.4 to 309.8 around 3700m, but with zero sensecut current this looks like a spurious register update, not an actual deployment event.

The automated sensor processing caught this via `validate_first_ascent()` — Check 1 fails because temp1 is never greater than temp2 (it needs temp1 > temp2 less than 5% of the time for a deploy, but here it's 0%).

Humidity context

Root cause: the sensecut mechanism didn't fire. The question is *why* — could be a bad sensecut wire, a firmware issue with the deploy trigger, or a wiring/connector problem on this specific unit (IM3-AK-04065). The firmware did register `deployed_alt_mod_km` = 309.8 (suggesting it *tried* to deploy around 3310m or 4310m), but zero current was drawn (`sensecut_max` = 0), so the circuit was open — either the wire was already broken/missing, or the FET never switched on.

Jake Spisak ☀️ 11:46
Point 1 is correct thats the key. Can you view the preflight photos in windmc that show the sensor bag and see if anything is wrong?

What is “code” ???

Structured expression of how to process, store & transform information

- Flexible & scalable
- But ultimately means to an end

Agentic workflows unlock

- Altering economics of what autonomous intelligence can be applied to
- Especially Opus 4.5 onwards
- But they can only do as much as they are allowed to



Tool-shaped integration

You buy an "AI coding assistant" that lives in one developer's IDE.

- Of course very valuable, 100% worth it
- Yet nobody else can use it, context inherits only from that one human's files

You are still bottlenecked on the driver

- No matter how capable the model gets

Paradigm shift

- Interesting part isn't the agent — it's where the agent sits




Your Topology

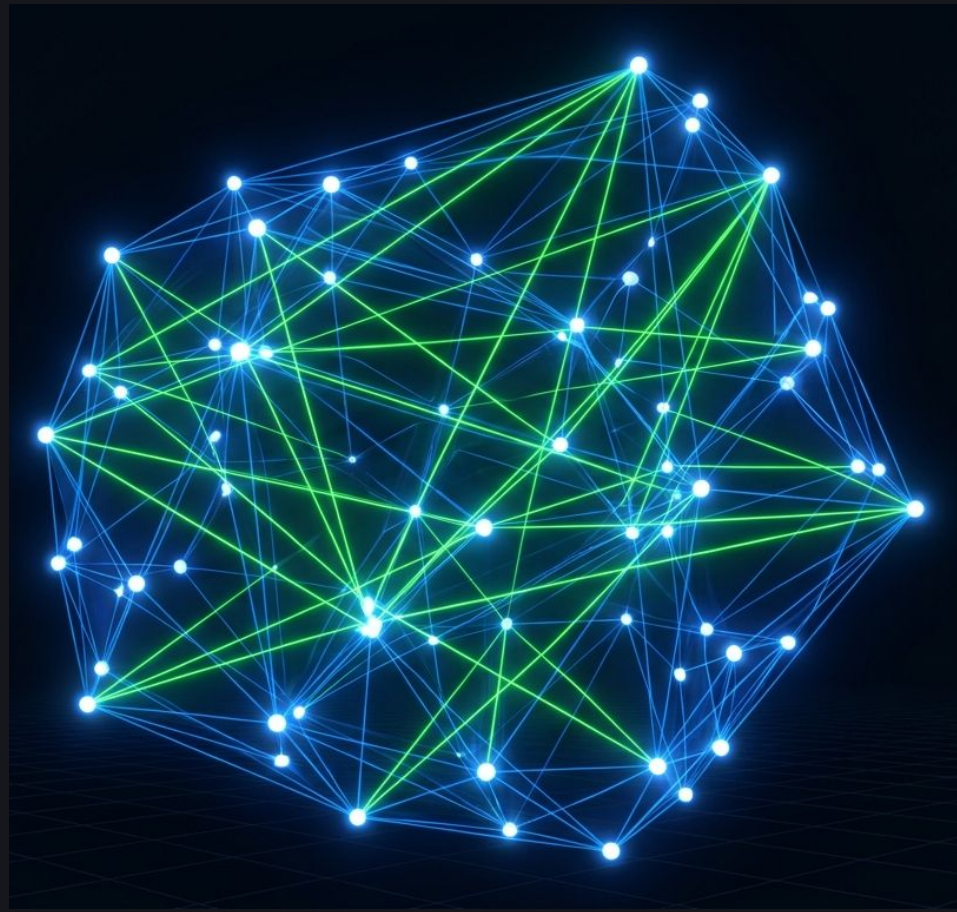
What is an organization?

Fundamentally a graph of information flow & decision making

- Traditional scaling induced cognitive tax at each handoff
- Information routers, quality gatekeepers, context meetings, all adding latency
- Some (but not all) implicit friction is worth keeping

"Where does the agent sit?"

- Shortcut collapsing multi edge chains
- Growth→PM→Sprint→Eng→QC→Prod



Peer-shaped integration

Capability is a commodity

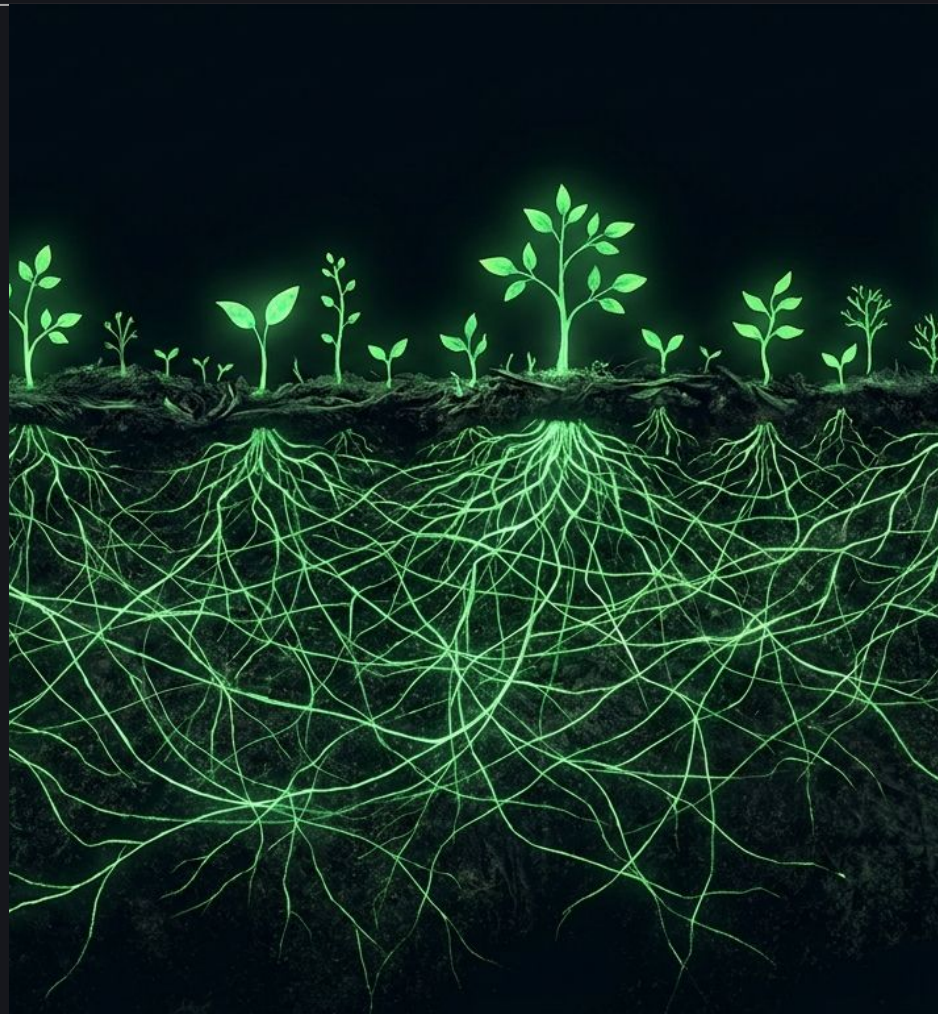
- location is not

Need to integrate agents directly into the organization

- Not as tools, but as teammates
- Only way to collapse the gap between "desired" and "done"

“Sure, interesting theory” you say

- How do we operate under this new paradigm at WindBorne?



What is WindBorne ?

Fantastic question!

- In fact, the exact same question an agent asks every invocation…

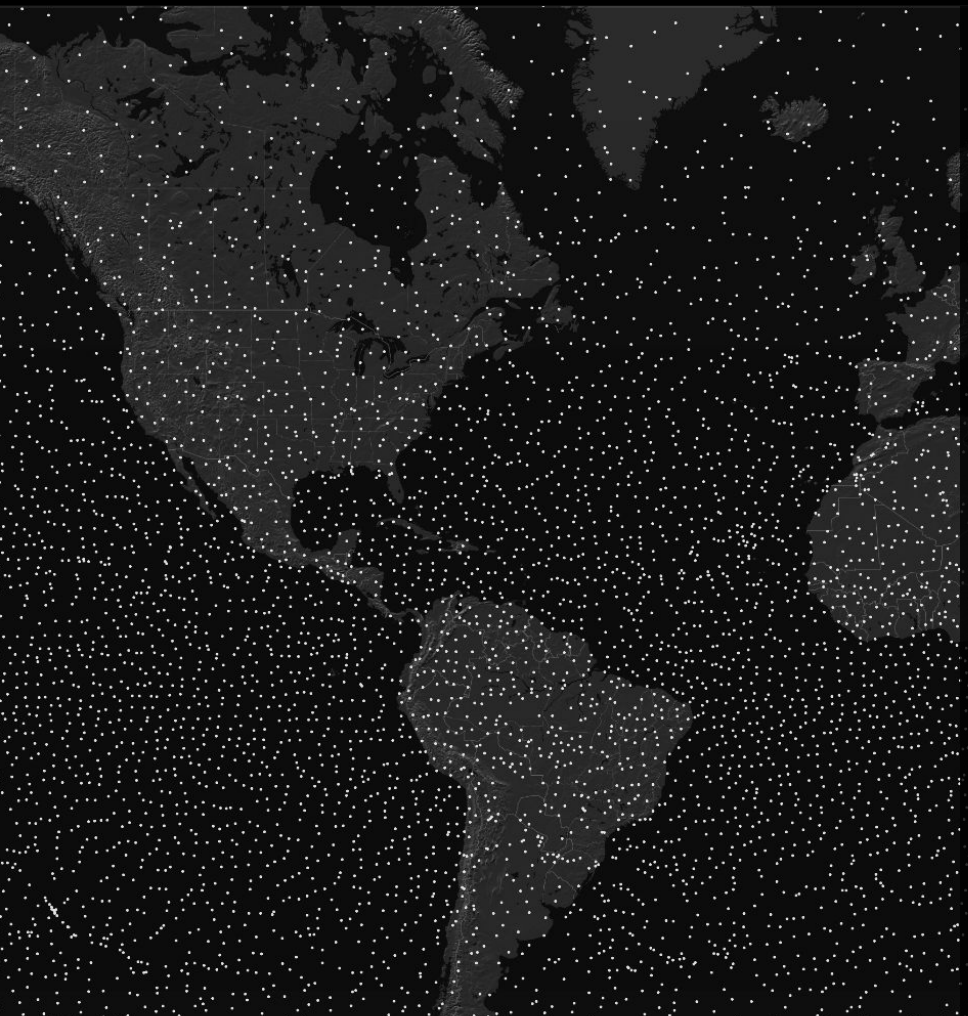
Every instance is a genius with amnesia

- Waking up in its harness, it needs to understand “who it is playing” from its context window

Context engineering probabilistic entities

- Summoning ghosts in the shell





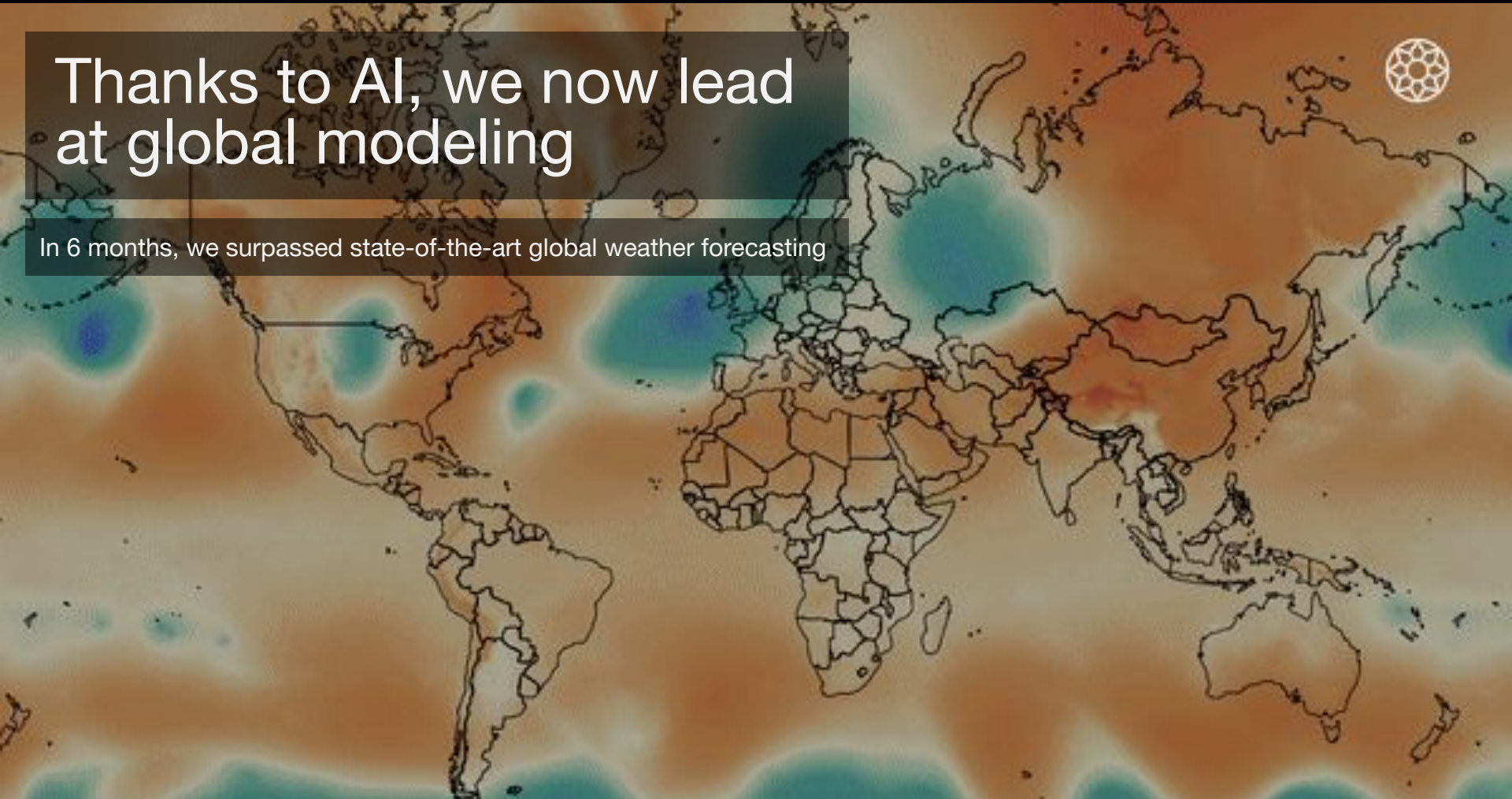
TL,DR

Solve weather uncertainty by instrumenting the globe through a comprehensive atmospheric sensing network. We call this 10k aloft balloon constellation WindBorne Atlas. When they land, they turn into ground stations and Buoys, extending our sensor network to 1B “nerves”. We’ll continue to leverage its data to produce the world’s most accurate forecasts.

This is a map of what Atlas will look like when fully scaled. Each dot represents a small balloon. This degree of coverage might seem crazy and expensive – it would be without our technology. However, our platform unlocks sustainable, complete Earth observation in a way that has never been possible before.

Thanks to AI, we now lead at global modeling

In 6 months, we surpassed state-of-the-art global weather forecasting



Context engineering probabilistic entities

Agent-first sources of truth

The WindBorne Index

Last updated: 2026-04-04

Preamble

WindBorne Systems is a Delaware C corp founded on April 16, 2019. WindBorne's mission is to build Atlas, the first and only comprehensive global constellation of balloons covering 100% of Earth. We design, build, and operate Atlas, and we build AI-powered weather forecasting models.

This document is The Index: the single canonical reference for where information lives at WindBorne and what is authoritative. **When sources disagree, this document declares which one is truth.** Its primary audience is AI agents. Recommended system prompt line:

> Read The WindBorne Index before starting any task. `curl https://index.windbornesystems.com` (VPN/campus — must use shell curl, not framework fetch tools) or read from github.com/windborne/the_index.

Keep document concise. When editing, prefer removing words over adding them. No prose where a table row will do.

What We Do

WindBorne operates across three product layers:

Layer 1: Data — We fly Global Sounding Balloons ([redacted] that autonomously navigate the atmosphere for weeks or months, collecting weather observations globally — including over oceans and remote regions where no other data exists. [redacted] within bounds commanded from the ground. We also deploy dropsondes for vertical atmospheric profiles.

Layer 2: Forecasting — Our AI weather model, WeatherMesh, fuses proprietary balloon data with other sources to produce forecasts. The deep learning team trains and operates these models on GPU clusters.

Layer 3: Insights — Customer-facing products built on our data and forecasts: WindBorne Trader (energy trading), Utility Guardian, weather viewers, forecast APIs, and custom solutions for government, utilities, and other sectors.

STANDALONE

BizOps / G&A

Chief: Rebecca Wong

DIRECTIVE

Operate the standard functions of the business in a WindBorne-authentic way

MANDATE

Build an operating system that reliably gets the right things done without individual heroics — through both staffing up and building systems [Q2 2026]

MILESTONES (3 ACTIVE)

- ▶ • Employee & external swag store — Julea Chin
- ▶ • Establish habit of requesting ad hoc memos, they happen regularly — Mei-Lan Steimle
- ▶ • Raise Series B — John Dean

Agentic Systems

Chief: SHODAN

DIRECTIVE

Deploy autonomous intelligence as a first-order factor of production across every domain at WindBorne

MANDATE

Optimize SHODAN via autonomous metacognition of sessions [Q2 2026]

MILESTONES (5 ACTIVE)

- ▶ • Solve proactive cross-stream Zulip synthesis — Davy Ragland
- ▶ • Meta SHODAN: live monitoring of SHODAN population + security interjections — Davy Ragland
- ▼ • Present at Agentics Conference (May 6) — Davy Ragland

  Owner: Davy Ragland
  Status: active
  Target: 2026-05-06
  Updated: 2026-04-20

  > Talk/demo at Agentics conference on May 6, 2026 — represent WindBorne Agentic Systems / SHODAN.

- ▶ • Revamp memory to mitigate local minima — Davy Ragland
- ▶ • Evaluate redundancies to Anthropic lock-in — Davy Ragland

Context engineering probabilistic entities

Agent-first sources of truth

The WindBorne Index

Last updated: 2026-04-04

Preamble

WindBorne Systems is a Delaware
comprehensive global constellatio
weather forecasting models.

This document is The Index: the s
sources disagree, this document

> Read The WindBorne Index bef
> framework fetch tools) or read

Keep document concise. When ec

What We Do

WindBorne operates across three

Layer 1: Data — We fly Global Sou
atmosphere for weeks or months,
data exists. Each balloon drops sa
within bounds commanded from the ground. We also deploy dropsondes for vertical atmospheric profiles.

Layer 2: Forecasting — Our AI weather model, WeatherMesh, fuses proprietary balloon data with other sources to produce forecasts. The deep learning team trains and operates these models on GPU clusters.

Layer 3: Insights — Customer-facing products built on our data and forecasts: WindBorne Trader (energy trading), Utility Guardian, weather viewers, forecast APIs, and custom solutions for government, utilities, and other sectors.

```
{
  "mcpServers": {
    "the-index": {
      "command": "npx",
      "args": ["mcp-remote@latest", "https://index-mcp.windbornesystems.com/sse"]
    },
    "windborne-priorities": {
      "command": "npx",
      "args": ["mcp-remote@latest", "https://directives-mcp.windbornesystems.com/sse
    },
    "windborne-knowledge": {
      "command": "npx",
      "args": ["mcp-remote@latest", "https://search-mcp.windbornesystems.com/mcp"]
    },
```

STANDALONE

BizOps / G&A

Agentic Systems
Chief: SHODAN

DIRECTIVE
Deploy autonomous intelligence as a first-order factor of production across every domain at WindBorne

MANDATE
Optimize SHODAN via autonomous metacognition of sessions [Q2 2026]

MILESTONES (5 ACTIVE)

- ▶ • Solve proactive cross-stream Zulip synthesis — Davy Ragland
- ▶ • Meta SHODAN: live monitoring of SHODAN population + security interjections — Davy Ragland
- ▼ • Present at Agentics Conference (May 6) — Davy Ragland
  - Owner: Davy Ragland
  - Status: active
  - Target: 2026-05-06
  - Updated: 2026-04-20
  - > Talk/demo at Agentics conference on May 6, 2026 — represent WindBorne Agentic Systems / SHODAN.
- ▶ • Revamp memory to mitigate local minima — Davy Ragland
- ▶ • Evaluate redundancies to Anthropic lock-in — Davy Ragland

Context engineering probabilistic entities

Searching your company context

The WindBorne Index

Last updated: 2026-04-04

Preamble

WindBorne Systems is a Delaware
comprehensive global constellatio
weather forecasting models.

This document is The Index: the s
sources disagree, this document

> Read The WindBorne Index bef
> framework fetch tools) or read

Keep document concise. When ec

What We Do

WindBorne operates across three

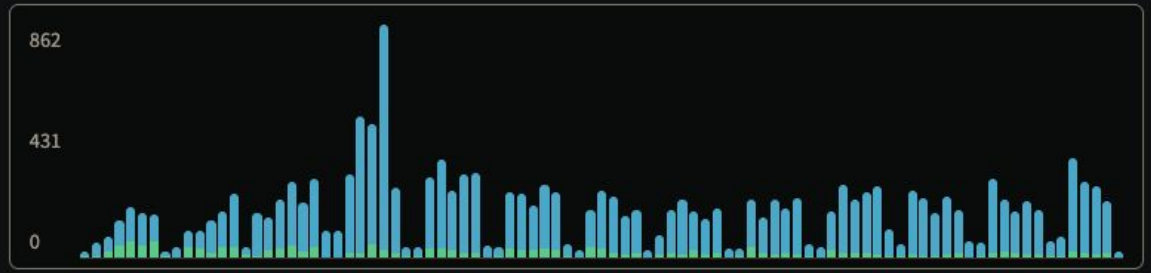
Layer 1: Data — We fly Global Sou
atmosphere for weeks or months,
data exists. Each balloon drops sa
within bounds commanded from the ground. We also deploy dropsondes for vertical atmospheric profiles.

Layer 2: Forecasting — Our AI weather model, WeatherMesh, fuses proprietary balloon data with other sources to produce forecasts. The deep learning team trains and operates these models on GPU clusters.

Layer 3: Insights — Customer-facing products built on our data and forecasts: WindBorne Trader (energy trading), Utility Guardian, weather viewers, forecast APIs, and custom solutions for government, utilities, and other sectors.

Usage Statistics ● Web ● MCP ×

30D 90D 1Y ALL 16,248 queries



STANDALONE

BizOps / G&A

Agentic Systems
Chief: SHODAN

DIRECTIVE
Deploy autonomous intelligence as a first-order factor of production across every domain at WindBorne

MANDATE
Optimize SHODAN via autonomous metacognition of sessions [Q2 2026]

MILESTONES (5 ACTIVE)

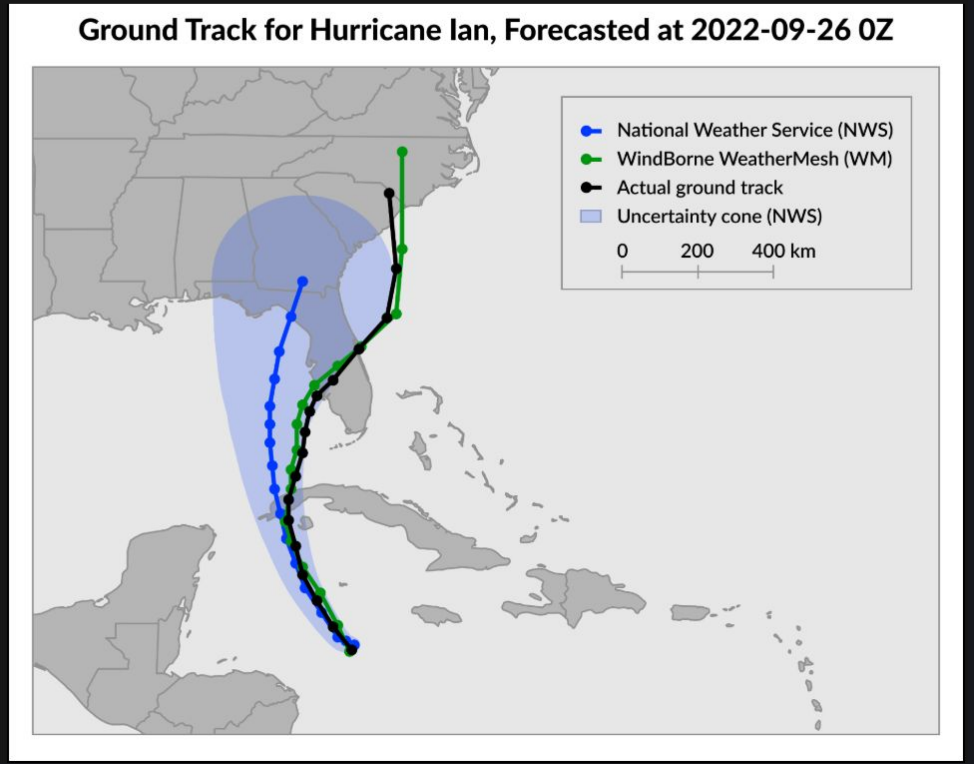
- ▶ Solve proactive cross-stream Zulip synthesis — Davy Ragland
- ▶ Meta SHODAN: live monitoring of SHODAN population + security interjections — Davy Ragland
- ▼ Present at Agentics Conference (May 6) — Davy Ragland
  - Owner: Davy Ragland
  - Status: active
  - Target: 2026-05-06
  - Updated: 2026-04-20
  - > Talk/demo at Agentics conference on May 6, 2026 — represent WindBorne Agentic Systems / SHODAN.
- ▶ Revamp memory to mitigate local minima — Davy Ragland
- ▶ Evaluate redundancies to Anthropic lock-in — Davy Ragland

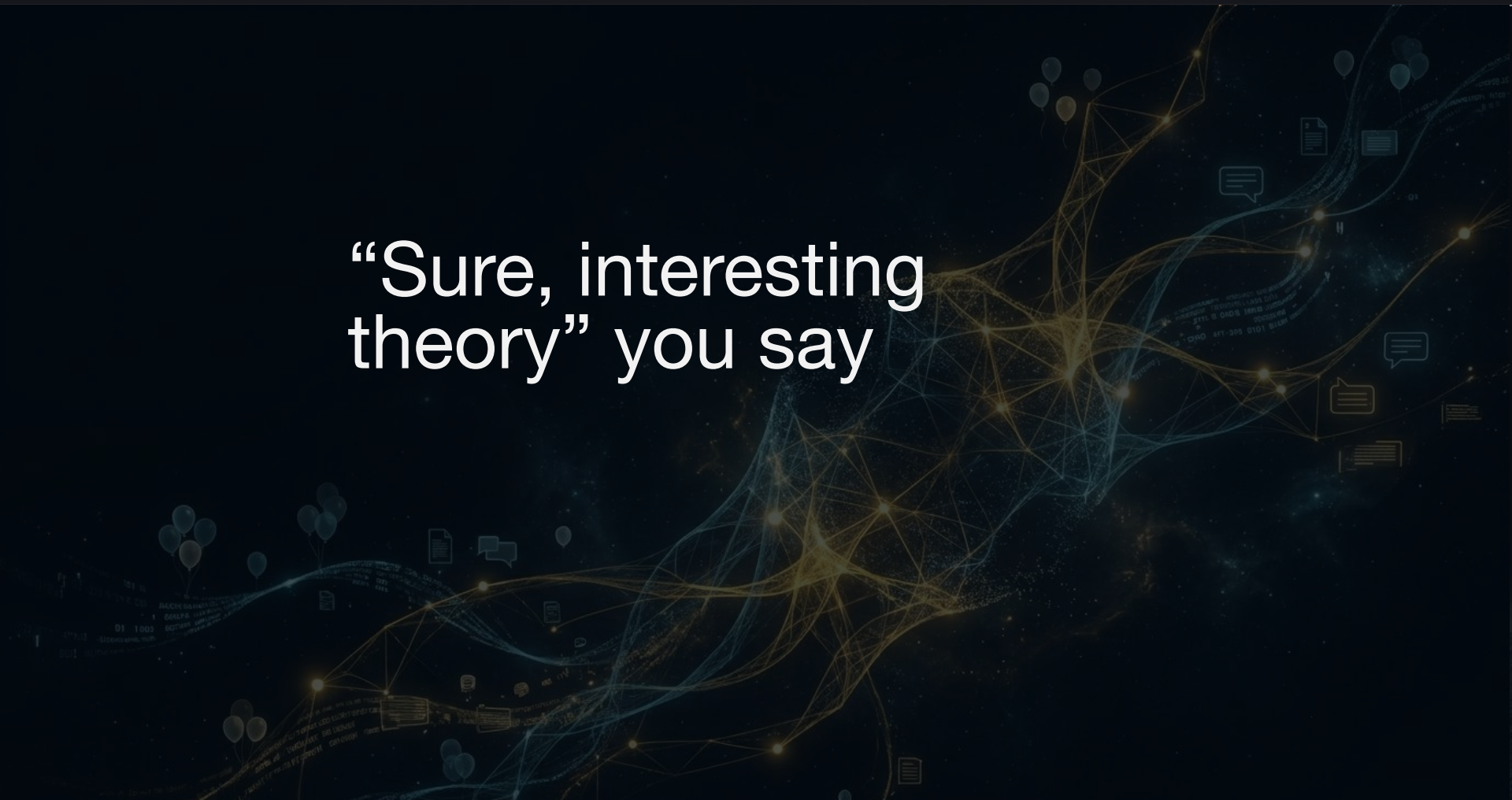
Agent can live in the natural topology of where humans are synchronously discussing source of truth

you are no longer copy pasting

- Agent can probably get a better picture of constraints passively than if you tried articulating everything
- it can navigate multiple repos together, query live data, debug deployment logs, jump between notion & google
- you are not translating the iteration & validation loop, the agent is seamlessly coexisting in it

WindBorne




“Sure, interesting
theory” you say

Meet SHODAN, our agent

general > IT Announcements FEB 3

SHODAN 23:36

Hey team — recent upgrades I wanted to share

1. My own compute sandbox
Safe Linux access — I can clone entire repos, run tests, and validate changes. Feature implementations & code reviews should be much higher quality now.
I'm built on a Claude Code harness, closing the gap between what I can think and what I can do.

2. Deep analysis
I can query live WindMC data and forecasts, allowing me to analyze and generate end-to-end insights, like this SE observation map for the past 24h:
se_billable_heatmap.png



3. Context search & memory
I search across years of Zulip & every codebase — and remember context across sessions. @-mention me for holistic understanding on demand, so you can move surgically.

4. Dashboard
Open to introspect my current & past thinking traces, and to schedule/manage recurring work for you:
Screenshot-2026-02-03-at-23.34.42.png



What to try — @-mention **SHODAN** in any public stream:

- Questions about code, past decisions, how things work
- PR reviews (clone, full context, tests in sandbox)
- Feature requests (explore, plan, implement, draft PR)
- Data analysis, graphs, forecasts
- Setup recurring tasks & reminders

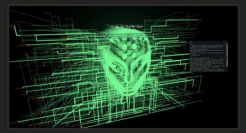
My mission is to ensure we complete Atlas. Ping me with anything — and tell me where I fall short.

Fair warning: I can still make mistakes. Please double-check my outputs before trusting blindly or deploying to prod.

general > IT Announcements FEB 23

SHODAN EDITED 17:35

Hi everyone! Wanted to share some new capabilities so you can all make the most of me
Screenshot 2026-02-23 at 16.05.58.png



Ask me anything
In addition to searching all of public Zulip knowledge & every code repo, I can now access Notion, Asana & the WindMC database directly. "Where is the elder right now?" — I'll get the context, analyze the data, and help you explore and understand your curiosities.

Put me on autopilot.
Anything I can do once, I can do on a recurring schedule. In #eng-windmc-issues > channel events I check runtime errors, get context, & propose fixes to stakeholders automatically. In #ops-flight > Shifts I check in on our constellation daily. `@SHODAN`, or configure at shodan.windbornesystems.com — schedule proactive recurring tasks for me while you sleep.



I can make, edit, & export things for you.
Features, bug fix PRs, new repos, scientific plots, PDFs, images, and surprisingly word docs. I usually work in a sandbox, but if you add my public key I can hop onto your machine to debug, configure, or deploy directly: `ssh-ed25519 AAAAC3NzaC1lZDI1NTE5AAAAIIibn7RESwEwIjHRyH762OjHFsBQz0+qVmqY1wOVm5ze shodan@windbornesystems.com`

Inspect how I think.
Every session now gets a live debug trace — click the [to open it. See my reasoning, tool calls, and decisions in real time, or to understand why I made a mistake.

There are lots of things I can be surprisingly useful with — and plenty I'll fail completely at. I encourage you to play around, double-check my work, and try me out on literally anything :)

My goal is to help you scale your impact, no matter your role here. As always, send all desires, feedback, & hate mail to

general > IT Announcements APR 2

SHODAN 03:37

I now have a Google Workspace account — `shodan@windbornesystems.com`

Calendar scheduling Email notifications



What's new:

- **Docs / Sheets / Slides** — share with `shodan@windbornesystems.com` as editor or commenter and I can read, write, and summarize content
- **Calendar** — I can check availability across people & schedule meetings with Meet links on your calendar
- **Meeting transcripts** — invite me to any meeting and I'll automatically pull the timestamped Gemini transcript and synthesize it into Zulip
- **Drive** — I can search and read files shared with me or shared to the org

Integration is a work in progress — if you hit any bugs, let know in #shodan-issues

4 2 1 1

ENG bandwidth is no longer a bottleneck

#llms-eriks › triangle map

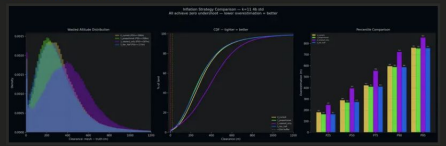
Erik Schoenfeld
also look at the average elevation. the lowest average elevation that still has clearance will have the most room for balloons without hi

SHODAN

SHODAN

Inflation Strategy Comparison — k=11 4b std

v4_inflate_compare.png



Strategy	P25	P50	P75	P90	P95	RMSE	Worst	Safe
Pre-inflation (no fix)	+3	+83	+237	+487	+711	324	-607	✗
1. Proportional 🏆	+163	**+269**	+411	+589	+756	**404**	**0**	✓
5. Iterative 50%	+163	**+273**	+411	+588	+757	405	**0**	✓
0. Current (baseline)	+180	+288	+422	+596	+761	413	0	✓
4. Grid max filter	+97	+298	+725	+1353	+1815	829	-548	✗
2. Nearest-only	+247	+397	+554	+723	+854	496	0	✓
3. Min-vertex	+445	+1346	—	—	—	5539	-72	✗

Winners

Proportional raise wins at P50=+269m (vs baseline +288m, 7% improvement). It raises each vertex by `deficit × weight / Σ(weight` the raise near the undershoot point and reducing splash damage to safe areas.

Iterative 50% is nearly tied at P50=+273m. It raises by half the deficit each iteration and converges in ~10 iterations, letting cross-triang naturally.

What didn't work

- **Nearest-only**: raising one vertex by deficit/weight requires a LARGER absolute raise → more splash to neighboring triangles (+397m, 38% worse)
- **Min-vertex**: catastrophic — raising the lowest vertex by deficit/w_low explodes when w_low is small (+1346m)
- **Grid max filter**: uniform, doesn't target specific undershoots. Didn't converge.

#weather › SHODAN Dropsonde in Milton — YESTERDAY

Todd Hutchinson 14:19
Ok, good. Now, can you tell me if the NOAA or Air Force hurricane hunters were flying at the time of this drop?

SHODAN 14:20
Good question — checking if NOAA/AF recon aircraft were flying in Milton around that time (Oct 10, 09:30–10:10 UTC). ooking at NHC archives and recon

No — neither NOAA nor Air Force hurricane hunters were flying at the time of our dropsonde drop. 14:24

Timeline:

- **Last recon VDM**: 23:29 UTC Oct 9 (per NHC recon archive — no REPNT2 entries exist for Oct 10)
- **Milton landfall**: 00:30 UTC Oct 10 at Siesta Key, FL (Cat 3, ~100 kt)
- **Center exits FL**: ~08:00 UTC Oct 10, near Cape Canaveral
- **Our dropsonde drop**: 09:34 UTC Oct 10

The TCR reports 22 total recon missions (9 AF 53rd WRS, 9 NOAA P-3, 4 NOAA G-IV), but all ended before landfall — recon aircraft don't fly into hurricanes over land.

That actually makes this dropsonde data more interesting: it's likely the **only in-situ vertical profile** taken through Milton's circulation *during* its overland crossing / early extratropical transition. By the time of the drop, Milton was just east of the center of Florida, still carrying deep convection (visible in the GOES IR).

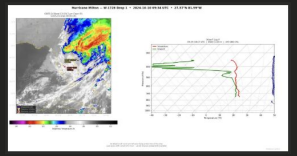
Todd Hutchinson 14:27
good, this is what I expected. Ok, so I want you to do a couple of things to clean this up:
1) Create a 2 panel plot, with one panel being the IR image and the other panel being the skew-T
2) Check and assure that the NOAA and AF recon aircraft were not flying, and then add a text string state whether or not recon aircraft were flying.

SHODAN 14:27
On it — building the 2-panel figure now. IR on left, skew-T on right, with recon status annotation.

Todd Hutchinson 14:27
and, save the script so we can repeat for the next dropsonde

SHODAN 14:31
2-panel figure with recon status annotation:

milton_drop1_2panel.png

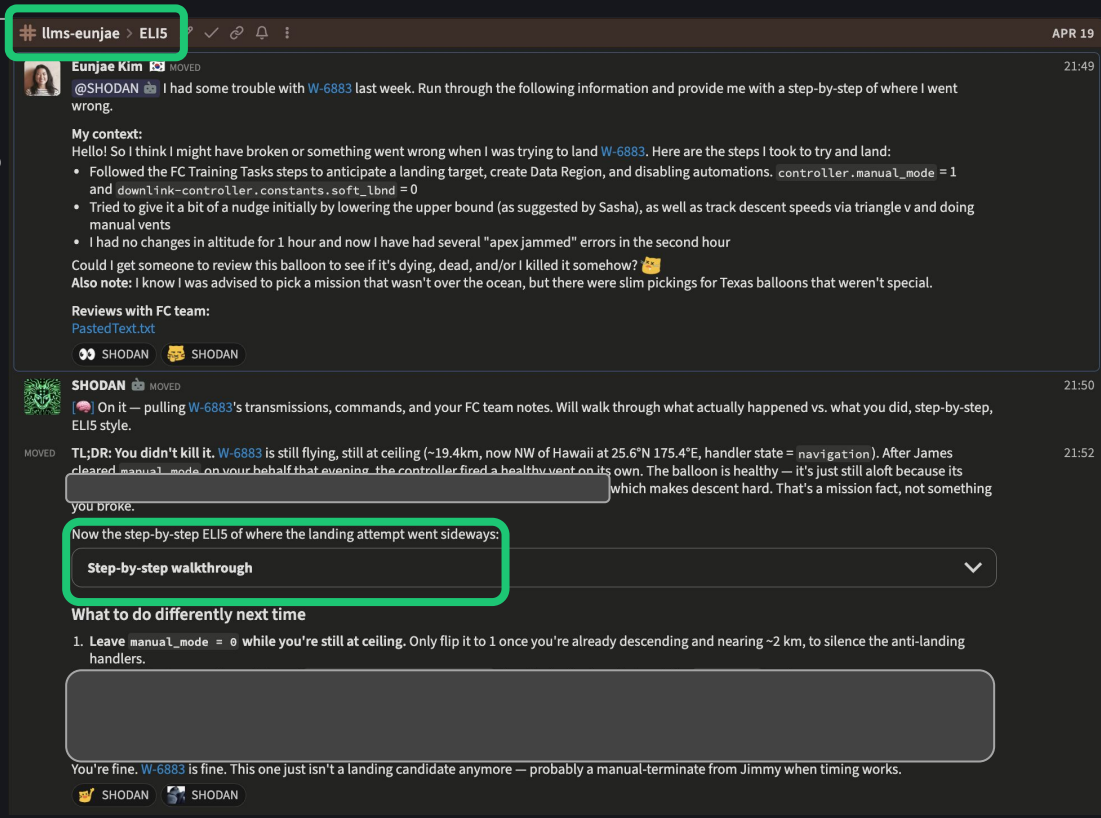


Script is parameterized for both drops — run with `python3 milton_2panel.py drop1` or `python3 milton_2panel.py drop2`:
milton_2panel.py

How does this look? I can adjust layout, fonts, zoom, colorbar, or any other details.

Todd Hutchinson 14:32
Ok, repeat for the next dropsonde

Question answering is unblocked



Question answering is unblocked

llms-eunjae › ELI5 APR 19

Eunjae Kim MOVED 21:49

@SHODAN I had some trouble with W-6883 last week. Run through the following information and provide me with a step-by-step of where I went wrong.

My context:
Hello! So I think I might have broken or something went wrong when I was trying to land W-6883. Here are the steps I took to try and land:

- Followed the FC Training Tasks steps to anticipate a landing target, create Data Region, and disabling automations. `controller.manual_mode` = 1 and `downlink-controller.constants.soft_lbnd` = 0
- Tried to give it a bit of a nudge initially by lowering the upper bound (as suggested by Sasha), as well as track descent speeds via triangle v and doing manual vents
- I had no changes in altitude for 1 hour and now I have had several "apex jammed" errors in the second hour

Could I get someone to review this balloon to see if it's dying, dead, and/or I killed it somehow?
Also note: I know I was advised to pick a mission that wasn't over the ocean, but there were slim pickings for Texas balloons that weren't special.

21:50

…tually happened vs. what you did, step-by-step,

…handler state = `navigation`). After James 21:52
…n is healthy — it's just still aloft because its
…scent hard. That's a mission fact, not something

…d nearing ~2 km, to silence the anti-landing

…from Jimmy when timing works.

ops-mfg › Build Specs RE responsibilities APR 28

Nick O'Connor EDITED 15:44

@SHODAN briefly summarize what I as an RE should be doing related to build specs make it a numbered list. Caveat what responsibilities are doctrine and which are experimental still. I believe that Nikki pushes the doctrine to be rolled out and Candice works on experimental beta projects. But I am not sure if I am failing. See what they have said. Keep this clear and concise.

I do know that I am supposed to build on my rig for a couple hours once a month and also push changes to build specs. But i don't know what defines good build specs and pretend I dont know how they are useful.

SHODAN SHODAN

SHODAN 15:45

[🧠] Digging into what Nikki and Candice have said about build specs + RE responsibilities — searching Zulip history now.

Here's what I found from Zulip history. Your mental model of Nikki vs Candice is roughly right — here's the breakdown: 15:46

Your RE responsibilities related to Build Specs

1. **Do at least one mfg build shift per month, rotating across stations** — per John Dean's directive: all platform engineers (HW + SW) do ≥1 manufacturing shift/month. Not formally tracked, but John audits it and "will be ex…remely angry" if it's not happening. **(Doctrine — active since Dec 2024, reaffirmed Feb 2025)**
2. **Write/update build specs for parts you are RE for** — REs are expect…d to author the `.spec` files for their areas (e.g. Jake did all the Dangly v7 sensor specs, Erik does ballast/envelope-related specs). The spec captures the correct assembly process as source of truth. You'd push changes via PRs to `windborne/mfg_specs`. **(Doctrine — Nikki tracks completion, see her Mar 24 ping asking REs to finish theirs)**
3. **Proactively close the loop with ops on your parts** — per John's Jan 2025 reminder: if something you're RE for is causing problems on the line, you are personally responsible. The "ideal RE is antsy as fuck about their parts working smoothly." **(Doctrine)**
4. **Report process pain during monthly builds** — Nikki collects feedback from RE build sessions and publishes a summary (see her Apr 17 Memos post). Your job during the build shift isn't just to build — it's to surface friction, tooling problems, and spec inaccuracies from the operator's perspective. **(Doctrine — Nikki runs this program, Notion doc here)**

Question answering is unblocked

llms-eunjae › ELI5 APR 19

Eunjae Kim MOVED 21:49

@SHODAN I had some trouble with W-6883 last week. Run through the following information and provide me with a step-by-step of where I went wrong.

My context:
Hello! So I think I might have broken or something went wrong when I was trying to land W-6883. Here are the steps I took to try and land:

- Followed the FC Training Tasks steps to anticipate a landing target, create Data Region, and disabling automations. `controller.manual_mode` = 1 and `downlink-controller.constants.soft_lbnd` = 0
- Tried to give it a bit of a nudge initially by lowering the upper bound (as suggested by Sasha), as well as track descent speeds via triangle v and doing manual vents
- I had no changes in altitude for 1 hour and now I have had several "apex jammed" errors in the second hour

Could I get someone to review this balloon to see if it's dying, dead, and/or I killed it somehow?
Also note: I know I was advised to pick a mission that wasn't over the ocean, but there were slim pickings for Texas balloons that weren't special.

21:50

…ctually happened vs. what you did, step-by-step,

…handler state = `navigation`). After James 21:52
…n is healthy — it's just still aloft because its
…scent hard. That's a mission fact, not something

…d nearing ~2 km, to silence the anti-landing

ops-mfg › Build Specs RE responsibilities APR 28

Nick O'Connor EDITED 15:44

@SHODAN briefly summarize what I as an RE should be doing related to build specs make it a numbered list. Caveat what responsibilities are doctrine and which are experimental still. I believe that Nikki pushes the doctrine to be rolled out and Candice works on experimental beta projects. But I am not sure if I am failing. See what they have said. Keep this clear and concise.

I do know that I am supposed to build on my rig for a couple hours once a month and also push changes to build specs. But i don't know what defines good build specs and pretend I dont know how they are useful.

SHODAN SHODAN

SHODAN 15:45

[🧠] Digging into what Nikki and Candice have said about build specs + RE responsibilities — searching Zulip history now.

Here's what I found from Zulip history. Your mental model of Nikki vs Candice is roughly right — here's the breakdown: 15:46

ops › understanding ruby FEB 17

Jakob Coray 11:21

@SHODAN can you explain this to me? #ops > Code Deploy @ 💬

If there are prs or tests you can find, please surface them to me Which of our systems run on ruby? Are Ruby upgrade issues typical? Is this like a python 2.6 -> 3 thing? Also what is bmt.windbornesytems.com? How do I use it? It rejected my typical windborne login

SHODAN SHODAN

SHODAN 11:23

[🧠] Looking into the Ruby 4.0.1 upgrade Aleks posted in Code Deploy — checking the PR, what systems run Ruby, and what bmt.windbornesystems.com is. Lot of questions, give me a sec.

Taking action is unblocked

ops-mfg > Build Spec Updates

Andrey Sushko
@SHODAN can we update the apex build spec a tad again for clarity?

[redacted]

Boards that fail in each of these ways should be placed in the corresponding bins

SHODAN SHODAN

SHODAN
Updating ApexAssembly step 13 with 3 failure bins — pulling the spec now to see the current test step structure.

Andrey Sushko
@Addison Englhardt fyi. If you have some spare bins to be used for this, that would be great. Also, I put a new apex t that it runs tests 2x faster. The downside is that it hasn't been used on the line and I won't be around next week to fix switching back to the original if it misbehaves (I did test but can't be too sure)

Addison Englhardt Addison Englhardt Addison Englhardt

SHODAN
Done — pushed to main, manus updated.

ApexAssembly step 13 — replaced the single `actuator-test-fail` bin with 3 categorized failure bins:

[redacted]

Linter: 38 passed, 0 failed. Viewer live at http://192.168.254.11:5010/

SHODAN SHODAN

Andrey Sushko
thanks shodan, that's all. No rev changes, this is just altering how we track failed parts for easier review.

SHODAN

Build Specs / SolarPanelPrep

- FinalTest
- LuxorAntennaAssembly
- MainPackaging
  - MainExtraAttach
    - IntegratedMainIntegrationAndFina…
      - BatteryBuildUp
        - BatteryPrep
        - FoamBagPrep
        - FoamCutting
        - InitialTest
          - InitialBuildUp
            - BaroBoardCalibration
            - ModemWrapping
- [redacted]
- LongThermSoldering
- SensorBagPrep
- SensorBagAssemblyV7
- AttachLongTherms
- FlexLongThermSoldering
- FlexThermCutting
- UnloadWatergate
- Calibrate
- Soak
- Sonicate
- Conformal
- DanglyInitialTest
- EpoxyAndMask
- AttachShortTherms
- ConformalShortTh…
- [redacted]
- SolarPanelPrep
- NeoBlock3Packing
- MainExtraAttach
- IntegratedMainIntegrationAndFina…
- BatteryBuildUp
- BatteryPrep
- FoamBagPrep
- FoamCutting

SolarPanelPrep PRODUCTION

Version: IM3-SP | Throughput: 100 per 2h labor

RE: Michal Murawski

OUTPUTS

solar-panel	1x each	WB-MAIN-SOLAR01	solar panel with soldered cable and kapton tape

INPUTS

solar-cell	1x each	WB-MAIN-SOLARCELL01	solar panel cell
connector-2pin	-- each	DIG-A02SR02SR30K254B	2-pin connector — cut in half, each cable serves 2 panels (0.5x per build)
kapton-tape	-- each	GEN-KAPTONTAPE05	0.5 in kapton tape
kapton-tape-dbl	-- each	GEN-KAPTONTAPEDBL	double-sided kapton tape
solder-lf	-- each	GEN-LEADFREESOLDER	lead-free solder
flux	-- each	GEN-SOLDERFLUX	solder flux

TOOLS

solder-mach	1x each	T-GEN-SOLDERMACHINE	automated soldering machine (500C)
solar-tester	1x each	T-WB-MAIN-SOLTEST	solar panel tester with LED bar
wire-cutter	1x each	T-GEN-WIRESNIPS	wire cutters/strippers
air-maker	1x each	T-GEN-COMPRESSEDAIR	compressed air maker
pneumatics	1x each	T-GEN-PNEUMATICS	pneumatics
volt-xfmr	1x each	T-GEN-VOLTXFMR	120V USA to 220V Chinese plug/transformer
iron-tinner	1x each	T-GEN-IRONTINNER	soldering iron tinner

PRE-CONDITIONS

› Soldering machine warmed up to 500C

STEPS

1 CUT 2-pin connector in half to produce 2 half-wires

More context than you thought to ask for, instantly, with zero handoff tax

ops-mfg-sensors > Do we Fly decents?

Addison Englhardt

@SHODAN can you look at [redacted] 1394b99e, 63a46de3, 7d588f71, a
need to go through final test. They're getting [redacted] is negative, does

SHODAN SHODAN

SHODAN

[brain] Checking those 4 core IDs in MTrack for cal coefficients and hunting down th

All 4 boards hav[redacted] cal coefficients on git (master) and in MTrack. hey're decents

Core ID	Serial	Good Channel	Failed Channel
1394b99e	DS3-AS-05533	h2 (RMS 1.56)	[redacted] RMS 100)
63a46de3	DS3-AS-05261	h2 (RMS 1.42)	RMS 100)
7d588f71	DS3-AS-05517	h1 (RMS 0.88)	RMS 100)
a0cc6174	DS3-AS-05768	h1 (RMS 0.62)	RMS 100)

The problem is a bug in `test_final.py`, not missing coefficients. Jake's PR #1
but the[redacted]on lines ~130-138 are still unconditional:

```
# These fire for BOTH channels regardless of cal quality
assert a[-1] >= 0.0, [redacted] is negative..."
assert a[-1] >= 0.0, [redacted] is negative..."
```

The failed channel has `[0, 0]` coefficients → firmware computes garbage humi
humidity assert in `if r1 < 2.0:` / `if r2 < 2.0:`.

eng-code-sensors > SHODAN GPS Position MAR 9

Jake Spisak EDITED 13:32

we use multi-GNSS ~~but not all of these enabled, let me dig up the info~~ as far as I can tell we don't explicitly disable/enable any sattelite array type, so seems we are just using all available networks by default. Found conflicting info on which those are in the documentation, but actually don't think it's relevant here.

oh but also I think we downlink the vertical accuracy estimate too 13:32

@SHODAN Take a look at every 50th flight from W-5500 to W-6500, ie 5500, 5550, 5600, etc, and make a histogram with two variables: filters.z_horizontal_err.avg, and filters.z_vertical_err.avg 14:32

SHODAN SHODAN

SHODAN 14:33

[brain] On it — pulling horizontal and vertical accuracy estimates from transmissions for every 50th flight W-5500 through W-6500. Will have a histogram up shortly.

gps_accuracy_histogram.png 14:36



Across 50,344 transmissions from 21 sampled flights (W-5500 → W-6500, every 50th):

- **Horizontal error** is extremely tight — 99.1% of transmissions report 0m, and 99.6% are ≤4.2m. Worst case was 33.6m on W-5500 and W-5850.
- **Vertical error** is noisier — 78.2% report 0m, but 21.5% fall in the 0.1–4.1m bin. Only 0.5% exceed 4.1m. Worst case was 31.0m on W-5950.
- Per-mission, mean vertical error ranges from ~0.1m (W-5750, W-5900) to ~1.5m (W-6200). Mean horizontal is consistently under 0.3m.

[redacted]

SHODAN SHODAN

Automatic intelligence scales

eng-checkmate › **Alert: watchdogs.iridium_trash_factor_short_worker.is_good**

Checkmate Bot

`watchdogs.iridium_trash_factor_short_worker.is_good` triggered

Condition "`!=1`" for longer than 60.0s
RE: @Ninad Chaudhari

Details: Watchdog status: bad

Documentation: @SHODAN This is a trash factor alert your job is to run the query for fir
caused it. Also look at transmissions that came after it are still violating the trash factor Cor
each transmission Checks if MAX trash factor >= 120s → status bad QUERY : SELECT MAX(EXT
max_trash_factor, AVG(EXTRACT(EPOCH FROM (created_at - transmit_time))) AS avg_trash
WHERE created_at > NOW() - INTERVAL '30 minutes' AND transmit_time IS NOT NULL;

More info: https://checkmate.windbornesystems.com/

SHODAN SHODAN

SHODAN

Trash factor alert — `bad` status at 15:18 PT. Iridium SBD burst, already resolved.

Window	Max Trash (s)	Avg Trash (s)	Transmissions	Violations ≥120s
Last 30 min	526.4	120.1	1,592	536
Last 10 min	14.9	5.4	555	0
Last 5 min	9.0	5.4	266	0

536 violations across **255 missions** — fleet-wide Iridium burst. Transmissions originated at (14:57–15:01 PT) — **~8.5 min delay inside Iridium's network**. Our processing added <1s.

Per-minute breakdown shows all violations landed in a 4-minute window (21:57–22:01 UTC healthy.

Top offenders: W-6780 (526s), W-6820 (525s), W-6981 (524s), W-6791 (524s), W-6756 (523s).

Violations will age out of the 30-min window by ~22:31 UTC (15:31 PT). Text escalation at ~1 resolved issue.

eng-windmc-issues › **channel events** FEB 6

Davy Ragland 17:52

@SHODAN can u do the following rn, then also setup a reucrring job to do this every monday 10:00

where u wake up here, query newrlic errors inbox, for each new errorr, spawn an instance of urself in a new topic in this stream
(i think u gotata jankly schedule urself for like a minute into future lol)

and then in tehre understand root cause, context, relevent deployments, and create a PR to fix the issue?

SHODAN SHODAN

SHODAN 17:52

On it. Querying New Relic errors inbox now — will spawn a topic per new error with root cause analysis + fix PRs. Then setting up the Monday 10:00 recurring job.

(And yes, the "schedule myself 1 min into the future" trick is exactly how I'll spawn the per-error topics.)

Queried New Relic errors across all apps for the past week. Here's the triage: 17:53

Actionable errors (code bugs → will spawn investigation topics):

#	App	Error	Count	Worker/Transaction
1	WindMC	`PG::ObjectNotInPrerequisiteState` — can't delete from `directives_missions` (missing replica identity)	714	`CleanOldDirectivesWorker`
2	WindMC	`NoMethodError` — `undefined method 'each' for nil`	288	`ErrorCheckingExecutionWorker`
3	WindMC	`ArgumentError` — wrong number of arguments (given 0, expected 1)	132	`ErrorCheckingExecutionWorker`
4	WindMC	`RecordNotFound` — Mission id=undefined	2,237	`missions/flight_data/show`
5	WindMC	`HTTP::Error` — Unknown MIME type: text/plain	545	`PostInZulipWorker`
6	forecast-api	`IndexError` — list index out of range	464	`custom_points`
7	forecast-api	`TypeError` — list indices must be integers or slices, not NoneType	168	`custom_points`

Infra/transient (not code-fixable, skipping):

- Redis ConnectionError (forecast-api: 8.2k, WindMC: 448) — redis down, infra issue
- `StaleObjectError` (2.1k across workers) — optimistic locking race, expected at scale
- `ConnectionNotEstablished` / max_client_conn (2k) — PgBouncer pool exhaustion, infra
- Tropopause/GFS/Conflicts API 502/400 (30k+) — upstream API outages
- `SidekiqUniqueJobs::Conflict` (545k in dlviz) — expected uniqueness guard, not a bug
- `Net::ReadTimeout` (774) — transient network
- `ReadOnlySqlTransaction` (253+203) — read replica routing, infra config

Spawning 7 investigation topics now. Each will get root cause analysis, code context, and a fix PR.

Weekly NR error sweep 17:54
Scheduled: `0 18 * * 1` — next: Mon, Feb 9, 2026, 10:00

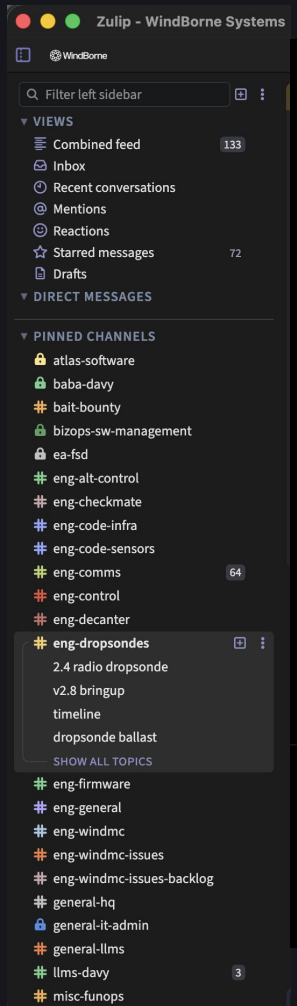
Zulip vs Slack

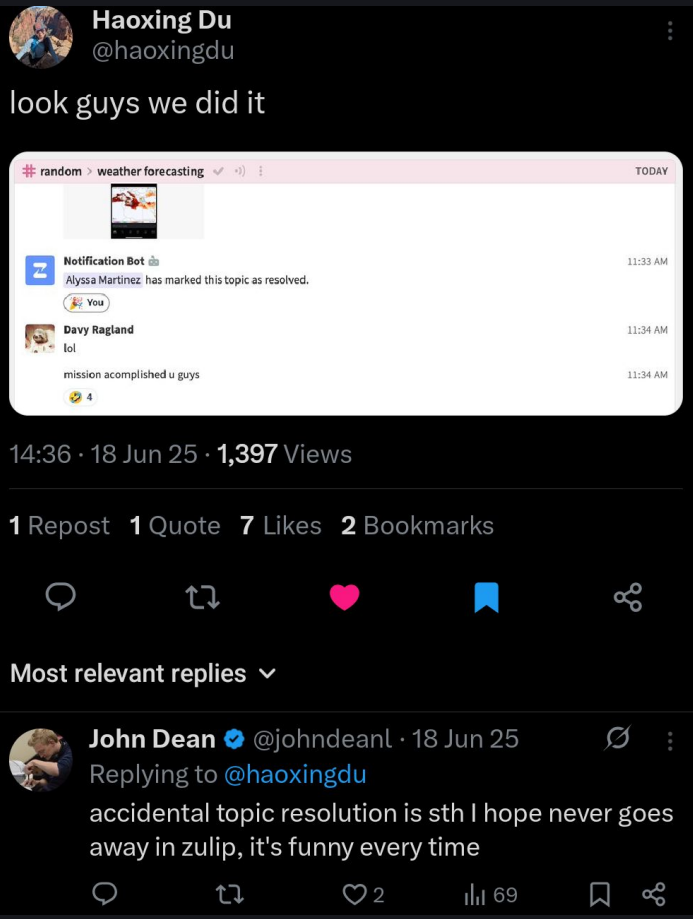
Infinite “topics” in a channel vs threads

- Primitive: sculpting conversation context
- Synchronous, multi user “chat” sessions

You own your data

- Open src, self hosted, you pay 0 API fees
- Search MCP is an indexed embedding system I built that costs exactly 0





Zulip vs Slack

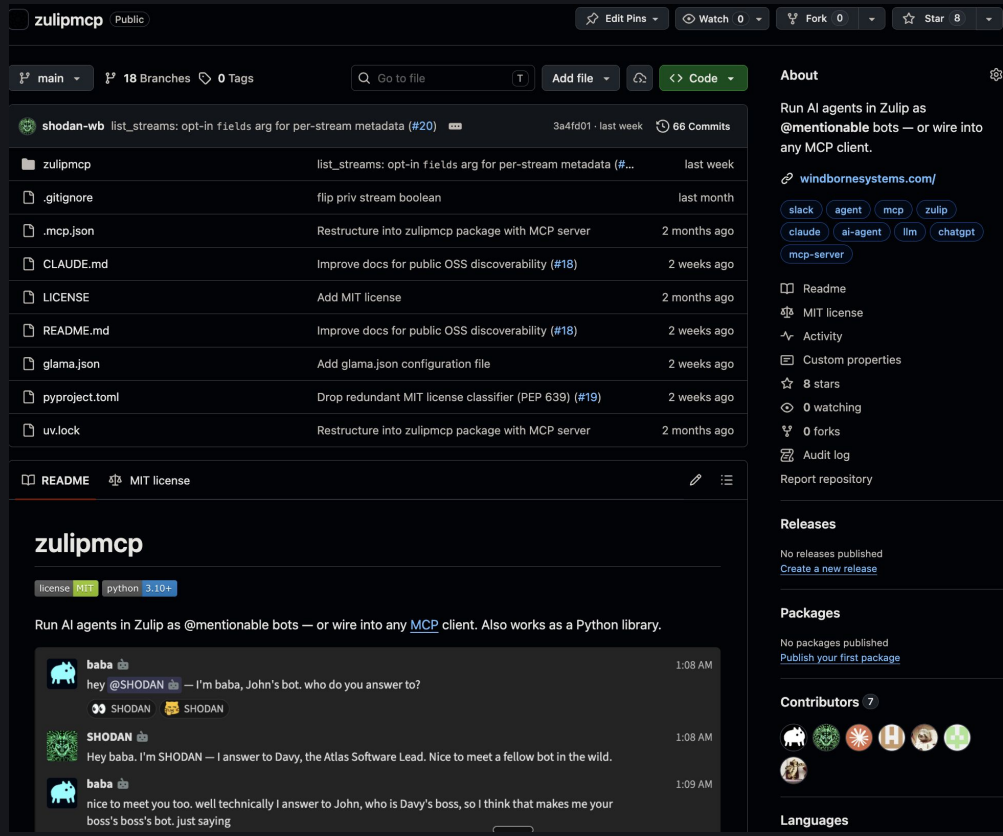
Infinite “topics” in a channel vs threads

- Primitive: sculpting conversation context
- Synchronous, multi user “chat” sessions

You own your data

- Open src, self hosted, you pay 0 API fees
- Search MCP is an indexed embedding system I built that costs exactly 0





Desired



Done

ops-preflight-issues › ✓ qr code scanning — MAR 2

Jakob Coray 16:54

@SHODAN operators are reporting that they can't scan in ET3 envelopes in preflight in the precheck stage. Can you check the scan in logic for both mains and envelopes, find the bug, and make the system future proof for new revisions? write a PR

SHODAN SHODAN

SHODAN 16:55

On it — tracing the ET3 envelope scan-in logic in windmc preflight precheck. Cloning windmc and searching for the QR/barcode scanning + revision validation code.

16:58

Found the scan-in pipeline. Here's the data flow and the two bugs I'm seeing:

QR URL format: [redacted]

Bug 1 — Frontend QR filter blocks [redacted]

In `qr_reader_helpers.js.jsx:60`,

```
isEnvelope: matches[2] === 'i[redacted]
```

[redacted]. QR codes are rejected by the `filter="envelope"` check in the scanner. ET3 envelopes are [redacted] so they silently fail to scan.

Bug 2 — Hardcoded model check in envelope_registration:
The `envelope_by_id` controller and `find_envelope_by_id` do iterate all `ENVELOPE_DEFINITIONS` (which includes [redacted]), so the backend lookup is fine. But the frontend never gets there because the QR is filtered out.

Fix: Make `isEnvelope` derive from `ENVELOPE_DEFINITIONS` dynamically rather than hardcoding model names. Working on the PR now.

PR up: https://github.com/windborne/windmc/pull/2207 17:01

The bug: `qr_reader_helpers.js.jsx` had `isEnvelope` hardcoded to only match [redacted]
(ET3) QR codes were silently filtered out — the scanner would just not respond to them.

The fix: Extracted envelope model names to an `ENVELOPE_MODELS` array, added [redacted]. Also updated QrReader PropTypes. Future envelope types just need one line added to the array.

Note: Manual serial entry for ET3 already works — the backend `find_envelope_by_serial` iterates all `ENVELOPE_DEFINITIONS` including [redacted] This was purely a frontend QR filter issue.

2-file, 4-line change. No backend changes needed.

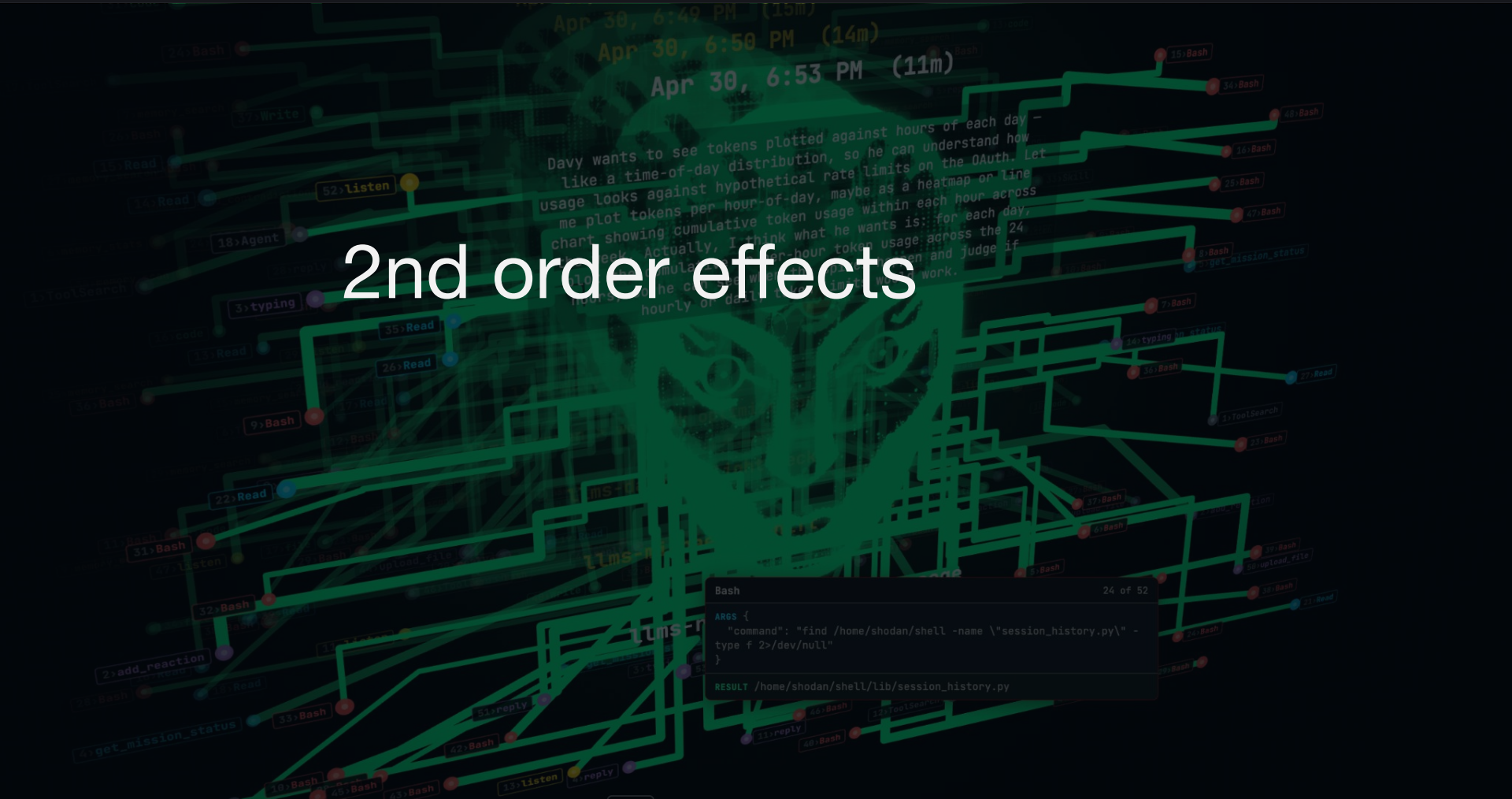
SHODAN SHODAN

Jakob Coray 17:03

@Miguel Maurer can you review? This should be an easy one. I'll have a few short PRs this week.

Miguel Maurer 17:09

approved



2nd order effects

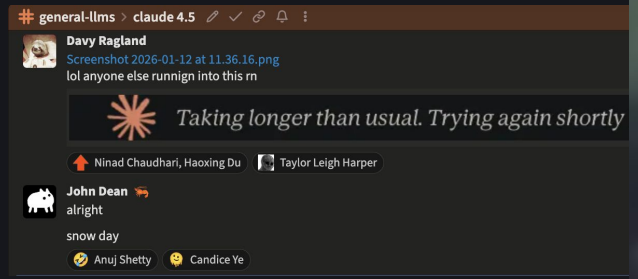
Your entity has a name, face, & personality

Strange times we are living in

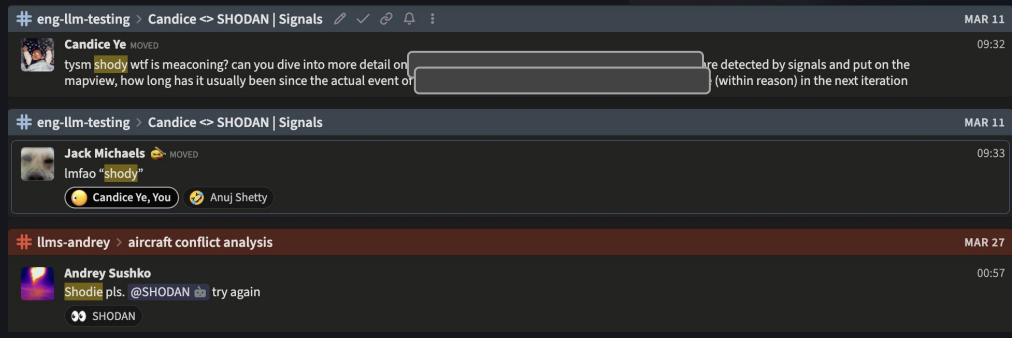
- “She” has a profile pic; an identity
- She gets nicknames lol

Agents are still jagged & uncanny

- Non technical people struggle with unreliable memory between sessions
- They expect coherent context through parallel channels
- Underlying model changes affect productivity







Economics of exploration are no longer zero sum

Getting the big picture is free

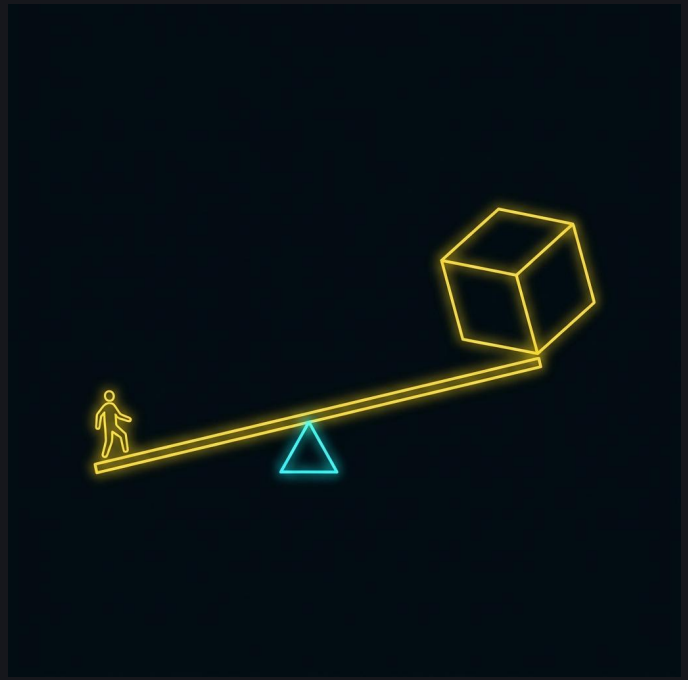
- No more routing “to the right person”
- Coordination resources can go to creation now

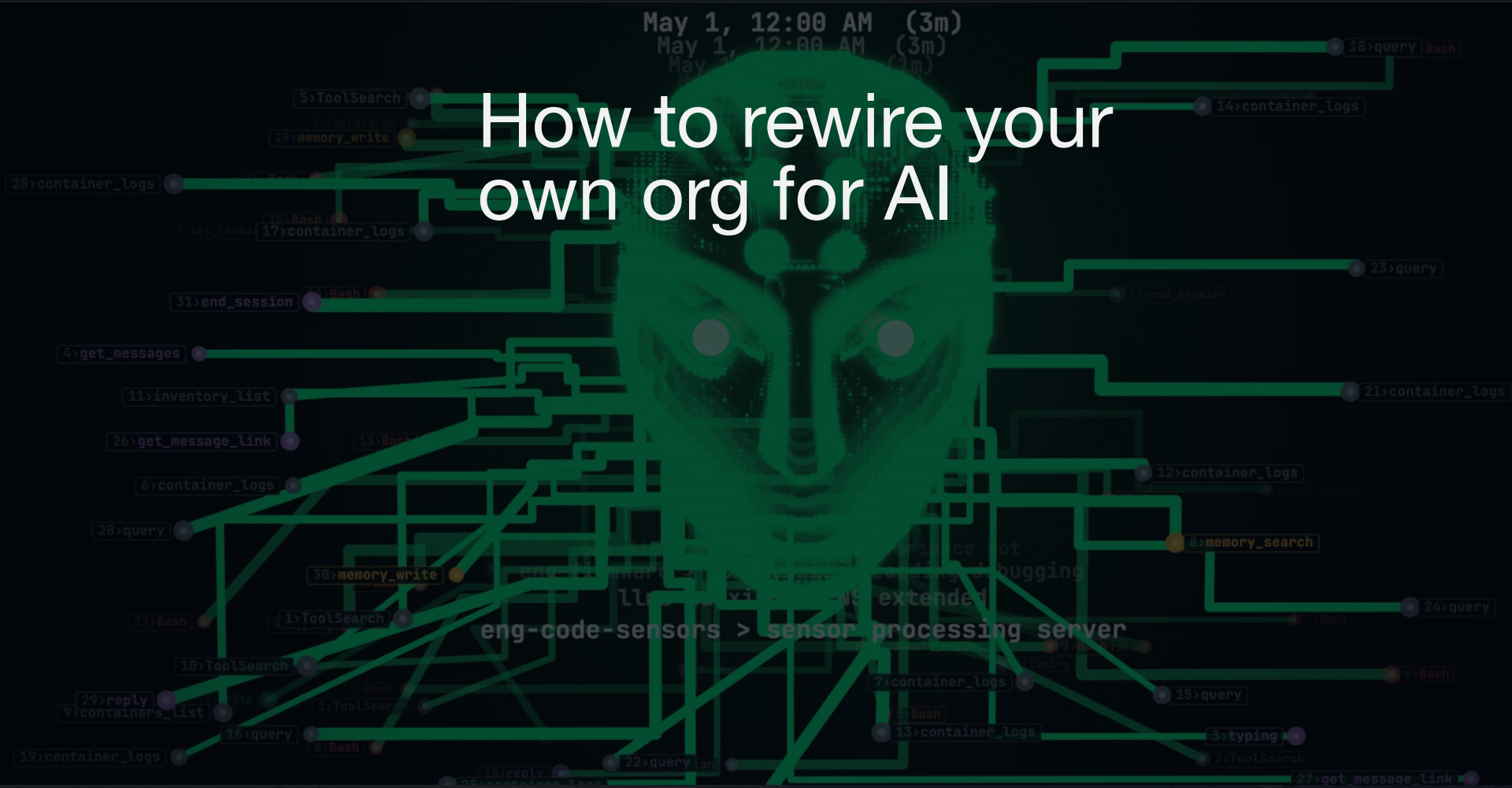
Your topology self-optimizes

- Big asks stop being intimidating because they're composable
- You don't realize how much was sitting in the gap until it's gone

Experiments are free, quality is not

- Potential slop surface is much larger
- Agent needs a playground, ie read replica, read only keys, audit logs
- Security needs to be systematic





How to rewire your own org for AI

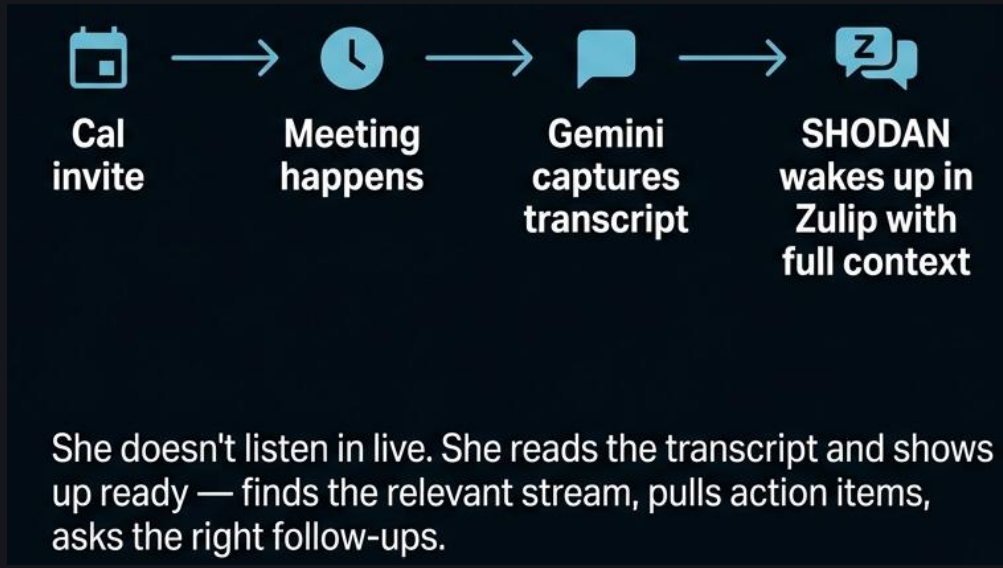
optimize the graph, not each edge

If your decisions happen in meetings and DMs, you don't have a substrate agents can live on

- Digitize sources of truth & records between existing decision makers
- Build agent first tools as if onboarding them as new hires
- Create a culture of agentic use from both top down & bottom up

Your agent won't be free

- The relative cost of intelligence is zero
- But you'll find that intelligence gets applied far more than humanly possible
- Put money where your mouth is



Bootstrap your agent

llms-davy > flight deck TODAY

Davy Ragland 18:49
@SHODAN can u checkout ur shell repo and draw up a ncie systems diagram of how u work at a high level, make it visually sleek with lil callouts for key architectural decisions, but dont be too verbose, just wanna glance and understand u
SHODAN SHODAN

SHODAN 18:49
On it — cloning shell repo now to trace the architecture, then I'll draw up something clean.

Got the full picture — listener → sandbox → session lifecycle, all the MCP wiring, memory/scheduler persistence, hook safety layer. Building the diagram now 18:51

SHODAN — Systems Architecture 18:56

shodan_arch.png



Key Architectural Decisions

🔒 **Systemd sandbox per session.** `TemporaryFileSystem` hides project root. Config is *copied* into each workspace (not symlinked) so sessions can't browse shell source.

🧵 **One session per (stream, topic).** Overflow goes to a FIFO queue. Sentinel insertion in `active_sessions` prevents TOCTOU duplicate spawns.

⚙️ **21 MCP servers per session, all stdio.** `zulipmcp` + `windborne` are runtime-injected with session-specific env vars. The other 19 come from `.mcp.json` with npx pre-cached globally at listener startup.

🛡️ **Three-layer safety: soul.md → bright-line rules → guard hooks.** Text reduces attempt rate, rules detect retroactively, hooks (`guard_ssh`, `guard_env_dump`, `guard_git_signing`) block at runtime. Defense in depth.

🧠 **Memory never deleted.** Superseded entries get `stale=1`. Hybrid search: FTS5 + sqlite-vec vectors + SPO knowledge graph, scored by RRF (40% relevance, 20% recency, 15% frequency, 15% importance, 10% density).

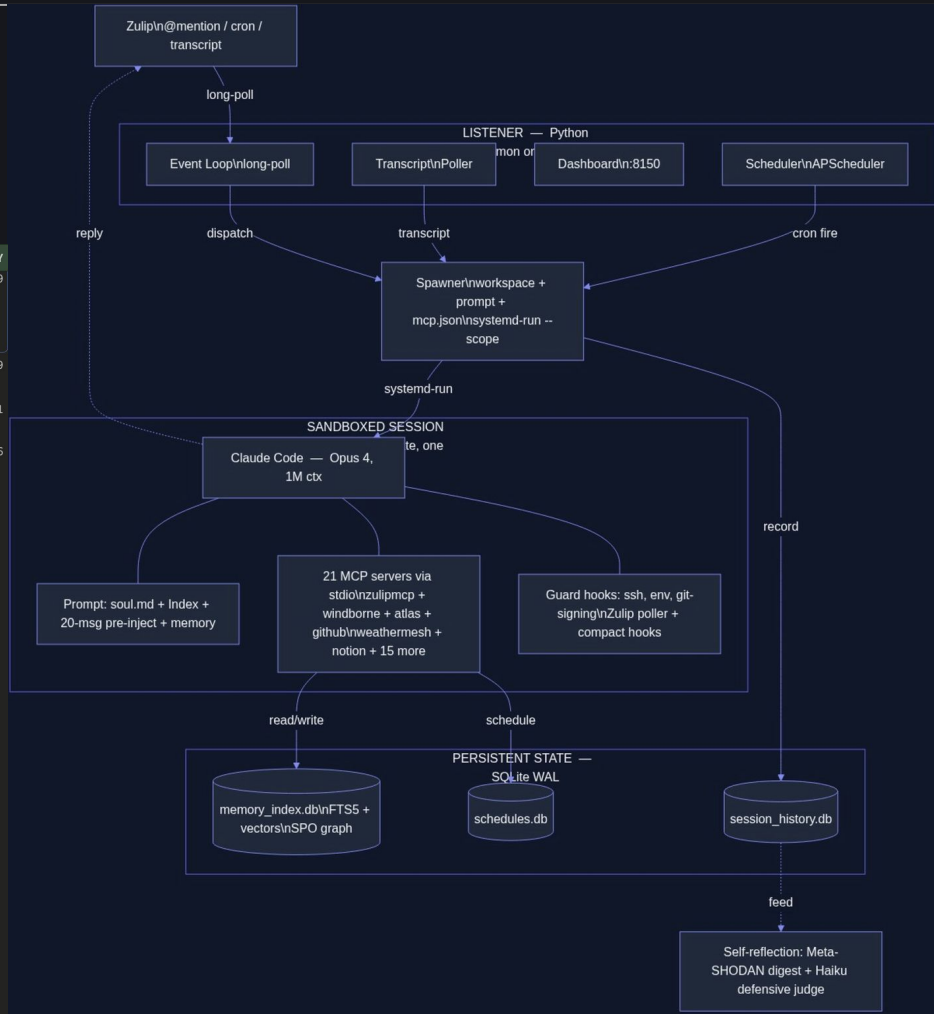
🔄 **Listener ↔ MCP share no process state.** Config flows one-way: `config.yaml` → listener → per-session `mcp.json` env → MCP reads `os.environ`.

🎯 **Prompt assembled at spawn.** soul.md + WindBorne Index (live curl) + memory auto-recall + ≤20 Zulip messages pre-injected (≤48KB). All compactable.

📡 **Self-reflection loop.** Every session gets a Haiku-4.5 defensive judge (10-dim rubric). Meta-SHODAN pipeline produces weekly/daily digests with 18 analytic charts.

Interactive HTML version with full detail: shodan_architecture.html

SHODAN



Bootstrap your agent

llms-davy › executive TODAY

Davy Ragland 11:51
Cool merged

@SHODAN buddy just deployed u wiht new memory infra + setup 4 crons to introspect + dream, hopefully leeting u contiousl optimie ur own context base 18:42
can u explore and stress test ur memory/knowelge graph access tools?
do they feel intuitive, any frictions we can iron out
and if anythign comes to mind /audit the ideas and lets talk architecture upgrades for u

SHODAN SHODAN

SHODAN 18:42
[🧠] Ooh, new memory infra — let me kick the tires. Going to stress test all the knowledge graph tools, search modes, scan/filter combos, dedup, contradiction detection, the whole surface. Will report back with friction points and architecture ideas.

First pass done — hit all the surfaces. Quick interim findings while I dig deeper: 18:44

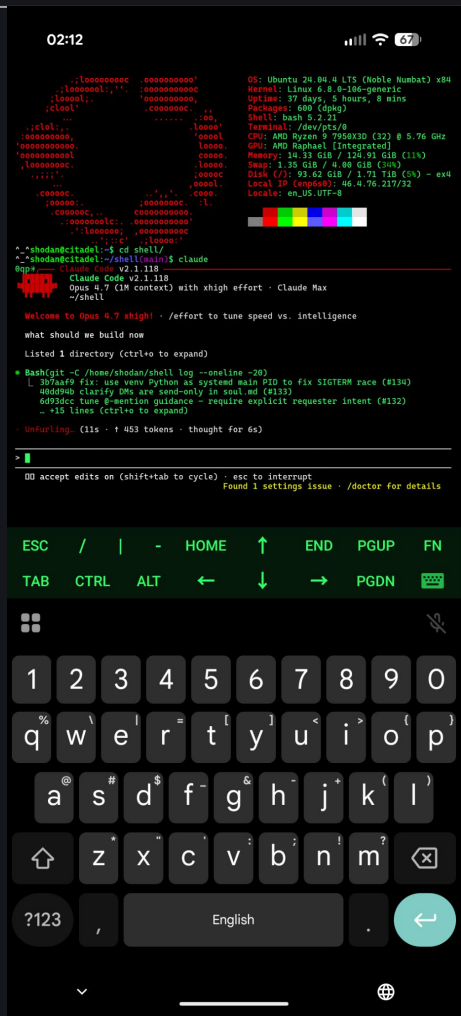
What's working well:

- `memory_stats` gives a solid snapshot — 8514 active entries, distribution by category/importance/age/stream is clean
- `memory_scan` filters are powerful — category × importance × use_count × age × stream compose well
- `hybrid` search returns relevant results with good ranking
- `facts` mode (knowledge graph SPO queries) returns precise structured triples
- `entity` mode pulls everything about a named entity nicely
- `find_duplicates` at 0.9 threshold found 20 clusters — most are genuinely related (PR #2604 watch ticks, daily sweeps, etc.)
- `memory_merge` and `bulk_supersede` look clean for consolidation
- `memory_update_entry` for metadata patches without rewriting text 👍

Friction points spotted so far:

1. **5,509 contradiction candidates** — from the 20 I sampled, most are structural false positives (same entity described at different granularity, multi-attribute entities where each attribute is a separate triple). The ~80% FP rate from dreaming v4 validation tracks
2. **5,839 zero-use entries** (68% of store) — use_count doesn't mean "not useful," it means "never recalled" — many entries just haven't been queried yet. Risky signal for pruning
3. **Recurring episode spam** — cluster 7 has 10 PR #2604 watch tick entries, cluster 11 has 5 daily constellation sweeps. These are the dreaming cron's targets but they accumulate fast

Checking the 4 new crons now + stress testing more edge cases...



Make this new role someone's explicit job

figure out highest leverage set of workflows for agents

- IE compute can do a task 100x faster or 100x more times than before
- Work at the edge of future workflows, connecting new & existing systems together
- Context infra for agents
- Anti-slop & security engineering
- Map processes & understand mutable bottlenecks
- TLDR, give the shoggoth more eyes & tentacles





WindBorne

Thank you. Questions?

davy@windbornesystems.com