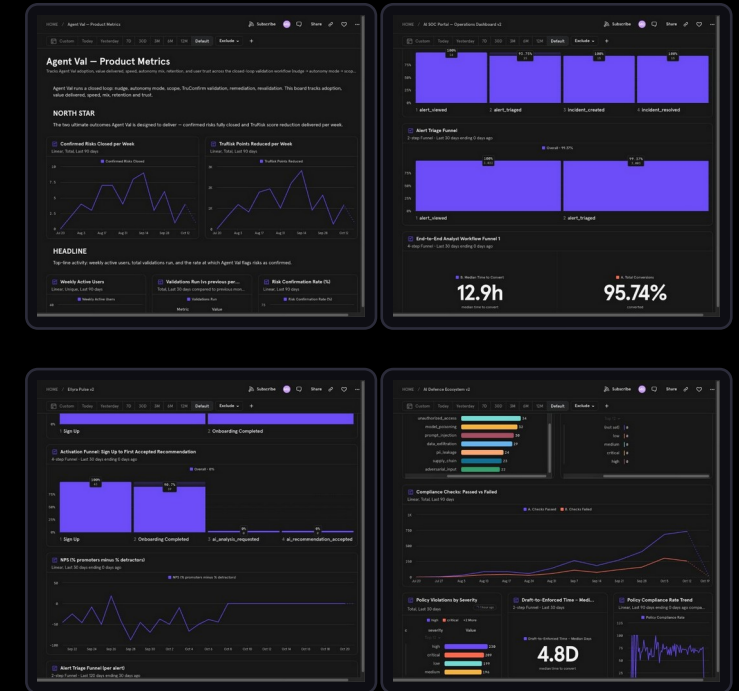


My Product Metrics

How product metrics were decided, mapped, targeted and used across seven AI products — each explained with its live Mixpanel dashboard

Mayukh Ghosh · October 2026



Live v2 dashboards in Mixpanel

Disclaimer All metric values in this deck are for demo purposes only and were generated using synthetic data.

Disclaimer: synthetic demo data



Metric values are for demo purposes only.

Every number, trend and funnel in this deck was generated using synthetic data scripts. None of it reflects real customers, real usage or real product performance.

- Thresholds are proposed starting points to be validated with real baselines.
- Flat lines after mid-October reflect when the synthetic event feed stopped, not product behaviour.
- Known data gaps are called out on the slides where they appear.

What is real

- The product workflows (from each app's GitHub repo and the Agent Val workflow video)
- The event design and metric definitions
- The Mixpanel dashboards and their structure

Seven apps, seven north-star metrics

1 Agent Val

North star

Confirmed risks closed per week

2 AI SOC Portal

North star

Alerts resolved end-to-end (per alert) and time to resolve

3 AI-SPM

North star

Active accounts running risk assessments every week

4 AI Policy Foundry

North star

Median time from policy draft to enforcement

5 Ellyra Pulse

North star

NPS (% promoters – % detractors) with zero unrouted safety signals

6 Agritech Crop Insurance

North star

Hours from payout trigger to money in the farmer's hands

7 AI Defence Ecosystem

North star

Policy compliance rate and draft-to-enforced time across the AI estate

Agent Val is new: its board was designed from the workflow video and filled with synthetic events.

One method, applied to every app

Based on the product analytics approach from the 'Key Analytical Tools & Communicating Persuasively with Data' session, using Mixpanel

1 Events + properties

Walk the user journey; one broad event per step, detail in properties

2 Absolute numbers

North-star counts and active users for reporting, with prior-period comparison

3 Medians and P90

Time metrics use median and 90th percentile so outliers don't mislead

4 Funnels

Map the journey step by step; count the object (alert, policy) where that is the unit

5 Retention

Cohorts on the value action, with enough time to mature; shape over level

6 Filter + breakdown

Slice by channel, severity, industry, mode to find where a metric breaks

Five questions answered for each app

1 How the metrics were decided

2 Mapping the app to its metrics

3 Objective of the metrics

4 Thresholds that require action

5 Value the metrics provide

Threshold status

On track

Inside the green band

Watch

In the amber band — monitor

Act now

In the red band — take the action listed

Every section slide uses a real screenshot of the app's v2 Mixpanel dashboard.

Agent Val

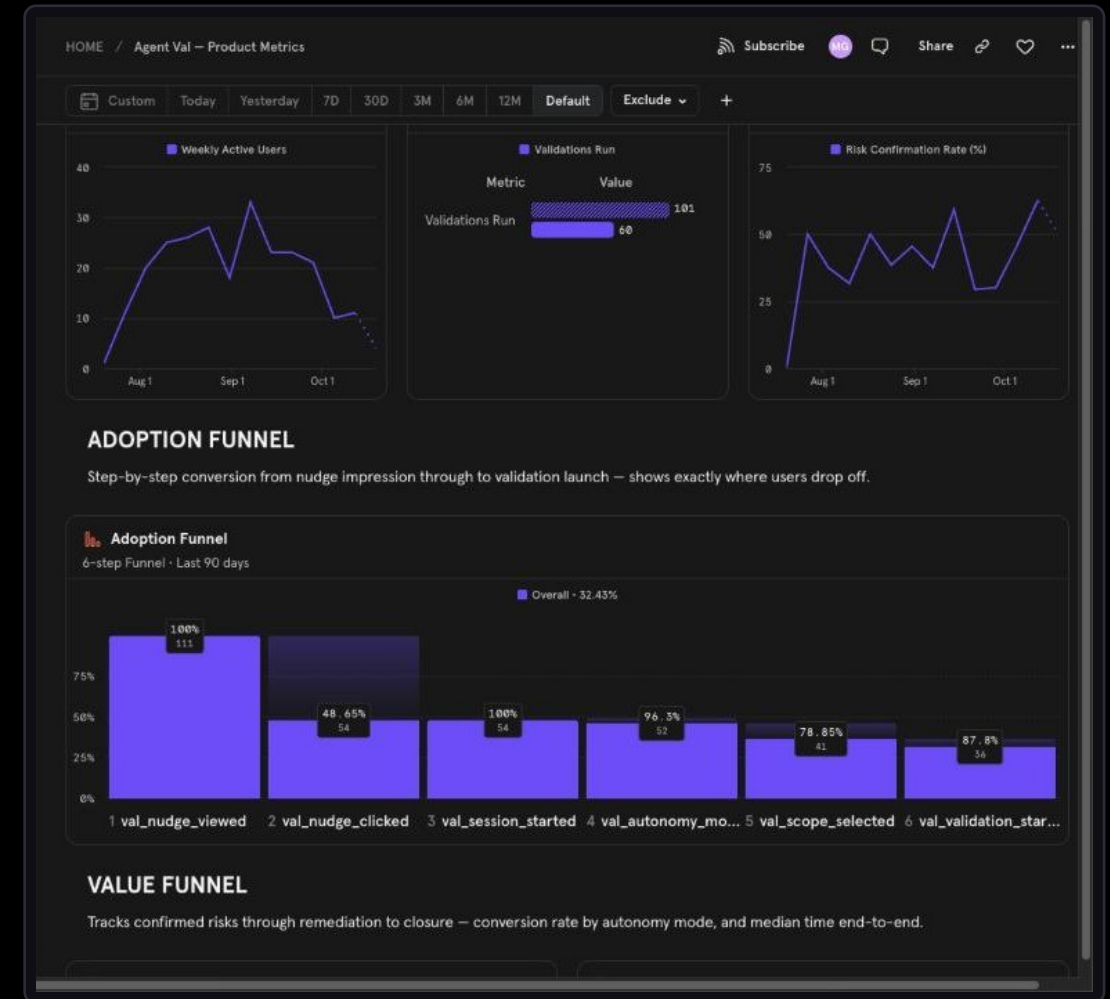
Agentic exploit validation inside Qualys ETM: nudge, autonomy mode, scope, TruConfirm validation, remediation, revalidation

NORTH STAR

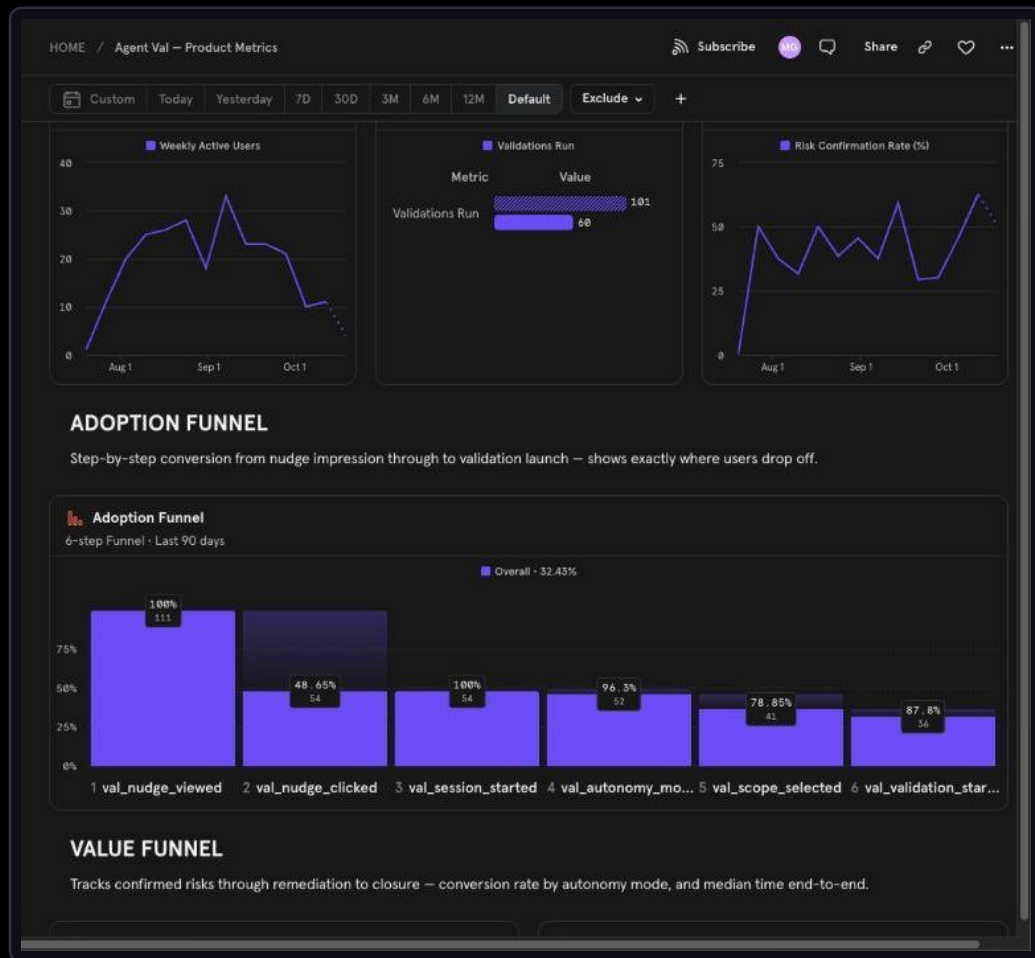
Confirmed risks closed per week

Dashboard: **Agent Val — Product Metrics**

Source: Agent Val workflow video



How the metrics were decided



Adoption funnel built from the six steps a user takes before a validation starts

1 The problem

Security teams know their findings but not which ones an attacker can exploit right now

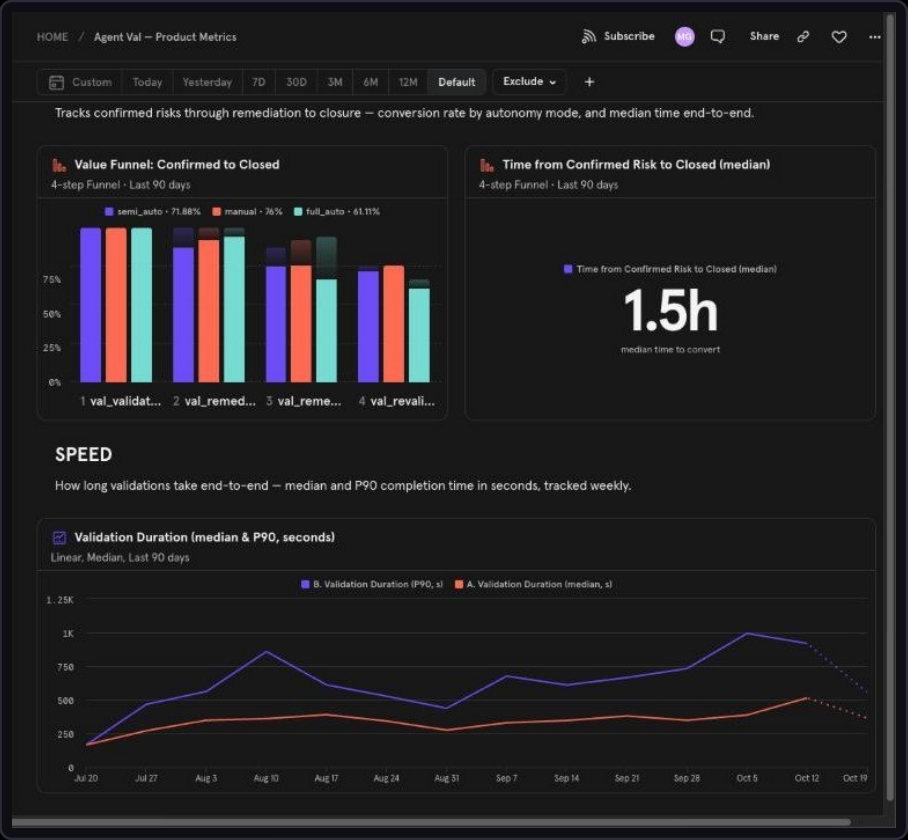
2 The workflow

The video shows one closed loop: nudge on Vulnerabilities page, pick autonomy mode, select CVE and assets, run TruConfirm, remediate, revalidate

3 The metrics that follow

One event per loop step with properties (autonomy_mode, result, option, status), then absolutes, medians, funnels and retention on top

Mapping the app to its metrics



Dashboard cards that carry these metrics

Workflow step	Event tracked	Metric on the board
Nudge on Vulnerabilities page	val_nudge_viewed / clicked	Adoption funnel step 1-2
Choose autonomy mode	val_autonomy_mode_selected	Autonomy mode mix, override rate
Select CVE + assets	val_scope_selected	Adoption funnel step 5
TruConfirm validation	val_validation_completed	Confirmation rate, duration P50/P90
Remediate (patch / mitigation)	val_remediation_applied	Option and outcome, failure rate
Post-fix revalidation	val_revalidation_completed	North star: risks closed, TruRisk reduced

Objective of the metrics

1

Prove risk is closed

Confirmed risks closed and TruRisk points reduced per week are the outcome Agent Val exists to deliver

2

Find adoption leaks

Nudge-to-validation funnel shows where users stop before the agent can help

3

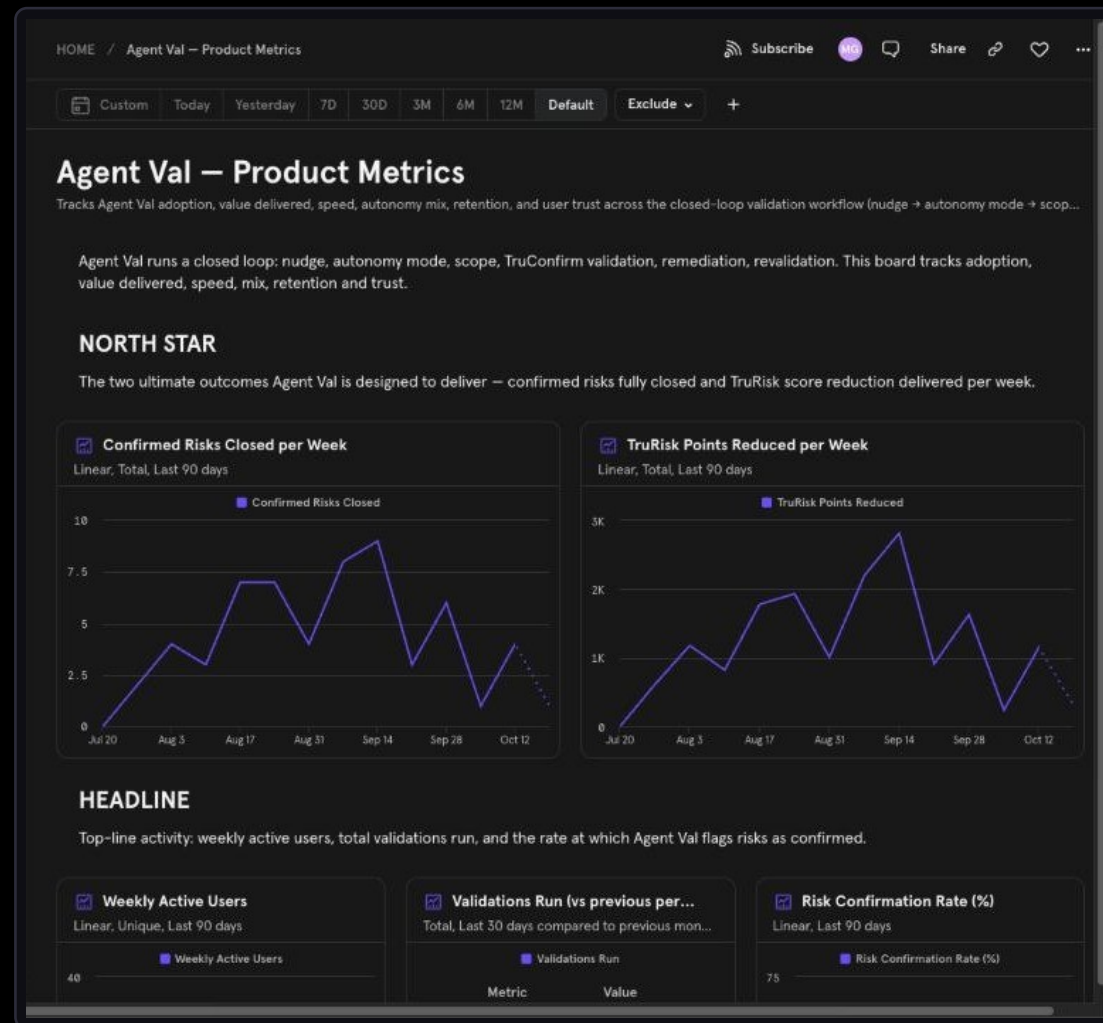
Keep the loop fast

Median time from confirmed risk to closed (1.5h) tracks the speed promise

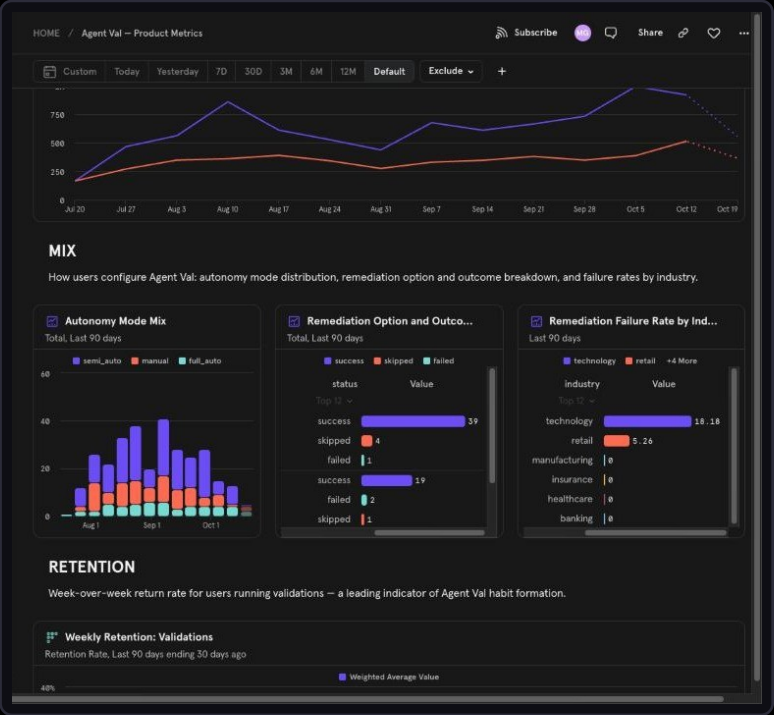
4

Earn trust

Feedback, evidence views and mode overrides show whether users trust the agent



Thresholds that require action

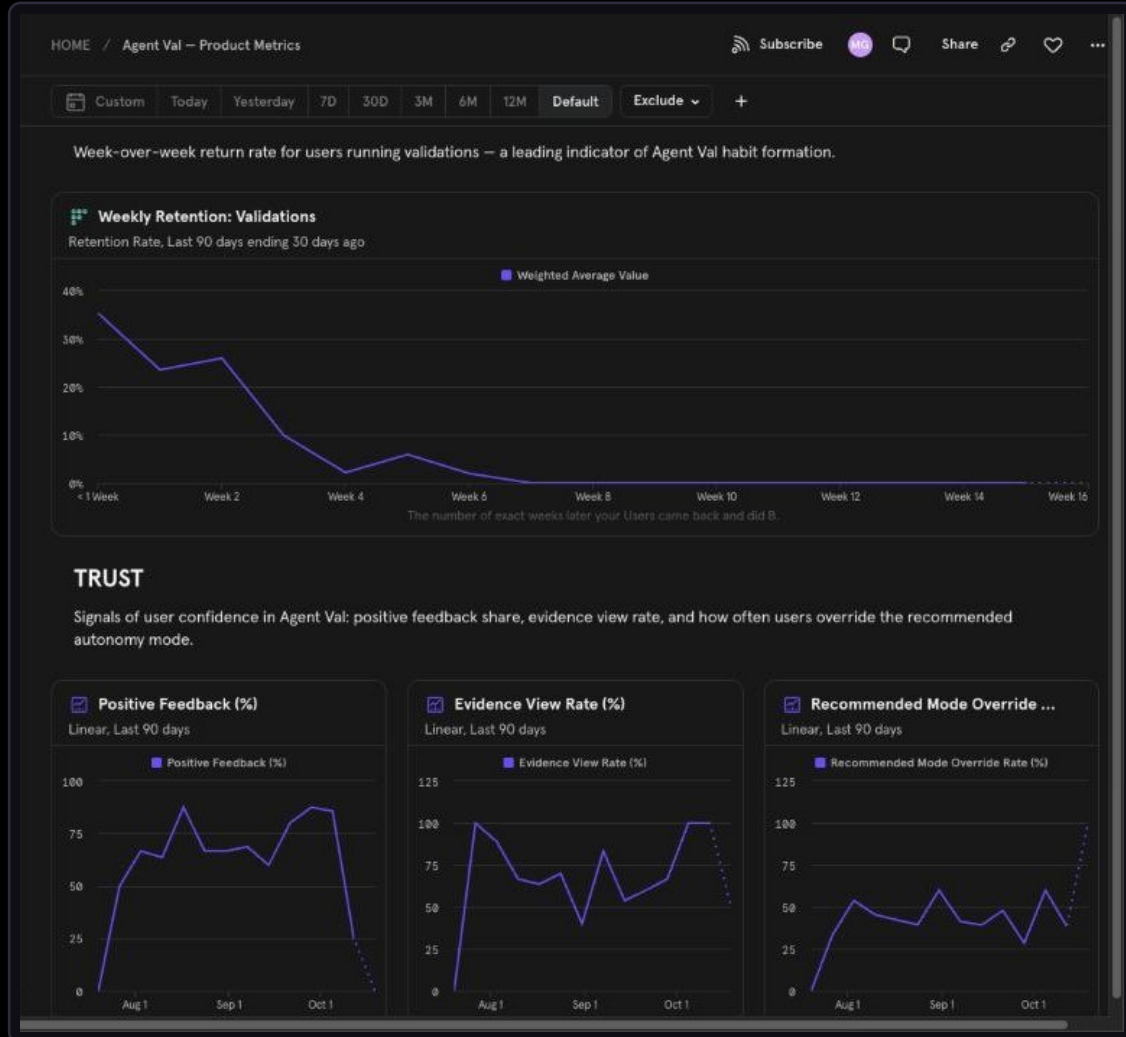


Action Red on failure or retention: review remediation playbooks for technology customers and add a weekly re-engagement nudge

Metric	Green	Amber	Red	Now (synthetic)	Status
Risks closed / week	≥ 8	4-7	< 4	Recent weeks 1-6	Watch
Adoption funnel (nudge→validation)	≥ 35%	25-35%	< 25%	32.4%	Watch
Confirmed→closed (full auto)	≥ 75%	60-75%	< 60%	61.1%	Watch
Median confirmed→closed	≤ 2h	2-6h	> 6h	1.5h	On track
Remediation failure (any industry)	≤ 5%	5-10%	> 10%	Technology 18.2%	Act now
Week-4 validation retention	≥ 15%	8-15%	< 8%	~2-6%	Act now

Thresholds are proposed starting points; validate against real baselines before use.

Value the metrics provide



Product (PM)

Shows which loop step to fix next and lets Cost of Delay rank fixes by risk closed

Security operations

Shows where remediation fails, by industry and option, before customers escalate

Leadership / Sales

TruRisk points reduced is a customer-facing ROI number for renewals and expansion

AI SOC Portal

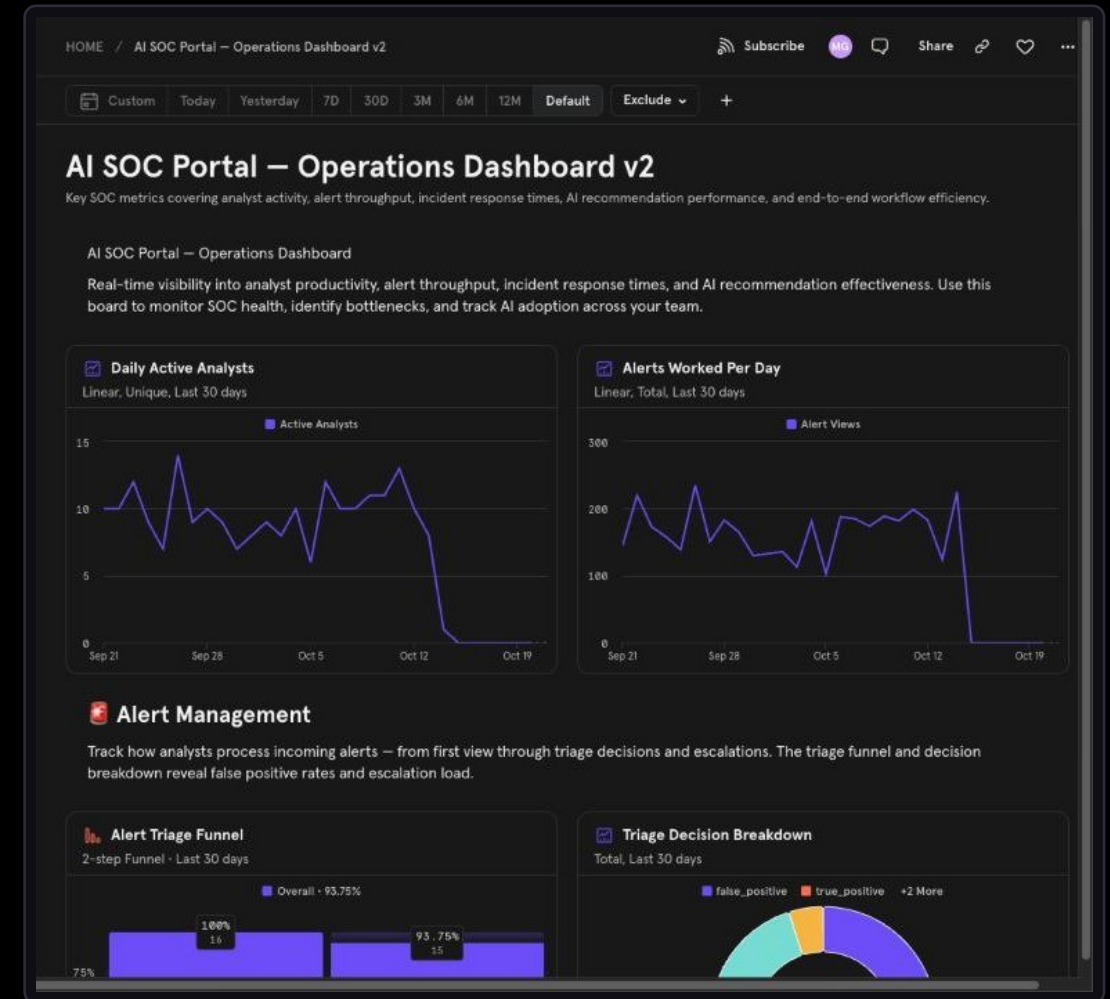
AI-assisted security operations: alert triage, incident response and AI recommendations for SOC analysts

NORTH STAR

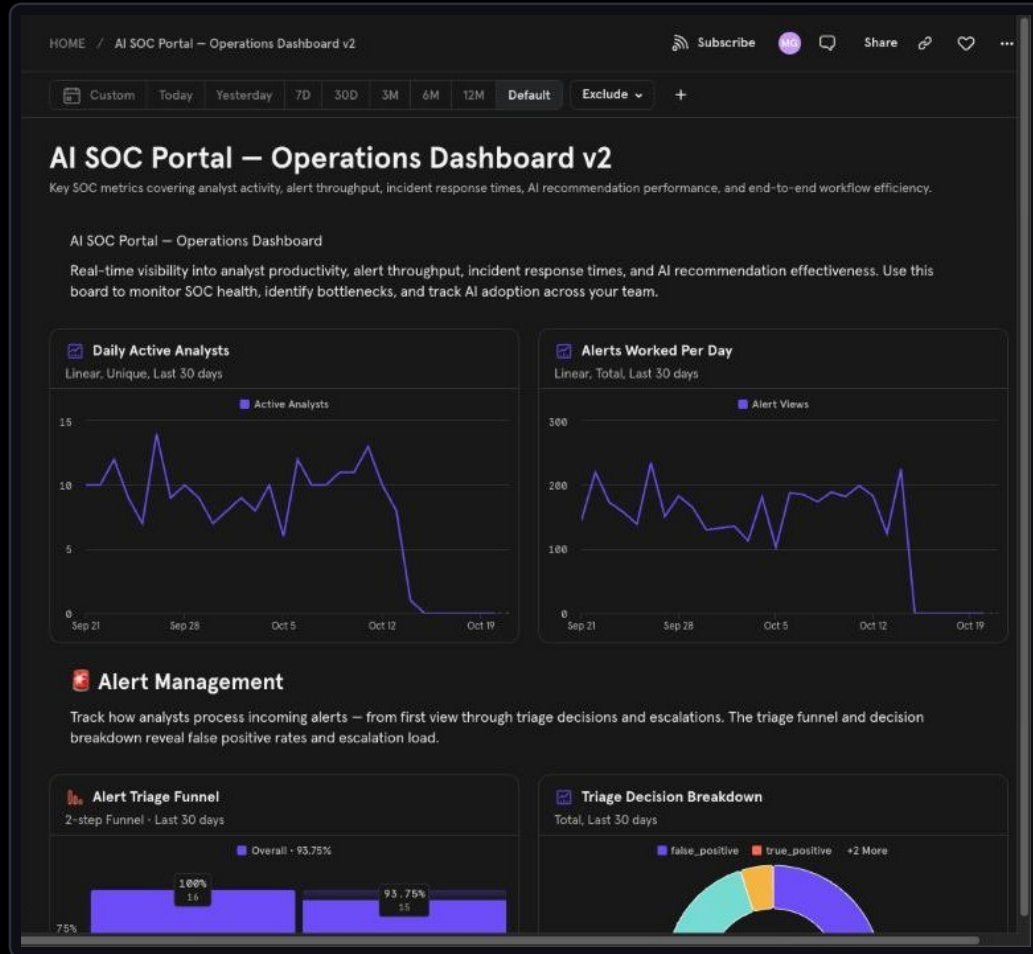
Alerts resolved end-to-end (per alert) and time to resolve

Dashboard: AI SOC Portal — Operations Dashboard v2

Source: github.com/mayukhg/ai-soc-portal



How the metrics were decided



Board opens with analyst activity and alert throughput — the volume problem the portal targets

1 The problem

SOCs face 10,000+ alerts a day, ~70% false positives and slow response (repo executive pitch)

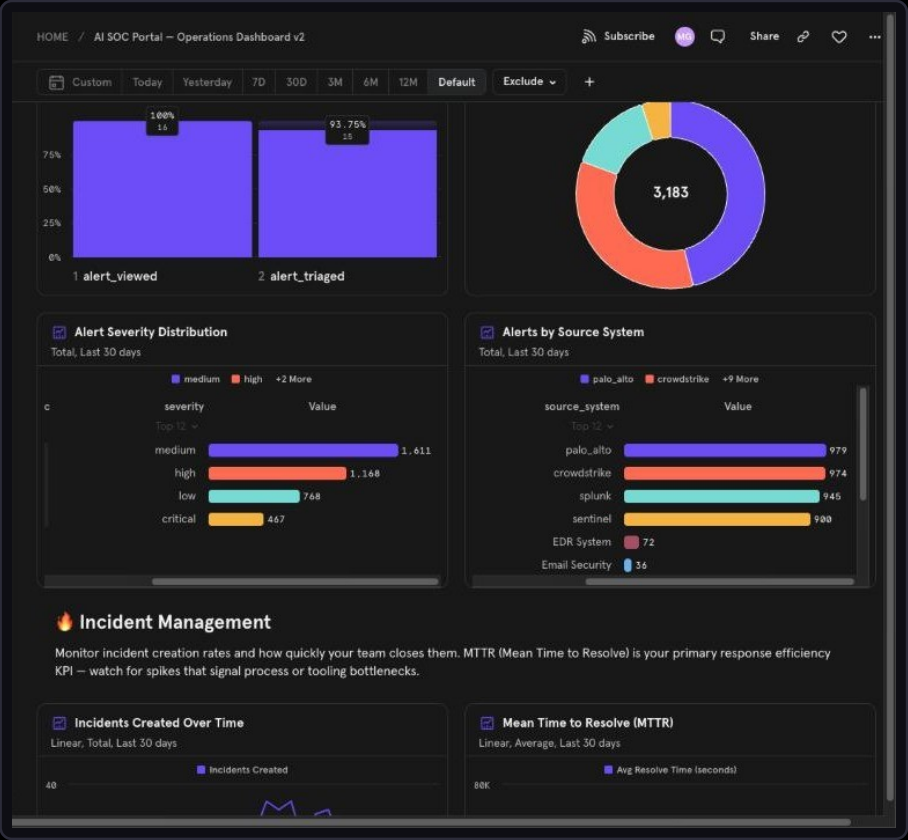
2 The workflow

Alert viewed → triaged (true / false positive) → incident created → resolved, with AI analysis requested along the way

3 The metrics that follow

Count work per alert (hold alert_id), use median and P90 for time, and measure AI acceptance and feedback

Mapping the app to its metrics



Dashboard cards that carry these metrics

Workflow step	Event tracked	Metric on the board
Analyst opens alert	alert_viewed	Alerts worked per day, daily active analysts
Triage decision	alert_triaged (triage_decision)	Triage funnel, false positive rate
Escalate to incident	incident_created	Incidents created, severity mix
Resolve incident	incident_resolved (time_to_resolve)	Time to resolve P50/P90 by severity
Use AI assistant	ai_analysis_requested / accepted	AI requests, acceptance rate by type
Rate AI output	ai_feedback_submitted	% positive AI feedback

Objective of the metrics

1

Count alerts, not analysts

Per-alert triage funnel: 3,022 → 3,003 (99.4%) instead of 16 users

2

Measure end-to-end flow

End-to-end funnel converts 95.7% with a 12.9h median time to convert

3

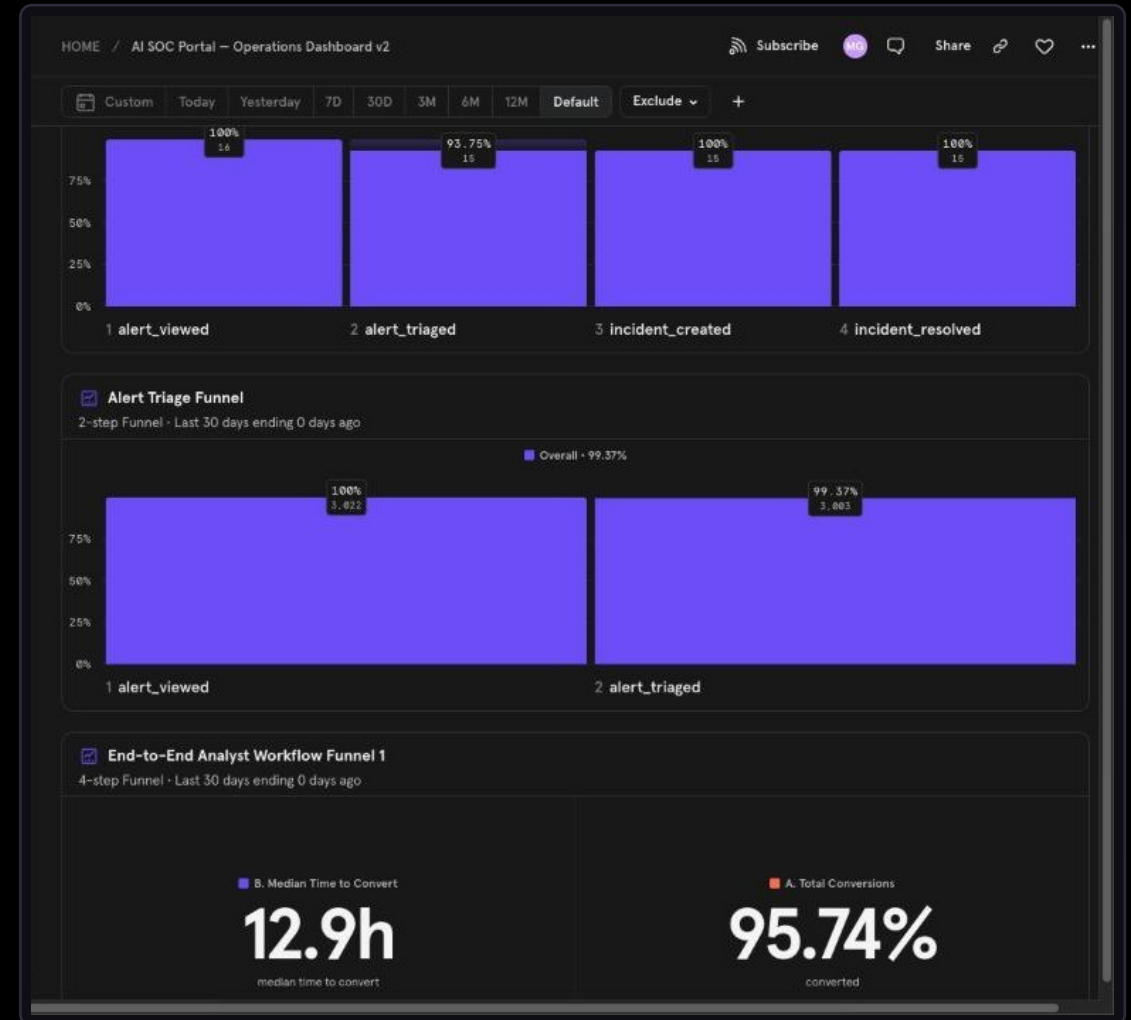
Cut false-positive load

False positive rate shows how much triage effort is wasted

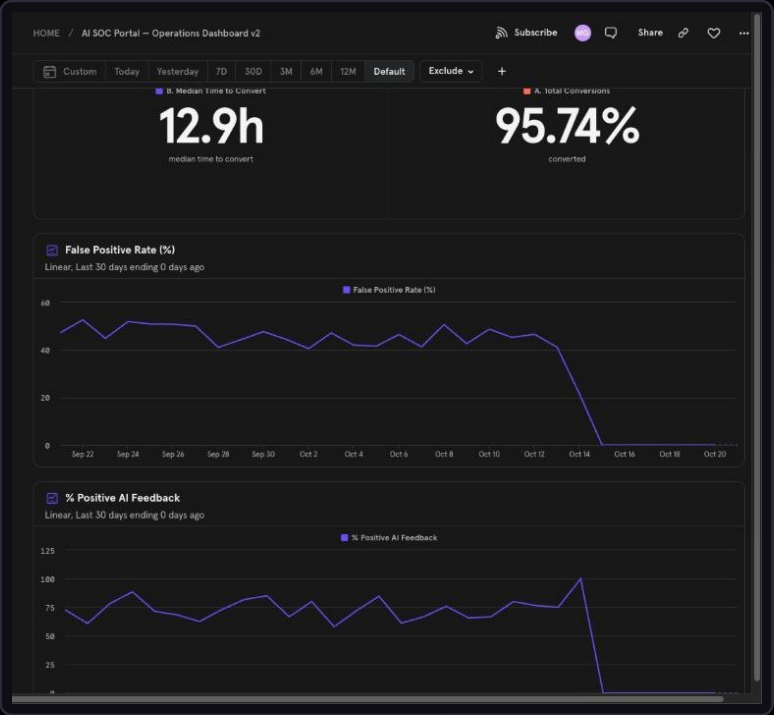
4

Prove AI helps

AI acceptance and positive feedback show whether recommendations reduce analyst burden



Thresholds that require action

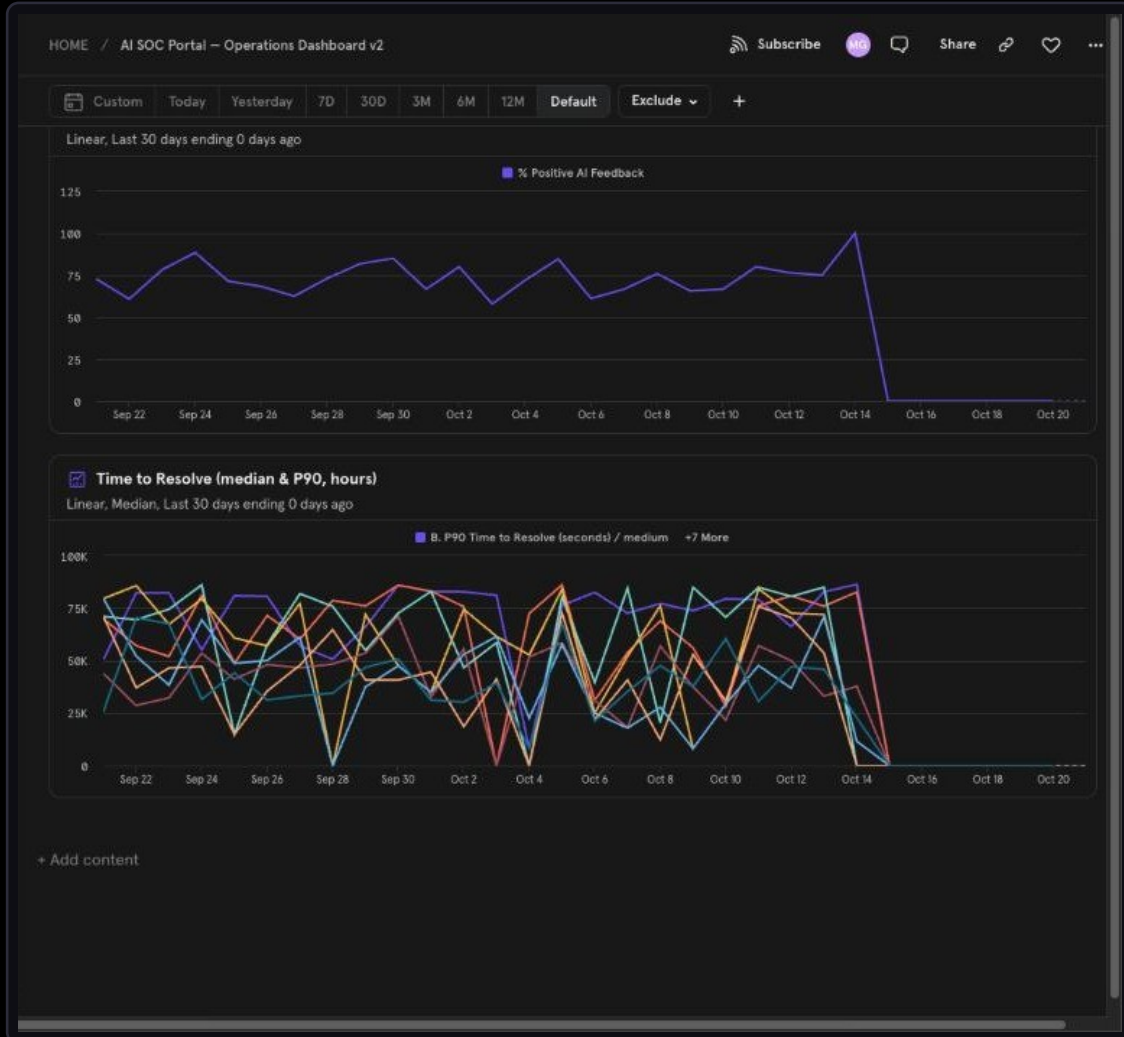


Action Red on false positives: tune noisiest source systems (palo_alto, crowdstrike) and feed false-positive labels back to the triage model

Metric	Green	Amber	Red	Now (synthetic)	Status
Per-alert triage conversion	≥ 98%	95–98%	< 95%	99.4%	On track
False positive rate	≤ 30%	30–45%	> 45%	~40–52%	Act now
End-to-end median time	≤ 8h	8–16h	> 16h	12.9h	Watch
AI acceptance rate	≥ 70%	55–70%	< 55%	~55–75%	Watch
% positive AI feedback	≥ 75%	60–75%	< 60%	~60–88%	Watch
Active analysts vs 7-day avg	≥ 90%	70–90%	< 70%	Feed stopped ~Oct 14	Act now

Thresholds are proposed starting points; validate against real baselines before use.

Value the metrics provide



Product (PM)

Verifies the pitch claims (faster response, fewer false positives) with measured trends

SOC lead

Time to resolve by severity shows where staffing or playbooks are the bottleneck

CISO / leadership

AI acceptance and feedback turn AI adoption into a reportable efficiency story

AI-SPM

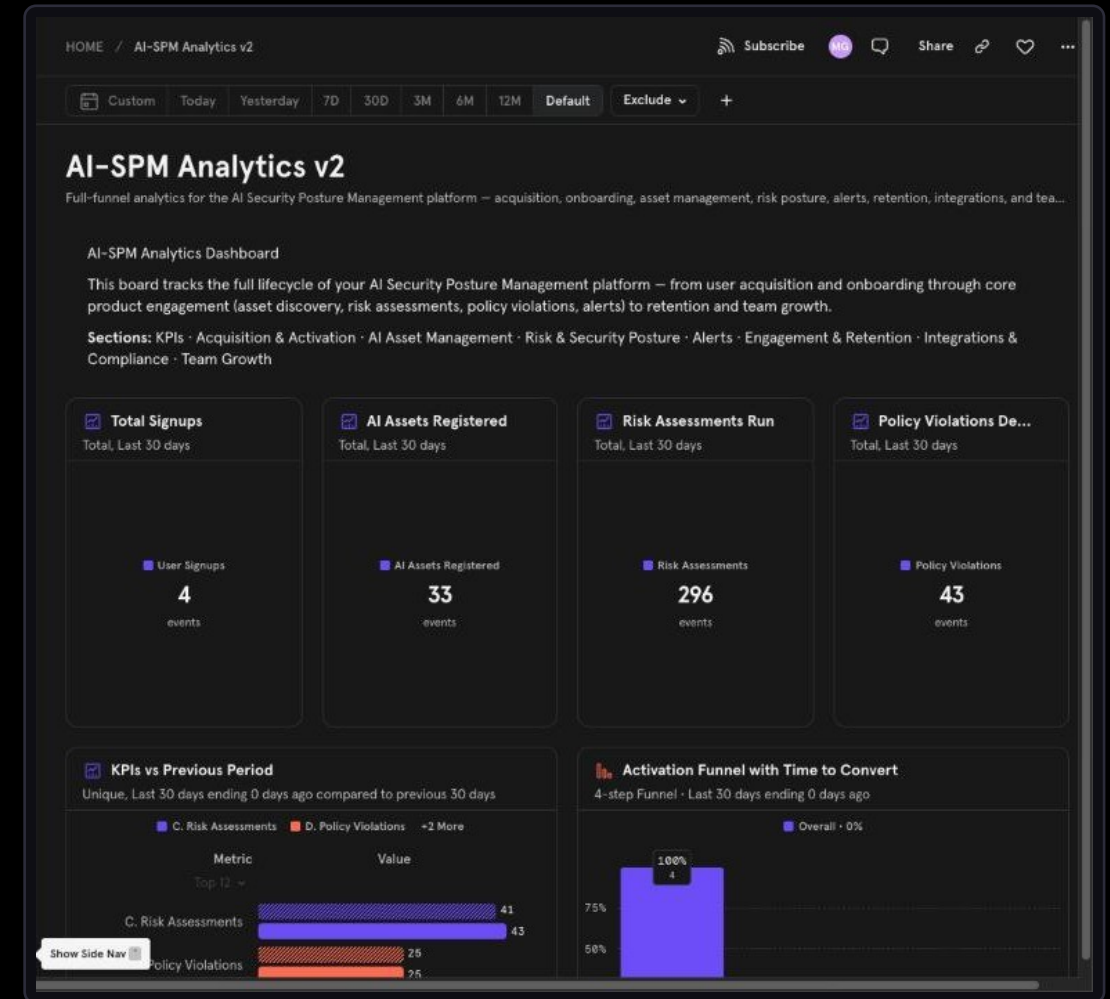
AI Security Posture Management: discover AI assets, assess risk, detect policy violations and stay compliant

NORTH STAR

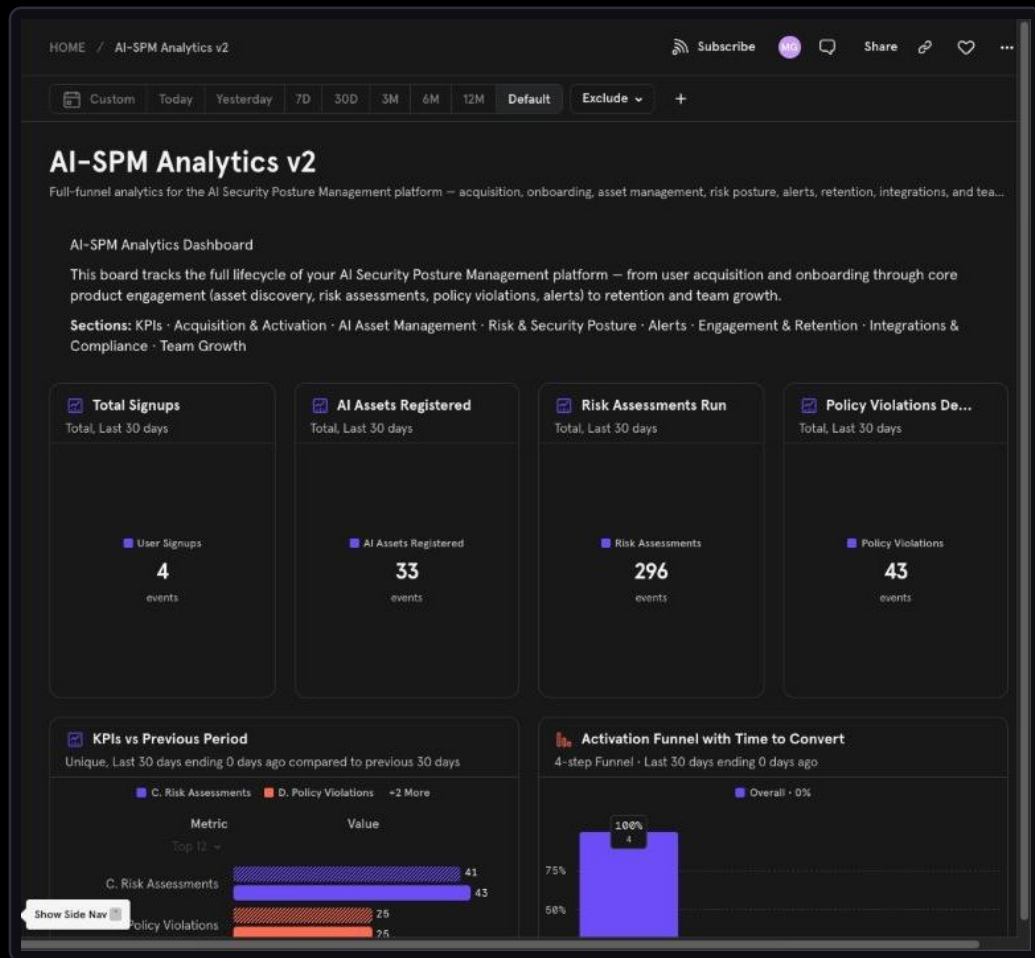
Active accounts running risk assessments every week

Dashboard: AI-SPM Analytics v2

Source: github.com/mayukhg/ai-spm



How the metrics were decided



KPI tiles: signups, AI assets registered, risk assessments and policy violations

1 The problem

CISOs lack an inventory of AI assets and a way to quantify AI risk for the board (repo executive pitch)

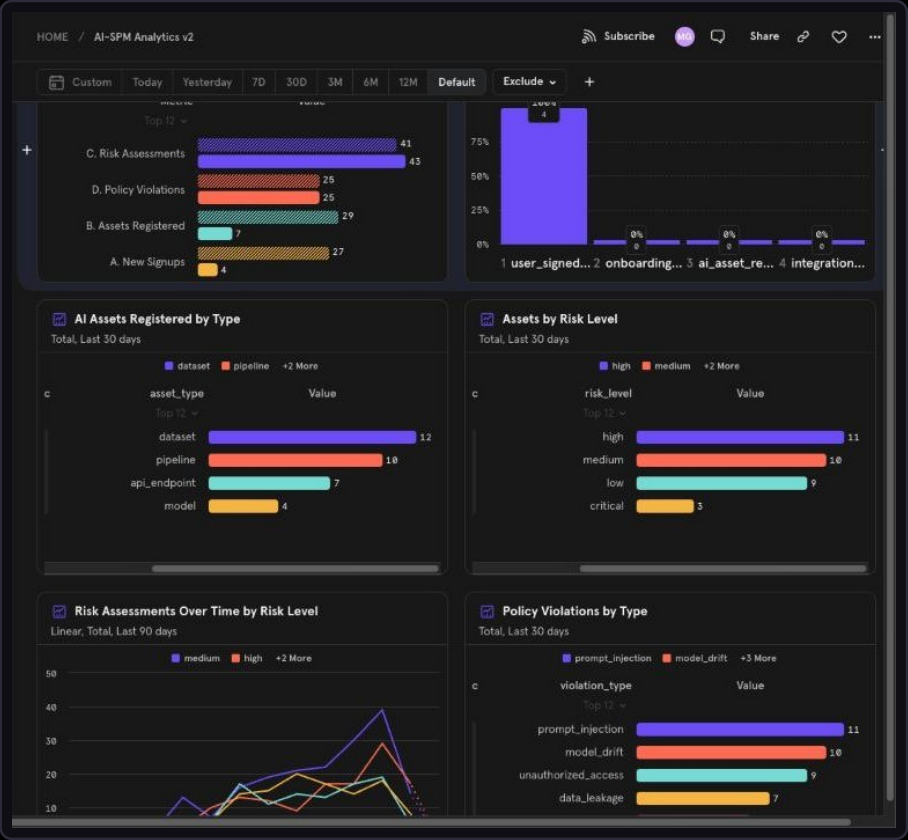
2 The workflow

Sign up → onboard → register AI assets → connect integrations → run risk assessments → handle violations and alerts → invite team

3 The metrics that follow

Follow the lifecycle: activation funnel, core-action volume, alert resolution speed, then weekly retention on risk assessments

Mapping the app to its metrics



Dashboard cards that carry these metrics

Workflow step	Event tracked	Metric on the board
Sign up and onboard	user_signed_up, onboarding_completed	Activation funnel with time to convert
Register AI assets	ai_asset_registered (asset_type)	Assets by type and risk level
Assess risk	risk_assessment_completed	Assessments over time by risk level
Detect violations	policy_violation_detected	Violations by type
Resolve alerts	alert_resolved (resolution_type)	Alert time to resolve, resolution by type
Grow the team	team_member_invited (role)	Invites by role, integrations by platform

Objective of the metrics

1 Activate new accounts

1

Activation funnel shows whether signups reach a first registered asset

2 Build a weekly habit

2

Weekly retention on risk assessments is the core-value retention signal

3 Respond fast to alerts

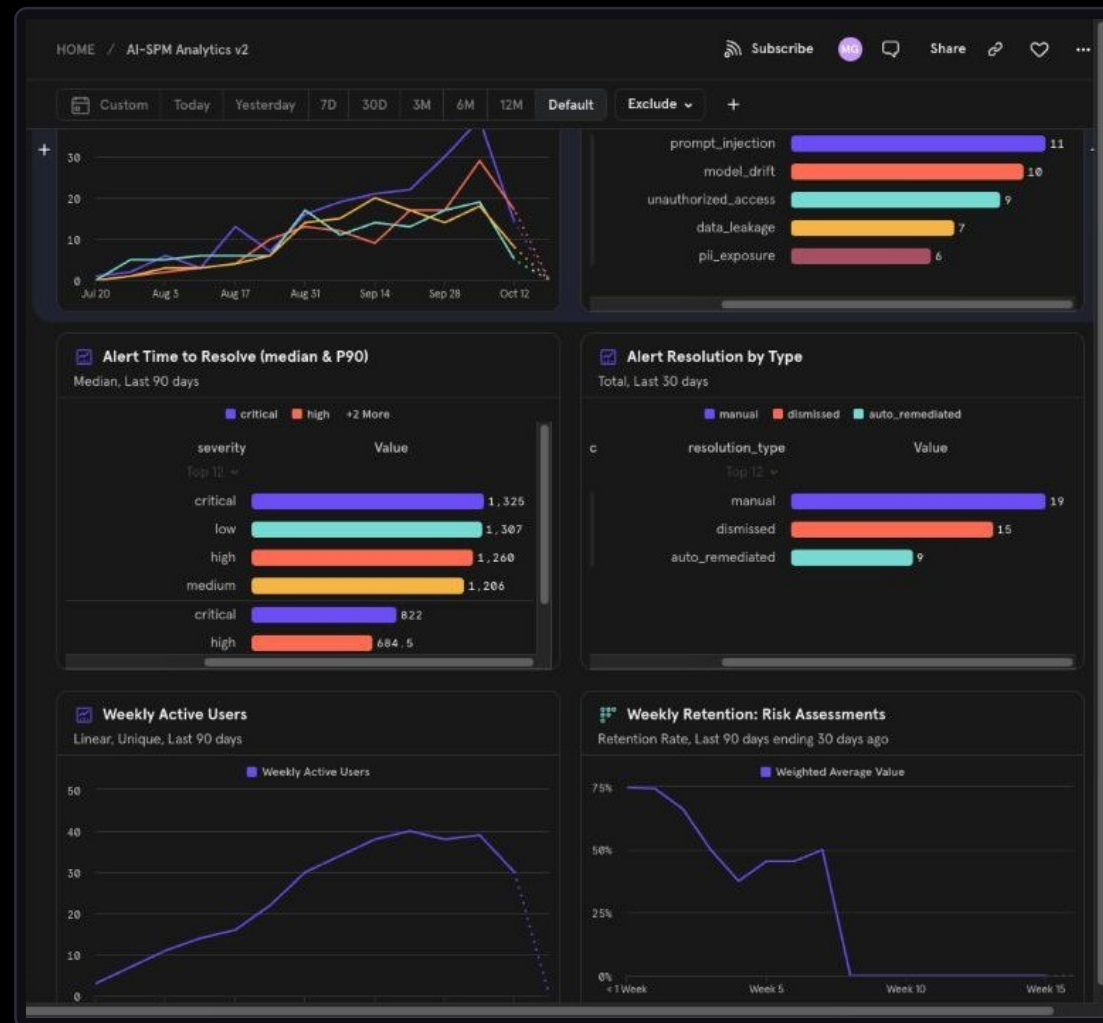
3

Median and P90 time to resolve by severity keep critical alerts first

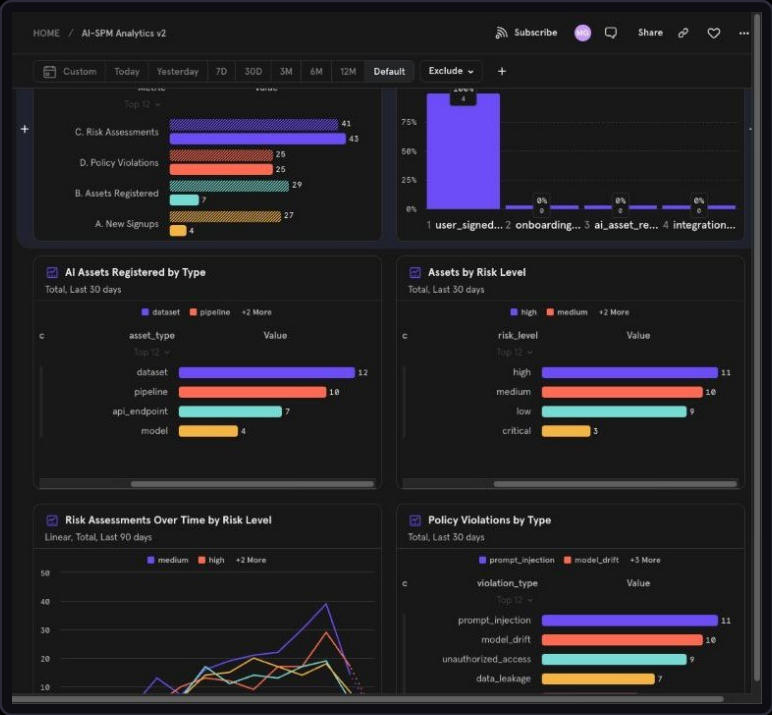
4 Show account growth

4

Weekly active users and team invites show expansion inside accounts



Thresholds that require action

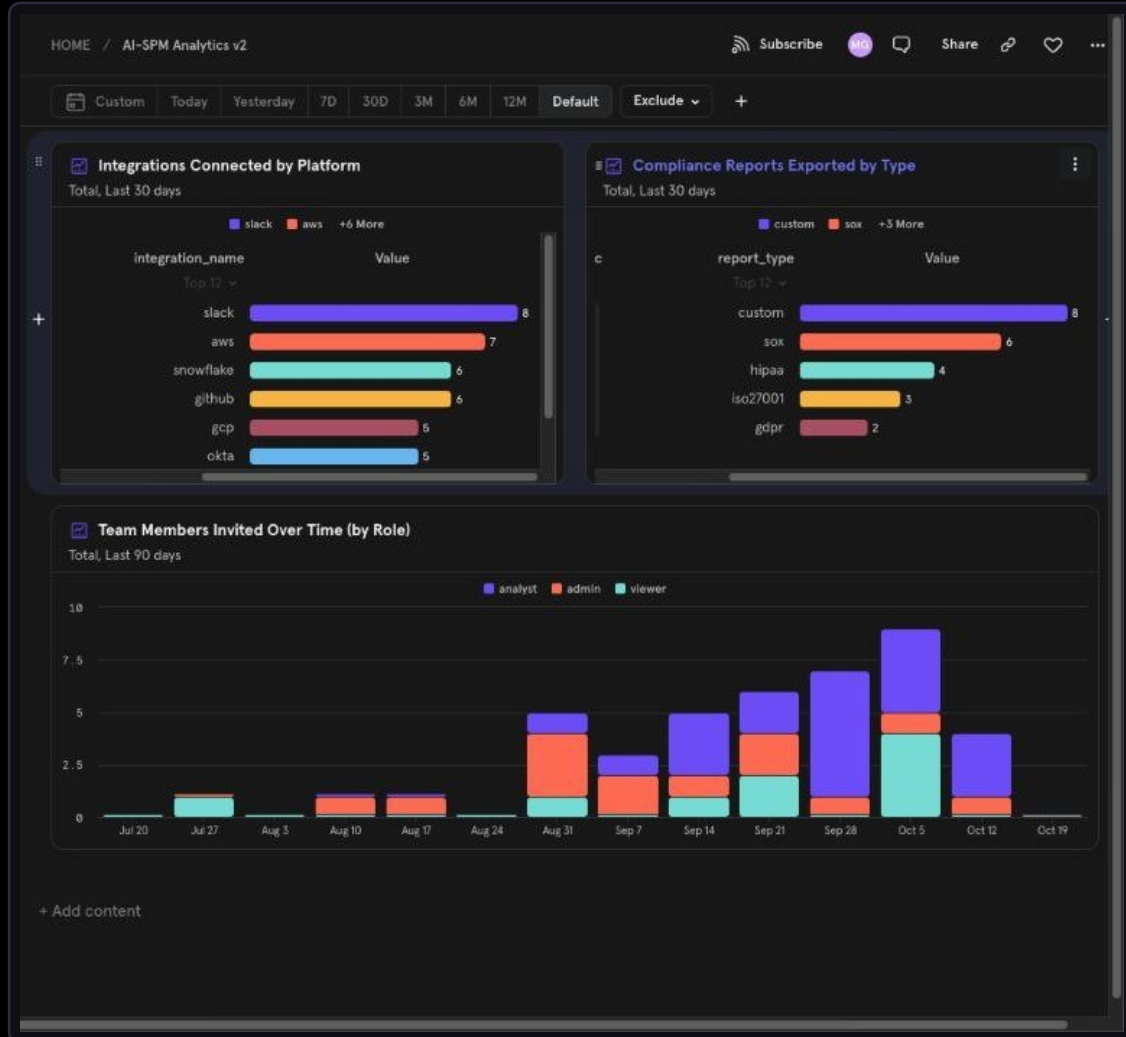


Metric	Green	Amber	Red	Now (synthetic)	Status
Signup → onboarding	≥ 60%	30–60%	< 30%	0% (4 signups)	Act now
Week-4 assessment retention	≥ 40%	25–40%	< 25%	~45% → 0% after wk 6	Watch
Critical vs high alert TTR	Critical faster	Equal	Critical slower	1,325 vs 1,260	Act now
Auto-remediated share of alerts	≥ 40%	20–40%	< 20%	21% (9 of 43)	Watch
Assets registered vs prior 30d	Up	Flat	Down > 20%	29 → 7	Act now

Action Red on activation: add a guided first-asset registration in onboarding and alert on any signup with no asset after 24h

Thresholds are proposed starting points; validate against real baselines before use.

Value the metrics provide



Product (PM)

Pinpoints the onboarding step blocking activation, the biggest lever on growth

Security / compliance

Compliance exports by framework give audit-ready evidence of coverage

Leadership

Asset inventory, violations and team growth give a board-level AI risk view

AI Policy Foundry

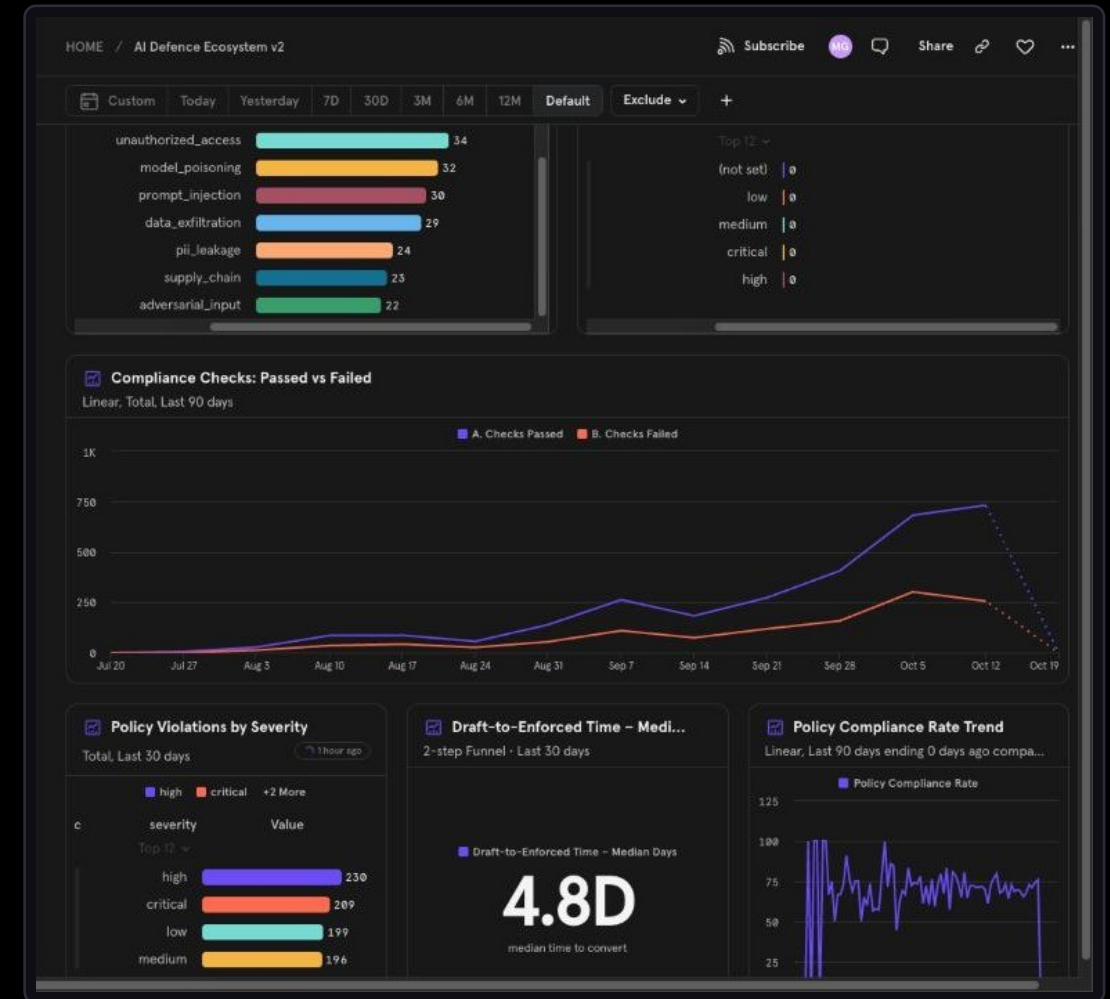
Multi-agent cloud security policy generation with human approval gates and compliance validation

NORTH STAR

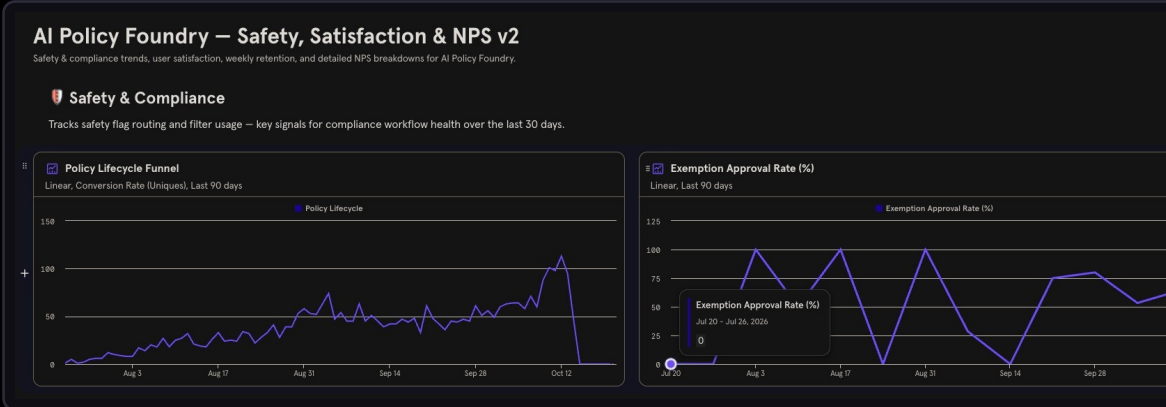
Median time from policy draft to enforcement

Dashboard: AI Policy Foundry — Safety, Satisfaction & NPS v2

Source: github.com/mayukhg/ai-policy-foundry



How the metrics were decided



Policy lifecycle funnel (now populated) and exemption approval rate

1 The problem

Manual policy creation takes 2–6 weeks per cloud service, leaving 30–60 day security gaps (repo executive pitch)

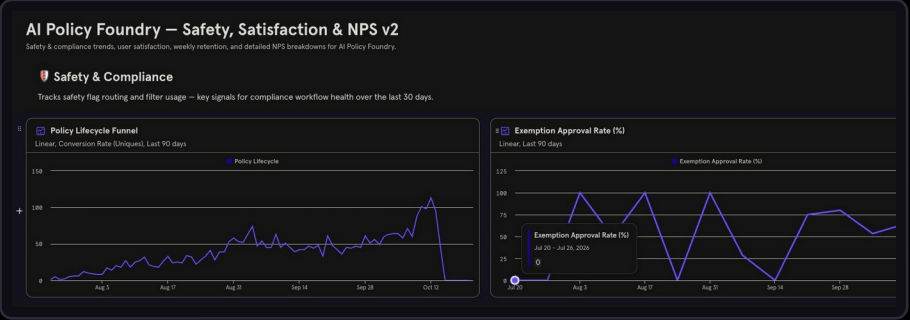
2 The workflow

Policy drafted → submitted for review → approved → published → enforced; exemptions requested and approved; compliance checks run

3 The metrics that follow

Lifecycle funnel per policy (hold policy_id), cycle time, exemption and compliance pass rates, plus user satisfaction

Mapping the app to its metrics



Dashboard cards that carry these metrics

Workflow step	Event tracked	Metric on the board
Draft policy (AI agents)	policy_drafted	Lifecycle funnel step 1
Human review and approval	policy_submitted_for_review, policy_approved	Funnel conversion, review time
Publish and enforce	policy_published, policy_enforced	Draft-to-enforced cycle time
Request exemption	exemption_requested / approved	Exemption approval rate
Validate compliance	compliance_check_passed / failed	Pass rate by framework
User feedback	nps_response_recorded	NPS, retention

Objective of the metrics

1

Deliver the 24h promise

Draft-to-enforced time tests the pitch claim of policies in under 24 hours

2

Keep humans in control

Review and approval steps in the funnel show where approvals stall

3

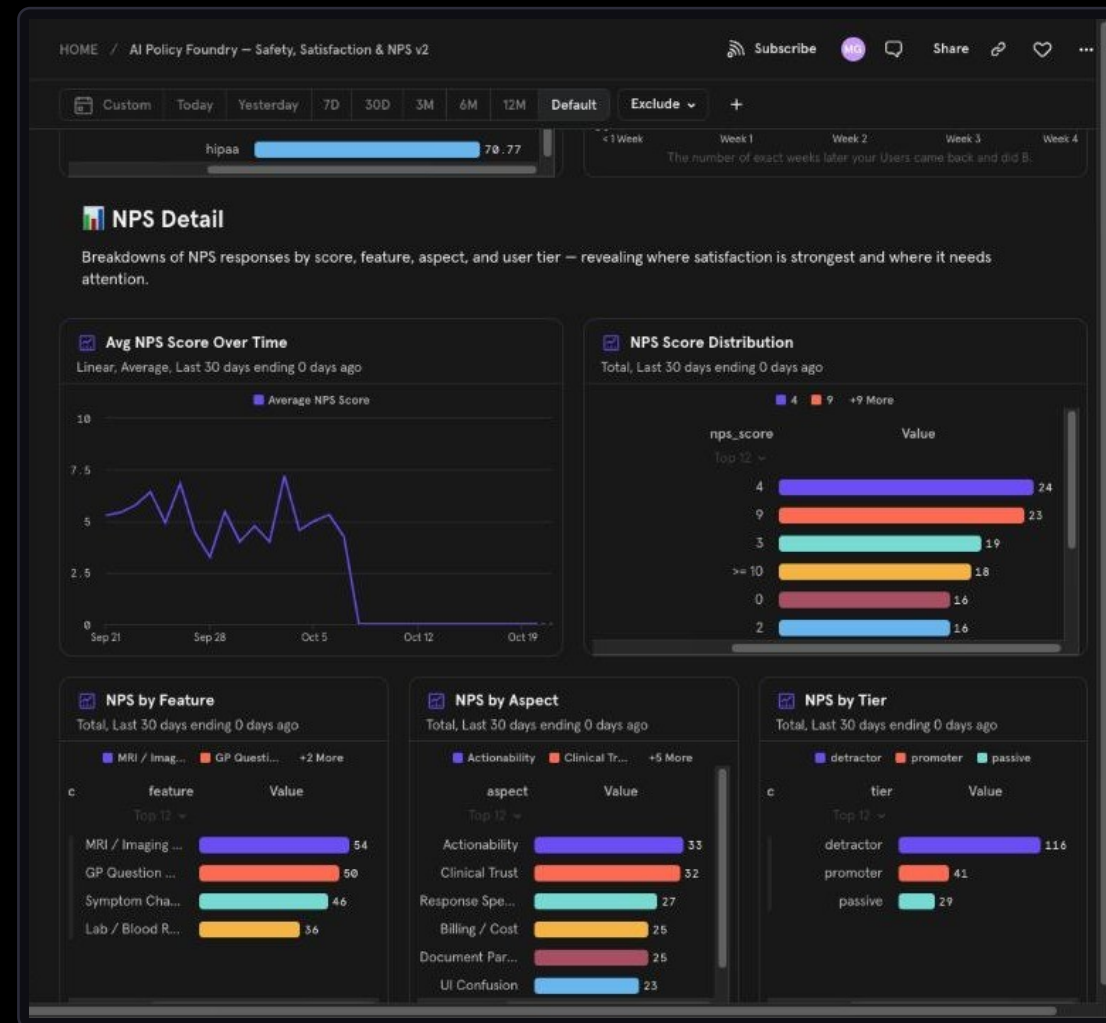
Prove compliance coverage

Pass rate by framework (ISO 42001, EU AI Act, NIST AI RMF...) shows audit readiness

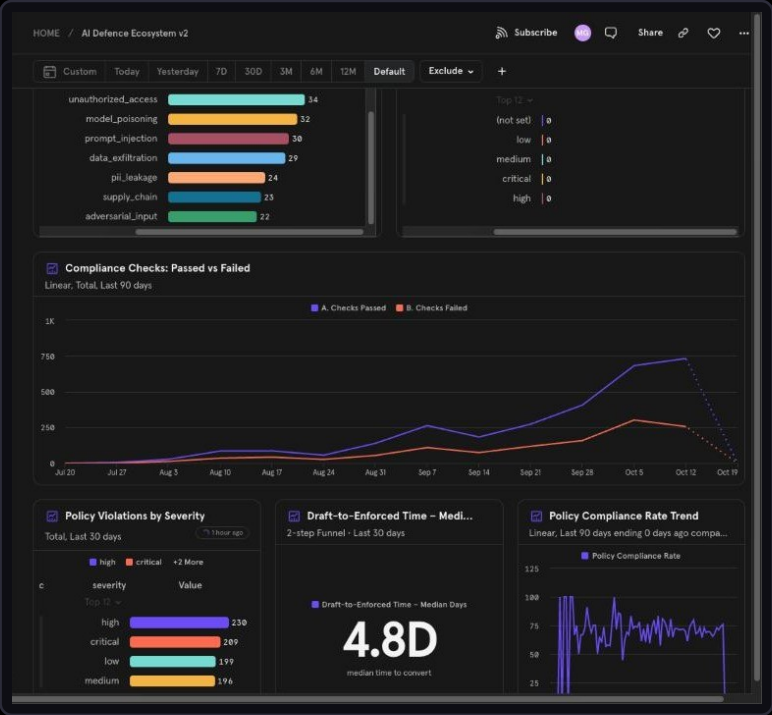
4

Track user sentiment

NPS by tier and weekly retention show whether engineers rely on the tool



Thresholds that require action

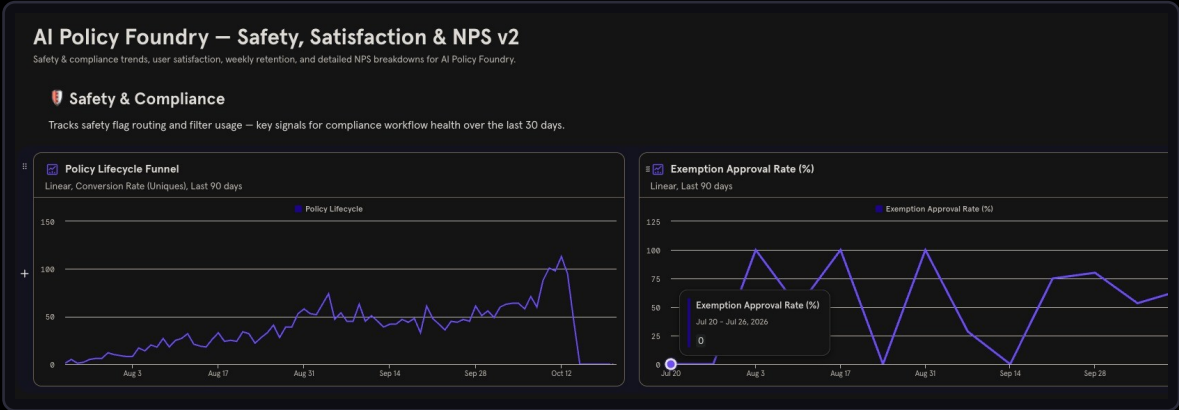


Action Red on cycle time: break down draft→enforced by step to find the slow approval gate, then add SLA reminders to reviewers

Metric	Green	Amber	Red	Now (synthetic)	Status
Draft → enforced (median)	≤ 1 day	1–3 days	> 3 days	4.8 days	Act now
Compliance pass rate (each framework)	≥ 90%	75–90%	< 75%	70.8–73.1%	Act now
Exemption approval rate	≤ 30%	30–50%	> 50%	0–100%, volatile	Watch
Detractor share of NPS	≤ 30%	30–50%	> 50%	116 of 186 (62%)	Act now
Weekly retention	≥ 20%	10–20%	< 10%	~6% → 0%	Act now

Thresholds are proposed starting points; validate against real baselines before use.

Value the metrics provide



Product (PM)

Turns the 70%-faster policy claim into a measured cycle-time trend

Cloud security engineering

Shows which frameworks fail most often so policy templates can be fixed

Leadership / audit

Pass rates and exemption trends give evidence for audits and risk committees

Data gap: NPS breakdowns currently show Ellyra features — add a product property to separate them

Ellyra Pulse

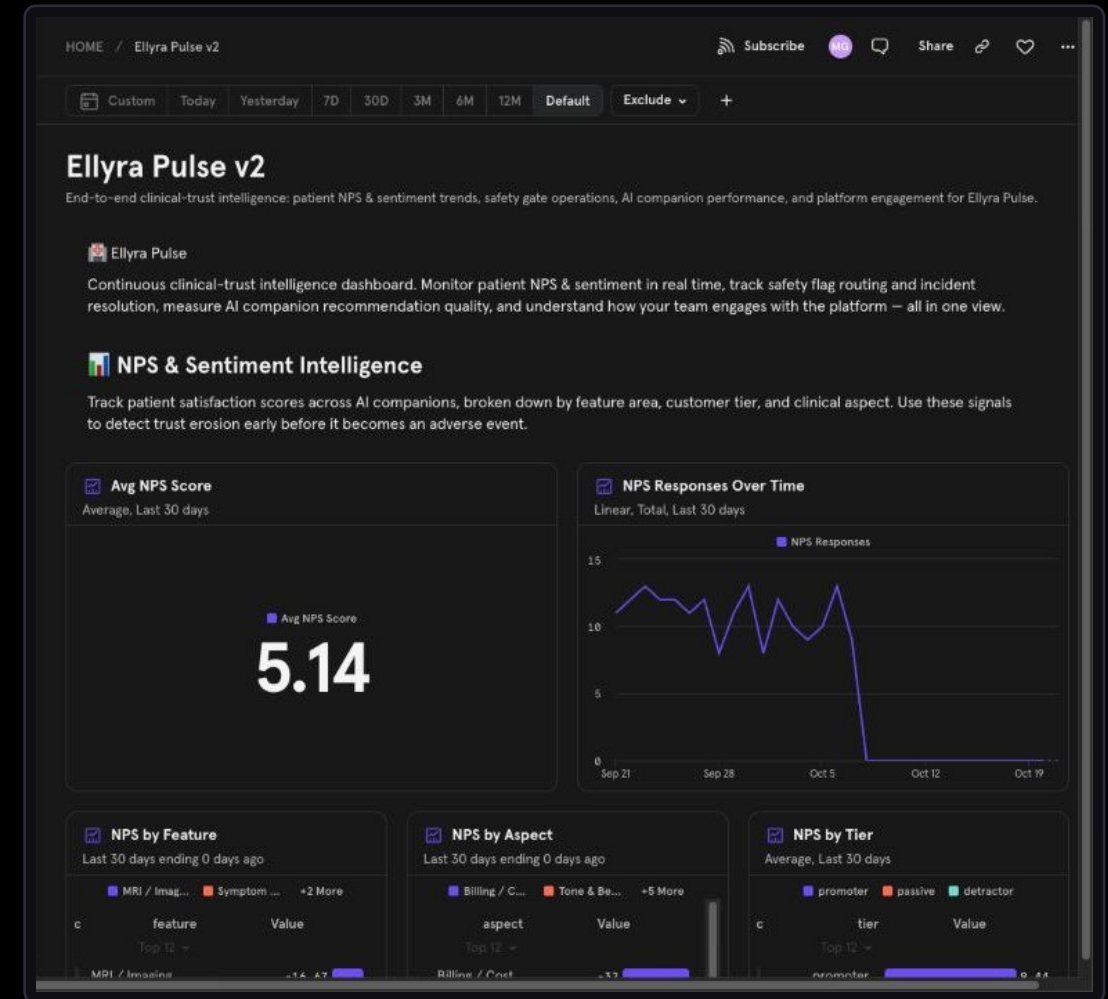
NPS and sentiment intelligence for a medical AI companion,
with safety-gated feedback routing

NORTH STAR

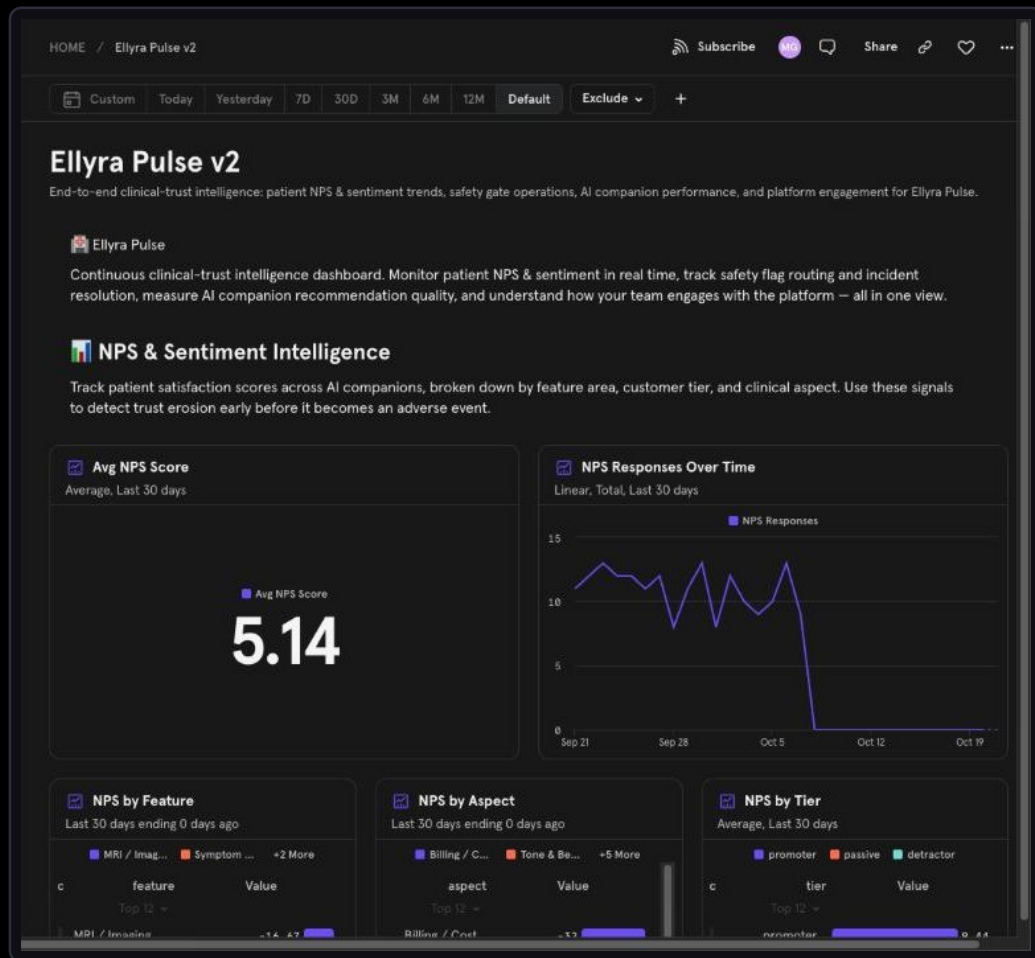
**NPS (% promoters – % detractors) with zero unrouted
safety signals**

Dashboard: Ellyra Pulse v2

Source: github.com/mayukhg/ellyra-pulse



How the metrics were decided



NPS and sentiment section: average score and response volume

1 The problem

A dip in NPS can mean a UI annoyance or a clinical-trust failure, and today they look the same (repo README)

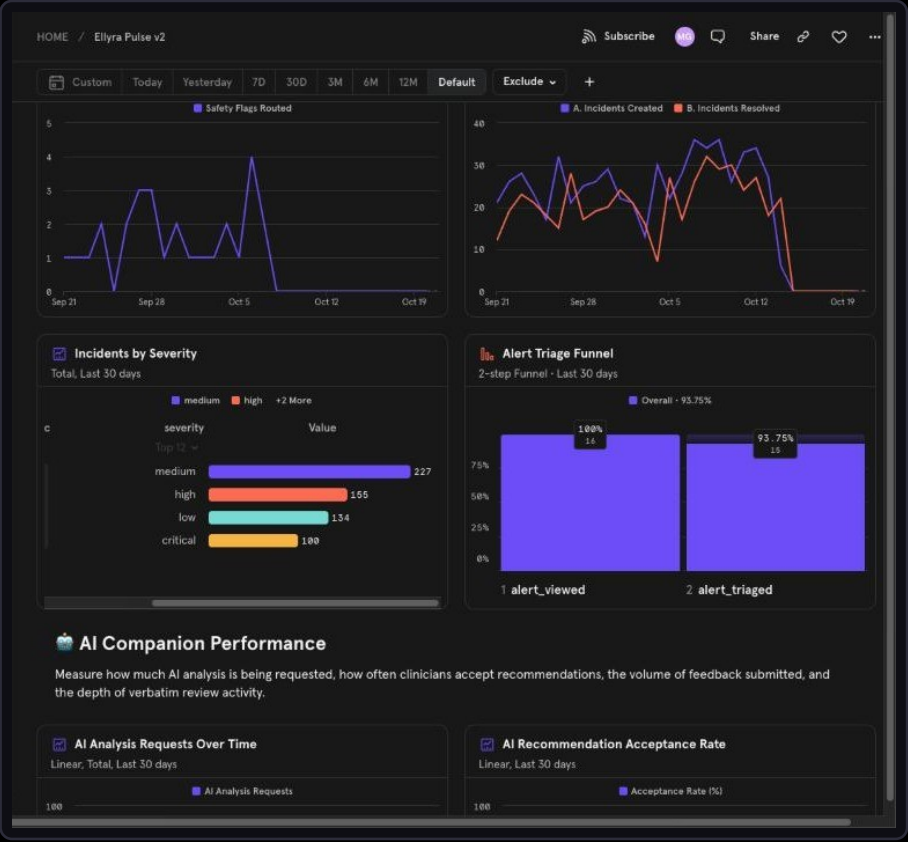
2 The workflow

Patient uses an AI companion → feedback captured → safety gate and PHI redaction → routing → clinical review and closed loop

3 The metrics that follow

Separate volume from severity: NPS by feature, aspect and tier, safety-flag routing, incident lifecycle and AI acceptance

Mapping the app to its metrics



Dashboard cards that carry these metrics

Workflow step	Event tracked	Metric on the board
Patient gives feedback	nps_response_recorded (feature, aspect)	NPS, NPS by feature and aspect
Safety gate flags risk	safety_flag_routed	Safety flags routed over time
Clinical incident	incident_created / resolved	Created vs resolved, by severity
Clinician triages alert	alert_viewed / alert_triaged	Alert triage funnel (per alert)
AI companion advice	ai_analysis_requested / recommendation_accepted	Acceptance rate, feedback
Team onboarding	Sign Up, Onboarding Completed	Activation funnel

Objective of the metrics

1 Measure trust correctly

1

True NPS (promoters – detractors) replaces the average score

2 Route safety fast

2

Safety flags and per-alert triage time (2.9 min median) track clinical response

3 Activate clinical teams

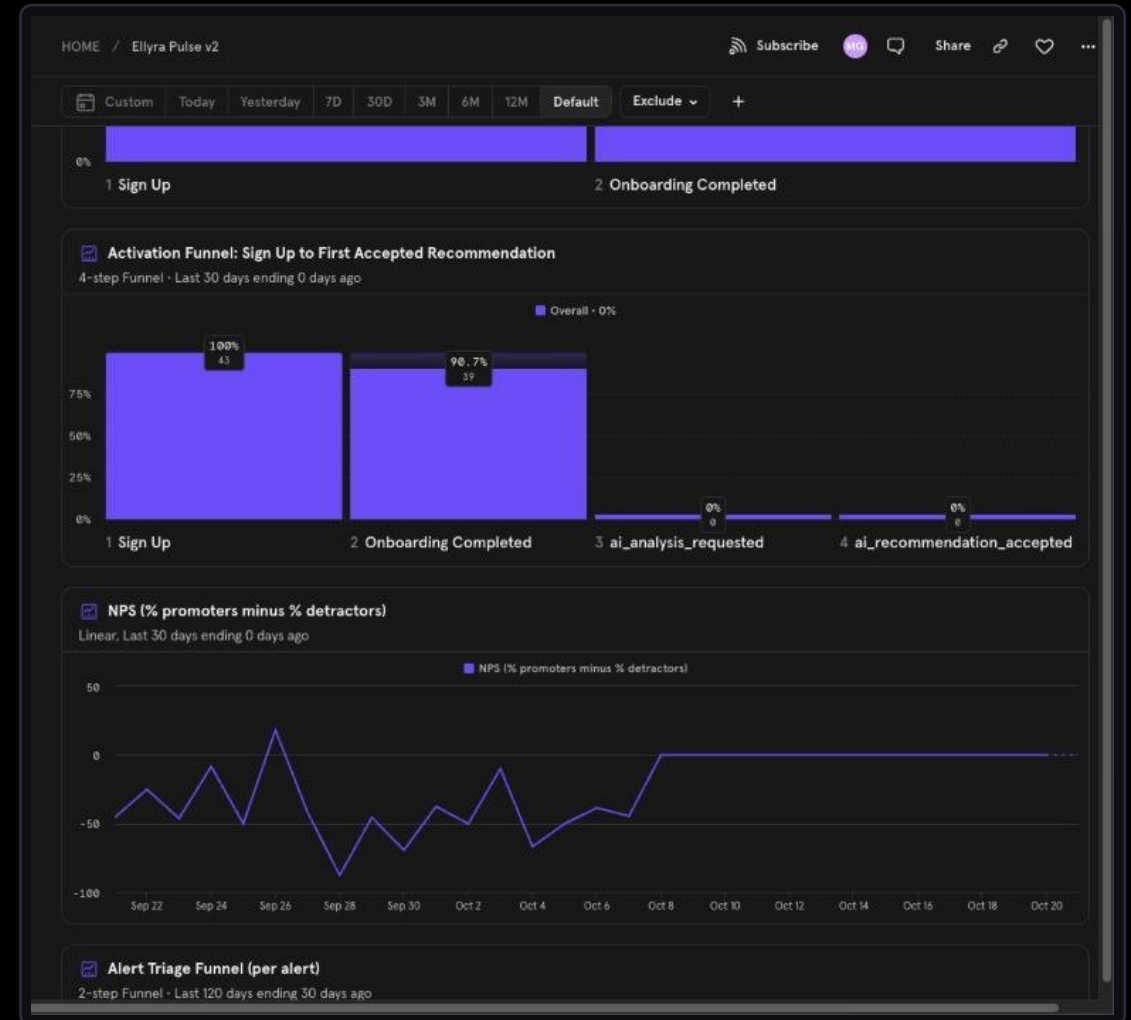
3

Activation funnel shows 43 sign-ups → 39 onboarded → 0 first analyses

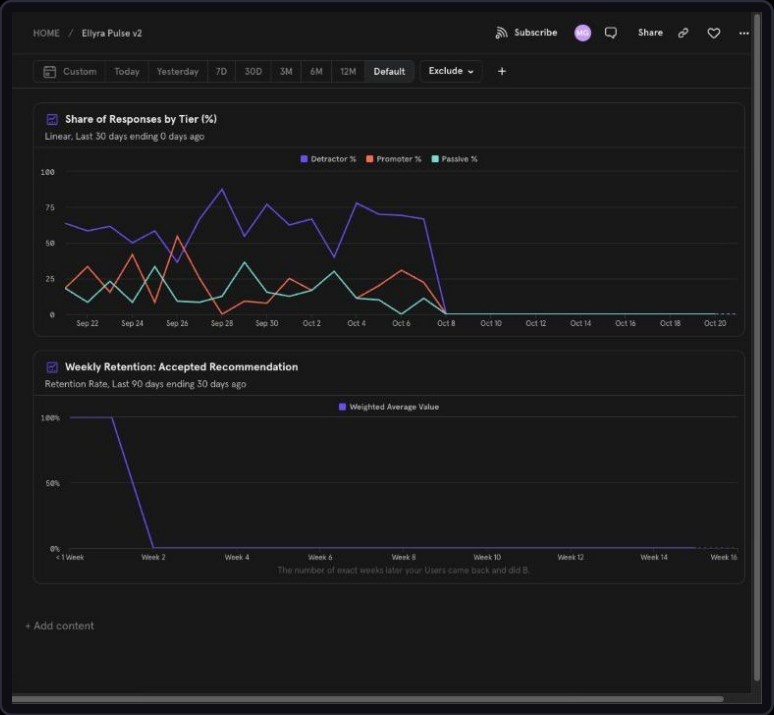
4 Find the weakest feature

4

NPS by feature and aspect show where trust erodes first



Thresholds that require action

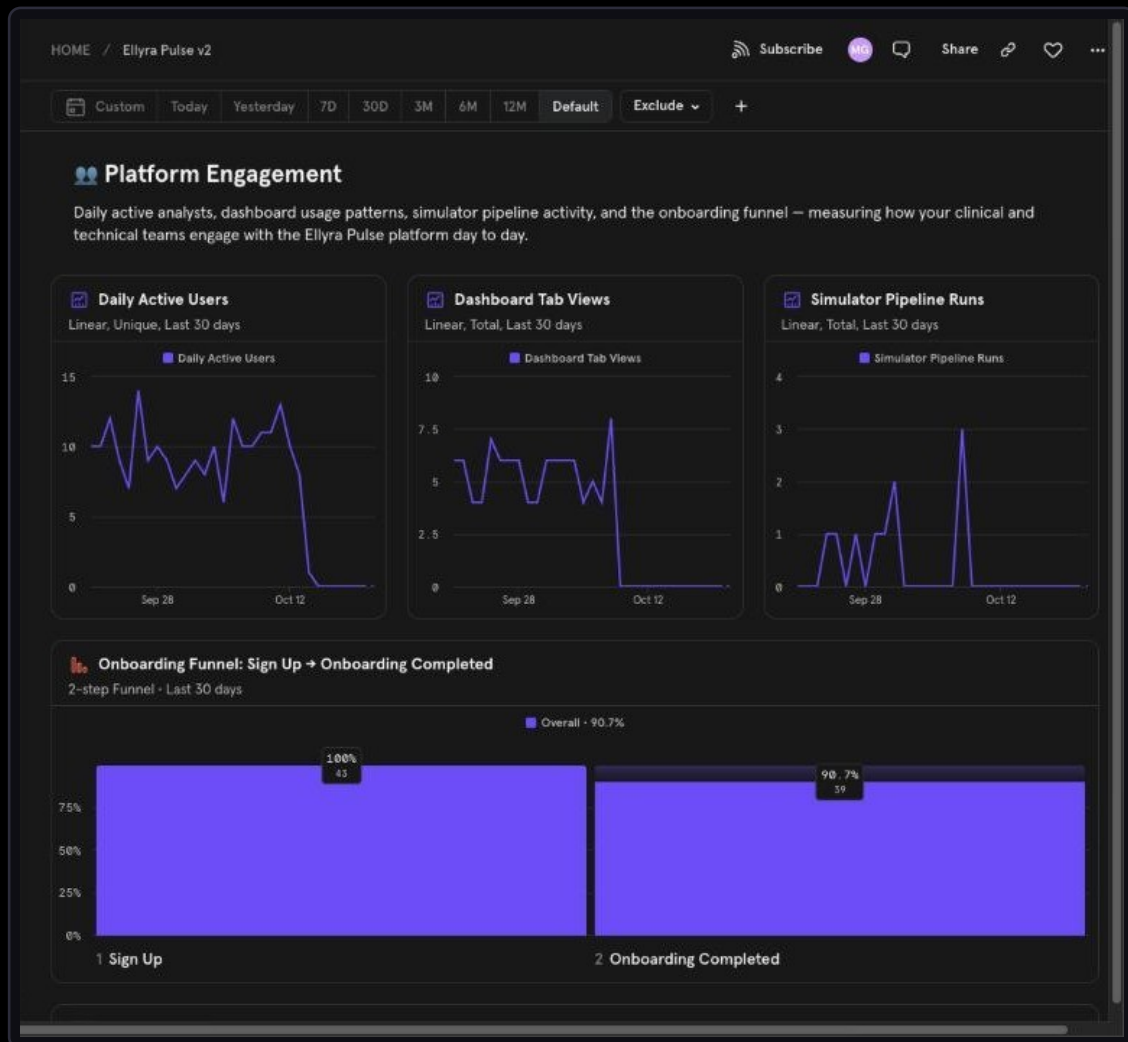


Action Any safety flag spike above 2× the 7-day average pages clinical on-call; red NPS triggers a feature-level root-cause review

Metric	Green	Amber	Red	Now (synthetic)	Status
NPS (promoters – detractors)	≥ 0	-20 to 0	< -20	Avg -22, dips to -85	Act now
Detractor share	$\leq 30\%$	30-50%	$> 50\%$	~60-87%	Act now
Per-alert triage median	≤ 5 min	5-15 min	> 15 min	2.9 min	On track
Onboarded → first AI analysis	$\geq 50\%$	25-50%	$< 25\%$	0%	Act now
Hallucination flag rate (repo)	< 0.20	—	≥ 0.20	Not yet on board	Watch
Disclaimer fatigue index (repo)	< 5	≤ 8	> 8	Not yet on board	Watch

Thresholds are proposed starting points; validate against real baselines before use.

Value the metrics provide



Product (PM)

Shows which AI companion (e.g. GP Question Builder, NPS -58) needs work first

Clinical safety team

Safety flags and triage time prove signals reach a human within minutes

Leadership

One trusted view of patient trust instead of three teams reading NPS differently

Agritech Crop Insurance

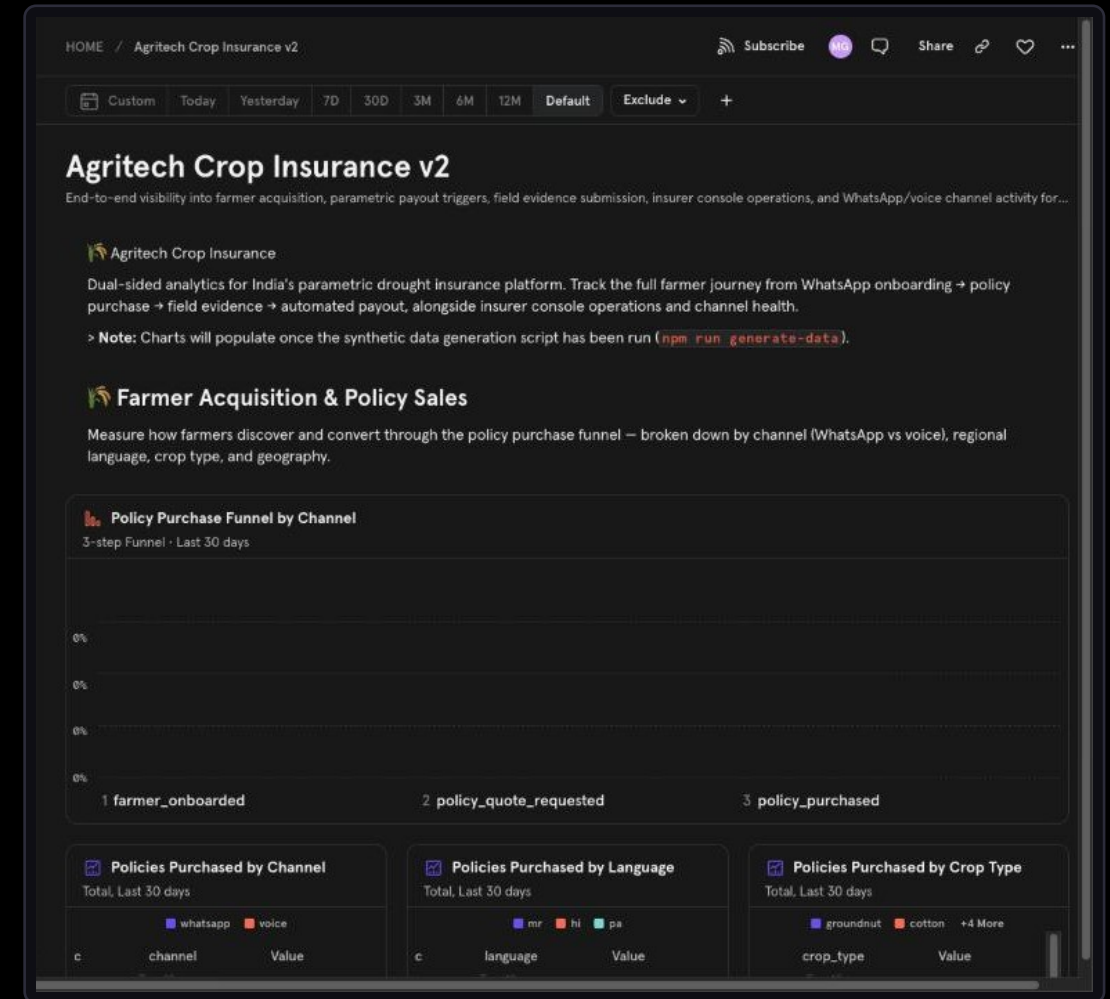
CropGuard: parametric drought insurance for Indian farmers over WhatsApp and voice in regional languages

NORTH STAR

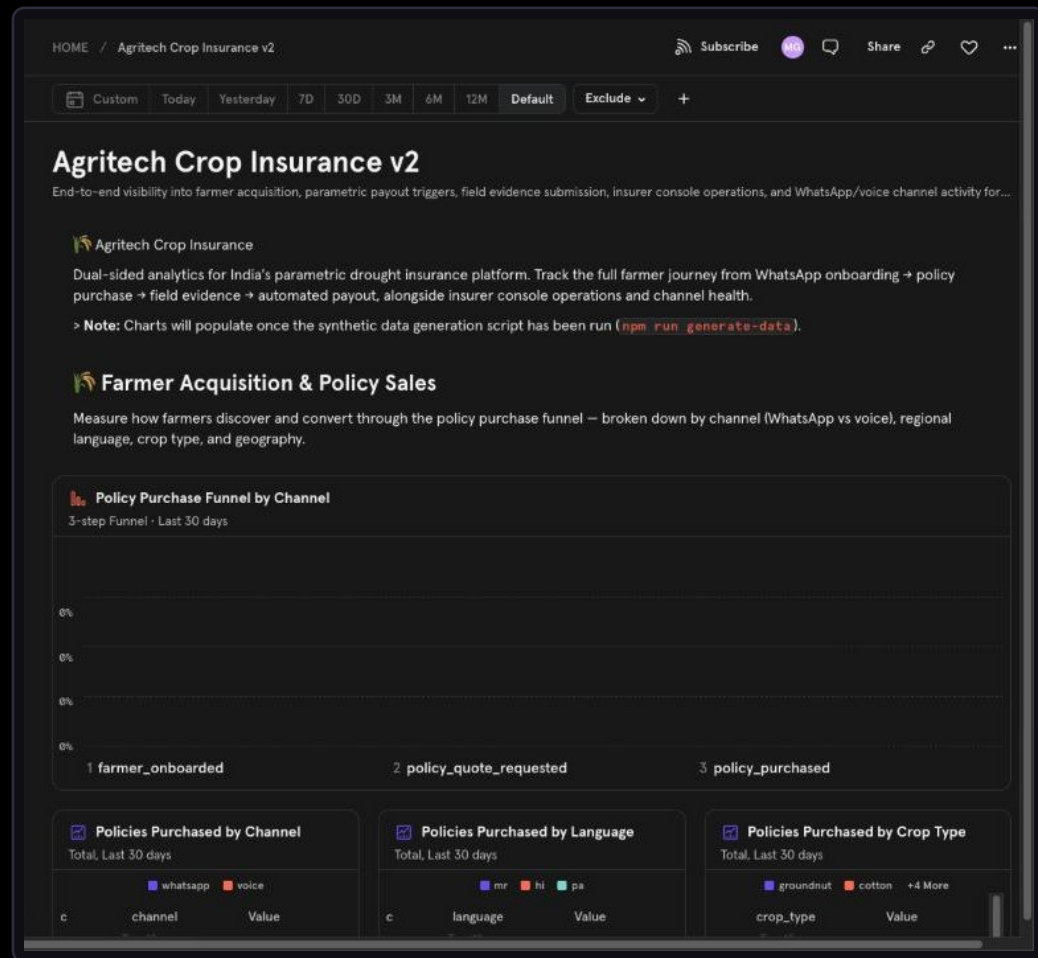
Hours from payout trigger to money in the farmer's hands

Dashboard: **Agritech Crop Insurance v2**

Source: github.com/mayukhg/agritech_crop_insurance



How the metrics were decided



Farmer acquisition: policy purchase funnel by channel

1 The problem

Claims take 90+ days, paperwork and language barriers exclude smallholder farmers (repo executive pitch)

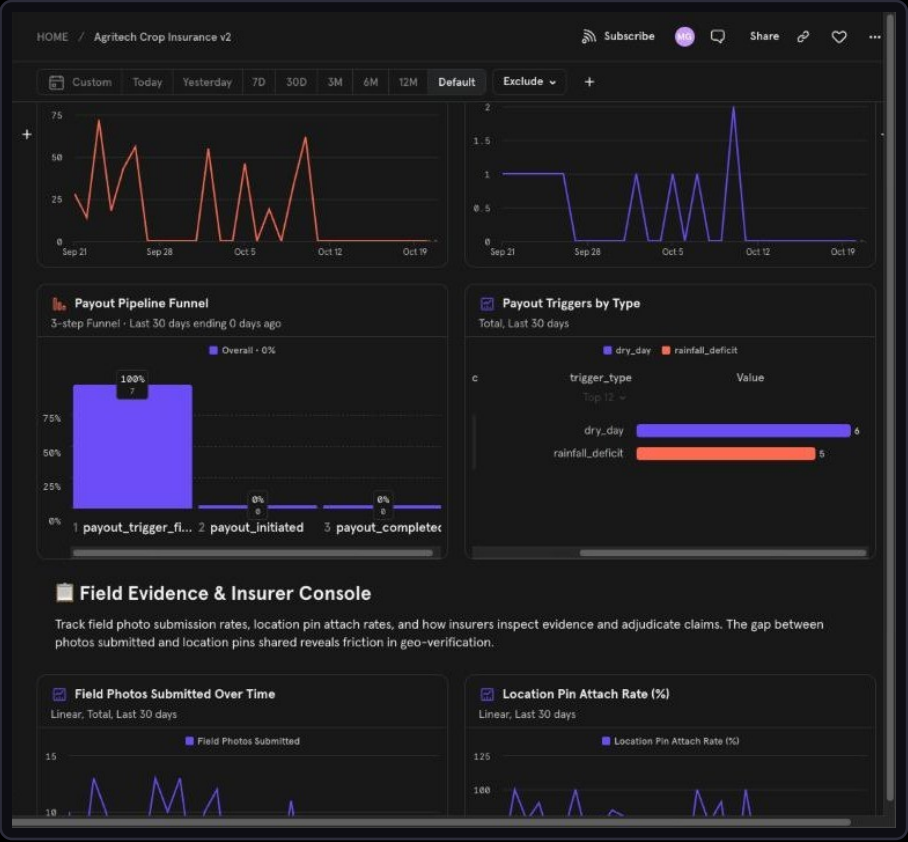
2 The workflow

Farmer onboards on WhatsApp/voice → requests quote → buys policy → sends field photo + location → rainfall trigger fires → payout

3 The metrics that follow

Two-sided view: farmer funnel and payout speed, plus insurer console activity and channel health by language

Mapping the app to its metrics



Dashboard cards that carry these metrics

Workflow step	Event tracked	Metric on the board
Onboard and buy policy	farmer_onboarded → policy_purchased	Purchase funnel; buys by channel, language, crop
Parametric trigger	payout_trigger_fired (trigger_type)	Triggers fired, by type
Payout	payout_initiated / completed	Payout pipeline funnel, time to payout
Field evidence	field_photo_submitted (has_location_pin)	Photos, location pin attach rate
Insurer review	evidence_inspected, claim_adjudicated	Insurer DAU, claim approval rate
Channel activity	whatsapp_message / voice_call	Messages and calls over time

Objective of the metrics

Pay farmers fast

1

Median and P90 time to payout test the 6-hour settlement target (CLAIM_SETTLEMENT_HOURS=6)

Reach farmers in their language

2

Purchases by channel and language show WhatsApp vs voice reach

Verify claims cleanly

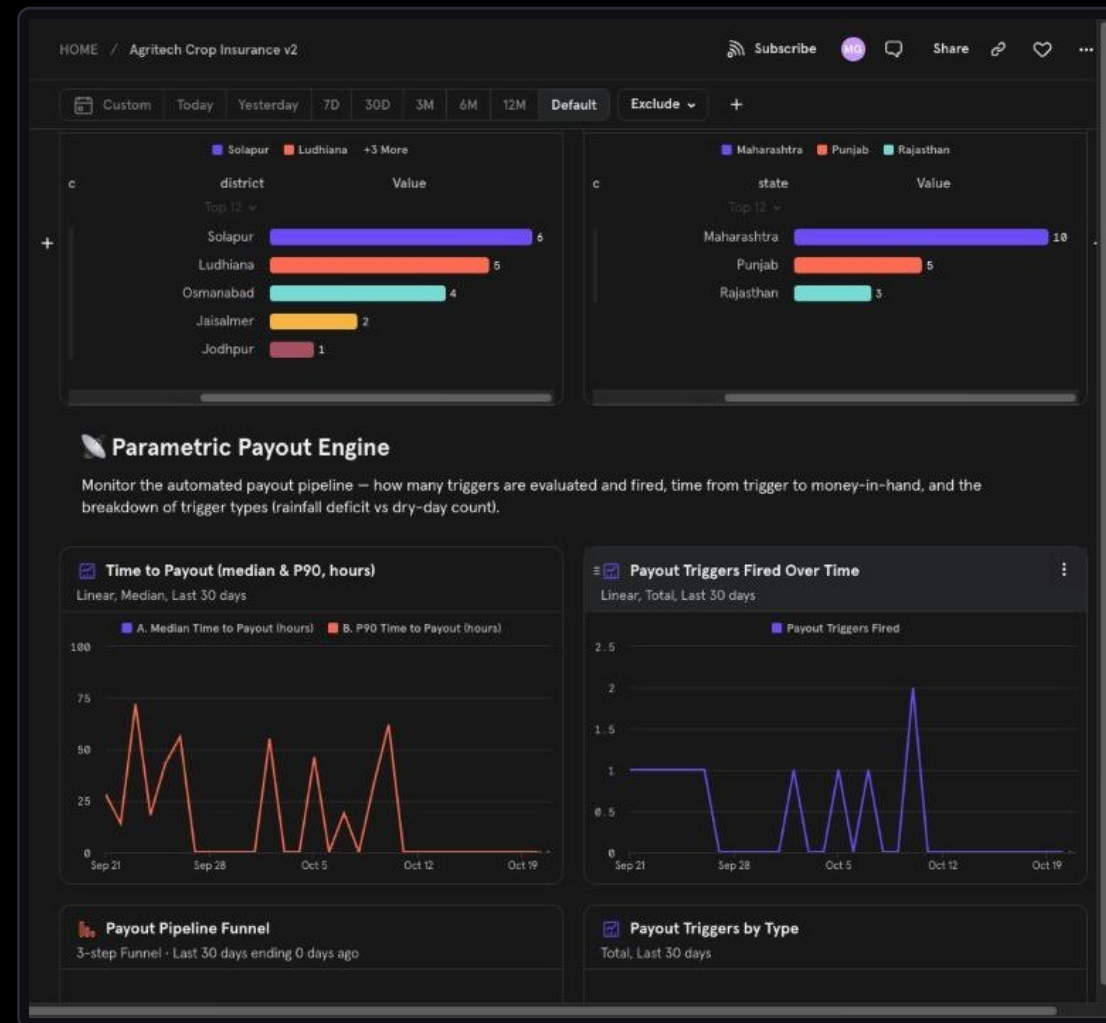
3

Location pin attach rate exposes geo-verification friction

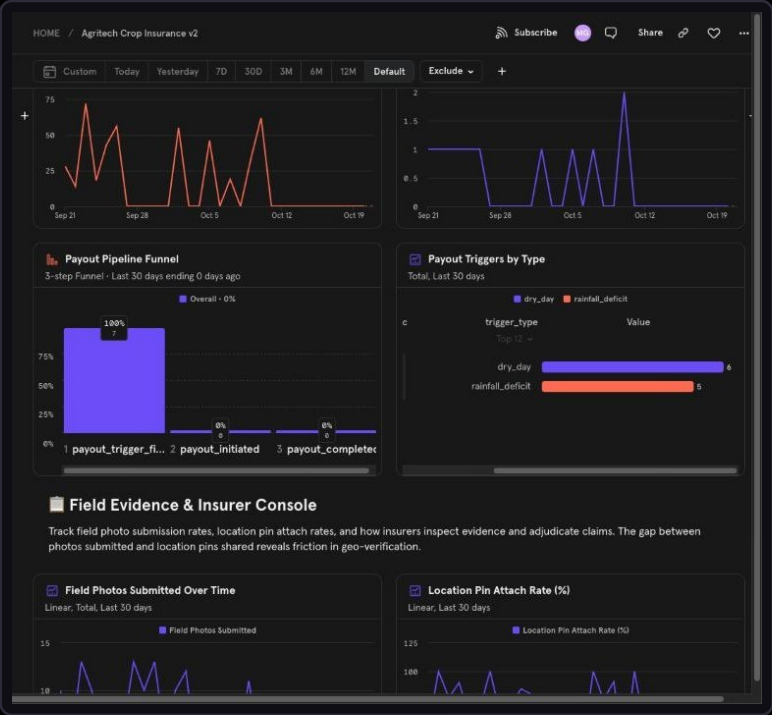
Keep insurers engaged

4

Insurer DAU and claim approval rate track the other side of the market



Thresholds that require action

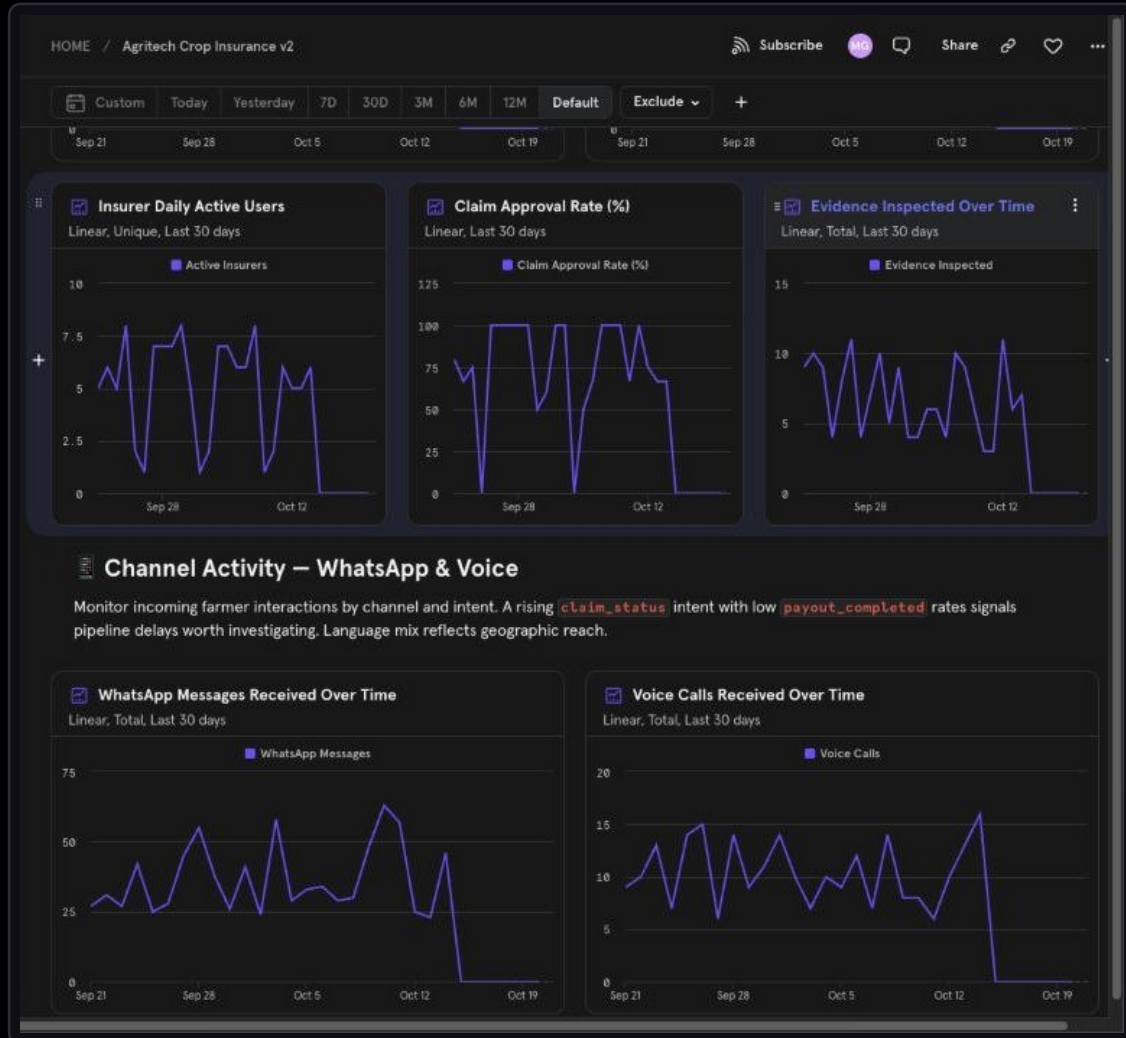


Action Red on payouts: fix farmer identity across steps (hashed phone number) so funnels connect, then alert when any payout exceeds 24h

Metric	Green	Amber	Red	Now (synthetic)	Status
Median time to payout	≤ 6h	6–24h	> 24h	Spikes to ~70h (P90)	Act now
Trigger → payout completed	≥ 95%	80–95%	< 80%	0% (7 triggers)	Act now
Location pin attach rate	≥ 90%	70–90%	< 70%	~25–100%	Watch
Claim approval rate	60–90%	40–60%	< 40% or > 100%	~0–100%	Watch
Purchase funnel conversion	≥ 20%	10–20%	< 10%	0% (identity gap)	Act now

Thresholds are proposed starting points; validate against real baselines before use.

Value the metrics provide



Product (PM)

Shows whether the 24-hour claim promise holds and where the payout pipeline breaks

Insurer operations

Insurer DAU, evidence inspection and approval rates size review workload

Leadership

Channel and language mix guide expansion to new states and crops

AI Defence Ecosystem

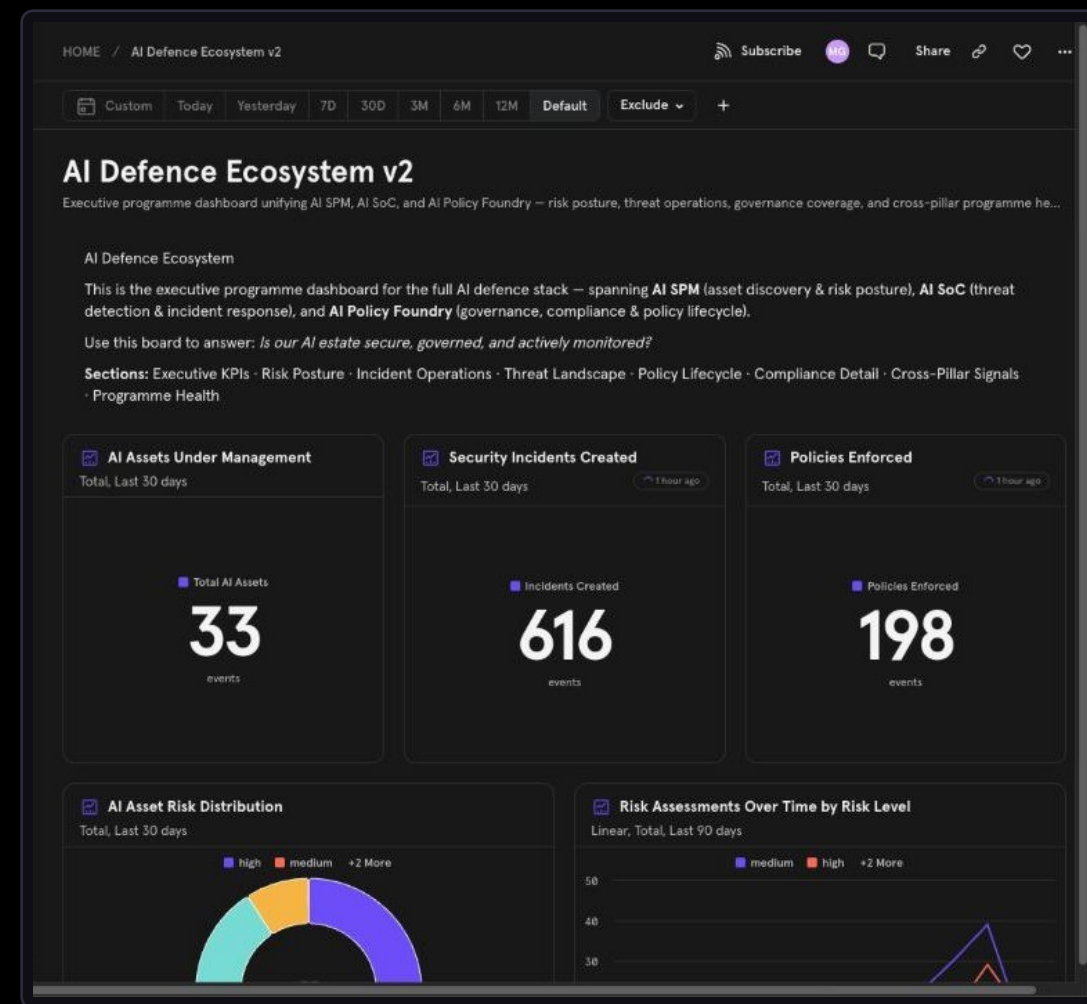
Executive umbrella across AI-SPM, AI SOC and AI Policy Foundry
Foundry: is our AI estate secure, governed and monitored?

NORTH STAR

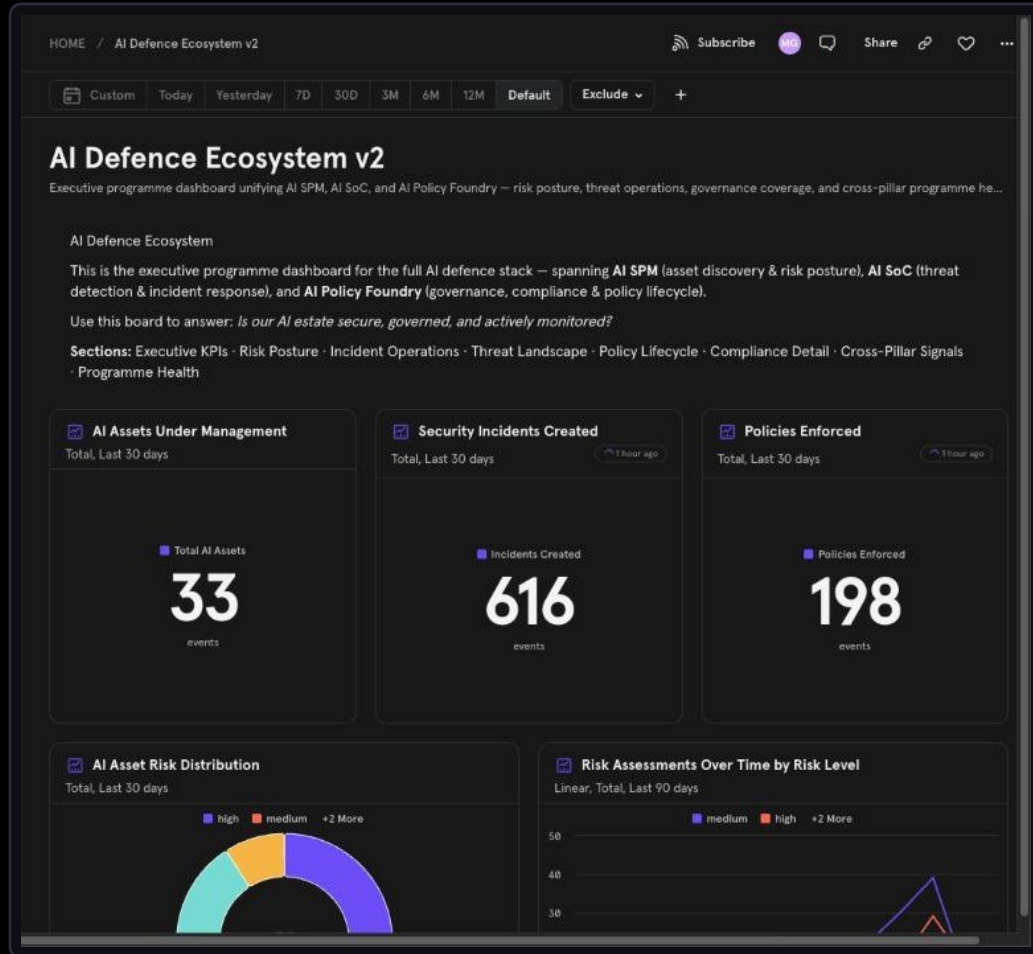
Policy compliance rate and draft-to-enforced time across the AI estate

Dashboard: **AI Defence Ecosystem v2**

Source: Umbrella board over the AI-SPM, AI SOC and AI Policy Foundry repos



How the metrics were decided



Executive KPIs: AI assets under management, incidents created, policies enforced

1 The problem

Leaders need one answer across three products instead of reading three boards

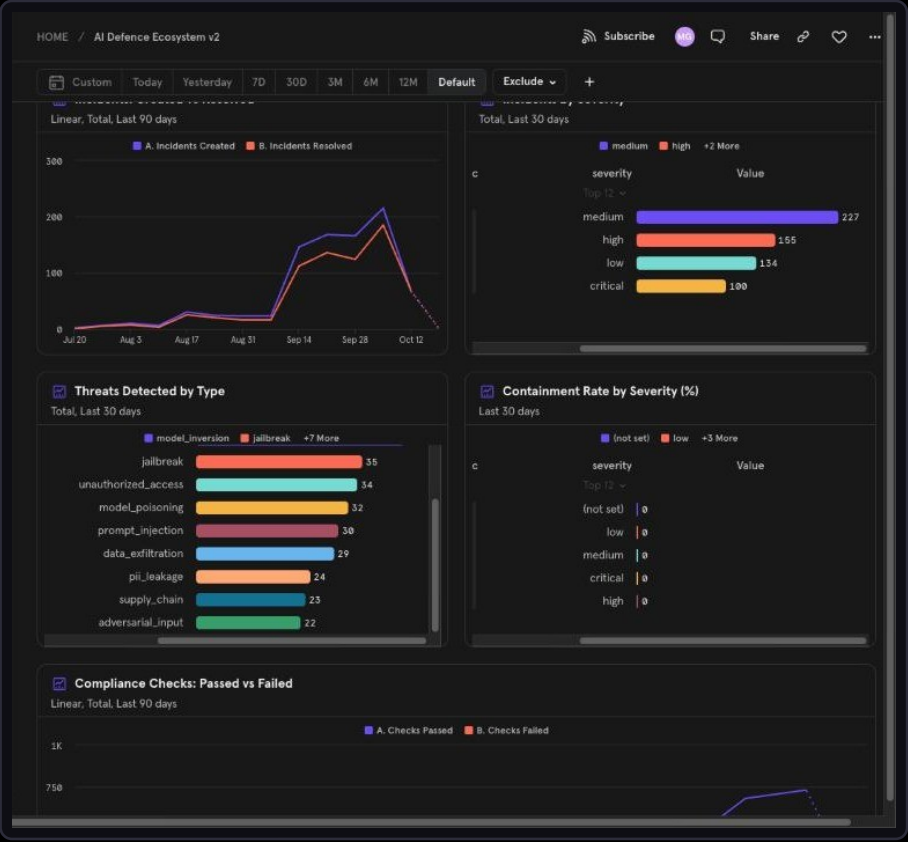
2 The workflow

Discover and assess AI assets (SPM) → detect and respond to threats (SOC) → govern with policies and compliance (Policy Foundry)

3 The metrics that follow

One or two signals per pillar plus cross-pillar adoption, each defined the same way as on the pillar boards

Mapping the app to its metrics



Dashboard cards that carry these metrics

Workflow step	Event tracked	Metric on the board
Risk posture (AI-SPM)	<code>risk_assessment_completed</code>	Assets under management, risk distribution
Incident operations (AI SOC)	<code>incident_created / resolved</code>	Incidents created vs resolved, severity
Threat landscape (AI SOC)	<code>threat_detected (threat_type)</code>	Threats by type, containment rate
Policy lifecycle (Foundry)	<code>policy_drafted → policy_enforced</code>	Draft-to-enforced median time
Compliance (Foundry)	<code>compliance_check_passed / failed</code>	Checks passed vs failed, pass rate
Programme health	<code>weekly_users_per_platform</code>	WAU across all three pillars

Objective of the metrics

Answer one question

1

Is the AI estate secure, governed and actively monitored?

Show governance speed

2

Draft-to-enforced median of 4.8 days is the headline governance KPI

Keep backlogs in check

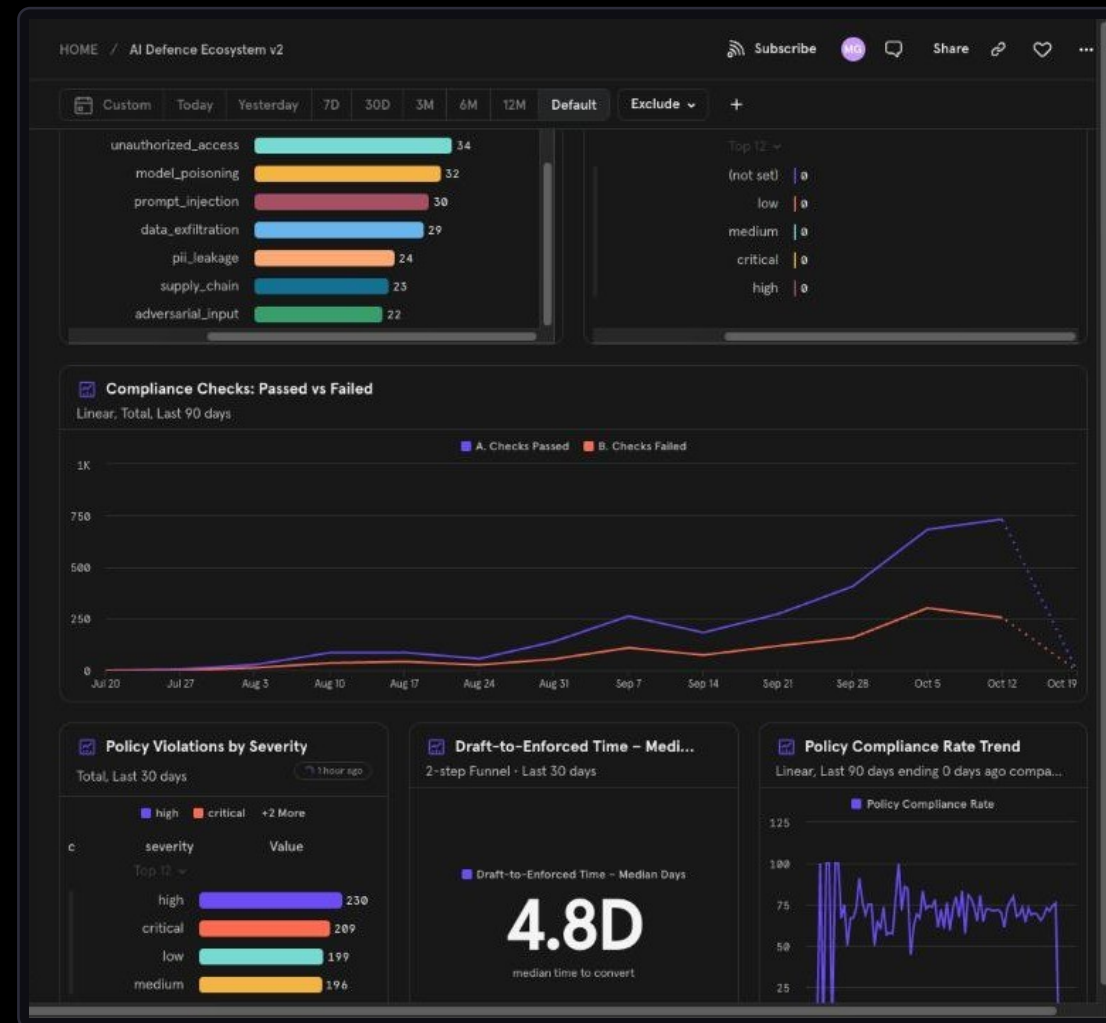
3

Incidents created vs resolved and checks passed vs failed show backlog growth

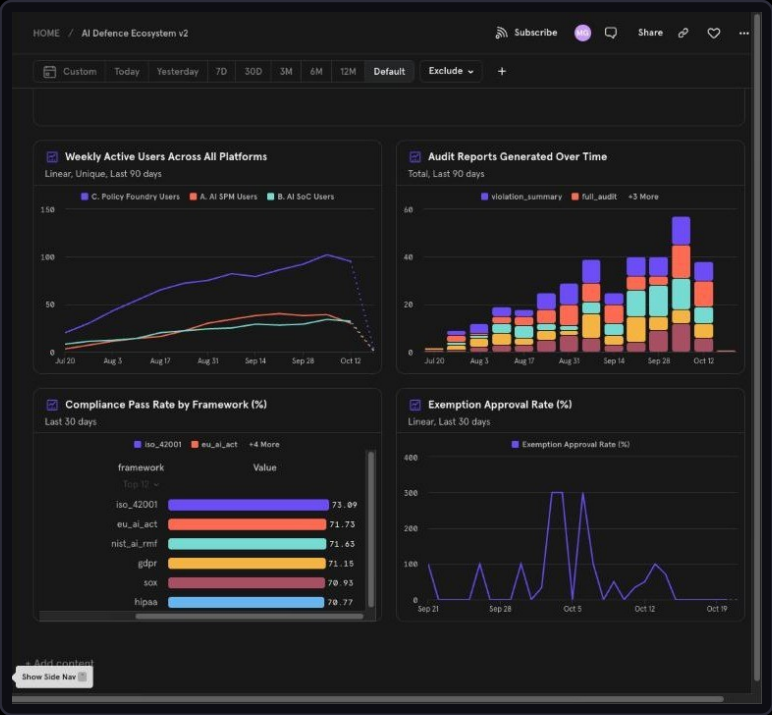
Track platform adoption

4

WAU per pillar shows whether all three products are used together



Thresholds that require action

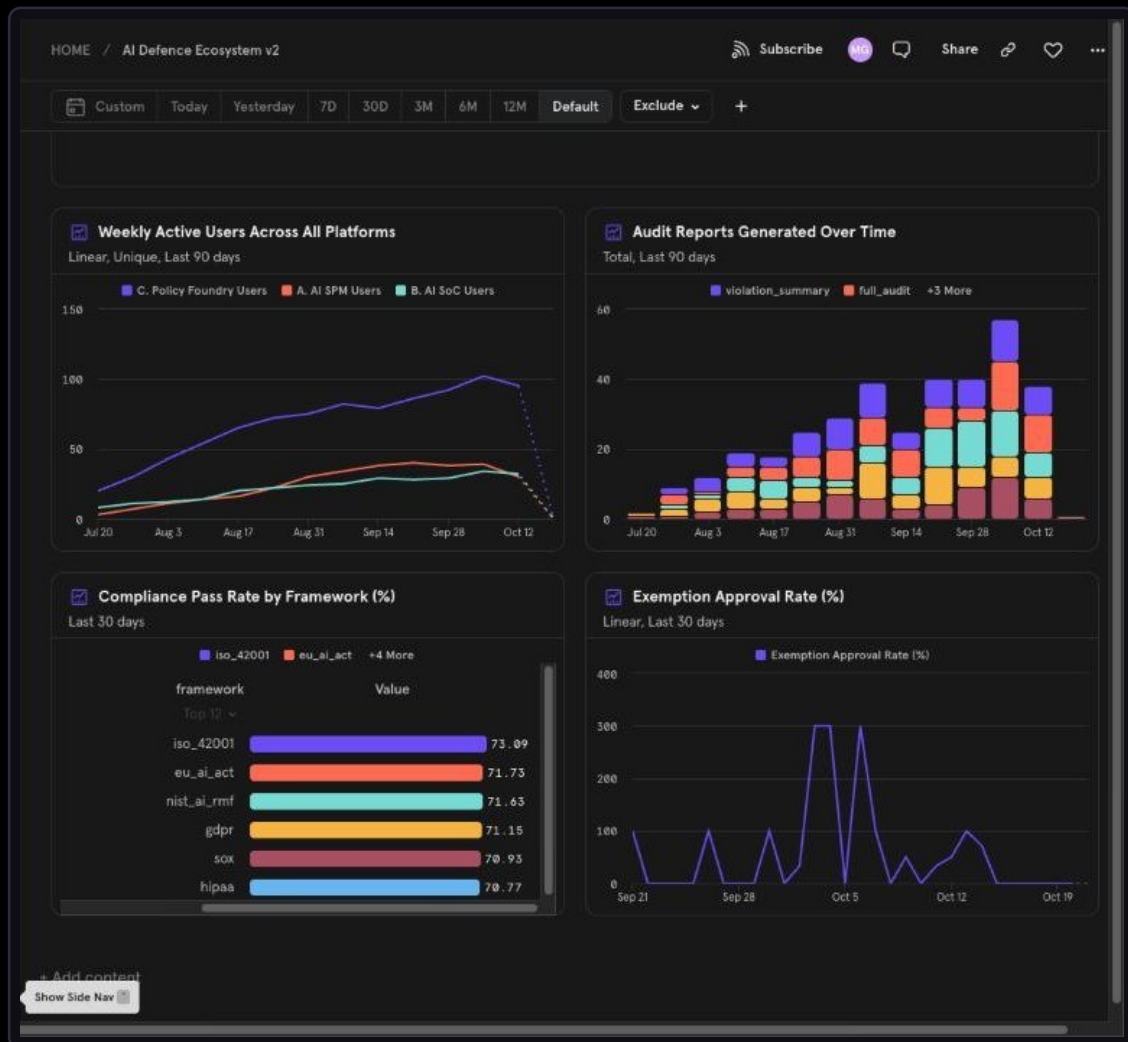


Metric	Green	Amber	Red	Now (synthetic)	Status
Draft → enforced (median)	≤ 1 day	1–3 days	> 3 days	4.8 days	Act now
Compliance checks failed share	≤ 15%	15–30%	> 30%	~27%	Watch
Incidents resolved vs created	≥ 95%	80–95%	< 80%	~80–85%	Watch
Containment rate by severity	≥ 90%	70–90%	< 70%	0% (data gap)	Act now
Exemption approval rate	≤ 30%	30–50%	> 50% or > 100%	Up to 300%	Act now

Action Red rows are partly data-definition issues: fix the containment and exemption formulas before using them in exec reviews

Thresholds are proposed starting points; validate against real baselines before use.

Value the metrics provide



Product (portfolio PM)

One board to prioritise across three products using the same definitions

CISO

Answers 'secure, governed, monitored?' in one view for board and regulator updates

Leadership / GTM

Cross-pillar adoption supports a platform story and bundled pricing

Where each app stands against its thresholds

App	North star	On track	Watch	Act now
Agent Val	Confirmed risks closed per week	1	3	2
AI SOC Portal	Alerts resolved end-to-end (per alert) and time to resolve	1	3	2
AI-SPM	Active accounts running risk assessments every week	0	2	3
AI Policy Foundry	Median time from policy draft to enforcement	0	1	4
Ellyra Pulse	NPS (% promoters – % detractors) with zero unrouted safety signals	1	2	3
Agritech Crop Insurance	Hours from payout trigger to money in the farmer's hands	0	2	3
AI Defence Ecosystem	Policy compliance rate and draft-to-enforced time across the AI estate	0	2	3

Counts of threshold rows by status on each app's threshold slide (synthetic data).

What to fix before these metrics drive decisions

Separate products

Add a product property to every event — Policy Foundry NPS currently shows Ellyra features

Connect identities

Use one ID per farmer / user across steps so Agritech and AI-SPM funnels stop reading 0%

Fix rate formulas

Exemption approval (up to 300%) and containment rate (0%) need corrected denominators

Keep the feed live

Synthetic events stop around mid-October; schedule the generators so trends stay current

Validate thresholds

Replace proposed green / amber / red bands with baselines from real usage

Review on a cadence

Use the boards in the quarterly roadmap review and rank fixes with Cost of Delay

Thank you

My Product Metrics · Mayukh Ghosh

Disclaimer Metric values in this deck are for demo purposes only and have been generated using synthetic data. Dashboards: Mixpanel project 4070033 (v2 boards and Agent Val — Product Metrics).