

Architecture and data flow ledger

A plain reading of how Sloane turns wearable and workspace signals into one useful prompt, and what the system deliberately refuses to keep.

01 Stateless ingestion

02 Computation core

03 Latent gatekeeper

04 Domain clustering

Four stages, one decision

Sloane reads two things: how your body is holding up, and how your working day is moving. Each is measured separately, on its own scale, and neither leaves a trail of raw content behind it.

The reasoning layer stays asleep. It only wakes when both readings agree that something has shifted, and even then it can suggest, never act on your behalf.

01

Stateless ingestion

Vendor data arrives on two lanes and is sealed under an anonymous key.

02

Computation core

A personal baseline and a work-pace reading, calculated apart.

03

Latent gatekeeper

The model stays dormant until both scores clear the line.

04

Domain clustering

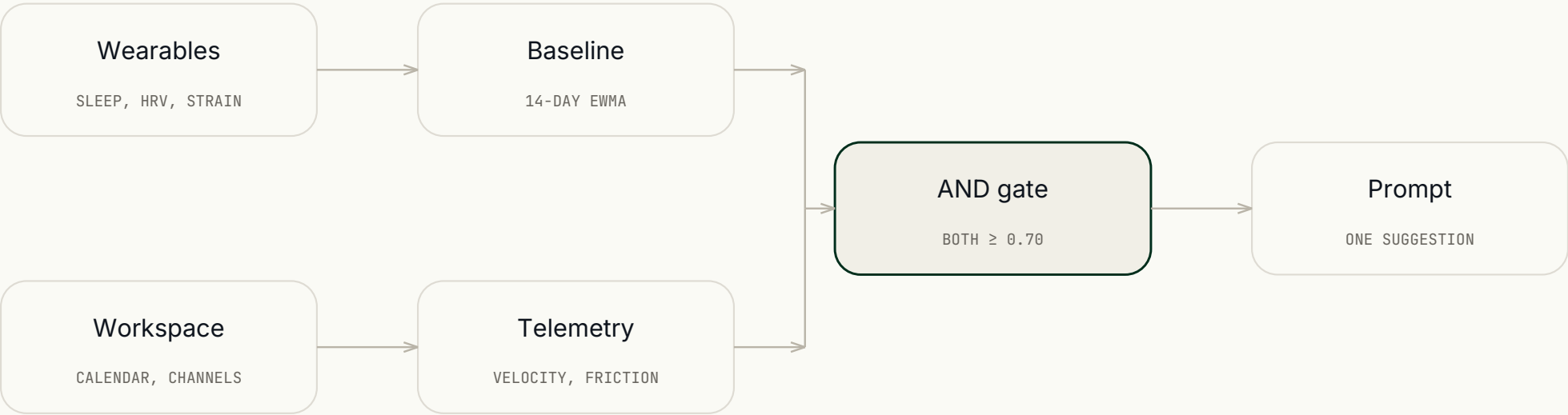
Individual use rolls up into anonymous team capacity.

MASTER STAGE MAP

How a signal moves

Two independent lanes converge on a single gate. If the gate stays shut nothing is generated and no tokens are spent.

TIER ONE / BIOMETRIC

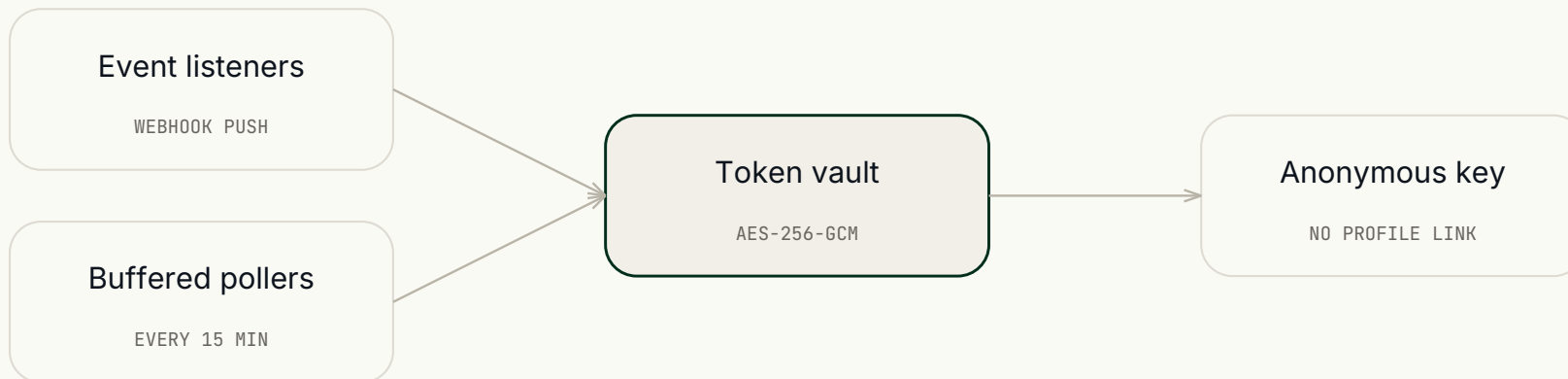


TIER TWO / EPHEMERAL TELEMETRY

01 / STATELESS INGESTION

Data arrives sealed and unnamed

Vendor connections run on two lanes. Some providers push updates as they happen, others are polled on a quarter-hour cycle. Both land in the same place: an encrypted token record that carries no identity.

**EVENT LISTENERS**

Endpoints that accept webhook payloads whenever a value changes, for example a daily sleep summary from a ring.

TOKEN VAULTING

Access and refresh tokens are sealed with AES-256-GCM under a master key held in a managed key service.

BUFFERED POLLERS

Scheduled workers that pull chunked readings every fifteen minutes from high-frequency telemetry sources.

IDENTITY SEPARATION

Each record is keyed by an anonymous identifier. No relation links a token block to a real name or profile.

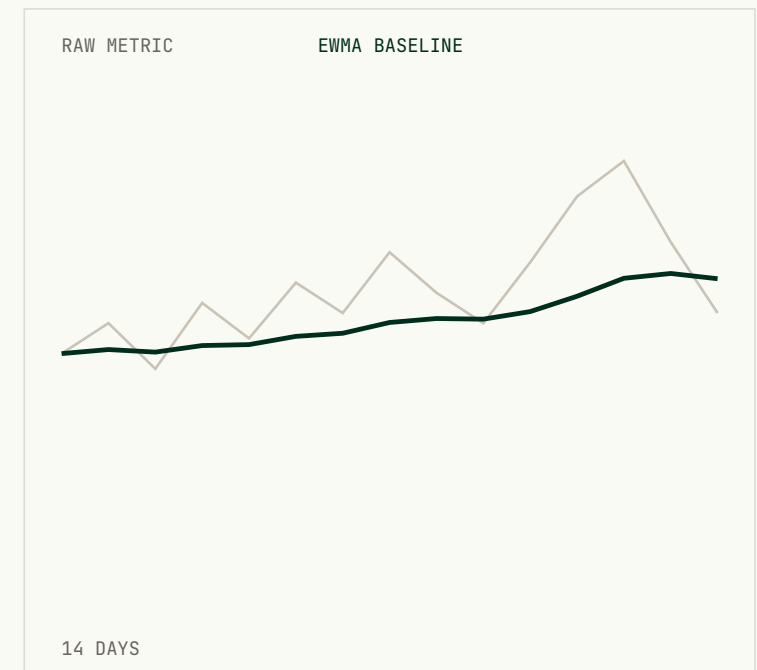
02 / COMPUTATION CORE – TIER ONE

A baseline that belongs to you

Raw metrics feed a time-series model that keeps a personal baseline, refreshed on a fourteen-day exponentially weighted average.

$$B(t) = a \cdot M(t) + (1-a) \cdot B(t-1), \quad a = 2/(N+1), \quad N = 14$$

The system compares the standard deviation of the last hour against that baseline and maps the result into a single unitless anomaly score, 0.00 for settled through to 1.00 for heavy strain. Raw readings are flushed from memory as soon as the score exists.



02 / COMPUTATION CORE – TIER TWO

Read in memory, then destroyed

Workspace hooks read calendars, channels, documents and repository events. The text itself never reaches storage. It lives in memory for a fraction of a second, long enough to count pace and friction, then it is destroyed.

Two factors survive: velocity, meaning interactions completed per minute, and friction, meaning lexical density. Both are folded into one anonymous score on the same 0.00 to 1.00 scale.

0 MS

Payload received

Message metadata enters volatile memory

12 MS

Counted

Velocity and lexical density measured

40 MS

Scored

One friction value on a 0.00 to 1.00 scale

100 MS

Wiped

Buffer overwritten, nothing persisted

03 / LATENT GATEKEEPER

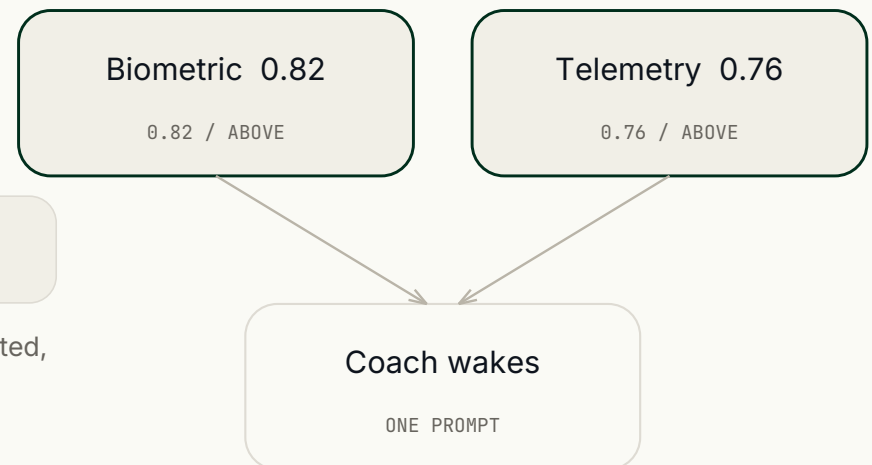
The model sleeps until both agree

Running a language model continuously is expensive and noisy. Sloane keeps the reasoning layer dormant behind a strict logic barrier, so it speaks rarely and only with cause.

`Trigger = True <=> (Ab >= 0.70) AND (Sc >= 0.70)`

Move either score below the line and the pipeline terminates quietly. Nothing is generated, nothing is shown, and no tokens are spent.

```
{
  "intervention_type": "recovery_window",
  "confidence": 0.86,
  "signals": ["hrv_below_baseline", "meeting_density_high"],
  "suggestion": "Move the 16:00 review to tomorrow morning.",
  "requires_confirmation": true,
  "write_permissions": []
}
```



IF EITHER SCORE DROPS BELOW 0.70 THE PIPELINE STOPS HERE

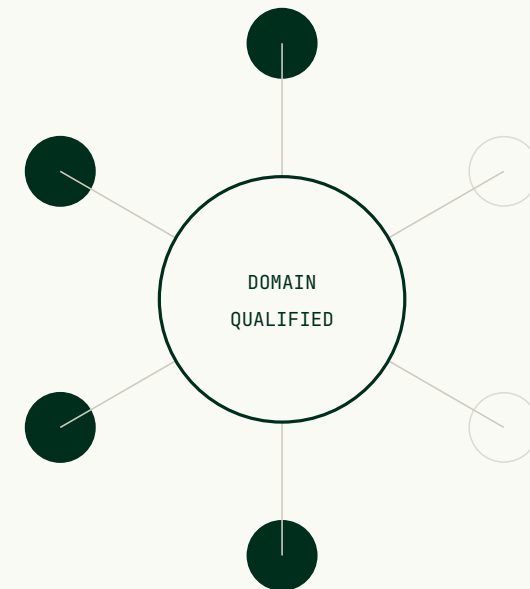
The agent holds no write permissions. It cannot change a record or touch a workspace endpoint until you confirm the action in front of you.

04 / DOMAIN CLUSTERING

One person at a time, then a team

People arrive on their own. When enough colleagues from the same organisation are active, the pattern becomes visible at a structural level, without exposing anyone individually.

Once a domain holds three or more active people it is flagged as qualified. That flag permits aggregated, high-level capacity signals for the group, and nothing more. Individual scores are never shared upward.



THREE OR MORE ACTIVE PEOPLE ON ONE DOMAIN

What the system refuses to do

These are constraints in the architecture, not policy promises.

No message content is stored	Workspace text exists in volatile memory only, and is wiped inside 100 milliseconds.
No identity in the token vault	Credentials are keyed by an anonymous identifier with no relation to a profile.
No silent action	Every suggestion requires explicit confirmation before anything is executed.
No manager-level surveillance	Only aggregated capacity signals are ever surfaced to an organisation.
No always-on inference	The reasoning layer runs only when both independent scores clear the threshold.

If this is the sort of system you want to help shape, we are running a small design cohort.

sloanelab.com/how-we-begin