

Proof-of-Concept Playbook

A 6-week operational pilot of Patrol Sight

HumanityAI'd — July 2026

CONFIDENTIAL

1. Purpose

This playbook lets a security organisation run a low-risk, time-boxed pilot of Patrol Sight at one site, and lets both sides judge success against numbers agreed in advance. It is deliberately small: one site, one shift pattern, a handful of glasses, six weeks. It is two weeks longer than a commercial-product pilot for a reason — a security deployment needs a legal and information-security gate before any real watchlist is loaded.

2. What the pilot proves

1. An officer on foot receives a correct, categorised identification in their lens fast enough to act on during the encounter — measured, in the customer's own building, on the customer's own wireless network.
2. The customer's watchlist team can enrol, categorise and activate people themselves, and see the change reach the fleet within seconds.
3. The system runs entirely inside the customer's network, with zero biometric egress verifiable by their own network audit.
4. The audit trail and governance controls satisfy the customer's data-protection function.
5. Accuracy in real operating conditions — including the false-positive rate, which we measure and report rather than avoid.

3. Scope

In scope	Out of scope (pilot)
1 site, defined patrol area	Multi-site or city-wide rollout
5-10 glasses units	Large fleet logistics
A test watchlist first, then a limited live watchlist after the legal gate	Full operational watchlist
On-prem GPU server (loaned)	Permanent server procurement
Control-console monitoring + alert stream	Integration with the customer's command-and-control (scoped separately)
Accuracy, latency and false-positive measurement	Third-party/national watchlist connectors
DPIA support and audit-log review	Enforcement policy and legal authority (customer's own)

4. Hardware & environment checklist

- [] **AR glasses:** 5–10 × RayNeo X3 Pro class (HumanityAI'd supplies for the pilot)
- [] **AI server:** 1 × GPU server. Minimum 8 GB GPU (the measured reference is an RTX 4060 Ti); **16–24 GB recommended** for fleet headroom. 32 GB RAM, 8+ cores, 512 GB SSD, NVIDIA driver 550+
- [] **Wireless: dedicated 5 GHz SSID** across the patrol area with power-save disabled, ≥ 2 access points, wired uplink to the server, server on a static LAN IP. Budget ~20 Mbps per streaming device. (*The reference glasses are 5 GHz only — no 6 GHz.*)
- [] **Network isolation:** glasses VLAN isolated with no internet uplink — a security control for the pilot, not a preference
- [] **Charging / provisioning:** charging station and a USB provisioning point at the equipment desk
- [] **Console access:** operator workstations on the staff VLAN
- [] **Power & rack space** for the server in the comms room

5. Timeline

Week	Activities	Owner
0 — Prep & governance	Kick-off; success metrics signed; patrol area and shift pattern chosen; RF site survey ; server delivered and installed; information-security review and DPIA started ; watchlist categories and colour scheme agreed	Both
1 — Stand-up on test data	Deploy stack; provision glasses; enrol a consented test cohort (volunteer staff); tune detection and recognition thresholds against measured false-detection rates; verify zero egress by network audit	HumanityAI'd + customer IT
2 — Legal gate & live watchlist	Information-security sign-off; DPIA conclusions applied (retention, access, audit); load a limited live watchlist with documented lawful basis per entry; operator console training	Customer legal/ security + HumanityAI'd
3 — Officer induction & shadow running	1-hour officer induction; glasses worn on shift with alerts to the control room only (no officer action) to establish the baseline false-positive rate; comfort feedback	Both
4–5 — Live operation	Officers act on verified matches per the customer's own procedure; daily alert review; weekly accuracy and false-positive report; threshold refinement	Customer operations
6 — Evaluate	Measure against the agreed KPIs; officer and operator survey; audit-log review with the data-protection function; readout workshop; go/no-go and rollout proposal	Both

6. Enrolment procedure (per person, ~2 minutes)

1. Obtain and record the **lawful basis** for the entry (this step is the customer's, and it comes first).
2. In the console, create the person record with name, reference number and provenance.
3. Upload one or more face images — multiple angles improve off-pose performance.
4. Assign the **category and alert colour** (e.g. wanted, banned, missing person, VIP, staff).
5. The index rebuilds and hot-reloads into the live engine — no restart, no retraining. The entry is active on every deployed device within seconds.

Reference point: bulk enrolment runs at **40–60 images per second** — about 15–25 seconds for 1,000 images.

7. Success metrics (agree & sign at kick-off)

KPI	Target	How measured
Glasses-to-label latency	≤ 150 ms p95	System metrics (reference: ~68 ms typical)

KPI	Target	How measured
Effective recognition rate per device	≥ 12 fps	System metrics (reference: 14–20 fps)
True-positive rate on the enrolled cohort at ≤ 5 m	$\geq 95\%$	Structured walk-through test, repeated per lighting condition
False-positive rate per officer-hour	Agreed ceiling, measured and reported	Shadow-running week + ongoing alert review
Recognition at range	Reliable ≤ 5 m; 6–7 m in good light	Marked-distance test — we publish this envelope rather than overclaim
Watchlist entry to live on fleet	≤ 2 min, unaided	Timed during Week 2
Biometric data egress	Zero	Customer's own network audit
Audit completeness	100% of searches and matches attributable	Audit-log export reviewed by data-protection function
Officer comfort over a full shift	$\geq 4/5$	Officer survey
Unattended recovery after reboot	Full stack live without intervention	Deliberate reboot test

8. Roles

- **Customer:** patrol area and shift selection, wireless network and power, information-security review and DPIA, lawful basis and watchlist content, officer and operator staffing, enforcement procedure.
- **HumanityAI'd:** server and glasses for the pilot, installation, threshold calibration, training, measurement and reporting, evaluation readout.

9. Risks & mitigations

Risk	Mitigation
Wireless quality dominates latency	RF survey in Week 0; dedicated 5 GHz SSID, power-save disabled; add APs where the survey shows gaps
Legal/DPIA review delays the live watchlist	Weeks 0–2 run on a consented test cohort so technical validation is never blocked by the legal gate
False positives erode officer trust	A full shadow-running week before any officer action; measured ceiling; per-site threshold tuning; category-colour display so officers treat labels as prompts, not verdicts
Officer discomfort or fatigue	Comfort survey; rotation; spare units; category-only display mode
Glasses-transport encryption not yet shipped	Isolated glasses VLAN with no uplink as a hard control for the pilot
Model-licensing item unresolved	Pilot proceeds under an evaluation arrangement; resolution is a precondition of the production licence and is stated in the agreement
Concurrent-device capacity unknown	Pilot fleet is deliberately small; we measure the ceiling here and contract on the measured figure

10. Exit & conversion

At the Week-6 readout: the KPI scorecard including the measured false-positive rate, the officer and operator survey, the audit-log review, and a rollout proposal (site-by-site plan, fleet sizing, GPU sizing against the measured device ceiling, licence and support quote). Watchlist content, thresholds and audit history carry forward into production — no rework.