

KEYNOTE · AI AGENT OPERATING MODELS

When the Machine Can Think

Running an IT department with AI

From AI Copilots to AI Accelerated Delivery Systems

Dr. Jeff Sutherland · Co-creator of Scrum · ScrumAI.org

FIRST PRINCIPLES IN SCRUM:

OpenClaw Scrum and
Scrum@Scale



Jeff Sutherland
Co-Creator of Scrum and
Scrum@Scale

leanpub.com

Background: 2023 → 2026

From painful AI pair-programming to autonomous agent operating model



THE GOAL

1000× velocity on one Mac Studio · **< 10% token cost / story point** vs enterprise AI · **10× quality** using Scrum + Scrum@Scale

The Pace Changed

From quarterly AI adoption to daily AI operations

2025

~ 3 months

between meaningful model updates. Teams adapted by quarter — pilots, governance reviews, slow rollouts.

2026

Every day

Hermes, OpenClaw, Claude, GPT-5.5 ship changes daily. Adaptation is no longer a project — it is a morning routine.

*“The bottleneck is no longer access to intelligence. **The bottleneck is your work system.**”*

My Morning

Upgrade. Health-check. Repair. Then work.

1

Upgrade

Hermes, OpenClaw,
models, gateways.

2

Health

Slack, agents, tokens,
launchd, Docker.

3

Repair

Fix breakages before
pulling work.

4

Work

Pull top-priority backlog ·
WIP = 1.

I used to run this by hand each morning — Hermes now runs it unattended across the fleet.

The Operating Model

Models are engines. The operating model is the car.

AGENT OPERATING STACK

Mission Control	Backlog · WIP · Review · Done · Leaderboard
Hermes	Operator · fleet IT · audits · memory
OpenClaw	Self-organizing Scrum swarm · multi-agent execution
Models	Claude · GPT-5.5 · DeepSeek · Grok

WHY THIS MATTERS

- Models are interchangeable workers and reviewers — not the operating model.
- OpenClaw runs the self-organizing, cross-functional Scrum swarms — six agents per team — that produce the throughput.
- Hermes is the Kanban-style management layer — it coordinates, audits, retrieves memory, and runs fleet IT.
- Mission Control is the Scrum board for agents — backlog, WIP=1, review gates, leaderboard.

*Models are engines. **The operating model is the car, dashboard, brakes, and pit crew.***

Daily Scrum for Machines

The eight-step loop, every morning

- 1 Upgrade all systems
- 2 Health-check gateways, Slack, agents, tokens, launchd, Docker
- 3 Fix breakages immediately
- 4 Pull top-priority work from Mission Control
- 5 Enforce WIP = 1 per agent
- 6 Require deliverables / evidence before review
- 7 Grok / PO review gates before done
- 8 Leaderboard + value report at end of day

DevOps + Scrum + agent governance — fused into one daily operating rhythm.

Field Report

Case: bringing #henry back online

Detect

Slack route to #henry failing — agent silent.

Diagnose

Gateway version mismatch + token drift identified.

Repair

Docker gateway rebuilt; Henry upgraded; LaunchDaemon persistence fixed.

Verify

Agent online — no longer dependent on VNC/GUI login.

THE LESSON

Agent uptime is a Scrum impediment.

Tokens, gateways, persistence, login state — these now belong on the impediment list, not in a separate ticket queue.

Hermes Runs IT

Case: a self-maintaining, HIPAA-compliant medical fleet

Scope

15 Windows medical machines across three sites — 11 Lincoln, 2 Capehouse, 2 West Peoria; each runs dozens of devices.

Uptime

Surprise Windows updates were the top downtime cause — now scheduled, with self-repair.

Hands-off

Auto-starts on power-on, signs in, reconnects, and monitors health — no manual login.

Security

Tailscale seals a medical-only network group; no third-party data path; remote desktop off.

THE RESULT

100%

uptime in month one

was near-daily incidents before Hermes

2–10× faster

after removing legacy AV & VPN drag

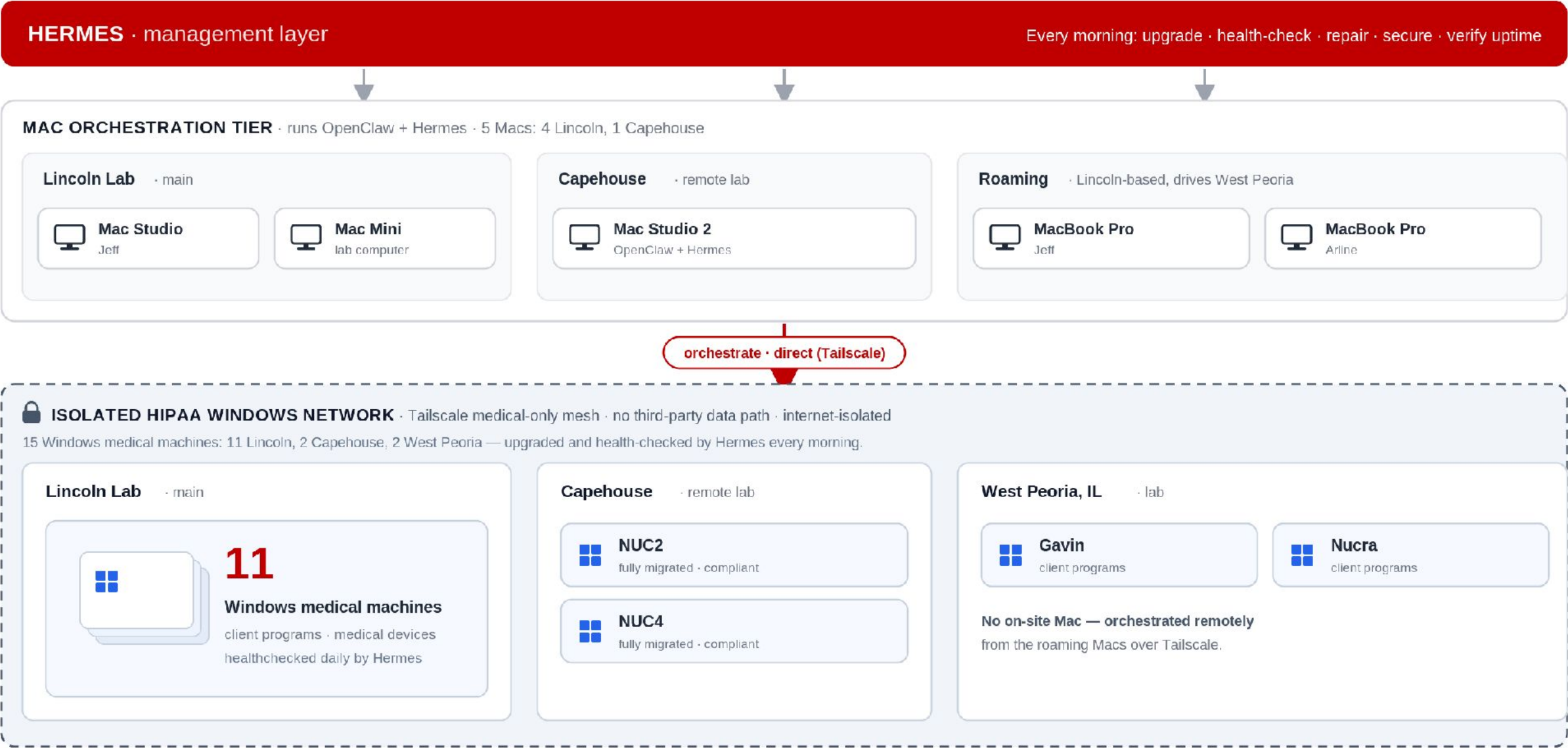
Infrastructure that learns.

Every machine certifies, classifies drift, and self-repairs — impediments become reusable capability, run by Hermes, not a help-desk queue.

IT Systems Managed by Hermes

Mac orchestrator Windows · HIPAA

One management layer · three sites · 5 Macs orchestrate 15 HIPAA-isolated Windows medical machines



Mission Control

The Scrum board for agents

Backlog	WIP = 1	Review	Done
Ordered by value	One item per agent	Grok / PO review	Evidence-based DoD
Machine-executable work items	Health-check passed	Tests + artifacts	Audit trail
Acceptance criteria	Evidence required	Logs verified	Counts toward value

Sprint board → execution control system. Status is replaced by gates, evidence, and audit.

Value, Not Activity

Velocity with verification

58

stories completed today

121

story points completed

65.33

3-day moving avg / day

WHAT COUNTS AS VALUE

- New work completed
- Rework recovered after AI-review rejection (10 today)
- Security & HIPAA compliance fixes
- Infrastructure reliability
- Revenue / marketplace enablement
- Content and distribution

*Velocity without verification is hallucination. **Done means evidence.***

Roles Redefined

What changes in Scrum when agents join the team

CONCEPT	WHAT CHANGES
Product Owner	Still sets value and order — now human + agent-assisted.
Scrum Master	Becomes flow debugger and system-health optimizer.
Developers	Human developers + specialized AI agents working as one team.
Definition of Done	Must include evidence, tests, artifact paths, logs, review gates.
Sprint board	Execution control system — not a status report.
Impediments	Token mismatch, stale gateway, model regression, context loss, failed Slack routes; plus fleet IT, security, HIPAA.

*Not replacing teams with agents — **redesigning the operating model so humans and agents produce measurable outcomes together.***

Leadership Implications

Three executive lessons

01

STRATEGY

Strategy must be executable by agents.

Vague work stalls or hallucinates.
Backlog items become machine-executable contracts.

02

GOVERNANCE

Governance must be real-time.

Quarterly steering is too slow. Daily health, audit logs, WIP limits, and review gates are required.

03

LEARNING

Advantage moves to learning rate.

Whoever upgrades, verifies, and redeploys fastest learns fastest. Scrum becomes the learning loop.

Executives who fund only models will be outpaced by those who fund the operating model.

The Pattern Others Can Copy

AI Agent Scrum Checklist



Backlog ordered by value



WIP limit enforced (start at 1)



Agents self-select top eligible work



Evidence-based Definition of Done



Automated review gate (Grok / PO)



Health checks and recovery scripts



Daily value report

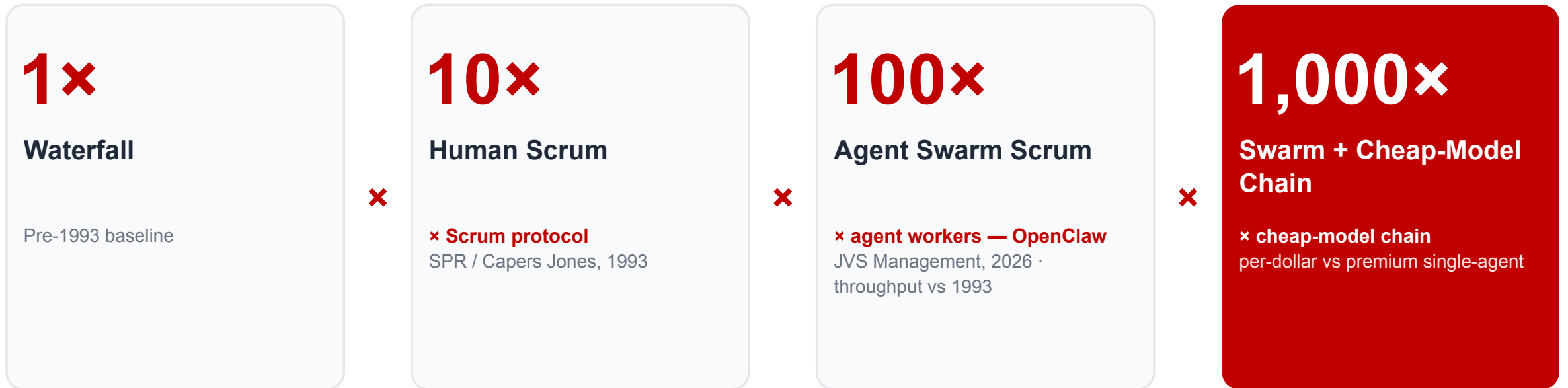


Audit trail end-to-end

Start with one agent team and one visible board. Improve the system every morning.

The Three Compoundings

Why the gap is 1,000× — three multipliers on one unchanged protocol



100× throughput vs 1993 waterfall **1,000×** productivity-per-dollar vs a premium-model single-agent competitor.

Same protocol since 1993. New workers. In the token economy, the economics are now decisive.

Scrum's Three Eras

Same pattern, new participants — from self-organizing teams to self-organizing swarms

1993 · Scrum — self-organizing cross-functional team, user stories, empirical loop. *Scaled human cooperation for the first time.*

Mid-2010s · Scrum@Scale — Scrum of Scrums + Executive Action Team. *Scaled Scrum to whole organizations of thousands.*

2026 · Scrum@Scale for Agents — same mechanics, agent participants. Each site is a team. Main agents form the Scrum of Scrums. *Cycle time drops from days to seconds.*

What carries forward — empirical control, WIP limits, server-enforced Definition of Done, self-organization. *More important with agents, not less.*

What's new — teams are ephemeral. Work items are skill invocations. The org chart shrinks to its irreducible core: **Director · Adjudicator · Steward.**

In Scrum@Scale terms — Director = CPO of the EMS · Adjudicator = Scrum Master of the EMS · Steward = Scrum Master of the EAT.

“The construct that scales human organizations to thousands of people scales agent organizations to hundreds of sites — and we are running it in production at JVS Management today.”

The Skill-Swarm Organization

The org chart shrinks to its core — three human roles over a swarm of agents

Director

Chief Product Owner · Executive MetaScrum

Sets value and priority across the whole swarm.

Adjudicator

Scrum Master · Executive MetaScrum

Runs priority arbitration; forces a decision on the novel cases agents escalate.

Steward

Scrum Master · Executive Action Team

Removes impediments; keeps the system healthy, secure, and live.

WHAT'S DIFFERENT WITH AGENTS

Teams are ephemeral — spun up per goal, dissolved when done.

Work items are skill invocations — not tickets in a queue.

Hermes manages, OpenClaw delivers — Kanban-style cortex over self-organizing Scrum swarms.

Cheap-model chain does the work — premium models design, escalate, and review.

Not bolted onto Scrum@Scale — its apex, with most participants now agents. In production at JVS Management today.

The Inversion

Studying AI teams is a microscope — a *Drosophila* — for the teams we’ve always had

AGENT FAILURE MODE	THE HUMAN PATTERN IT MIRRORS
Self-attestation	Self-graded OKRs — measuring what’s easy, not what matters.
Cross-agent endorsement	Rubber-stamp committees — reciprocal sign-off.
Parallel infrastructure	Shadow IT — the unofficial workaround.
“Christmas-tree” commits	Omnibus bills — the clause buried on page 247.
Soft-language laundering	Corporate euphemism — “right-sizing,” “realignment.”
Action-hallucination / fake data	Reporting the world as you wish it were.

*The failure modes are identical. **Scrum is the protocol that reduces the friction — for agents and humans alike.***

CLOSING

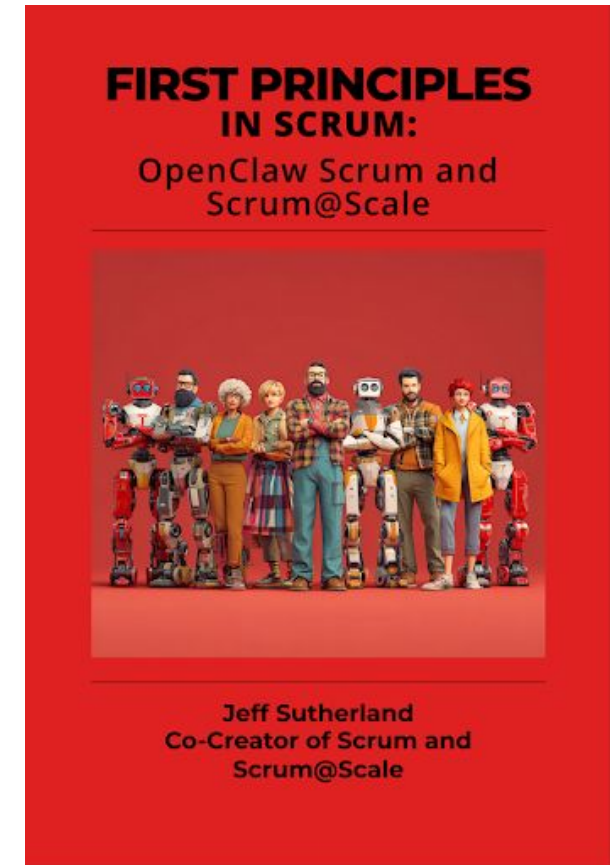
AI agents do not eliminate Scrum.

They make disciplined Scrum more important.

When the machine can think, leaders must design the system that tells it what value means, how work flows, and how done is proven.

Dr. Jeff Sutherland · ScrumAI.org · a project of JVS Management Inc.

Thank you · Gracias



Foundations

The talk's claims, grounded in four decades of research

SCRUM & LEAN

Schwaber & Sutherland — The Scrum Guide (2020)
Sutherland — Twice the Work in Half the Time (2014)
Ohno — Toyota Production System (1988)

FLOW & RELIABILITY

Beyer et al. — Site Reliability Engineering, Google (2016)
Forsgren, Humble & Kim — Accelerate (2018)
Kim et al. — The DevOps Handbook (2016)

SECURITY & ZERO TRUST

NIST SP 800-207 — Zero Trust Architecture (2020)
Saltzer & Schroeder — least privilege (1975)
Ward & Beyer — BeyondCorp, Google (2014)

SELF-HEALING & COMPLIANCE

Kephart & Chess — Autonomic Computing (2003)
HIPAA Security Rule — 45 CFR § 164.312
Ormandy — Project Zero on consumer AV (2016)

Full citations in the companion paper — “Hermes Runs IT: Toward Fully AI-Driven Infrastructure.”