

Building trust through responsible AI

Julian Vido

AI Safety Lead, MYOB

Mona Saadeldin

Manager of Strategic Initiatives, National AI Centre

Responsible AI?

Responsible AI is the set of technical and human practices that make AI trustworthy. It spans considerations into how a system is designed and built, to how it is governed and overseen

Building Trust in Three Parts

01

Understanding AI

What AI is, why we need to build trust in it, and the role responsible AI plays in meeting this moment.

02

Responsible AI: The building blocks of trust

National AI Centre, Australia's National AI Plan, and the Guidance for AI Adoption.

03

Trust beyond frameworks

Responsible AI beyond the Guidance for AI Adoption.
The role of human systems and relationships in building trust in AI.

PART 01

Understanding AI

The origins of today's AI and why trust is a central question.

Technology and trust across human history

The idea behind AI: giving knowledge and intent to an inanimate object, has preoccupied humans for millennia.

In Greek myth, Hephaestus forged golden assistants to help run his forge. Autonomous helpers made to do open-ended work.

“In each of them was a mind... a voice and strength, and they knew the works of the immortal gods.”

Homer, Iliad 18.417–21

Across history, our stories have carried the same concern: that human inventions must be handled carefully and wrangled into trustworthiness.



The Sorcerer's Apprentice illustration by F. Barthe c.1882

A rapid-fire history of AI

1956

Dartmouth

John McCarthy coins “artificial intelligence”, deliberately departing from cybernetics.

70s–80s

AI Winters

Two periods of defunding as the rule-based approach stalls and promises go unfulfilled. They can't handle edge cases and scale poorly.

→ 1990s

Symbolic AI

Expert systems encode hard rules by programming.

2010s

The Neural Network Revolution

Neural networks +
Abundant data +
GPU compute +
Backpropagation algorithm

The result: instead of teaching a computer the rules by hand, we trained them to build probabilistic models from data.

Today's AI probabilistically calculates the most likely answer.

Today's models aren't retrieving facts from a database. They work out their own mathematical model from vast amounts of data, which they use to generate the most statistically likely output.

That is what produced the ChatGPT moment of 2022, and the systems have improved dramatically since. But at their core they still produce a best guess — an extraordinarily accurate one, but a probabilistic guess all the same.



Today, these systems are genuinely capable

The best AI can now do methodical, multi-step work that was thought to be beyond it just a few years ago:

- **Categorise transactions and extract data from documents**
- **Reconcile accounts against the bank feed, flagging mismatches**
- **Forecast cash flow**
- **Surface patterns and make observations we might have missed**

The advances in a short time have been genuinely remarkable.

But capability is uneven — and understanding how is the key to using it well.

‘Navigating the Jagged Technological Frontier’

A Harvard Business School study of management consultants using AI, published in Organization Science.

See: <https://pubsonline.informs.org/doi/10.1287/orsc.2025.21838>

~25%

faster, with quality graded
significantly higher

1 person + AI

=

the output of a team

The largest gains went to those who had been performing worst.
AI lifted the whole field.

Used in the right place, and with the right controls, the
capability is seriously impressive.

The sting in the tail: on tasks that fell outside what the AI was good at, the people using it did worse than those with no AI at all.

MBABench: Evaluating LLM Agents on End-to-End Spreadsheet Tasks in Finance

MBABench (Columbia Business School, 2026) tests AI on building complete financial models end to end. The complex, daily work of finance professionals.

See: <https://arxiv.org/html/2605.22664v4>

69

/100 for the best system, against
professional standards

53

/100 once tasks reach real-world
difficulty

The outputs looked polished and the figures correct.
However, the workings were fragile: models broke when
assumptions changed, and the logic was often too tangled for a
reviewer to follow.
Performance is unreliable beyond the frontier – use with caution!

Knowing where to lean on AI, and where to check it



INSIDE THE FRONTIER

Lean in here

Constrained, high-volume tasks with a checkable answer:

Categorising transactions
Reconciling the bank feed
Extracting document data



AT THE EDGE

Keep a human in the loop

Open-ended, judgement-heavy work:

Models built from scratch
Analysis under uncertainty
Final advice a client relies on



RESPONSIBLE AI

How you navigate

You can see the shape, not every edge.
RAI is the compass:

Know who is accountable
See where the risks sit
Monitor how it performs
Decide: trust vs check

MYOB's Bi-Annual Business Monitor

SMEs see the AI opportunity and are chasing it

~40%

Of ANZ SMEs are now using AI

×2.8

Faster growth for SMEs using AI

50%+

Of AI users report real time savings

MYOB's Bi-Annual Business Monitor confirms the impact: adoption is reaching critical mass. But we're also seeing value lag adoption — and it isn't the tools that are the blocker. It's trust.

See here: <https://info.myob.com/hubfs/INTL-1253%20MYOB%20Bi-Annual%20Business%20Monitor-compressed%201.pdf>

Trust, attitudes and use of artificial intelligence: A global study 2025

It's not the tools. It's trust.

University of Melbourne & KPMG, 2025: nearly 50,000 people across 47 countries. Trust, not capability, is what gates adoption.

3 in 5

People in advanced economies are unsure about trusting AI

4 in 5

Would trust AI more with assurance mechanisms in place

1 in 10

Australian SMEs have a responsible-AI policy

People's biggest concern is safety, security and responsible use. Assurance mechanisms (**human oversight, accountability, monitoring, recognised standards**) are exactly what close that gap.

See here:

https://figshare.unimelb.edu.au/articles/report/Trust_attitudes_and_use_of_artificial_intelligence_A_global_study_2025/28822919?file=54013232

Sources: University of Melbourne & KPMG, Trust, attitudes and use of AI, 2025; MYOB Bi-Annual Business Monitor; ASIC, 2024.

Responsible AI is the thread through both halves of the story



Navigate the frontier

Know where to lean on AI, and where to check it.



Close the trust gap

Build the assurance that unlocks adoption.

Building Trust Through Responsible AI

AFAANZ Conference 2026

Mona Saadeldin | National AI Centre

7 July 2026



Australian Government



The opportunity

1

More time for what matters

AI takes on the routine work, freeing you up for deeper client conversations and higher-value advice.

2

Sharper insights, faster

AI helps you spot patterns and trends you'd otherwise miss.

3

A real competitive edge

Practices that adopt AI well move faster than those that don't.

Trust unlocks that advantage!

The trust gap

50% of Australians already use AI regularly — but only 36% trust it.

~⅔ of Australian SMBs use at least one AI tool- but 78% are concerned about negative outcomes.

1% have “complete trust” AI will be used responsibly

85% support government regulation of AI

70% would feel more comfortable with AI if regulation was in place

Trust, not technology — is the real adoption barrier

Sources: NAIC AI Adoption Tracker, ai.gov.au (50%/78% figures); KPMG/University of Melbourne, Trust, Attitudes and Use of AI 2025
Sources: [Deloitte Access Economics; The AI edge for SMBs- Amazon 2025](#)

The National AI Centre

1

The Australian Government's lead AI body

Helping organisations understand and use AI safely.

2

A practical focus, not a regulator

Moving businesses from curiosity to real, value-add use — especially SMEs, where AI adds the most value but confidence is often lowest.

3

Delivery arm of the National AI Plan

Capture the opportunity. Spread the benefits. Keep Australians safe.

New: ai.gov.au — Australia's platform providing , guidance, tools & free templates

Essential practices

1

Decide who is accountable

Every AI tool needs a named human owner — before something goes wrong, not after.

2

Share essential information

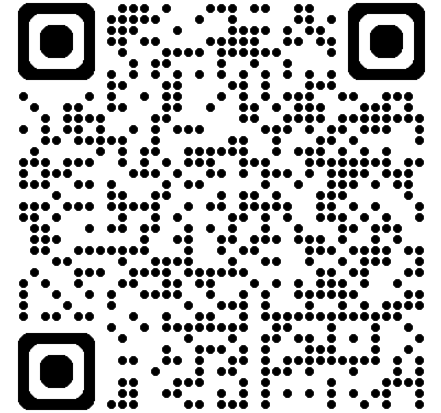
Team, clients, regulators — different depth, same transparency about where AI is used.

3

Understand impacts, plan accordingly

Map who's affected before you deploy — not after.

Example: A firm's AI tax tool consistently over-recommends deductions for sole traders — a gap in the training data.
Fix: retrain the model, add mandatory human review for that cohort, and flag existing clients to double-check past assessments.



Essential practices

4

Test and monitor

AI trained on last year's rules can quietly drift when the rules change.

5

Measure and manage risk

Not every AI use case is equally risky — don't treat them the same.

6

Maintain human control

Match oversight to how high the stakes are.

Example:

Submitting a tax return to the ATO? A human signs off, every time.

Booking a client meeting through a chatbot? Let AI handle it end to end. Same firm, same AI — different level of control because the stakes are different.

Where to go from here

1

Explore ai.gov.au

ai.gov.au has the free policy guides and templates, the AI Adoption Tracker and the AI Directory.

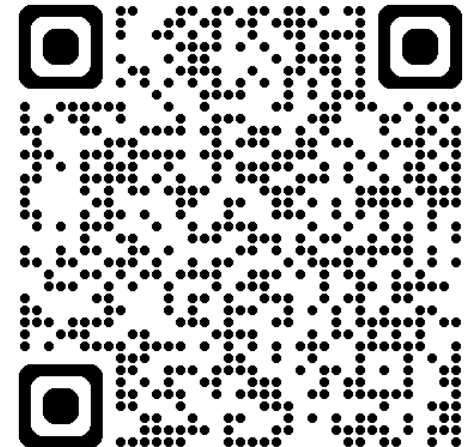
2

Upskill your team

Upcoming AI literacy program for bookkeepers (Pilot) program. Watch this space.

3

Connect with AI adopt centers



Responsible AI is a team sport

We treat the Guidance as the baseline governance and engineering practices that let us ship AI we can stand behind, because it pulls in expertise from across the company.

Accountability sits with a named executive.

Impact is mapped with design researchers, in-house accounting and bookkeeping experts, and voice-of-customer teams.

Risk is prioritised against a bespoke taxonomy built for our context and customers.

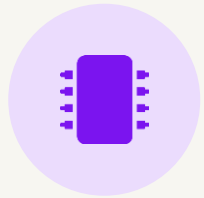
Monitoring draws on data scientists, AI engineers and customer support, so products perform across our diversity of customers.

Transparency means standing in the customer's shoes — with communications, marketing and design — so people don't just know AI is there, they use it well.



**Every principle feeds
the sixth: keeping a
human in control.**

Responsible AI in the context of the AI supply-chain



The engine

Google, Anthropic, OpenAI

Frontier models built to be as powerful and general as possible. Extraordinary raw capability — but the makers, based in California, don't know the reality of a café in Fitzroy North or a butcher in Karori.



The vehicle

MYOB

We're where the rubber hits the road: grounding those models in applications that understand our customers' data, stay compliant with tax and accounting obligations, and get the job done. Raw power made genuinely useful.

Grounding frontier capability in the reality of local business is where the SME upside becomes real.

Making the AI opportunity real for local business

Through a responsible-AI lens, Responsible AI means extending the opportunity to ANZ businesses in a form that genuinely meets their needs — raising AI literacy, giving SMEs cutting-edge tools built for their reality, and never compromising the data-security standards we've held since day one.

We all have a role in shaping the transformation AI is bringing about.

If you have an idea for a responsible-AI collaboration in your teaching or research, come and find me.

FIVE THOUSAND YEARS AGO

The first writing and the first ledger

The first thing humans wrote wasn't a poem or a prayer. It was a receipt: so many measures of grain, so many head of cattle.

The the mark in the clay was only a claim, worth no more than the trust built around it.

Qualified scribes

Trained experts whose mark carried the weight of a qualified hand.

Named parties

Records bound to identifiable individuals, answerable for them.

Held securely

The most important records centralised in temples and palaces.

Certification, expertise, accountability, secure custody.

Across history, we see new human systems of trust being established around new information systems.

Trust always lived around the tool, not in it



Medieval England

The first 'auditors'

Henry I sent auditors across the country to hear barons account, out loud, for the crown's money — the word 'audit' from the Latin for 'to hear'.

Through the industrial age auditing became essential to the modern corporation.



The Italian cities

Double-entry bookkeeping

Double-entry built a check into the books, so entries agreed with one another. But that confirmed accuracy, not honesty.

What let merchants send fortunes across the sea was everything around it: notaries, merchant courts, reputation, and accounts reconciled against a shared method.

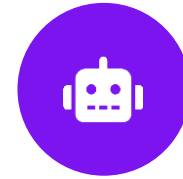
The technique made the books checkable. The people and their practices made them trustworthy.

AI is a technology that *makes* claims, not just records them

Every tool before it recorded a claim a person had already made.

AI is different. It generates the claim itself.

For this technology, the human systems matter more now, not less.



Responsible AI has its own antecedents — and it's how we put a person back behind the number.

Four things you can do



Put a name against every AI output that matters

RAI fails most when a wrong output had no one accountable. Decide who stands behind the number. If it's "no one," that's your first priority.



Set the boundary before you deploy, not after it stumbles

Specify what AI may do alone, what it may only recommend, and what it must never decide alone. Map your frontier.



Collaborate across the profession

Decide together what standards you expect for the use of AI.



Introduce RAI to your students

Graduate them with the judgement to ask the write questions

Thank you

Please get in touch if you have any questions, or wish to discuss.

Julian Vido

AI Safety Lead, MYOB

Contact via: Shailan.Patel@myob.com

Mona Saadeldin

Manager of Strategic Initiatives, National AI Centre

Contact via: naic@industry.gov.au