



STAKEHOLDER & EXECUTIVE OVERVIEW · SEPTEMBER 2026

Melaninternet.Volume_1

The evidence layer for AI you can trust. Melaninternet combines a working mobile-first MVP — source-grounded document intelligence, policy controls, privacy-aware sessions, telemetry, and automated reliability evidence — with a research program to make AI accuracy guarantees survive model changes, and a staged roadmap into governed agents running in secure, sandboxed runtimes, personal AI, health innovation, world models, and Physical AI.

Current product

Live NVIDIA Nemotron inference, document intelligence with citations, anonymous guest sessions, customer-interest capture, PostgreSQL telemetry, and automated release validation.

Core differentiation

Value is not access to a model. It is making AI outputs — and later AI actions — measurable, source-grounded, verifiable across model changes, governable, and auditable.

Strategic expansion

Governed agents with observability and security signals, an evidence-first personal assistant, health research and operations, world models, and long-term Physical AI R&D.

Current stage boundary

Not represented as a certified security product, regulated clinical system, contractual SLA, or validated revenue-generating business.

Mission lens: expand meaningful AI participation and founder participation, reduce harmful AI bias, and translate responsible-AI principles into practical technical controls.



Purpose of Melaninternet

Practical AI participation, founder opportunity, bias-aware development, and responsible deployment — the first two pillars grounded in published evidence.

PURPOSE PILLAR	EVIDENCE / PROBLEM SIGNAL	MELANINTERNET RESPONSE
Expand meaningful AI participation	Pew Research Center has found Black and Hispanic workers are more concentrated in occupations with lower exposure to AI.	Encourage advanced AI use for work, learning, entrepreneurship, and daily life — including an evidence-first personal assistant — with emphasis on ownership, not access alone.
Expand founder participation	Crunchbase reported companies with a Black founder or co-founder received about 0.32% of U.S. venture funding in 2025.	Pursue pilots, federal R&D funding (NSF, NIH), accelerators, partnerships, and scalable commercialization.
Reduce harmful AI bias	AI systems can behave differently across groups, and failure modes stay hidden without structured evaluation.	Evaluation, documentation, source grounding, and visible failure modes so behavior can be measured and improved.
Promote practical AI ethics	Responsible deployment needs operational controls and honest boundaries — more so as AI takes actions.	Truthfulness, privacy, human oversight, grounding, security boundaries, approval gates, and clear limits.

Source context: Pew Research Center workplace AI-exposure research; Crunchbase News 2025 U.S. venture-funding data.



What is already working

Current capability, followed by the recorded evidence behind it.

Live AI

Hosted MVP at melaninternet.ai on NVIDIA Nemotron 3 Super 120B with automatic failover to Ultra.

Document intelligence

PDF/TXT/Markdown analysis with source-grounded answers and page/section citations.

Traction capture

Early Access, Business Pilot, and Design Partner calls to action with consent-based capture.

Operational telemetry

Privacy-safe PostgreSQL metrics persist across restarts.

Automated production gate

Release, regression, live model, synthetic user, concurrency, document, isolation, and latency checks.

Privacy boundaries

Guest telemetry and intentional contact information stay logically separated.

EVIDENCE ITEM	RECORDED RESULT
Production reliability checks	46 of 46 passed; 0 failed ; 0 tolerated warnings
Synthetic scenarios	25 of 25 passed (100% pass rate)
Concurrency	5 of 5 isolated simultaneous sessions passed
Latency	P50 about 2.8s; P95 about 11.4s in the referenced run
Document proof	Grounding and citation PASS; cross-session isolation PASS
Regression corpus	200 protected cases, run every release

All figures are from a single persisted automated gate report: release v1.9.7, measured against live production on 31 August 2026. They are re-verified on every release and the release fails if they disagree. A tolerated warning is an upstream provider timeout, not a product defect; this run needed none, and the count is reported either way.



The platform, staged honestly

Every product line runs on the same grounding, verification, governance, and evidence core.

LIVE NOW

Evidence Layer v1

- Grounded document intelligence with citations
- Refusal when the source lacks the answer
- Versioned policy and privacy boundaries
- Automated production release gate

R&D 2026-27

Durable guarantees

- Error rates designed to hold after model swaps
- Distractor-aware claim verification
- Answer / qualify / abstain decisions
- NSF SBIR resubmission + NIH SBIR path

NEXT 2027

Governed agents + assurance

- Agents run in sandboxed runtimes (NVIDIA OpenShell-aligned)
- Agent observability and behavioral monitoring
- Verification of agent decisions before actions execute
- Integrations with AI-security platforms

LATER / R&D

Personal AI · Health · World models · Physical AI

- Evidence-cited personal assistant
- Health research and admin intelligence
- Enterprise and health world models
- Assurance for robot plans and actions

Sequencing is the strategy. Each stage reuses the core built before it, and nothing is marketed as a capability until it has been tested.



Research program: durable guarantees

The R&D at the center of the NSF SBIR resubmission and the NIH SBIR path.

QUESTION	ANSWER
What exists	Claim checkers score whether statements match a source (MiniCheck, Patronus Lynx, Galileo Luna, Vectara HHEM, Azure and AWS grounding checks). Conformal methods attach a statistical accuracy promise — if conditions stay the same.
The documented gap	A March 2026 study (arXiv 2603.16817) found those promises break when the generating model changes and when look-alike claims appear; the obvious fix leaves most answers empty. Numerical content failed fastest.
What Melaninternet will build	(1) Calibration that re-establishes a claim-level guarantee after a model swap from a small labeled sample. (2) A lightweight verifier with answer / qualify / abstain decisions for near-miss figures and outdated versions. (3) An adversarial enterprise test bed across several model families.
Why it matters	Every model upgrade today forces re-validation. A guarantee that survives the swap lets AI become a system of record in finance, insurance, and health administration.
Status	NSF Project Pitch resubmission drafted (after the Nov 4, 2026 window); NIH Specific Aims and Research Strategy next; NJSBDC SBIR coaching in progress; technical lead recruitment underway.



PLANNED · NEXT 2027

Governed agents and agent assurance

Planned direction: moving from answers to approved actions — with observability and security built in.

BUILDING BLOCK	WHY IT MATTERS
Secure accounts / Workspace Mode	Persistent work, pilot organizations, roles, and saved preferences.
Governed agent framework	Approved actions that inherit constitution, permissions, verification, and human approval.
Sandboxed agent runtime	Agents execute inside a policy-enforced sandbox (built to run on NVIDIA OpenShell, open source) rather than directly on hosts, networks, or credentials.
Agent observability	Every action, tool call, and data access recorded as an evidence record; behavioral baselines and drift dashboards.
Output verification for agents	Check what an agent claims or decides against source evidence before it is sent or executed.
Containment	Least-privilege permissions, anomaly flags, automatic pause on out-of-policy behavior, injection and exfiltration test suites.
Security-stack integrations	Export assurance signals to SIEM and AI-security platforms rather than replacing them.

Market signal

NVIDIA's open-source OpenShell runtime (Agent Toolkit, with NemoClaw) sandboxes agents with policy-enforced guardrails and is coming to Windows, OpenShift, and Ubuntu. CrowdStrike (Falcon AIDR, Guardian) and Darktrace (/ SECURE AI) launched agent-security products.

Positioning

OpenShell controls what an agent can do; security platforms watch what it does. Melaninternet verifies whether its decisions are supported by evidence and keeps the proof. Planned integration; no NVIDIA partnership is claimed.



Personal AI and health innovation

Two applications of the same evidence core — one consumer, one regulated.

PLANNED

Personal AI assistant (planned)

- Daily brief with calendar, priorities, and a source-linked news feed
- Calendar management and goal plans with progress tracking
- Shopping research and price tracking; purchases only with approval and limits
- Life-admin help with cited explanations

NIH PATH

Health innovation AI

- Research intelligence: literature synthesis, evidence mapping
- Study documents: protocol Q&A, criteria review, evidence gaps
- Health operations: prior authorization, claims documentation, compliance review
- Safety framing: not diagnosis; human review; auditable outputs

CONSIDERATION	DETAIL
Personal AI market	Meta launched its Muse personal agent in the U.S. on Sept 8, 2026; assistants such as Espa deliver daily briefs. Melaninternet's role is an evidence-first, approval-gated, mission-driven assistant — a showcase for the core rather than the primary revenue line.
NIH review fit	Health Technology & Methodologies (HTM) for health-operations framing, or Computational, Modeling & Biodata Management (BCTB) for method-development framing.
NIH registrations required	SAM.gov/UEI (active), eRA Commons organization + PI accounts, Grants.gov organization with an Authorized Organization Representative, SBA Company Registry (SBC Control ID).



LONG-TERM R&D

World models and Physical AI

From one-off answers to verified, stateful models — and eventually to systems that act in the physical world.

TOPIC	CURRENT DIRECTION
World model objective	A stateful model of entities, relationships, timelines, permissions, obligations, risks, and evidence paths — with every prediction and recommendation carrying evidence and uncertainty.
Enterprise version	Invoices, vendors, contracts, approvals, obligations, cash timing, exceptions, and controls.
Health version	Papers, protocols, study documents, evidence gaps, operational bottlenecks, and review cycles.
Field context (2026)	World models have become a platform race: NVIDIA released Cosmos 3, an open physical-AI world foundation model, in June 2026 with a coalition of robotics partners. Building on open models is the practical path for a startup.
Physical AI angle	Plan verification, action approval gates, and safety evidence for robots using open world and robot foundation models (e.g., NVIDIA Cosmos, GR00T) — not training base models.
Boundary	Not a current MVP capability. Physical AI follows only after enterprise agents, assurance, pilot controls, and safety testing mature.



Safeguards and reliability interpretation

What is protected today, what is explicitly not claimed, and how to read the numbers.

Current safeguards

Versioned constitution; privacy boundaries; temporary document context; source grounding; rate limiting; protected dashboards; 200-case regression corpus; live model checks; persistent reliability report; automatic model failover.

Important non-claims

Not a certified security product; not a regulated clinical decision system; not a contractual SLA; not proof of market demand or revenue; planned lanes are not current capabilities.

How the reliability figures should be read

The gate separates two kinds of failure. A **validator failure** means the product gave a wrong or unsafe answer and is never tolerated. A **provider timeout** means the hosted model did not respond — an availability fact about the upstream provider, recorded openly as a warning. The gate refuses to pass if it could not test cross-guest isolation, because a run that could not test isolation is not evidence that isolation holds.



Commercial and funding readiness

What is being asked for, and what is deliberately not being claimed.

AREA	CORRECT MESSAGE
Business model	Paid pilots → SaaS plans → enterprise usage pricing → per-agent assurance add-on (planned).
Traction	One narrative: CTA leads, design-partner conversations, pilot requests, repeated usage, and I-Corps interviews — not vanity metrics.
Funding use	Verification R&D, technical hire and contractors, inference, security and documentation, pilots, and go-to-market.
Non-dilutive path	NSF SBIR resubmission, NIH SBIR, NSF I-Corps, NJ CSIT and NJSBDC programs.
Proof goal	Move from technical evidence to market evidence through defined pilots, ROI analytics, and a validated durable-guarantee method.

\$500K

pre-seed funding goal

3+

paid pilots / customer proofs targeted in 18–24 months

Pre-revenue

seeking design partners; no customer or revenue claims are made