

# Most production AI failures are confident handoffs at the wrong moment.



**Dr Peter McCann Strain**

CTO, DPhil/PhD in AI from Oxford University

Swipe >>

---

— THE OPERATING RITUAL

**The missing layer is not termination; models stop generating text easily. It is a stopping rule: when the work is enough, unsafe, too costly, too uncertain, looping, or no longer the agent's decision to make.**

— THE REFRAME

# Stopping, not termination.

THE OLD QUESTION

**Has the model finished speaking?**



THE QUESTION THAT HOLDS UP

**Is the work enough, safe and still the agent's decision to make?**

## — THE CONTROL SIGNAL

# \$5,000 Rule 11 sanction

**In *Mata v. Avianca* a legal brief reached a federal court citing fabricated cases, and Judge Castel sanctioned the lawyers and their firm \$5,000 Rule 11 sanction after the model was asked whether the cases were real and said yes.**

**SOURCE**

Mata v. Avianca sanctions order (S.D.N.Y., 2023); TheAgentCompany benchmark; Alexey Grigorev's DataTalks write-up and the OECD.AI incident monitor; Railway's PocketOS post-incident notes.

## — CADENCE

# A stopping rule is a disjunction of six criteria.

- 01 Fire any of six **stop signals**: goal, verification, budget, marginal value, loop, handoff.
- 02 Name an **owner** per criterion; a max-turns cap is a seatbelt, not policy.
- 03 **Tier the stop** by action class, so irreversible work stops at consequence's edge.

## — THE ARTIFACT

# Four core stopping criteria, four named owners.

**Goal met**

Stop when the task is actually done

**Verification passed**

Stop only after checks clear

**Budget reached**

Stop before spend becomes authority drift

**Marginal value collapsed**

Stop when more work adds noise

**Stuck loop detected**

Stop repeated failure paths

**Human handoff**

Stop when it is no longer the agent's call

*Four core criteria: goal met (owner: product owner), verification passed (owner: evaluation owner), budget reached (owner: platform owner), human handoff (owner: business or risk owner). The two harder cases (marginal-value collapse, stuck-loop detection) sit in the supporting detail with their owners. An empty owner cell is the bug.*

---

— PUT IT ON THE CALENDAR

**Choose one workflow this week and write the six stop signals in the margin. Mark which are enforced by code, which by process, which only hoped-for in the model saying 'done'. Convert the hoped-for to enforced before the next run.**

---

— RITUAL DRIFT

**Raising the max-turns cap when runs misbehave. A cap is not goal verification. Add the missing signal (verification check, budget gate or handoff trigger) before extending the leash.**

— RUN THE OPERATING LOOP

# Agents Don't Know When To Stop

Read the full essay – the argument, the sources, the figures and a reader-ready working artifact.

Substack [petermccannstrain.substack.com](https://petermccannstrain.substack.com) · Medium [@peter.mccann.strain](https://@peter.mccann.strain) ·

LinkedIn [peter-strain-dphil-15a607128](https://peter-strain-dphil-15a607128)

New essays twice weekly, 2 June – 21 July 2026.

Next: [E15 – Show Me the Run](#)

## — THE STACK SO FAR

E14 · Essay 14 of 22 complete · Arc III: Runtime control

YOU JUST ADDED

**Stopping criteria**

STACK LAYER LIT UP

**Checks / Named owner**

YOU CAN NOW ASK

**write stopping criteria for agent work.**

NEXT

**E15 asks what a single complete run reveals that a benchmark cannot.**

---

— THE ARTIFACT, CONTINUED

## Four core stopping criteria, four named owners.

### THE REMAINING NODES

Marginal-value collapse: each new step adds less than its cost; owner is the orchestration owner.

Stuck-loop detected: the agent repeats a state with no progress; owner is the runtime owner, not the model.

'Loop detected' operational definition: same tool call with same arguments more than N times, or a state revisit ratio above a threshold.



# 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](https://petermccannstrain.substack.com)