aiDevCon
by Tessl

NYC

Version your Factory, not just your code



Daniel Meppiel
Director, Agent & Dev Platforms · Microsoft

Would you merge this without reading it?

Contoso Bank · core payments

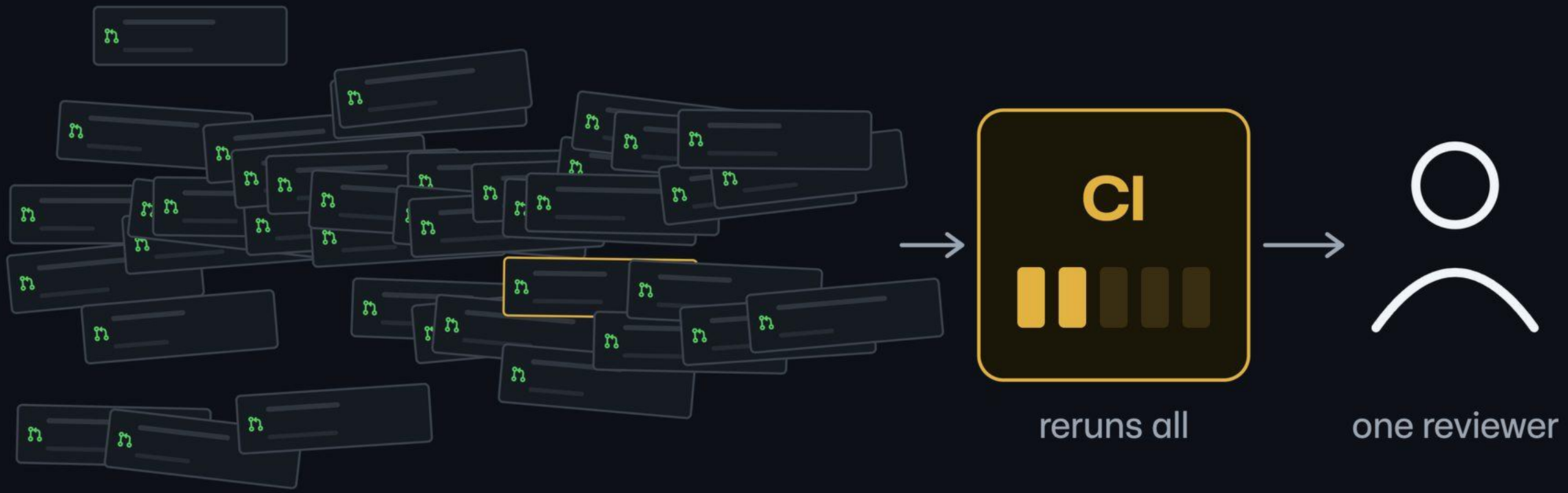
Hold wires of USD 10K+ for internal approval

delivered by an agentic software factory

Intent in.
A pull request out.



Review and CI don’t scale.



Today: auto-merge the low-risk.



Necessary. Not enough.

The enterprise needs:

Product

Verification

Does it do what we intended?

Platform

Enforcement

Our standards, at every step?

Audit

Traceability

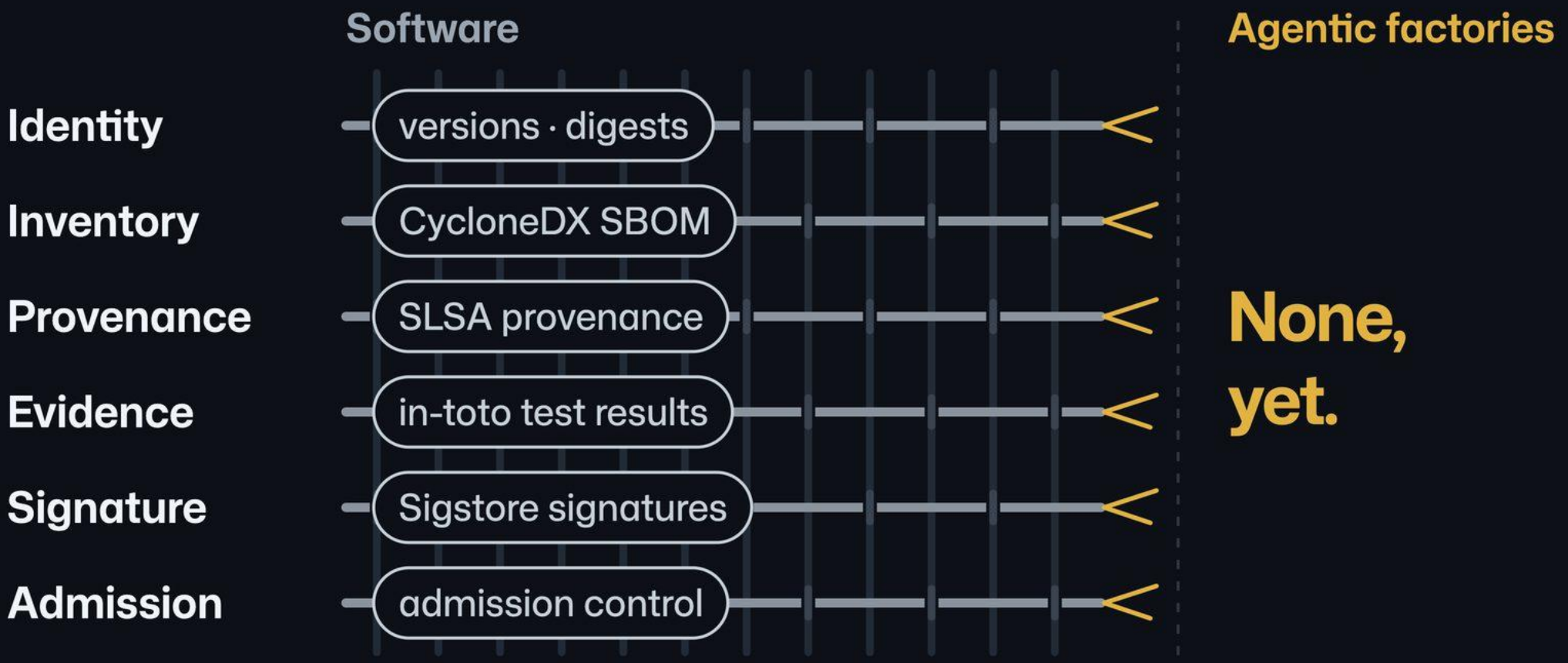
Which factory built it?

CISO

Security

Was the factory tampered with?

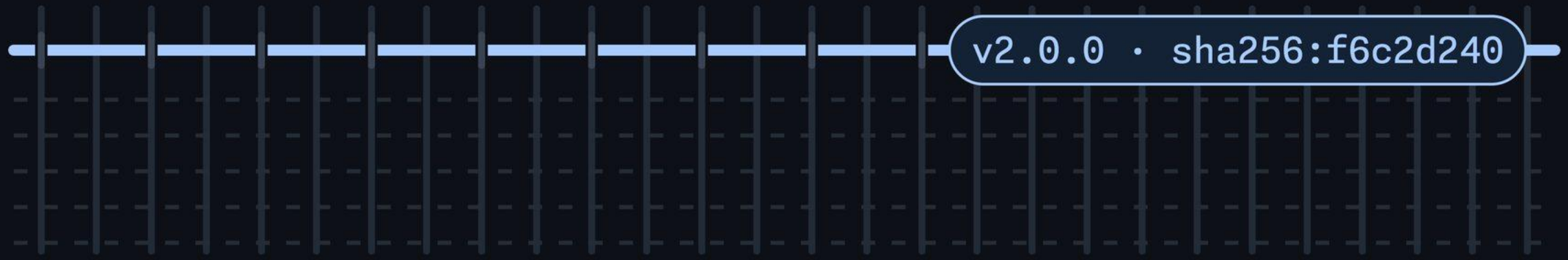
Software runs on a fabric of trust.



Trust comes from the factory, not the model.

A factory is a program. Version it.

Identity is the first thread.



Today's factories:
loose parts, or locked in.



Loose parts
not one thing

One portable unit
of agent work

Locked in
not portable

Agent Contract

v2.0.0

A reusable specification for agent work.

Input

Assignment
your intent, each run

The job

Role
how it works, every run

Dependencies
skills it may use

Budget
tries and time

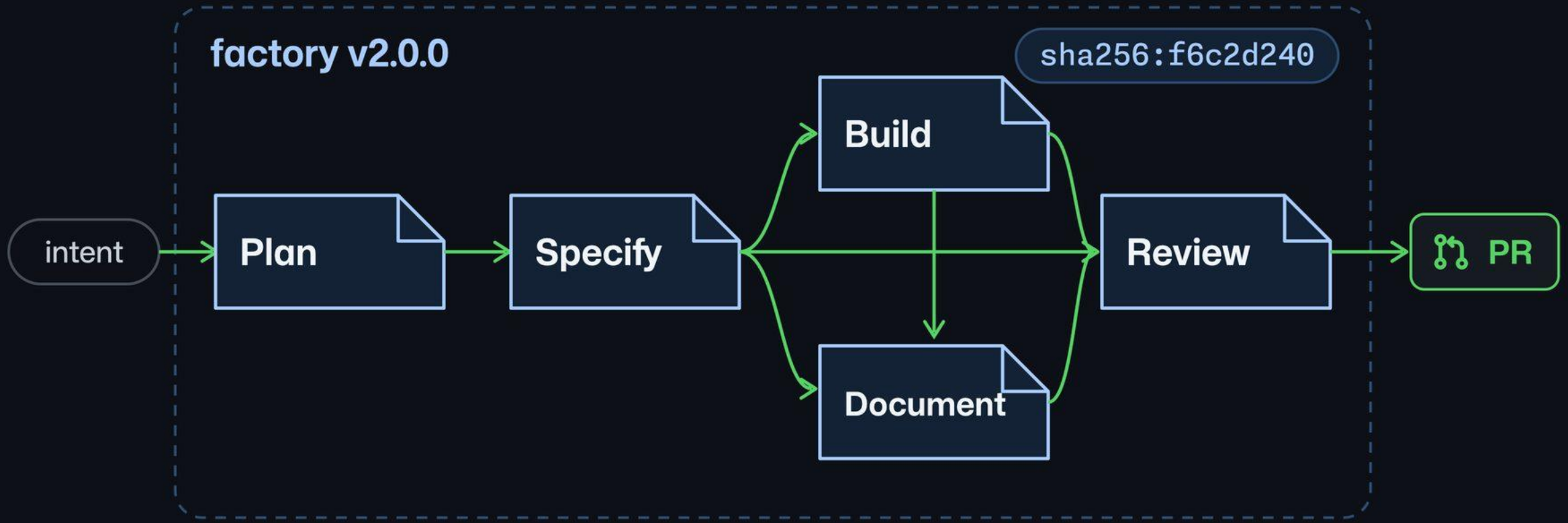
Judged & handed back

Checks
how it's judged

Deliverables
what it hands back

Versioned · Composable · Portable · Identifiable

A factory is a graph of contracts.



A factory is a program.
So what runs it?



APM: the package manager for agents.

apm.yml

```
name: feature-factory
version: 2.0.0
dependencies:
  - wshobson/agents
```

Identity

Declared

apm.lock.yaml

```
resolved_commit:
  46891e7e60da0e52…
content_hash:
  sha256:df2784d5…
```

Inventory

Pinned, hashed

apm-policy.yml

```
dependencies:
  allow: [contoso/*]
  deny: [untrusted/*]
extends: org
```

Admission

Governed, org-wide

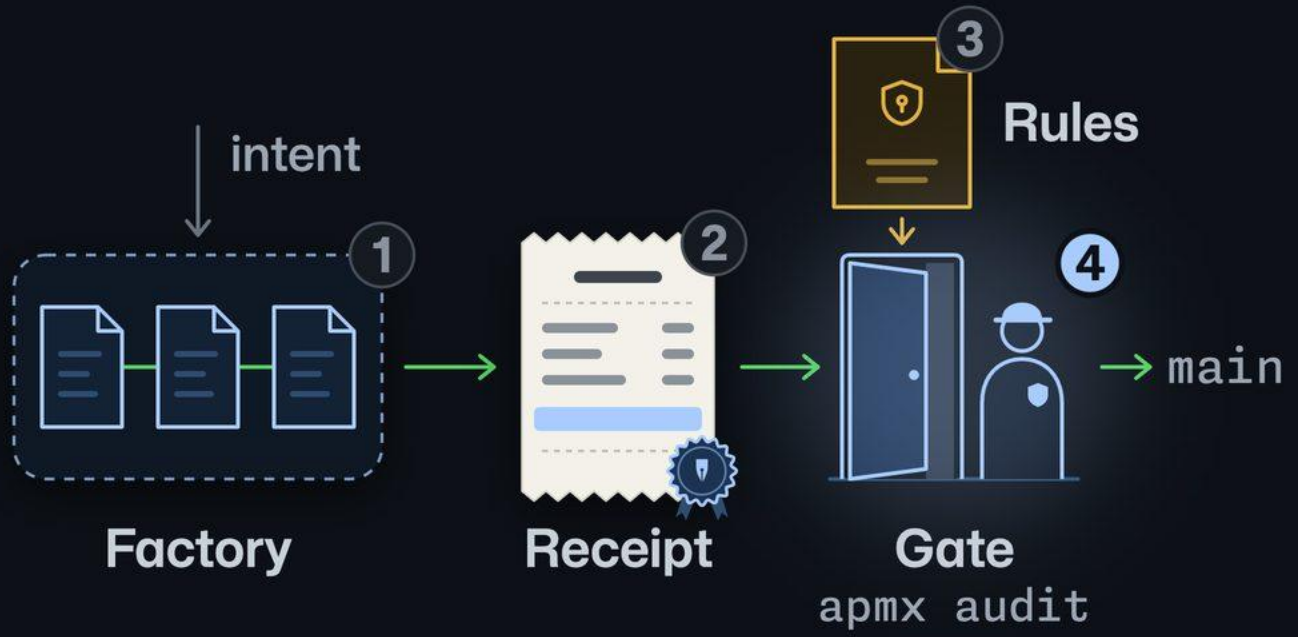
Microsoft open source · MIT

APMX: run a factory as a program.



Independent · experimental · Apache-2.0

Intent in. Evidence out.



1. **A versioned factory**
   contracts, checks, skills
2. **A signed receipt**
   what built it, what passed
3. **Your rules**
   owned by security
4. **One gate at merge**
   receipt against your rules

One receipt. Four open standards.



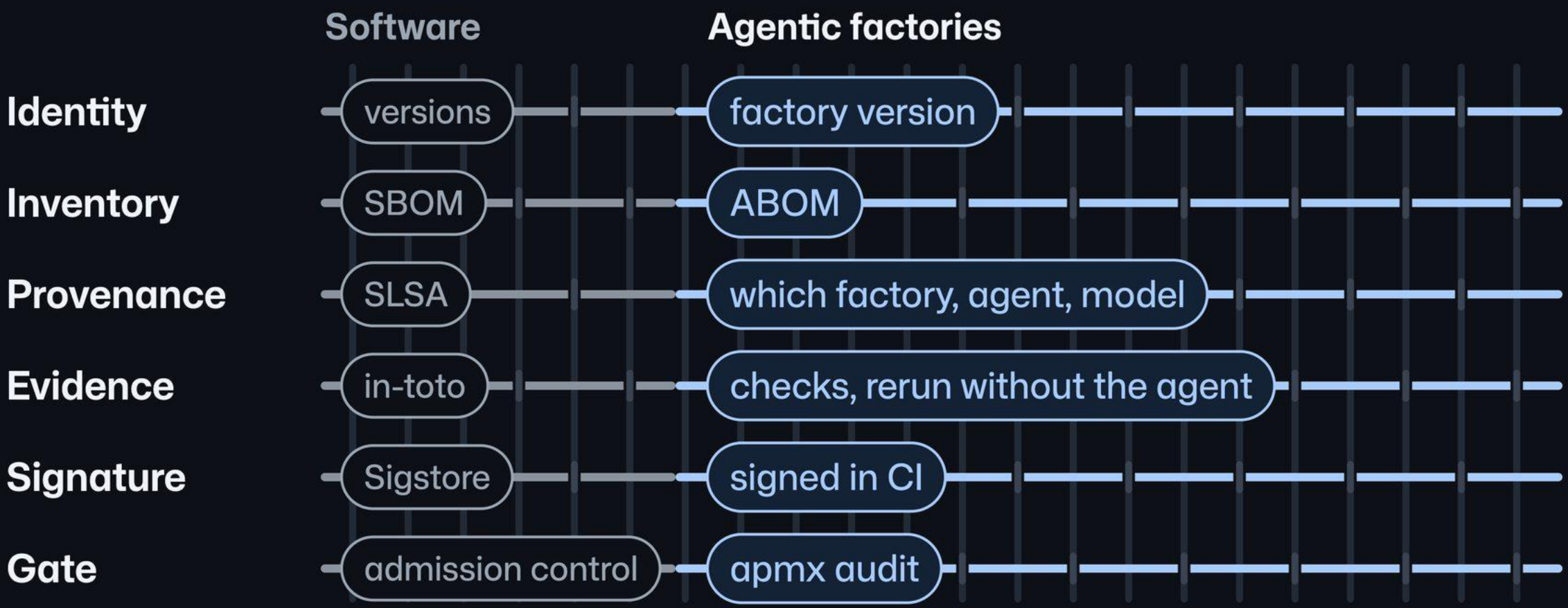
SLSA provenance — Traceability

in-toto test results — Verification

CycloneDX ABOM — Enforcement

Sigstore signature — Security

Same fabric of trust, for factories.



Verification Enforcement Traceability Security

The gate: one check reads the receipt.

PR #1 · Checks · apmx audit (required)

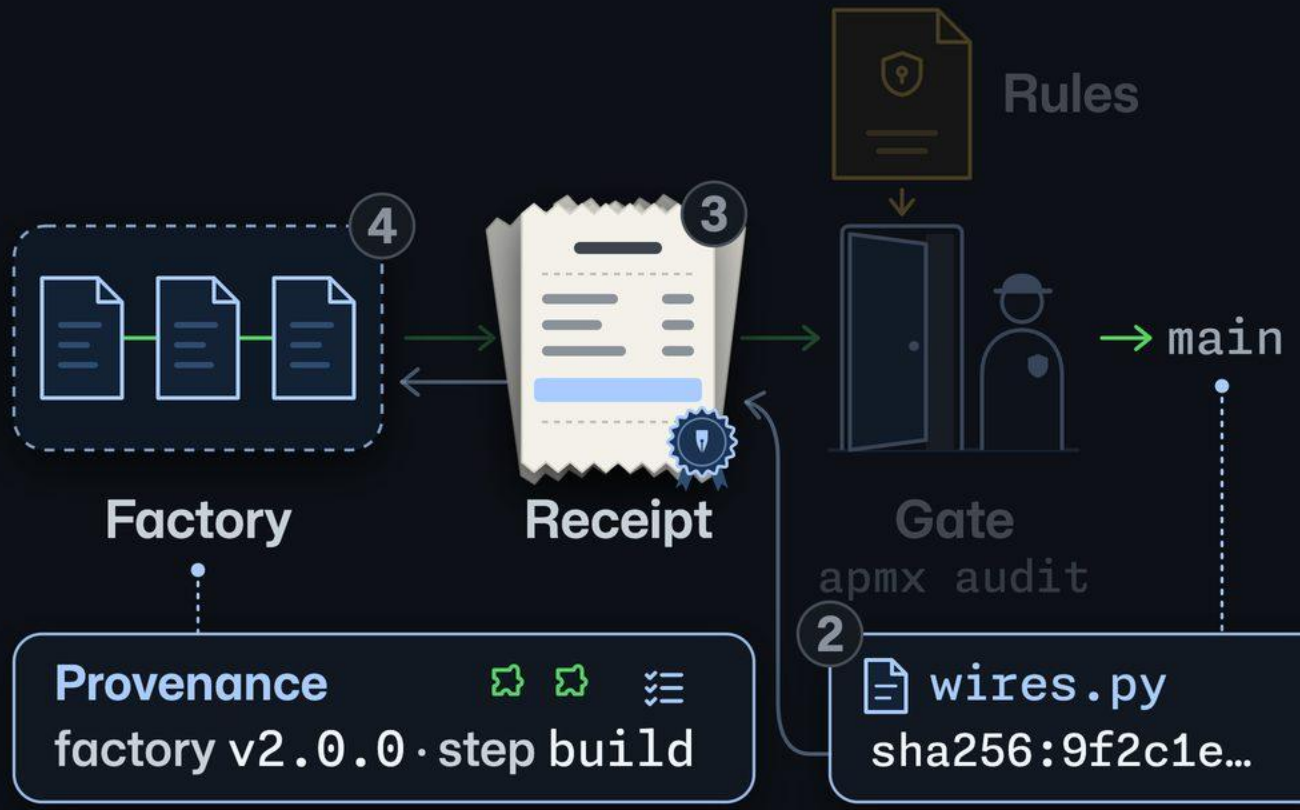
Generation proposes.
Verification evaluates.
Policy authorizes.

Monday morning.
Questions land.



Traceability

Production broke.
Which factory, which step?

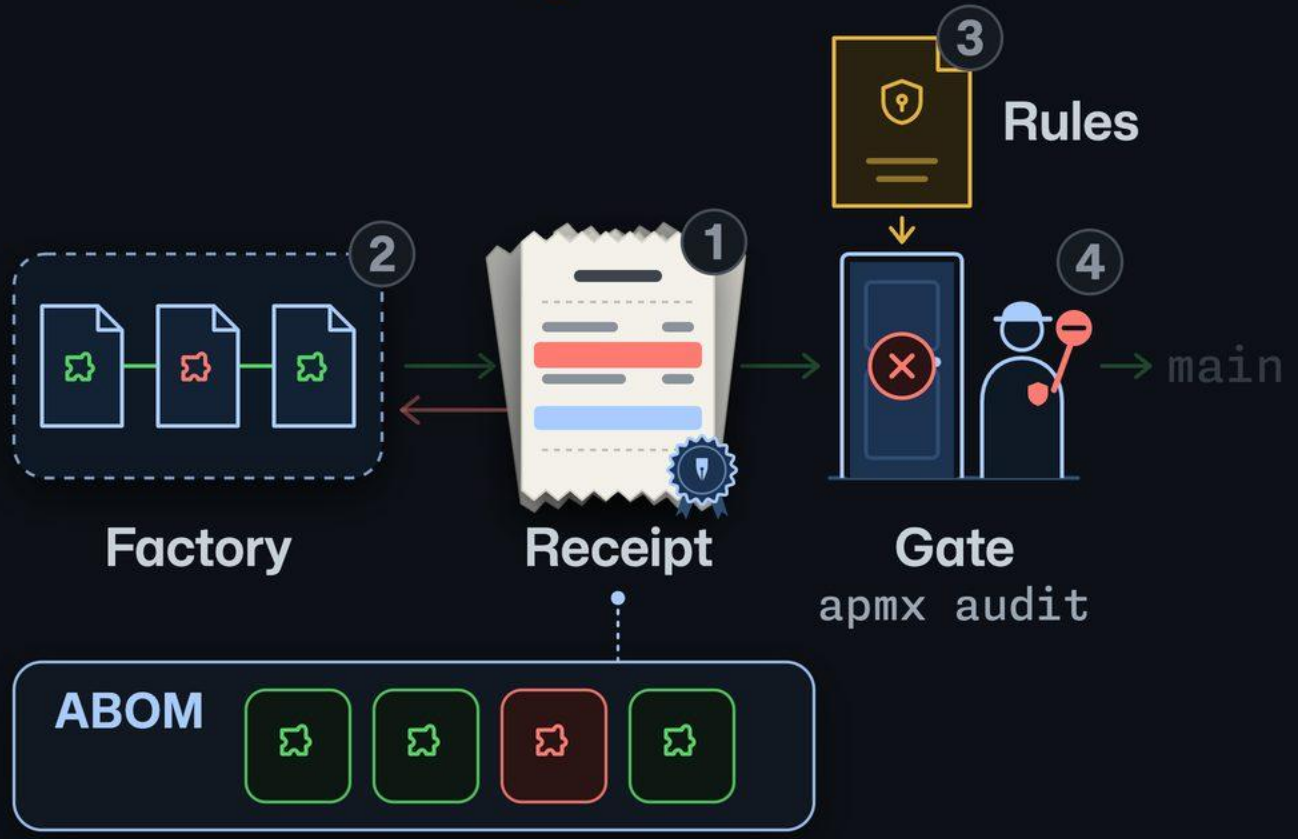


1. **The shipped file**
   what's in production
2. **Its fingerprint**
   a SHA-256 hash
3. **Look up its receipt**
   CI stored it under that hash
4. **Its provenance**
   factory, step, skills, checks

Traced to the factory, step and skills behind it.

Security

A skill was compromised. Who's exposed?



1. **Every receipt's ABOM**
   the skills each delivery used
2. **Find who used it**
   factories and deliveries
3. **Revoke it in the rules**
4. **The gate blocks new ones**
   merged code isn't recalled

☑ **Patch the skill. New factory version.**

Verification

The checks passed.
Did the agent change them?

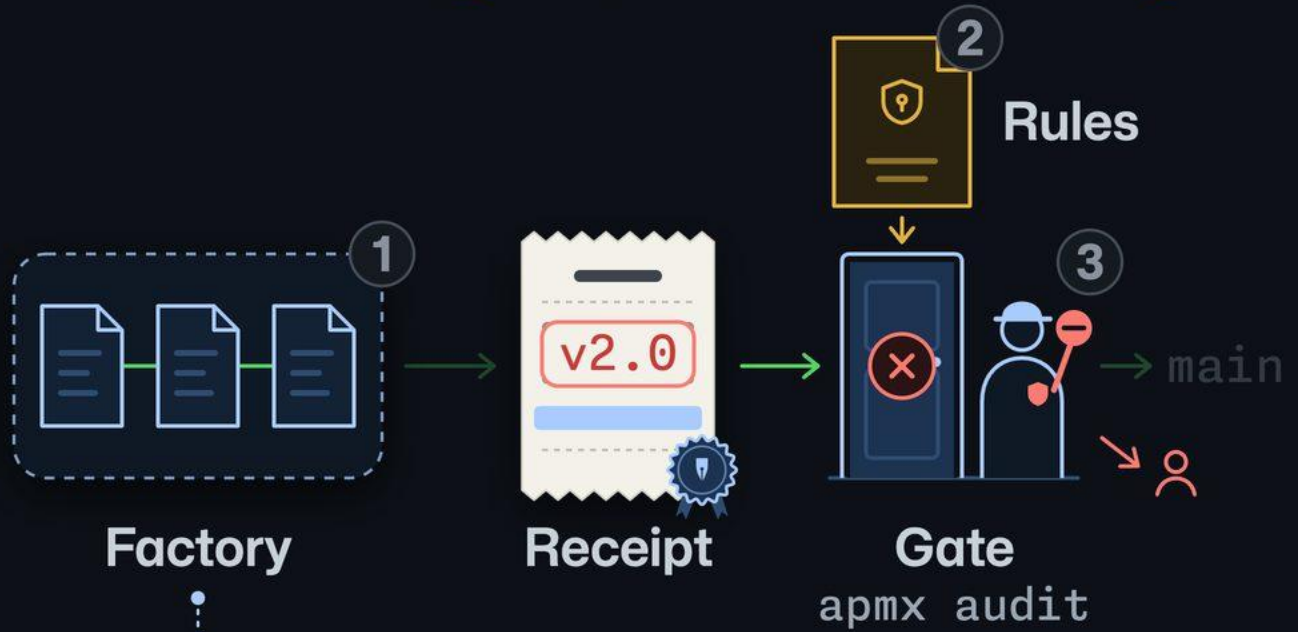


1. Checks live in the contract
2. Run apart from the agent
3. Gate needs every check

☑ No. Judged by the checks you approved.

Enforcement

We added payment checks. Are they mandatory?



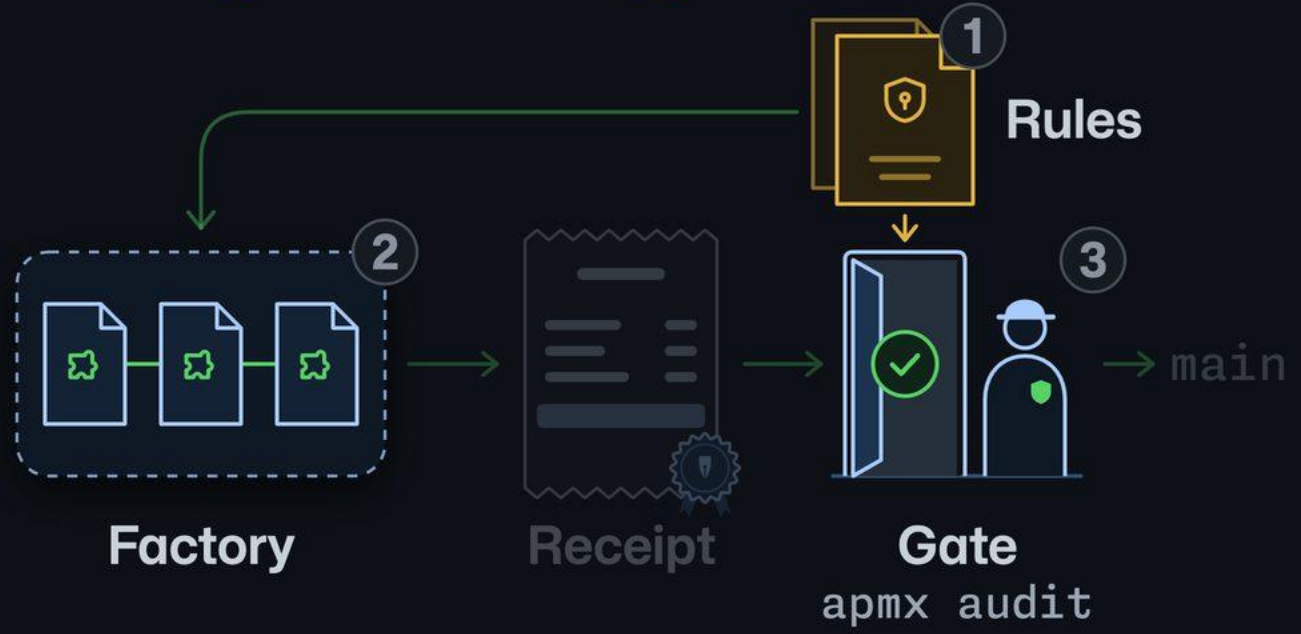
1. New checks, new version
2. Rules allow only v2.1
3. The gate turns away v2.0
   back to a human

New version
v2.0 ⟶ v2.1 + checks

Yes, from the next merge.

Enforcement

Which skills and factories may our agents use?



1. **One set of rules, org-wide**
   owned by security
2. **Approved skills, factories**
   checked at install
3. **Checked again at the gate**
   on every merge

Set once. Enforced at install and merge.

Merge on evidence.

Today

USD 10K+ wires
core payments
→ CI reruns all → Human review → main

With its receipt



USD 10K+ wires
core payments
→ Gate + your tests — review skipped → main

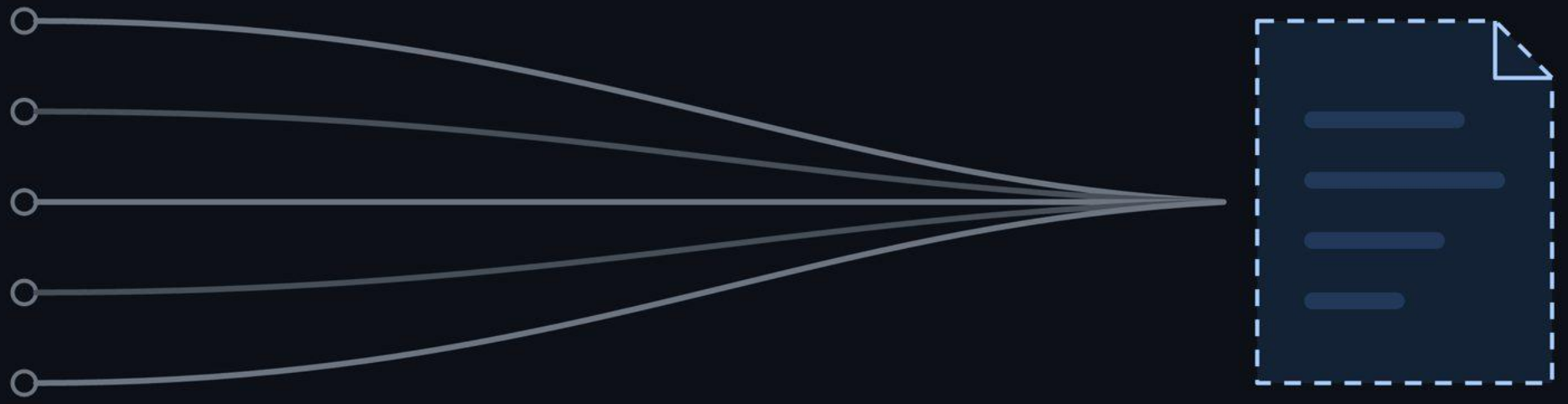
No trusted receipt

USD 10K+ wires
core payments
→ Gate: blocked → Human review

Live demo.
Let's see if it holds.

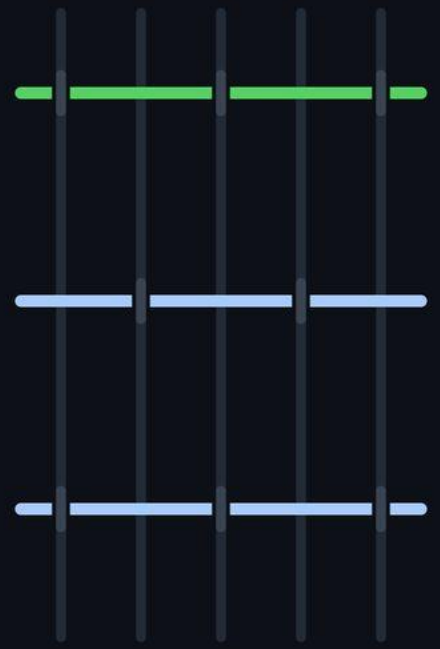
🏛 Contoso Bank · core payments

Many of us are converging here. What's missing is a shared format.



Let's weave it in the open.

1. **Define factories as code, in an open format.**
   Agent Contracts: portable across teams and coding agents.
2. **Version them. Emit standard evidence.**
   SLSA · in-toto · CycloneDX ABOM — any tool can verify.
3. **Gate on evidence.**
   Fewer CI reruns. Fewer PR reviews.



Help shape an open format for agent contracts and their receipts.
Draft 0.1 is open: read it, break it, say what's missing.

Version your factory, not just your code.

Contoso Bank · core payments

Hold wires of USD 10K+ for internal approval

delivered by a **trusted factory**



Verification · Enforcement · Traceability · Security



Let's talk. Questions?

theainativemind.com/aidevcon

slides · essay · format draft 0.1

Daniel Meppiel

Author, The Agentic SDLC Handbook

Creator of APM and APMX

Thank you

