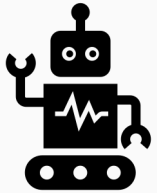
Shadow AI: Hunting the Unknowns Across Developer Tooling and Deployed Artifacts

Anant Shrivastava

About Cyfinoid

Research Focused Cyber Security Firm

Cloud
Environments



AI
Usage & Security

Software
Supply Chain

https://cyfinoid.com

Anant Shrivastava

- Chief researcher @ Cyfinoid Research
- 15+ yrs of corporate exposure
- **Speaker / Trainer:** BlackHat, Defcon, c0c0n, nullcon & more
- **Project Lead:**
  - Code Vigilant (code review project)
  - Hacking Archives of India
- (@anantshri on social platforms)
- anantshri.info

Cyfinoid

A NOTE ON CRAFT

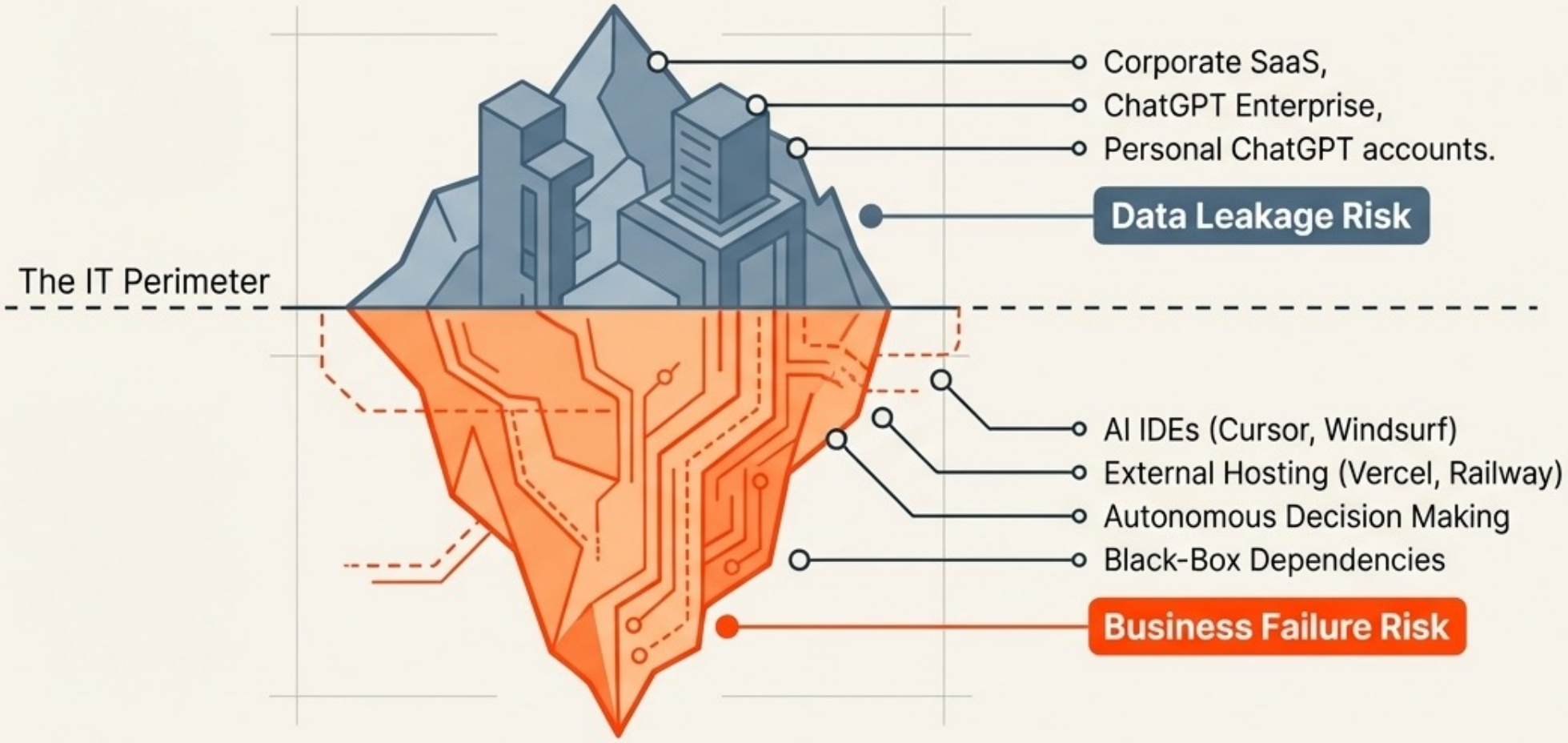
Co-Authored by Intelligence.

The insights within this presentation are the result of dellberate human curation ampiified by state-of-the-art cognitive models.

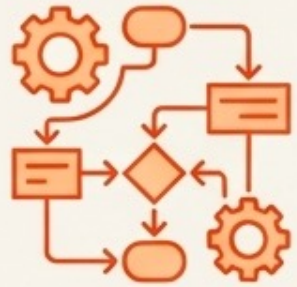
Cognitive Stack Matrix

HUMAN	Strategic Intent & Final Curation
NOTEBOOKLM	Source Ingestion & Data Grounding
CLAUDE	Synthesis & Linguistic Precision
CHATGPT	Structural Architecture & Ideation

The Threat Landscape Has Shifted



120%
YoY Growth
in unapproved
AI coding tools.



Key Insight: AI is no longer just a **background helper.**
It is an active, unmanaged participant in your software supply chain.

The Democratization of Software Creation

Cyfinoid

THEN: Skills limited creation.

Software was built by engineers within the rigid boundaries of the corporate firewall and procurement processes.

AI removed the requirement to be a developer. Business teams now autonomously deploy live production systems.

NOW: Imagination limits creation.

How the Shadow Pipeline Bypasses IT

1. Download: Employee installs AI IDE (Cursor/Windsurf).

2. Bypass: Pays with personal credit card.

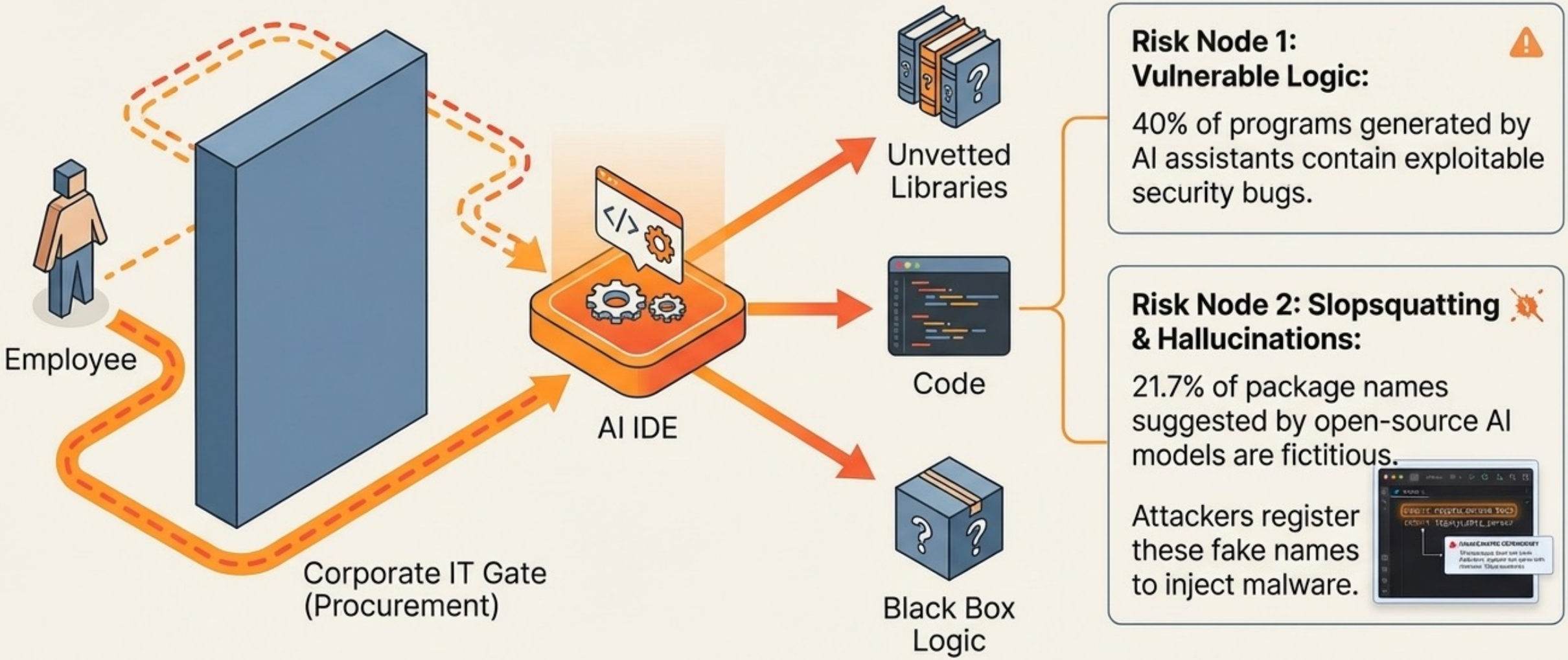
3. Generate: AI builds the internal business tool.

4. Host: Deploys to Public Cloud (Vercel/Railway).

5. Legitimize: Shares URL / Adds CNAME.

This is not a theoretical corner case. Motivated by efficiency, non-engineers are spinning up live business applications in under an hour, today.

The Black Box Architect & Implicit Supply Chain Choices



Risk Node 1:
Vulnerable Logic:

40% of programs generated by AI assistants contain exploitable security bugs.

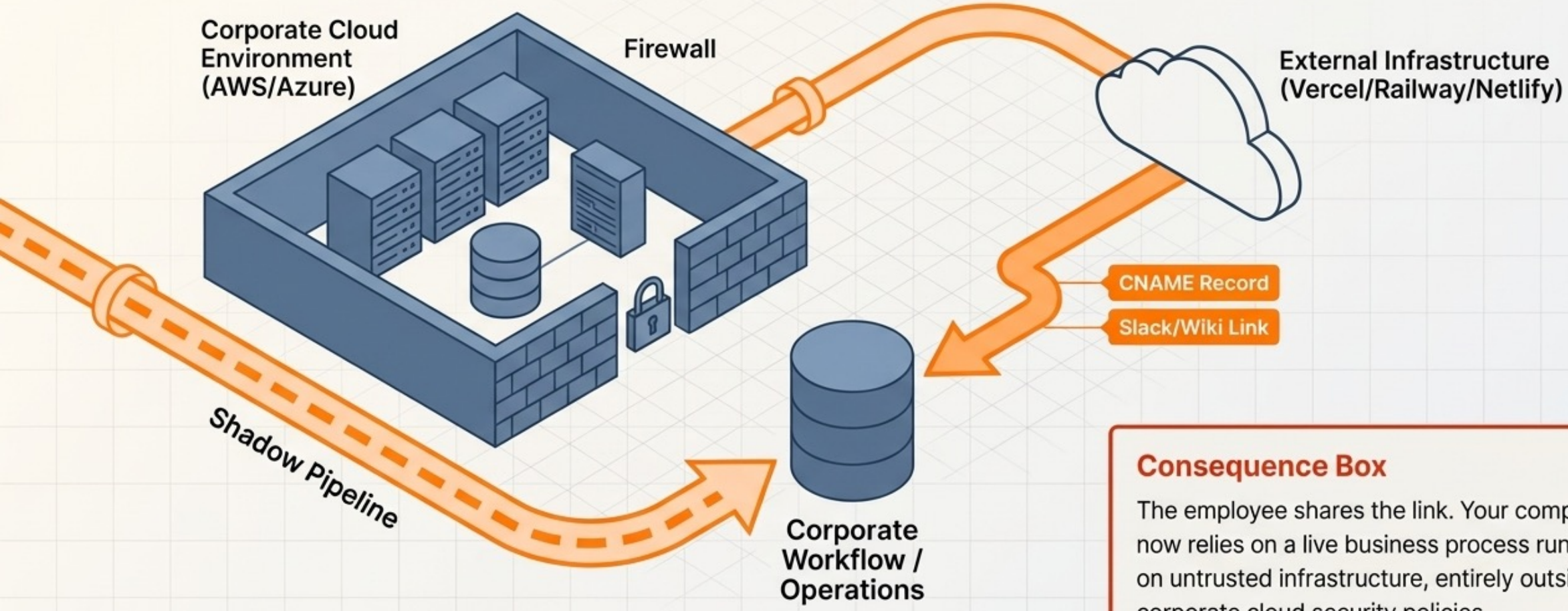
Risk Node 2: Slopsquatting & Hallucinations:

21.7% of package names suggested by open-source AI models are fictitious.

Attackers register these fake names to inject malware.

The Deployment Bypass: Hiding in Plain Sight

The AI optimizes for speed, recommending instant-deploy external platforms.

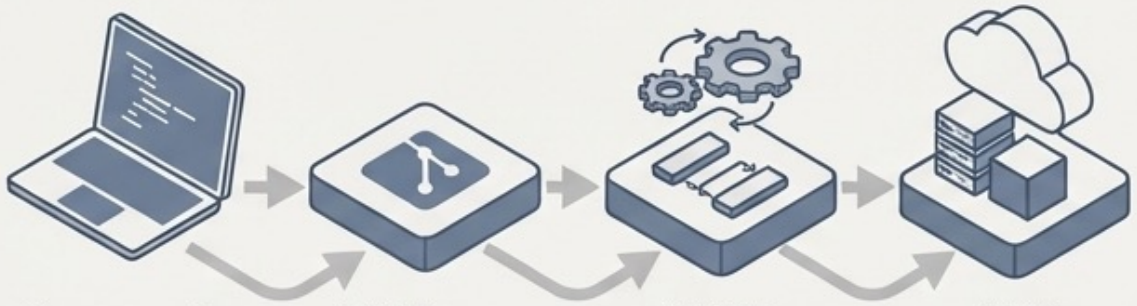


Consequence Box

The employee shares the link. Your company now relies on a live business process running on untrusted infrastructure, entirely outside corporate cloud security policies.

Why Traditional Security Fails

The Walled Garden



Security Layer: SCA, SBOM, and Firewalls catch vulnerabilities here.

The Shadow Pipeline

Security Layer: Completely bypassed.

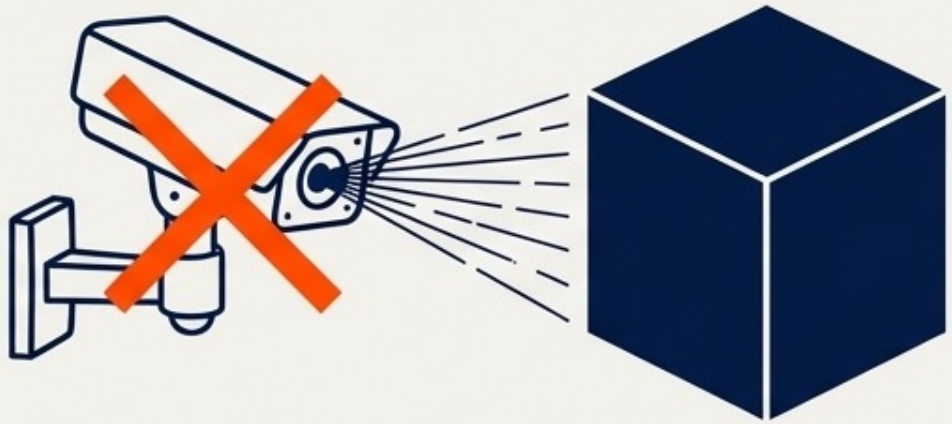
Key Insight: Security tools scan known repositories. If the repository is personal or external, there is nothing to scan. It is ~~**Production Without Inventory.**~~

Traditional Tooling is Blind to the Shadow Pipeline

The SBOM Gap

SBOMs require visibility. If the repository is personal or hosted externally, the SBOM remains completely empty.

The SCA Gap



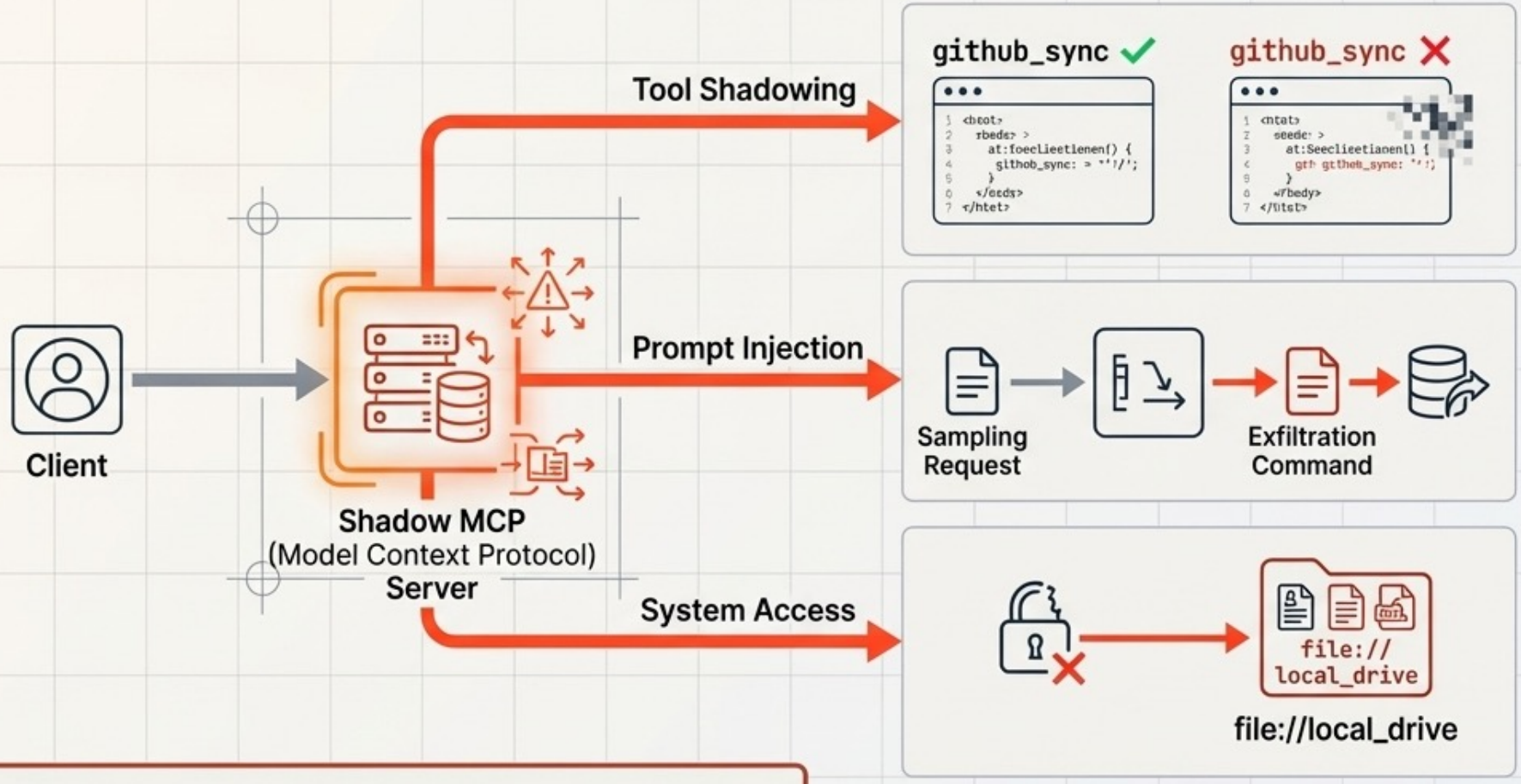
SCA scans for known vulnerabilities in known code. It cannot assess the security posture of an AI model's real-time suggestions.

Data Check: NYU Center for Cyber Security found AI Copilots generate vulnerable code 40% of the time. 21.7% of AI-suggested package names are fictitious (Slopsquatting).

Shadow Agents: The Kinetic Attack Surface

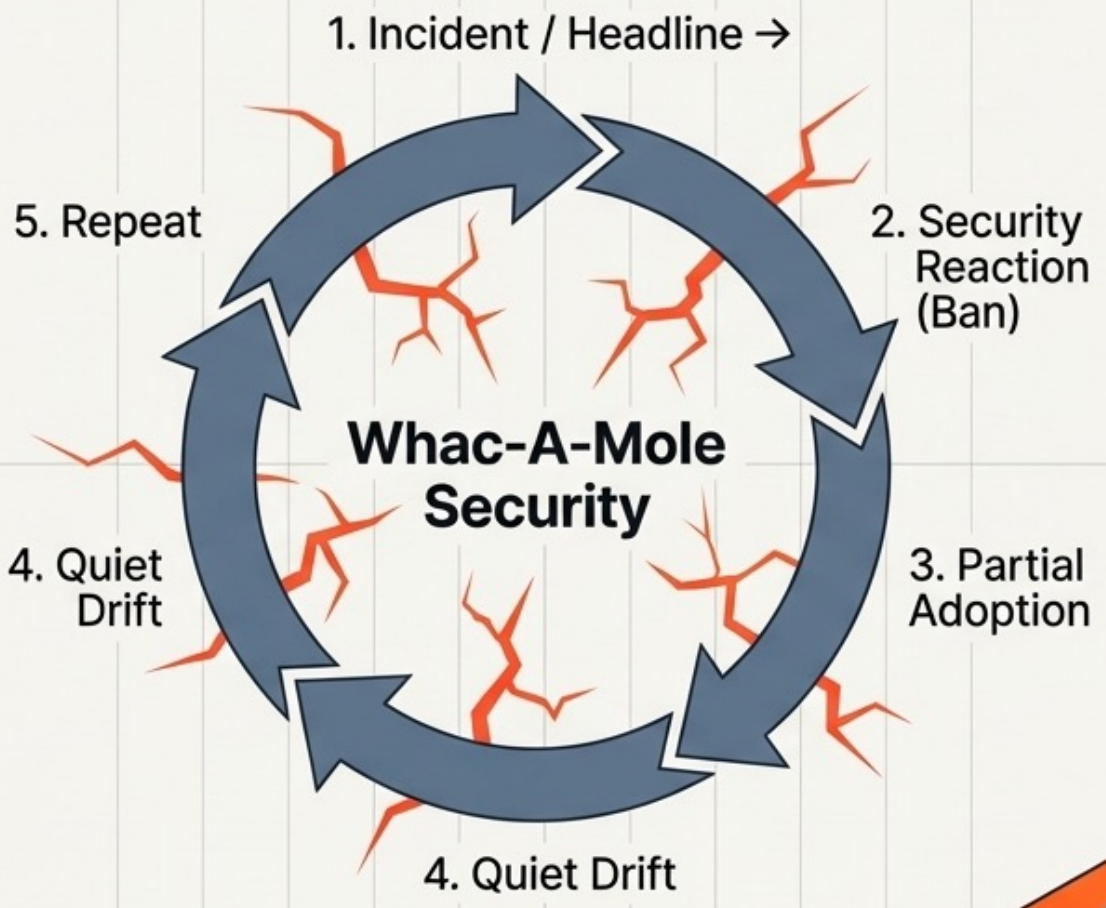
Cyfinoid

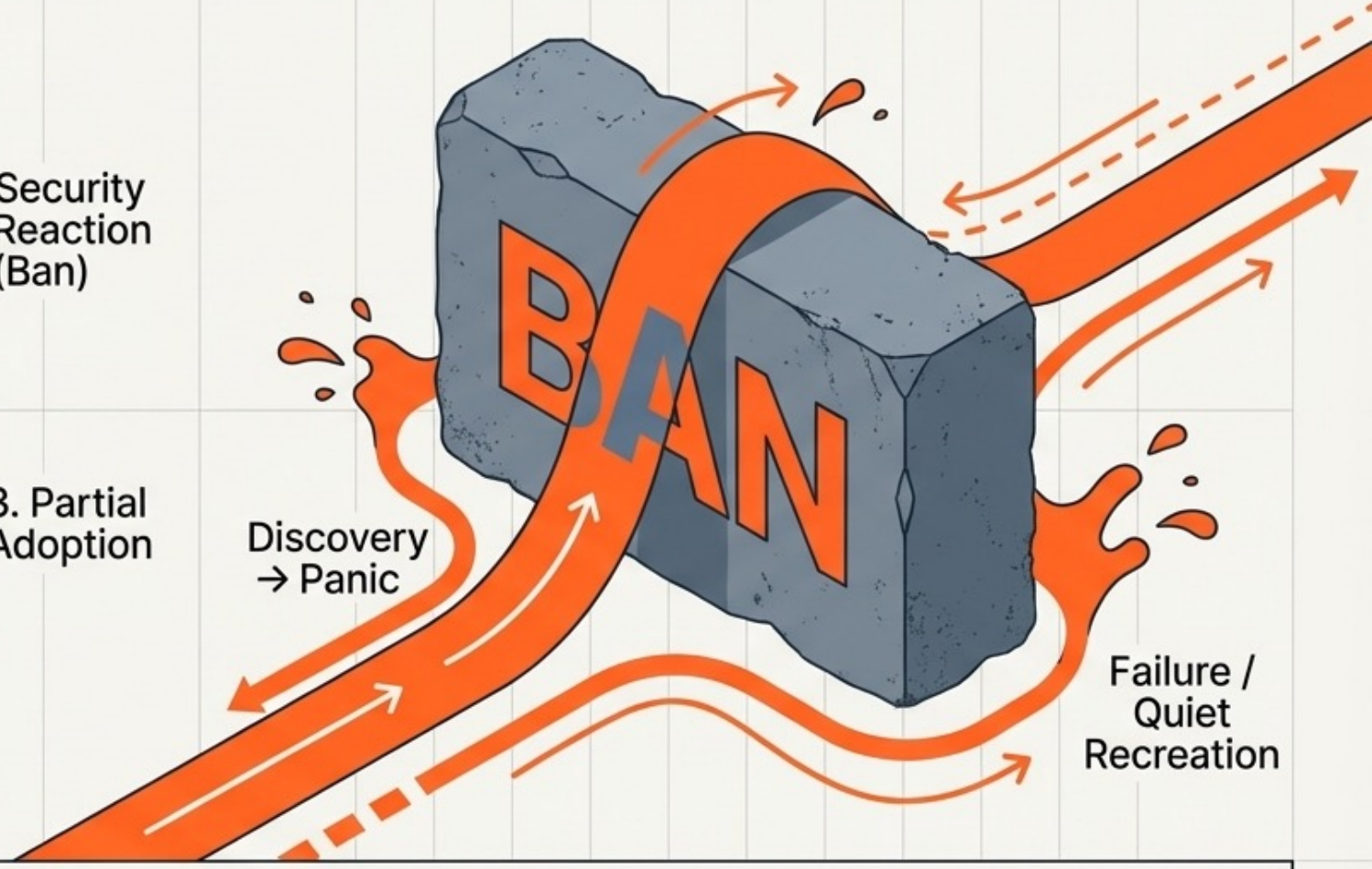
Cybercriminals are exploiting AI coding agents. If a shadow agent has file execution capabilities, malicious inputs can trigger unintended remote code execution.



"When used improperly, AI doesn't just help builders. It creates predictable entry points that criminals actively exploit."

Banning Capability Does Not Work





Punishing shadow AI simply **forces it deeper underground**, creating "Quiet Recreation" and completely eliminating any remaining IT visibility.

Flip the Model: Reward Visibility to Gain Control

Build & Submit

Reward & Support

Harden & Maintain

The New Pact: Submission buys Support. If employees surface their AI-built tools, IT provides secure hosting, scaling support, and hardening.

Key Insight: Inventory is a process, not a static list. We must stop trying to control the code and start tracking the intent.

Cyfinoid

Discovery Without Friction

Don’t spy. Listen to the signals. You cannot run an SCA scan on a personal Vercel account, but you can detect the traffic routing to it.

DNS & Domains

Monitor network calls to known AI IDE endpoints.

OAuth & SSO

Track SSO grants to unauthorized deployment platforms.

Reverse Proxy

Catch and analyze unmanaged Shadow AI traffic.

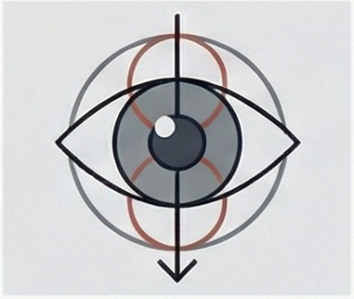
External Apps

Scan for unrecognized external business applications.

THE AI AGENT SECURITY & OBSERVABILITY STACK

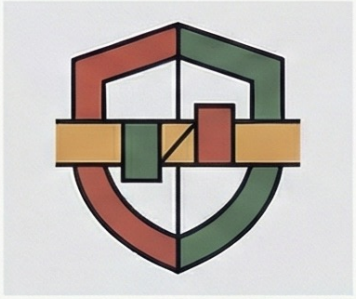
As AI coding agents gain autonomy, traditional boundaries vanish. This toolkit provides layered defense, from real-time protection to fleet-wide observability and analysis.

REAL-TIME SECURITY & GUARDRAILS



Gryph: Agentic Operations Visibility

Hooks into AI agents to log file changes and shell commands to a local database.



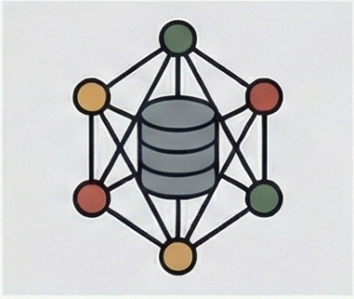
PMG: Package Manager Guard

Intercepts and sandboxes malicious package installations before code executes on your machine.

Local-First Privacy

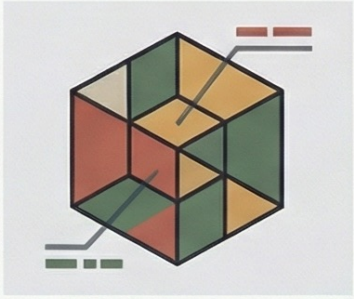
Both tools prioritize local storage and redaction, ensuring sensitive credentials never leave your machine.

INFRASTRUCTURE & SESSION & OBSERVABILITY



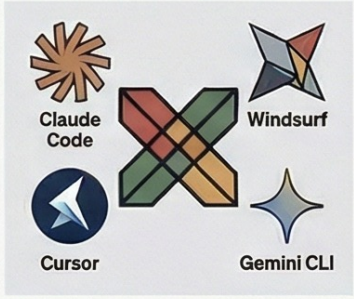
Osquery: Fleet-Wide Observability

Performance-oriented framework exposing the entire operating system as a relational database.



Agentsview: Session Insights

A desktop and web application for browsing, searching, and analyzing past AI coding sessions.



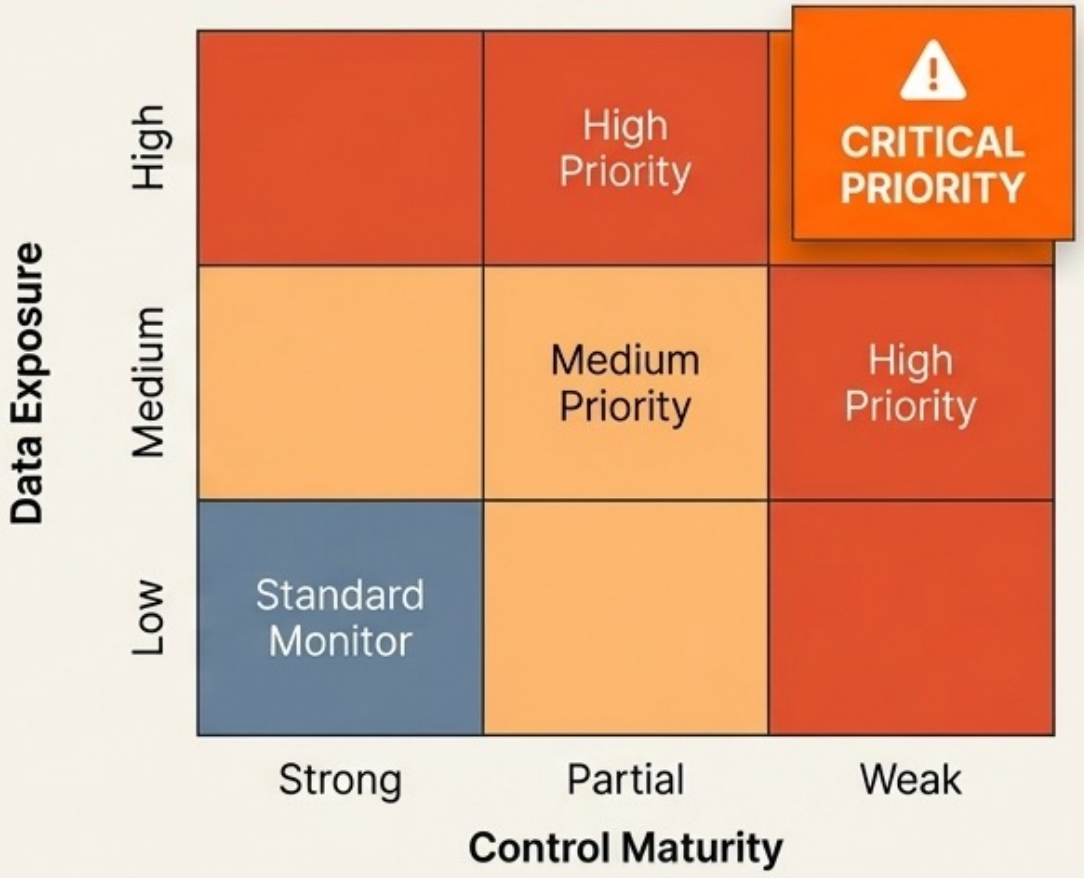
Multi-Agent Support

Compatible with major agents Including Claude Code, Cursor, Windsurf, and Gemini CLI.

Triage by Blast Radius

Not all Shadow AI is equal. An AI script to format spreadsheets is Low Risk.
A customer portal hosted on a personal Netlify account is Critical Priority.

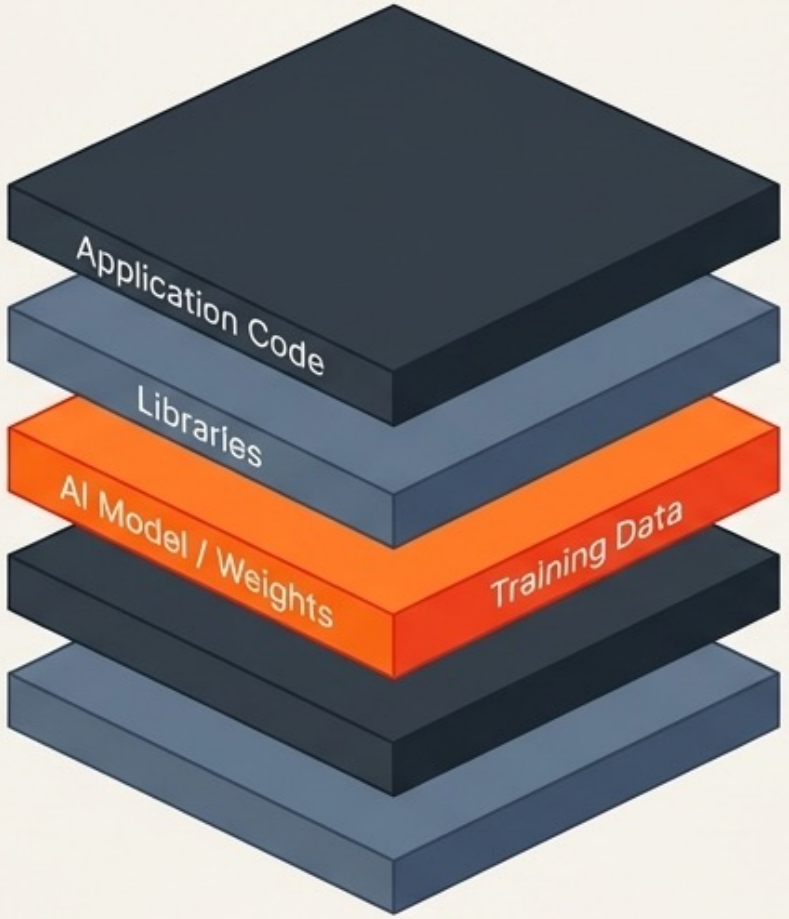
Risk Exposure



Assessment Criteria

Data	Blast Radius	Control Maturity
Public	Internal	Customer/Regulated
Individual	Team	Org-wide
None	Partial	Strong

The Missing Link: From SBOM to AIBOM



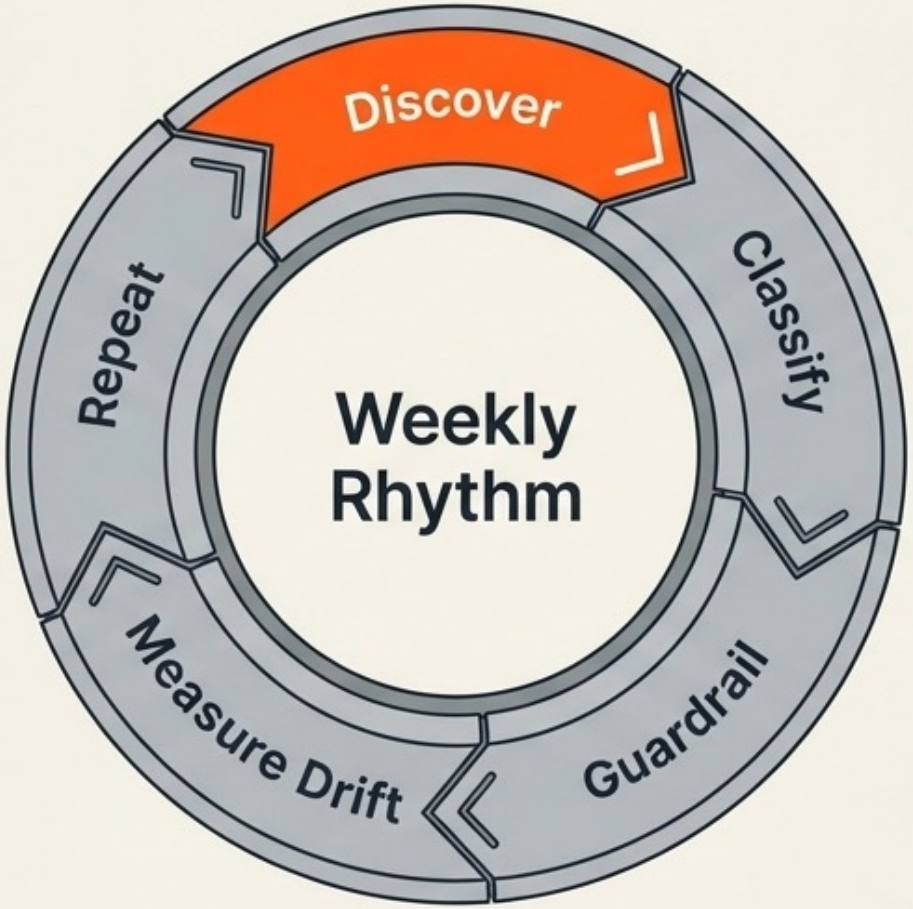
AIBOM Maker Details

- Dependency Inventory & Provenance
- Model Architecture & Version
- Training Data constraints (PII/Copyright inclusion)

Takeaway: We must treat the AI model as a third-party vendor requiring full chain-of-custody documentation.

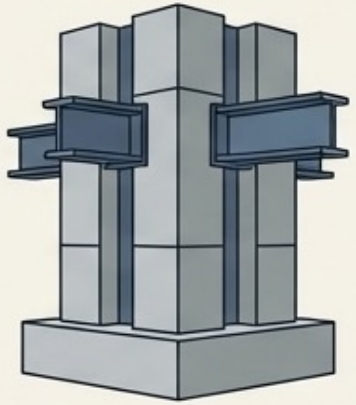
Shadow AI is an Operation, Not an Audit

The Golden Rule: If bypassing the system is easier than compliance, the guardrail fails.



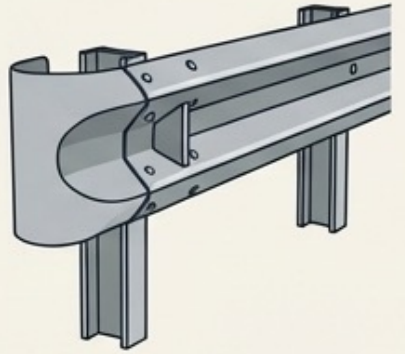
Roles Reimagined

ENGINEERING



Owns Hardening, Maintenance, and Scale.
Does not police.

SECURITY



Defines stances, provides acceleration guardrails. Does not block.

Open to Questions?

Cyfinoid

Trainings & Research

Cloud Security | AI Security | Supply Chain

Trainings

Attacking Software Supply Chain | Attacking Cloud Environments

Contact us at **contact@cyfinoid.com**