

By Ethan Lawrie

# Session 6 - Staying Current

PRACTICAL AI LITERACY FOR WORK



# Things Will Move

---

## What this means

- You need a owner who is named for each workflow
- You need a cycle to recheck the workflow or prompts.
- You need permission to pause the workflow when it stops being safe



## Models will change

- The same prompt can give a different answer next month
- Format can drift, tone can drift, and facts can drift
- You cannot assume it will behave the way it did in training



## Policy will change

- HR rules, privacy rules, pricing rules, disclosure rules
- Something that was allowed last quarter can become blocked with no warning
- You are still responsible

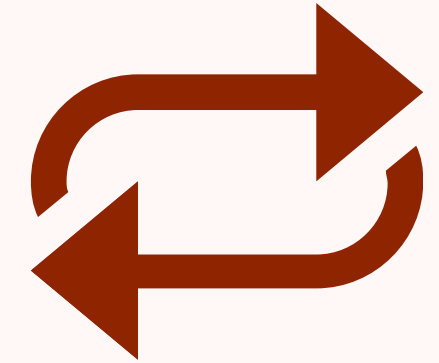


## People will change

- New staff will run the workflow
- They will not know the guardrails unless they are written into the workflow card

# The Maintenance Loop

---



- 01 Sample**
- Take 5 recent runs
  - Use real inputs and real outputs

- 02 Check**
- Accuracy. Facts match the source
  - Safety. No private or restricted content leaked
  - Format. Output still matches the contract

- 03 Record**
- Write down any defects you found
  - Write how they were fixed
  - Keep this log

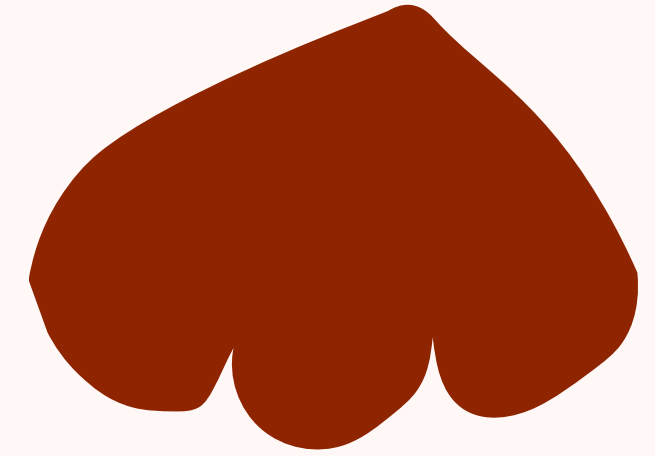
- 04 Decide**
- Keep running
  - Update the prompt or the workflow card
  - Pause the workflow

# Where to get Updates on AI Advancements



Where to get Updates on AI Advancements

# Inside Your Workplace



## **Your workflow plans**

- These are the source of truth. If someone changes a step, that plan must change.

## **Your policy or compliance lead**

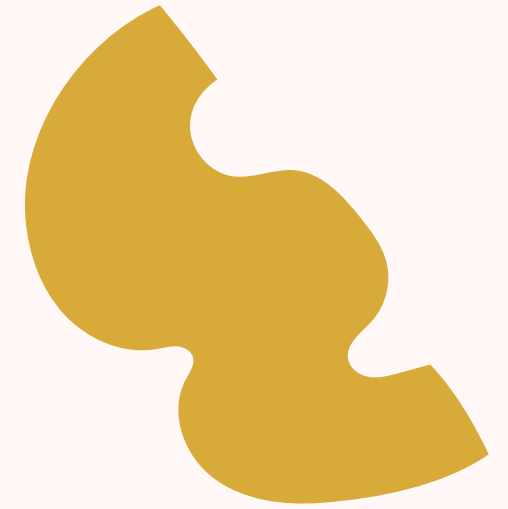
- They know what data is allowed to be used and what is now restricted.
- Policy and privacy rules can change and you are still responsible.
- Trusted guidance treats human oversight, safety, privacy, and accountability as ongoing duties, not one time checks.

## **The workflow owner**

- Each live workflow needs an owner. That owner decides when to pause the workflow if risk appears.

Where to get Updates on AI Advancements

# External & Reliable Sources



## National standards bodies

- NIST publishes the AI Risk Management Framework which is useful to know.
- It focuses on keeping systems trustworthy by managing risk through four functions: govern, map, measure, manage.
- It is updated and supported with a public playbook and guidance, and is meant for anyone using AI, not just engineers.
- Use this to justify audits, stop rules, and pause conditions.

## International policy groups

- OECD AI Principles set expectations for transparency, safety, human oversight, and accountability.
- Governments including Australia reference these to justify things like human approval before anything leaves the organisation.



Where to get Updates on AI Advancements

# What Not to Trust



- Unofficial tips that tell you to skip human review
- Content that suggests guessing missing info instead of stopping
- Hype about AI speed with no mention of traceability or accountability
- NIST and OECD both stress traceability, human review, and the ability to override or shut down unsafe behaviour.

# Final Activity: Capability Check







Thank you!