— Insight

Would your AI work stand up?

Five checks before it leaves your organisation

Quality assurance for AI-assisted knowledge work — A practical guide for reviewers, subject-matter experts and sign-off holders.

ALLEN + CLARKE

AI didn't invent weak analysis, inaccurate claims or missed assumptions.

What it changes is the speed, volume and confidence with which poor-quality work can be produced, making strong quality assurance more important than ever.

Would your organisation's AI-assisted work stand up to the same scrutiny?

The issue: Output is faster (and easier) than quality assurance

Across the organisations we work with, we see a common pattern. An enterprise AI tool is made available. A policy is published. Staff attend training sessions and the expectation is set that a person must check the output. How the work gets done often stays the same. We call that individual value.

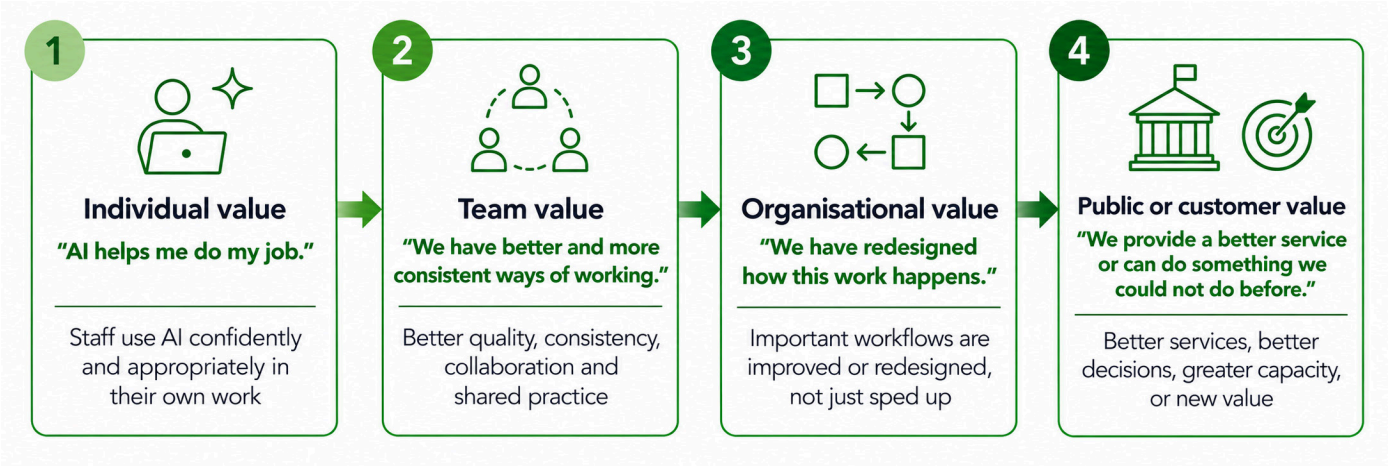


Figure 1. The AI value chain. From individual use through to public or customer value.

This approach ignores one of the main implications of org-level AI rollout: producing seemingly good work becomes fast and requires very little thought. Because the first draft is no longer the time-consuming hard part, the limitation on total output moves to the subject-matter expert or the department head who is responsible for sign-off.

If you're in one of those positions, you will already know the old quality assurance model no longer works because traditional indicators of poor work, bad phrasing, inconsistent logic, etc are hidden by polished AI-assisted writing.

In a recent AI Questions Answered webinar, 85% of poll responders said "people stop checking or thinking" as their biggest worry about AI at work. That's a small sample but a good indicator of how people are feeling.

85% are concerned people stop checking or thinking

Source: Allen + Clarke AI questions answered poll

Remember, AI can produce a plausible answer instantly. Your job is to build a defensible answer deliberately.

Apply the principles of traditional QA to your AI-assisted work

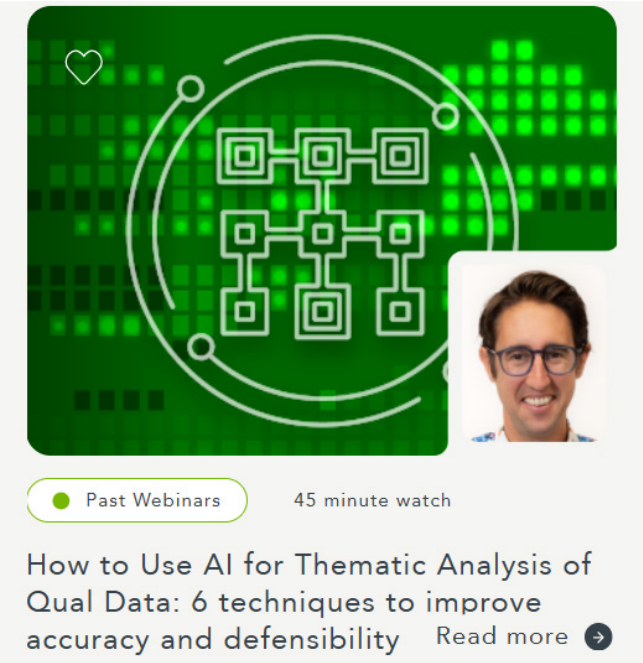
Before AI, quality assurance (QA) happened naturally throughout the work. Writing required people to return to sources, check details, test assumptions, consider limitations, and discuss difficult points with colleagues. Those activities took time, but they also made the final product better.

Take thematic analysis as an example. Traditionally, an analyst would read interview transcripts, code or label the important ideas, group those codes into themes, compare the themes with other evidence, and then write the report. Each stage created a natural opportunity to question the work. AI can seemingly complete all those stages quickly. You can give it the transcripts and ask for a finished report. The result may look polished, but the final reviewer must work backwards to determine whether the interviews were coded correctly, whether the themes reflect what people said, whether contradictory views were lost, and whether the conclusions are supported.

A better approach is to use the principles of traditional QA in your AI-assisted work. Review the AI coding before using it to identify themes. Get the interviewers to review the themes against their perception and check the original interview data where themes feel incorrect and so on.

The final quality check then becomes much more manageable. The reviewer is confirming that the report reflects inputs and analysis that have already been tested; they're not trying to reconstruct the entire analytical process from polished text.

Quality assurance therefore needs to be deliberate and iterative. The five checks below will help you decide what needs review, where it belongs, and who should be involved.



Before AI-assisted work leaves your organisation, put it through these five checks.

1 Don't check everything equally: match reviews to the consequence of being wrong

Brainstorming for an internal meeting doesn't need the same review as advice that could affect a person's rights, a funding decision, or the future of a role. The more serious the consequence of being wrong, the more independent and expert the review should be.

The best way to implement this successfully leads into check two, but you shouldn't just stop at redesigning workflows, training your team on when to apply the higher level of review and when it's not necessary is also required.

Figure 2. Matching the level of review to the rising cost of being wrong.

2 Show where AI was used in the workflow

Start with what process was used and show where AI contributed and how.

AI is not suitable for all stages of knowledge work and, in some instances, be used with extreme caution or not at all. Here you want to create a simple map of the steps that went into creating the output and showing where AI was used and where it wasn't. This includes the tasks, inputs, outputs, decisions, and hand-offs.

A reviewer or sponsor shouldn't have to be an AI expert, but they should be able to easily see where and how AI was used throughout the drafting process. This helps them narrow in on sections or datasets where AI use is material to the findings or recommendations and therefore more accurately apply the appropriate level of review.

The best way to do this is to redesign how the work gets done. This is painstaking work that often requires real AI expertise, but it is how you move up the value chain from individual value to team or organisation value. If your process is still "use AI, create a draft, send it to an expert at the end", you haven't redesigned the workflow, you've made one part faster and pushed the pressure upstream.

For more information on how to identify and redesign workflows, our AI Risk Management webinar and resources will help.

3 Are the people using AI ready?

Access to an AI tool and a short demonstration are not capability-building.

Learning to use AI safely is like learning to drive – theory gets you behind the wheel, but practice is what gets you safely on the road. People using AI need to understand more than how to access it and how to prompt (though we advise not to learn how to prompt and do this instead). They need to know what tasks are best for AI assistance, what information they can provide, and where its answers are likely to look good but be wrong.

Workflow redesign is a great hands-on way to do this; so is access to AI superusers or AI team members who have more understanding and capacity to support real work. That's the instructor in the passenger seat we all need when we start learning to drive.

So, how can you tell whether you or someone is ready? Ask them to walk you through one real example of how they used AI using our SAFE framework:

SAFE framework	Four questions to ask about one real piece of work
S: SAFE Where did this information come from? Can you verify it independently?	This is still your critical check even with enterprise AI. The AI doesn't cite sources accurately just because it's running on your tenant. Verify every citation against the original source.
A: Accuracy Is this factually correct?	Cross-check figures, quotes, claims. Enterprise AI reduces privacy risk but it doesn't fix hallucinations. Every AI output still needs human verification. For policy work, check against official documents. For data analysis, verify against source data.
F: Fairness Who could this harm or exclude?	Your enterprise setup doesn't solve bias. You still need to ask equity questions. Whose perspective is missing? Could this entrench discrimination? For community-facing work, this check is non-negotiable.
E: Exposure What data went into AI?	With enterprise AI and proper permissions, this risk is lower but not zero. Still review what information was shared. If it required special permission to release to the AI, document that decision.

You don't need to be an AI expert yourself; the point is to understand to what extent the person (or you) understands the work and can explain the decisions they made.

Readiness is not measured by how often someone uses AI. It is measured by how well they can use it on real work while retaining their understanding and judgement.

4 Check the accuracy of claims and numbers

"It takes me just as long to check as it does to do the work" is a common statement and is what causes most people to stop using AI.

We've already covered some of the mitigation for this above, but most of the time advice relies on accurate analysis of data and checking that does take time because, for every material source, you're asking three questions: does it exist, is it the source being claimed, and is it true to what's written?

We've all had that correct-looking citation that goes nowhere, an AI summary that misses the outliers that make the data useful, or a set of numbers that may or may not be accurate.

The best solution we've found is to use an AI-powered QA tool to help speed this process up. For security reasons, we've created our own, but there are many on the market; one example is Wauldo. These tools tend to work by taking your final document and extracting all the claims and numbers. It then checks if the same claim or data point exists in the associated source material and provides a visual of where the claim or data point is supported. This, of course, does not remove the need to review the judgement applied and the recommendations made, but it does help the reviewer understand to what extent the supporting data exists within the background material.

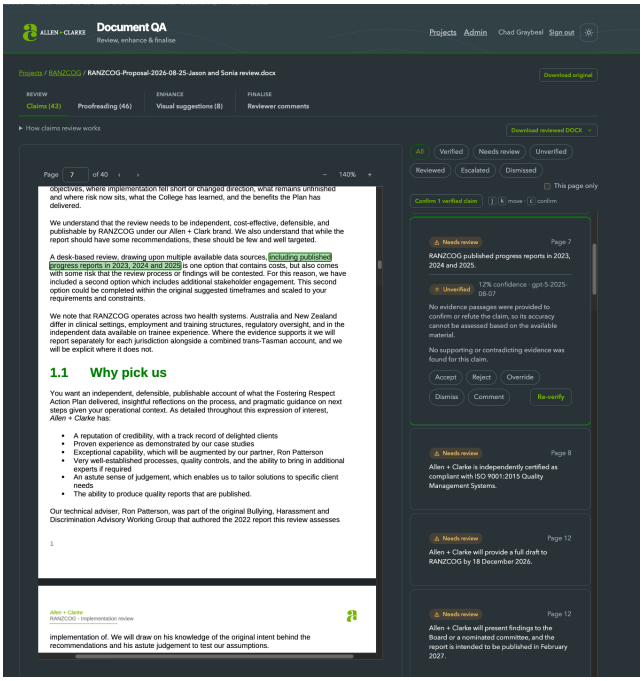


Figure 3. An AI-powered QA tool extracting claims from a document and showing where each is supported in the source material.

5 Ownership and accountability: Whose work is it really?

"Human in the loop" and disclosure of AI are not enough.

We've found the most effective ways to ensure quality come down to ownership and accountability. The question is: how do you support and, in some cases, enforce this?

We apply three levels:

Level 1

Names and positions within the disclosure

When your name is in the disclosure box at the top of the deliverable, you're suddenly more likely to own its contents. You should be proud of the work you delivered and your use of AI is not something to hide, so putting the names of those on the team and the role they played should be standard practice.

Level 2

Record who did what

This goes back to the workflow map. If you know what work was done, then you can also record who did it, helping the reviewer with the next level. This also speeds up the scenario where a mistake was made and rework is required.

Level 3

Use open questions to check understanding

AI use will let you outsource thinking, but it cannot outsource your understanding. Using open questions in person is the easiest way to do this. Start by asking: What options did you consider and reject? What evidence would change your view? Which part of this conclusion are you least certain about?

Then plan for the possibility that something still goes wrong. Who can stop the work? Who corrects the record? Who needs to be told? How will the workflow change so the same failure is less likely next time?

This doesn't mean an audit trail for every casual prompt. Match the record to the consequence. If the work could materially affect people or decisions, you should be able to explain why it was trusted.

Things will go wrong; are you ready?

AI did not invent inaccurate work; people have always been able to (and have) misunderstood sources, used poor assumptions, made spreadsheet errors, and approved work that wasn't ready. The difference is AI makes bad work harder to detect at face value.

That's why individual checks must be part of an organisational response. Make sure you've added AI-related risks to your existing risk register. Give your teams a clear escalation path and permission to pause work when something feels wrong, and create a culture where reporting an AI concern is as ordinary as reporting any other quality issue.

Your incident response plan needs enough detail so everyone can spot and contain the problem, including who must be notified, how the organisation will respond, and what needs to change afterwards. That might mean correcting the work, contacting affected people, or changing the workflow, safeguards, or training.

Most importantly, thank people who raise concerns. If reporting an AI mistake feels like admitting misconduct, problems will stay hidden and the errors will compound.

If you'd like more information, free frameworks, and guides, here are some helpful links.

Past Webinars 45 Minute Watch

AI Risk Management: How to Avoid Things Going Wrong

Read more

Past Webinars 45 minute watch

Leading through AI disruption: the friction holding back your AI adoption

Read more

Past Webinars 1 hour watch

Stop Learning AI Prompting (And Do This Instead)

Read more

Guides Less than a minute to read

Are AI chats subject to the Official Information Act?

Read more

AI use statement: This article was written by Adam Emirali with assistance from OpenAI Codex. Codex was used to search internal source material, help structure, edit, and proofread it. Adam wrote, reviewed and revised the content and is responsible for the final article.