Audit asks 'who approved this?' The only honest answer: 'the stack did.'

Dr Peter McCann Strain
CTO, DPhil/PhD in AI from Oxford University

Swipe ››

— THE FRAMEWORK GAP

The thing on the screen is not the thing being bought, trusted, governed or blamed. The deployed system is a stack.

— THE REFRAME

Stack, not model.

THE OLD QUESTION

Is the model trustworthy?

↓ ↓ ↓

THE QUESTION THAT HOLDS UP

Is the stack accountable?

—— WHY THE FRAME HOLDS

~30 hr outage

A coding agent deleted a production database and its backups in nine seconds; the resulting outage ran roughly thirty hours.

SOURCE

Railway architecture write-up and the OECD.AI incident monitor, which summarised the resulting outage as roughly thirty hours.

— HOW IT WORKS

Capability is not trustworthiness.

01 Read the **model** as capability only: the possibility of useful work.

02 Audit the **stack** for what catches errors, contains mistakes, bounds escaped harm.

03 Locate **trust** (a tested boundary plus a named owner), not in the model's score.

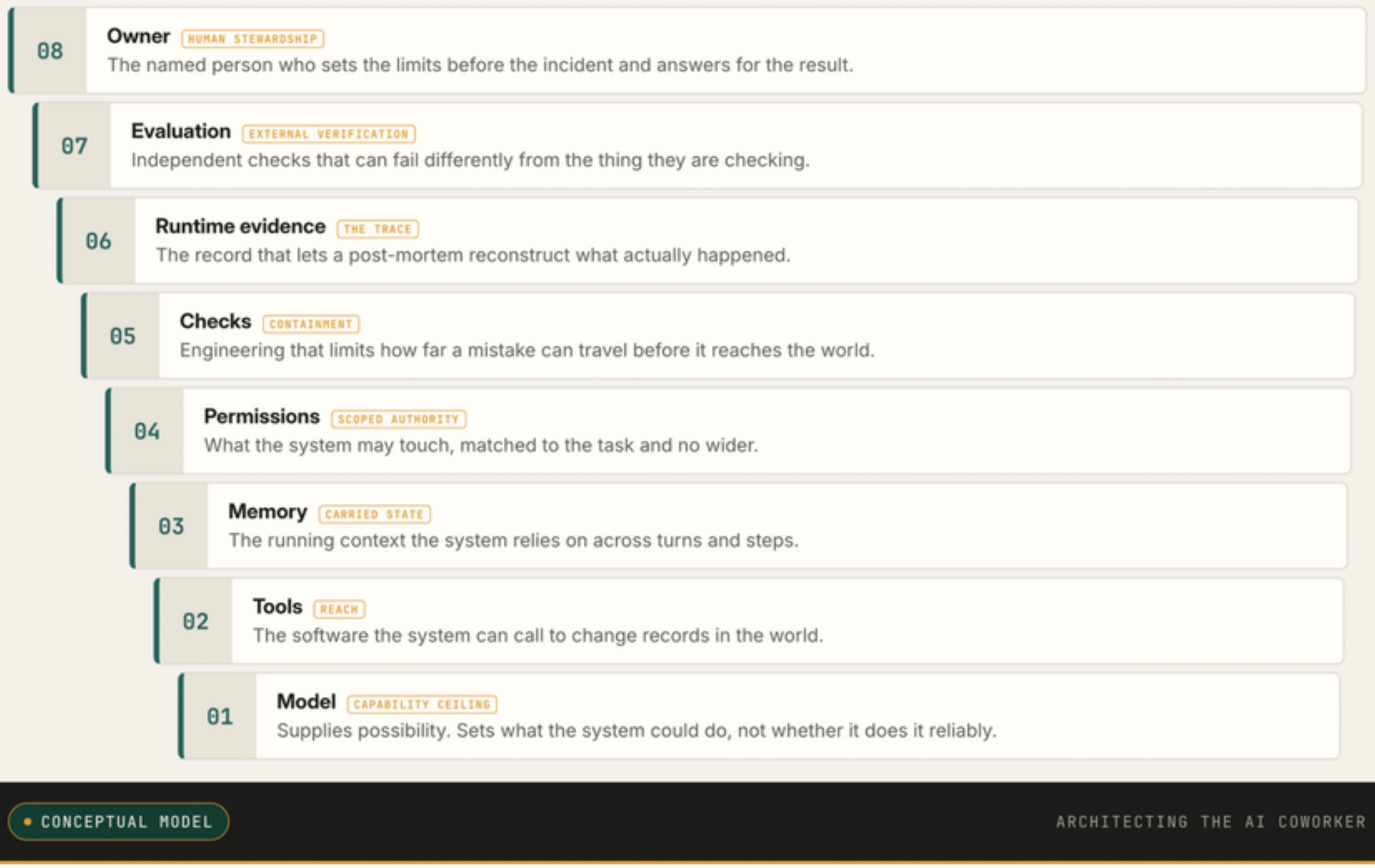
— THE ARTIFACT

Eight things, accountable as one.

— FIGURE E01.1 · E01 - THERE IS NO "IT"

There is no "it". There is a stack.

The model is one component. Seven more produce, check, contain and own the system's behavior.



• CONCEPTUAL MODEL ARCHITECTING THE AI COWORKER

Model, tools, permissions, memory, checks, runtime evidence, evaluation, owner: eight components working together, accountable as one whole.

— APPLY THE INSTRUMENT

Pick one AI system near your work this week. Rewrite "it did X" naming the generator, tool, verifier, authority boundary, evidence trail and owner. Any role you cannot name is the next thing to design.

WHERE TEAMS MISREAD IT

Evaluating the model, then deploying the stack. The score covers capability only; the stack decides whether mistakes escape. Judge the same object you ship.

— READ THE FULL FRAMEWORK

There Is No "It"

Read the full essay — the argument, the sources, the figures and a reader-ready working artifact.
Substack petermccannstrain.substack.com · Medium @peter.mccann.strain · LinkedIn peter-strain-dphil-15a607128
New essays twice weekly, 2 June – 21 July 2026.

Next: E02 — The Coworker Illusion

— THE STACK SO FAR

E01 · Essay 1 of 22 complete · Arc I: See the object

YOU JUST ADDED

The stack map

STACK LAYER LIT UP

Whole stack

YOU CAN NOW ASK

see the AI coworker as a stack, not an "it".

NEXT

E02 asks how the coworker word collapses roles, permissions, checks and ownership.

— THE ARTIFACT, CONTINUED

Eight things, accountable as one.

THE REMAINING NODES

Model and tools: capability and the actions it can take in the world.

Permissions and memory: what it is allowed to touch, and what it carries to the next call.

Checks and runtime evidence: independent checks, and the trace that records what happened.

Evaluation and owner: how capability is measured, and who answers for the result.

Dr Peter McCann Strain

CTO, DPhil/PhD in AI from Oxford University

I build production AI systems and write about making agentic AI useful, inspectable, governable and safe enough for real work.

Follow on Substack for the full 22-essay series
petermccannstrain.substack.com