An agent bought the wrong shoes. The charge cleared. Who acted?

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

Swipe ››

— THE BUYER RISK

When an agent can buy, send, file or delete on your behalf, "did the user click?" stops being enough. The system needs an identity chain, and user permission is not the same thing as platform authorisation.

— THE REFRAME

Authority, not clicks.

THE OLD QUESTION

Did the user's account produce the instruction?

↓ ↓ ↓

THE QUESTION THAT HOLDS UP

Whose authority travelled through the action, link by link?

— WHAT TO ASK FOR

1 Jan 2026

California AB 316 took effect 1 January 2026: a defendant in a covered AI-harm case may not assert that the AI autonomously caused the harm. The identity-chain stops being a curiosity and becomes a litigation defence.

SOURCE

California AB 316 (signed 13 October 2025, effective 1 January 2026, leginfo.legislature.ca.gov); Amazon v. Perplexity preliminary-injunction order (N.D. Cal., 9 March 2026, stayed pending Ninth Circuit appeal, docket No. 26-1444).

— CHECKLIST LOGIC

User permission and platform authorisation are separate principals.

01 Separate user **permission**: the account, data and payment instrument the user lends.

02 Verify platform **authorisation**: the receiving system can still refuse the agent.

03 Refuse architectures that fuse the two. One fused **principal** (the named party acting under its own authority) becomes a **hidden principal** (an unnamed party whose authority is silently spent).

— THE ARTIFACT

The Agent Identity Chain: five visible links, two behind them.

— FIGURE E18.1 · E18 - WHEN AN AGENT ACTS, WHO ACTED?

The Agent Identity Chain

Seven links every consequential agent action must answer from credentials and logs, not memory. Break any one and you cannot trace authority across it.

01 · LINK 1
Who authorized it
A human or organizational principal, bound to a credential and a time - not a vague "the account did it".

→

02 · LINK 2
What it was allowed to do
A scoped mandate: the amount, merchant, tool, time window, data classes, and limits on anything irreversible.

→

03 · LINK 3
Which tool acted
The exact integration, API call, browser session, sub-agent or workflow step - identified specifically enough to point at.

→

04 · LINK 4
What data it touched
A record of the inputs, retrieved context, sensitive fields, and anything disclosed onward.

→

05 · LINK 5
What decision led to the action
A decision record reconstructable from logs and evidence - not only the agent's own self-report.

→

06 · LINK 6
Was it reversible
A refund, undo, recall, delete or escalation route - or an explicit marker that the action was final.

→

07 · LINK 7
Who is accountable
The named party that owns the loss, the dispute handling, the compliance response and the remedy to the user.

PROPOSED OPERATING TEMPLATE · not a formal standard ARCHITECTING THE AI COWORKER

Five core links: authorising human, scope, acting tool, data touched, decision. Two more links (reversibility and consequence owner) sit in the supporting detail; together the seven links match the Monday worksheet. Amazon v. Perplexity is cited only as a preliminary-injunction order, stayed pending Ninth Circuit appeal.

— ASK THIS ON MONDAY

Pick one consequential agent action this week. Answer the seven identity-chain questions: who authorised it, what scope was allowed, which tool acted, what data it touched, what decision led to it, was it reversible, who owns the consequence.

VENDOR TRAP

Pasting OAuth scopes into the design doc and treating that as the identity chain. User consent does not bind the receiving platform. Get the platform's authorisation token before the agent crosses the boundary.

— USE THE CHECKLIST

When an Agent Acts, Who Acted?

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: E19 — Compliance Is Not a PDF

— THE STACK SO FAR

E18 · Essay 18 of 22 complete · Arc IV: Proof and accountability

YOU JUST ADDED

The Agent Identity Chain

STACK LAYER LIT UP

Permissions / Runtime evidence / Named owner

YOU CAN NOW ASK

trace who acted through authority, tool, data, decision, reversibility and owner.

NEXT

E19 asks how compliance becomes runtime evidence rather than a PDF.

— THE ARTIFACT, CONTINUED

The Agent Identity Chain: five visible links, two behind them.

THE REMAINING NODES

Reversibility: was the action recoverable, and through what mechanism? Consequence owner: which named party answers if the action was wrong? Cross-platform record: which receiving systems can confirm or deny the agent's standing?

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