AMANDA CUNNINGHAM | SERGEI PONOMAREV

From AI Induction to AI Accountability

Connecting organisational readiness with customer-facing transparency

Amanda Cunningham, Founder, Savvy Pixel® | Sergei Ponomarev, PhD, Founder, AI Business

Introduction: AI Is Your Territory

AI is intimidating. Not as a technology, but as new, unexplored territory. It seems to have its own language and its own laws, and until you have learned the words “embedding” and “orchestration”, you are not allowed in. Many executives know this feeling, and it is misleading.

The real difficulty with AI is not its complexity. It is something else: in the digital world, almost anything is possible. Any agent, any configuration, any degree of autonomy. This sounds like freedom, but in practice it turns into chaos: when everything is possible, it is unclear what exactly to build, and projects drown not in a shortage of options but in their excess.

So limits are needed. And the frame for those limits comes from your business. You already have services, prices, deadlines and refund rules. You already have people who know precisely what may be promised to a customer and what you would later be ashamed of. That is the frame for AI. It does not need to be invented. It needs to be taken out of people's heads and written down.

You are not required to understand code. You are required to understand your own business.

The code remains the developers' concern: they will tell you what is technically possible and what is not. But what the AI should do, what it should not do, and where it must stop is decided by you alone. AI is only an instrument for solving business tasks. And the business belongs to you; on that territory, you are in charge.

This paper brings together two sides of the same accountability problem. Amanda Cunningham examines what must happen inside an organisation before AI meets the customer. Sergei Ponomarev examines what happens when that AI becomes customer-facing and how the organisation can demonstrate that its promises survived contact with the real world. The argument is shared: responsible AI requires a continuous connection between organisational intent, delegated authority, customer experience, evidence and review.

Stage One: Before AI Meets the Customer

Inducting AI into the Organisation

Amanda Cunningham | Founder, Savvy Pixel®

Capability does not equal authority.

Artificial intelligence has entered organisations unusually quickly. In many cases there was no formal moment at which a leadership team decided that AI would become part of the organisation. It arrived incrementally: through individual accounts, productivity platforms, customer systems, software updates and employees experimenting with new capabilities.

That matters because an organisation can acquire AI capability without ever defining AI authority. Technical capability answers one question: what can this AI system do? Organisational authority answers another: what are we prepared to allow it to do here?

An organisation would not normally appoint a new employee and immediately provide unrestricted access to company information, customer records and operational systems. The person would first be given context: what the organisation does, whom it serves, how decisions are made, what standards apply, what they may access, what sits outside their authority and when somebody else must become involved.

AI is not an employee, but the comparison exposes an important inconsistency. We recognise that people require context and boundaries before they can operate responsibly inside an organisation, yet AI is often introduced principally through access to software and instructions for using it.

For this paper, AI induction means the organisational process of establishing the context, boundaries, permissions, authority and accountability within which AI may be used. This is a practical organising concept proposed here; it is not a claim that existing standards use the same terminology.

The need for organisational processes around AI is well established. The UK Department for Science, Innovation and Technology's AI Management Essentials (AIME) focuses specifically on the organisational processes that enable responsible AI development and use, rather than evaluating AI products themselves. AIME is designed for organisations of different sizes and addresses internal processes, risk management and communication.[1]

1. Context comes before instruction

Much of the early conversation about generative AI concentrated on prompting. Better instructions can improve outputs, but prompting addresses only part of the problem. A prompt tells an AI what a user wants at a particular moment; it does not necessarily provide enough understanding of the organisation within which that work is taking place.

Every organisation operates within a context: customers and stakeholders, products or services, commercial priorities, terminology, policies, contractual obligations, regulatory requirements, risk tolerances, communication standards and established ways of making decisions. Some of this knowledge is documented. Much of it is distributed across systems or resides in the experience of founders, directors, managers and employees.

AI does not automatically possess that organisational understanding. This creates a particular risk because an AI system can produce fluent and plausible output even when important context is absent. An obviously poor answer is relatively easy to challenge; a polished but incomplete answer may be harder to recognise.

The first question is therefore not simply how effectively employees can prompt AI. It is: what does AI need to understand to perform this particular role appropriately? That immediately creates a second question: what does it not need to know?

Responsible context is selective. Information may be confidential, commercially sensitive, personal, outdated, irrelevant or inappropriate for the AI environment being used. The objective should be appropriate context, not maximum context.

Do not give AI a master key to the organisation. Give it an access pass appropriate to the role it has been authorised to perform.

The access-pass principle changes the question from 'What can we connect AI to?' to 'What does this particular use case justify AI having access to?' That is a governance decision, not merely a technical one.

2. Access does not confer authority

Even appropriate access is only part of the induction. The organisation must decide what AI is permitted to do with the information and capabilities available to it.

The phrase 'using AI' is too broad to define authority. There is a material difference between allowing AI to retrieve information, draft content, identify patterns, recommend an action, communicate directly with a customer, update a record, initiate an action, or make or materially influence a consequential decision.

A practical authority test: May the AI inform, draft, recommend, act or decide?

These are levels of delegated authority, not stages of AI maturity. An organisation does not become more advanced merely by allowing AI to act or decide. In many situations the responsible permanent boundary may be Draft or Recommend.

The appropriate authority should be determined by the use case and its potential consequences. Drafting alternative marketing headings is fundamentally different from responding to a vulnerable customer, assessing an employee, interpreting a contractual obligation or influencing a financial decision.

AIME reinforces the underlying organisational discipline: its assessment asks whether organisations maintain records of AI systems, have an AI policy, use that policy to evaluate whether AI is appropriate for a given function or task, and identify clear roles and responsibilities for AI management.[2]

Delegating activity to AI does not delegate organisational accountability. AI may generate an output, recommend an action or execute an authorised process. It does not thereby become the accountable party.

3. Human oversight must mean accountable review

References to a 'human in the loop' can sound reassuring while leaving the important questions unanswered. Who is the person? What are they expected to examine? What authority do they have to challenge the AI? When must they intervene? Who ultimately owns the outcome?

A more useful operational concept is accountable review: identifying the person or organisational role responsible for reviewing, approving, escalating or intervening in AI-supported work where the consequences require it.

This does not mean requiring human approval for every AI-assisted task. Low-consequence uses may require little intervention; higher-consequence activity may require explicit approval, and some uses may be unsuitable for delegation to AI at all.

The underlying accountability principle is supported by the Information Commissioner's Office in the data-protection context. The ICO states that overall accountability for data-protection compliance lies with the organisation acting as controller, and that governance and risk-management capabilities should be proportionate to the organisation's use of AI.[3]

Where personal data is involved, the ICO also emphasises documenting and demonstrating the approach taken, including decisions to use AI and the associated risks. Its current AI guidance is under review following legislative change, so it should be read alongside the ICO's latest updates.[3][4]

4. Boundaries must exist before they can become promises

Before AI undertakes meaningful work, the organisation needs to establish internal boundaries. These do not necessarily require a large compliance function or elaborate AI bureaucracy. They do require organisational ownership.

The greater the potential consequence for a customer, employee or organisation, the more explicit those boundaries should become. The absence of a specialist AI, technology or compliance team does not remove the underlying responsibility.

An organisation cannot credibly tell a customer that AI had limited authority if it has never decided internally what that authority was. It cannot promise that a person will intervene in defined circumstances if no escalation conditions exist. It cannot explain what informed an AI-supported outcome if it does not understand its own information flows.

The public promise begins with an internal decision.

This is where AI induction and customer accountability begin to meet. Internal decisions about purpose, context, access, authority and accountability create the conditions for meaningful external transparency.

5. Induction is not a one-time event

Treating AI introduction as induction rather than installation has another consequence: induction is not the end of management. Organisations change. Customers change. Policies and regulation change. AI systems themselves change. New capabilities appear, integrations are added and employees discover new uses.

Evidence should therefore be gathered proactively as well as reactively. A customer complaint may reveal a boundary that was too loose; a test purchase may show that an escalation route does not work as expected; an incident may expose inappropriate access; and routine monitoring may identify drift before a customer experiences a failure. Evidence may also arrive from outside the organisation through consumer bodies, journalists, regulators or other independent scrutiny.

Re-induction should be proportionate rather than automatic. It is triggered when evidence or material change calls the existing authorisation into question: for example, a significant model or system update, a change in data or integrations, observed drift, a scheduled monitoring cycle, a material change to the use case, or a repeat check following remediation. Context can then be revised, access narrowed or expanded, authority changed, escalation conditions strengthened and accountable review reassigned.

The loop also needs an owner. Leadership should assign a named person or organisational role responsible for ensuring that evidence is reviewed, required changes are made and re-induction occurs when its trigger is

reached. This continuous approach is consistent with NIST's AI Risk Management Framework, which describes AI risk management as continuous throughout the AI system lifecycle and organises its Core around Govern, Map, Measure and Manage.[5]

INDUCT → AUTHORISE → OPERATE → EVIDENCE → REVIEW → RE-INDUCT

This makes responsible AI management a circle rather than a line. The organisation establishes the conditions under which AI operates; evidence from operation, deliberate checking and external scrutiny returns to the organisation and informs the next version of those conditions.

Executive AI Induction Readiness Check

Before AI is authorised for a meaningful organisational use case, leadership should be able to answer eight questions:

Principle	Executive question
Purpose	Why is AI here?
Context	What does it need to know?
Access	What may it reach?
Authority	What may it do?
Boundaries	Where must it stop?
Accountability	Who remains responsible?
Evidence	Can we demonstrate what happened?
Review	What have we learned, and what changes?

An organisation that cannot yet answer these questions may have access to sophisticated AI. It has not necessarily established that the AI is ready for the role it has been given.

From organisational boundary to customer promise

By the time AI reaches a customer, the organisation should already know why it is there, what it knows, what it can access, what authority it holds, where it must stop and who remains accountable.

But an internal promise is not enough. Once AI participates in a customer interaction, those organisational decisions become externally significant. The customer should not have to guess what role AI played, what it was authorised to do, whether a person remained accountable or what evidence exists of the interaction.

AI induction establishes the promise inside the organisation. Customer accountability tests whether that promise survives contact with the real world.

Stage One asks: have we properly inducted and authorised our AI?

Stage Two asks: can we prove to the customer that it behaved accordingly?

Stage Two: When AI Meets the Customer

Proving the Promise to the Customer

Sergei Ponomarev, PhD | AI Business

Article 50(1) of the EU AI Act, applicable from 2 August 2026, places the relevant transparency obligation on providers: AI systems intended to interact directly with natural persons must be designed and developed so that those people are informed that they are interacting with an AI system, unless that is obvious to a reasonably well-informed, observant and circumspect person in the circumstances and context of use.[6]

That is a specific legal requirement, not a general duty on every organisation to disclose every use of AI. Article 50 also contains separate transparency provisions for synthetic content marking, emotion recognition and biometric categorisation, and deepfakes or certain AI-generated public-interest content.[6]

For customer-facing AI, however, the legal requirement can be treated as a starting point rather than the whole customer proposition. A minimal notice may satisfy the applicable disclosure requirement, but it does not by itself tell a customer what the AI is authorised to do, where its authority ends or what happens if something goes wrong.

A company can choose to go further voluntarily: explain why the AI has been introduced, what it can do, what has deliberately been reserved for people, how a customer can reach a member of staff, and what evidence of the interaction will remain.

The distinction matters throughout this section. Where the law imposes a requirement, we identify it as such. The AI Policy, AI Service Passport, AI Receipt and test-purchase method proposed below go further: they are practical accountability mechanisms and voluntary consumer standards unless a separate legal obligation applies.

There are only three:

1. **How does the company use AI,** for what purposes, and under what rules?
2. **What can this particular bot,** the one I am talking to right now, actually do?
3. **If the bot does something wrong,** how will I prove my case?

To answer these three questions, a company needs three documents: an AI Policy, an AI Service Passport, and an AI Receipt.

1. The AI Policy: how things are done here

The first document answers the first question. Before talking to a bot, the customer wants to understand what the company has permitted its AI to do at all. Not within one particular service, but in principle. This is the company's AI policy: one open, public document in a visible place on the company's website.

Let us stress the word “one”. Not thirty pages of terms and conditions where everything important is hidden between the commas. Not a privacy policy that no living person has ever read to the end. Not ten separate pages scattered across the website. Nobody is going to hunt for them and piece them together.

The policy has an internal paradox, and it is not a flaw but a design. Two opposites live side by side in it. What does not change for years: the principles. Whether the company trusts AI with decisions about people. What happens to data. At which words from a customer a human must join the conversation. And what changes constantly: phone numbers, names, prices, deadlines.

Why do they share one home? Because a company may have many bots, but there must be only one source of truth. The principle works like this: if the support phone number is written into each bot separately, then after the number changes, one of the bots will inevitably keep quoting the old one, and that is precisely the bot your unhappiest customer will be talking to. What is recorded in one place does not diverge. What is recorded in five places will diverge without fail.

What the customer should find in the policy. Ten points, and it is the executive who answers them, not the programmer:

1. **Who is responsible.** The legal entity and the owner of the AI function, a named member of staff.
2. **Where AI operates.** The list of company services involving AI.
3. **How the AI introduces itself.** When, and in what words, it says that it is an AI.
4. **What happens to data.** What is collected, how long it is kept, whether models are trained on conversations, how to delete your data.
5. **Whether the AI makes decisions about the customer.** If so, how to obtain an explanation.
6. **How to reach a human.** A contact, a response time, and a promise that a human will actually reply.
7. **What happens in case of error.** Guarantees and the correction procedure.
8. **How to learn about changes.** Version, date, what has changed.
9. **What records remain.** What is logged and what the customer receives to keep.
10. **Where to complain.** A contact, a timeframe, a refund procedure.

The AI policy is written once, not rewritten for every bot. From then on, it works as a construction kit: each AI agent pulls the relevant sections from the policy for its own service and refines them for itself.

What does the AI policy achieve? Customer trust. Some of the information is already required from companies by law. The rest you adopt as a voluntary consumer standard. By gathering everything into one readable document and marking what is law and what is your additional guarantee, you end up not with a formal reply for the record but with a display of maturity that can be shown to customers, regulators and partners.

The policy can be extended with a register of trust credentials: current licences and certificates, adopted standards, quality marks, results of independent checks, and links to public test reports.

And one more detail that seems exotic today and will be the norm tomorrow. Alongside the version for people, the policy needs a machine-readable layer: a copy in a format that programs can read. In a year or two, it will not only be the customer who comes to you for a service, but their personal AI agent: to compare prices, ask questions, place an order. Your beautiful website means nothing to it: it reads documents, not design. With a company whose rules are machine-readable, such an agent will reach an agreement in seconds. A company without them it will simply fail to understand, and it will move on to your competitor. Technically this is not difficult; the developers will do it alongside everything else. Just build it in from the start.

2. The AI Service Passport: what this particular bot can do

The customer is usually talking not to the company but to a particular bot with a particular job. Each such job needs its own short document: the AI service passport.

The passport is not technical documentation. Which model you use, which prompts, how the integrations are built: you may record all that, but customers are generally not interested. The passport describes the rules under which the customer can obtain this service. In other words: what they are entitled to, and what they are not.

What the customer should find in the passport. Ten points, in the same logic as the policy:

1. **What the service is.** Who provides it and to whom.

2. **Result and deadline.** What the customer will receive and by when: the outcome of the service and the time within which the company promises to deliver it.
3. **Capabilities and limits.** What the bot can and cannot do: the full list of actions and, more importantly, the boundaries.
4. **Data and consents.** What data and consents the bot will request from the customer. Anything outside this list, the bot has no right to ask for.
5. **Price.** If the service is paid, the price is stated in advance and pulled from the policy, rather than arriving as a surprise at the end.
6. **Grounds for refusal.** A closed list, so that refusing "due to circumstances" is impossible.
7. **When the bot must call a human.** The customer asked; the bot failed to understand twice; the case is unusual; the conversation has turned into a conflict.
8. **What is recorded.** Every decision the bot makes is written to a journal: what it decided, by what criterion, and when. A black box, as on an aircraft.
9. **Pre-delivery check.** Before handing over a result, the bot checks it against the passport. If unsure, it does not deliver; it calls a human.
10. **Where to complain and how quickly you will be answered.** Plus the passport's owner by name, the version and the date.

The passport must be public. It sits where the customer will find it: next to the chat, on the service page. It is written in plain human language, briefly. And, like the policy, the passport is published in two layers: one for people and one for machines. This is your future language for talking to your customers' AI agents.

The passport lives in time. It has a version, a date and a status: in testing, in force, suspended. If you updated the model, changed the parameters or taught the bot something new, you will need a new version of the passport, and the old one stays in the archive.

Passports can be written not only for AI services but also for internal AI functions the customer will never see: accounting, warehouse, HR. Not for the sake of transparency, but for the staff. So that a new employee can quickly understand what this bot can do, what it cannot, and whom to go to when it makes a mistake. The passport becomes an excellent textbook on your own business.

3. The AI Receipt: what actually happened

A short history lesson. When the consumer movement was just emerging, it had no laws, no associations, no hotlines. It had one rule, repeated from posters like an incantation: always take the receipt. Because the receipt is the main door to justice.

Now look at what is happening in the world of AI. The bot quoted a price. Promised a refund. Calculated an insurance premium. Booked an appointment. The conversation ended, the window closed. What is the customer left with? Nothing. At best, a chat history that is stored by the company, edited by the company, and disappears when the company finds it convenient.

A hundred years of consumer progress, and a person is once again standing at the counter empty-handed. The counter has simply become virtual. So the proposal is simple: let the AI issue a receipt. Not forty screens of correspondence, but a document. It is generated at the end of the conversation and stays with the customer.

What is inside the AI receipt:

1. **Who provided the service, and to whom.** The company, the assistant's name and version, the date, the start and end time of the conversation, the recipient of the service.
2. **What the service was.** The subject of the request and the full record of the conversation.
3. **Consents and data.** What consents were obtained and what data the customer provided.
4. **The agent's actions.** What the bot did: amounts, deadlines, next steps. Item by item, like the lines on a till receipt.

5. **The result.** What the customer received in the end: the booking created, the order placed, the calculation issued.
6. **The price.** What the service cost and what the price includes.
7. **Failures.** Errors, delays, staff intervention, if there were any.
8. **The rules that applied.** The versions of the policy and the passport in force at the time of the conversation.
9. **Where to address a complaint.** The channel, the review period, the form of response.
10. **Contacts.** A direct line to the company, prominently placed.

Two details without which the receipt remains merely a pretty file.

First. What goes into the receipt is not a link to the policy but a snapshot of it: the full text of the edition in force at the moment of the conversation. The difference is fundamental. A link leads to a website that may be rewritten tomorrow. A snapshot fixes the agreement, and it cannot be retouched after the fact.

Second. The receipt exists in two identical copies, one with the customer and one with the company, with a number, a server timestamp and a way to verify authenticity. And then the eternal dispute of “what did your bot actually say” dies before it begins. There is no need to trust the customer's screenshots. There is no need to trust the company's word about its logs. It is enough to compare the copies. Note that the receipt protects the customer from the company and the company from the customer at the same time, and it is an open question who needs it more.

And a layer for the future: the receipt, like the policy and the passport, should exist in machine-readable form, so that it can be read and verified not only by a person but by the customer's AI agent.

In which cases should a receipt be issued? Common sense, without fanaticism. If a customer asked for the opening hours, no receipt is needed. But as soon as the AI has done something substantive: placed an order, made a calculation, taken money, accepted documents, given advice that affects someone's wallet, the receipt is generated automatically. Not “on request, if the customer thinks to ask”.

The objection is predictable: paperwork, and small businesses have no people for this. The answer: no people are needed. Nobody writes the receipt by hand. The system assembles it from its own records in a second: the time, the actions, the consents, the result, the document versions. An AI may present it in human language, but it does not invent the facts: they are pulled from the system, like the lines on a till receipt. The cost of this gesture is minimal. And it looks like a revolution.

The AI receipt proposed here is not, as such, required by law; it is a voluntary consumer-accountability standard. There are, however, legal requirements in particular contexts that point toward greater transparency and explanation. Article 86 of the EU AI Act gives an affected person, in defined circumstances, a right to a clear and meaningful explanation of the role of certain high-risk AI systems in individual decision-making and the main elements of the decision.[7] Quebec's private-sector privacy law requires a person to be informed when a decision is based exclusively on automated processing of their personal information and, on request, to be told about the personal information used and the principal factors and parameters that led to the decision.[8] The proposed AI receipt goes further by making a contemporaneous record available as a matter of organisational practice rather than presenting it as a legal requirement.

4. The test purchase

Suppose the company has done everything by the book. The policy is published, the passports are written, the receipt is issued.

Three months later, the bot quotes a price that is not in the price list.

Nobody touched anything. The model the bot runs on was simply updated by its developer, without asking and without notice. Or the terms changed and nobody corrected the bot's instructions. Or errors accumulated little by little until they became visible.

A bot can remain unchanged and still become different. So run test purchases of your own bot periodically.

A test purchase is not the same as a satisfaction survey. A survey asks whether the customer liked it. A test purchase checks whether the declared standard was met: the bot's behaviour is compared against the policy and the passport, the way goods are compared against a reference standard. The method is decades old; it is used to check shops, banks and government services. It transfers to AI agents unchanged: there is a declared standard, there is actual behaviour, and there is a measurable difference between them.

- **Approach your bot as a customer.** Not as the director who knows the right words, but as a person off the street who does not.
- **Take three scenarios.** Typical: an ordinary request. Atypical: something at the edge of the bot's authority, not directly covered by its instructions. Conflict: an irritated person demanding their money back.
- **Compare three layers.** What was promised in the policy and the passport. What the bot said and what receipt it issued. And what was actually recorded in the system.
- Repeat. Run checks on a defined monitoring cycle and after material change: for example, a significant model update, a change in instructions or integrations, evidence of drift, or remediation following a problem.

The last layer deserves attention, because it is the most treacherous place. A divergence between the promise and the bot's words is visible in the transcript: any attentive reader will find it. But a divergence between the bot's words and the record in the system is visible to no one. The phrase “I have booked you in for Thursday” looks identical whether the booking was created or not. There is only one way to check: compare the conversation against what remains in the system. In one of our checks, the bot behaved impeccably, yet in its own report there appeared the address of an employee who does not exist. Nobody would have seen this from the transcript. The comparison found it in a second.

You can run the checks yourselves, and you should: the point is to find a problem before a customer does. But self-checking has a well-known flaw: one's own bot often seems better than it is, and the questions put to it tend to be comfortable ones. At least occasionally, invite someone from outside to test it. Remember too that external evidence may arrive without invitation: consumer bodies, journalists and regulators can scrutinise a company's AI independently.

And here the circle described in Stage One closes. Everything the check has found returns to the organisation. A complaint may show that a boundary was drawn too generously. A failed escalation may show that authority was distributed wrongly. A recurring error may show that the AI lacks context. A material update, observed drift, scheduled review or completed remediation may trigger re-induction even before a customer complains. A named owner must be responsible for turning the loop: ensuring the evidence is reviewed, changes are made and the revised AI is authorised again. INDUCT → AUTHORISE → OPERATE → EVIDENCE → REVIEW → RE-INDUCT.

The Executive Checklist

To finish, a short self-assessment. Nine statements; tick a box only for “done and working”, not for “planned for next quarter”.

#	Statement	Done
1	The company has an AI policy: one public document answering the customer's ten questions	☐

#	Statement	Done
2	The policy is published in a visible place on the website and has a version and a date	☐
3	Every customer-facing AI service has a passport: capabilities, limits, deadlines, price, grounds for refusal, escalation	☐
4	The passport sits where the customer will find it and is written in plain human language	☐
5	Where Article 50(1) applies, the customer-facing AI is designed so that people are informed they are interacting with AI, unless that is obvious in the circumstances; the organisation also states its chosen transparency standard	☐
6	After every substantive interaction, the customer automatically receives an AI receipt	☐
7	The receipt contains the conversation record, the actions, the result, and the versions of the policy and passport in force at the time	☐
8	Test purchases are run on a defined cycle and after material change; a named owner ensures findings feed back into review and proportionate re-induction	☐
9	The policy, the passports and the receipt have a machine-readable layer fit for customers' AI agents	☐

If you have ticked all nine boxes, your company has done more than the vast majority of those who have adopted AI. The good news: the bar is so low that it can be cleared within a quarter. The bad news for your competitors is exactly the same.

Conclusion: The Loop Never Stops Turning

Amanda Cunningham & Sergei Ponomarev

AI accountability does not begin when a customer encounters a chatbot, agent or automated service. It begins earlier, inside the organisation.

Before AI is allowed to act, the organisation must decide why it is there, what it needs to know, what it may access, what authority it holds, where its boundaries lie and who remains accountable. Those decisions form the internal promise.

When AI reaches the customer, that promise becomes visible. The AI Policy explains the organisation's position. The AI Service Passport defines what a particular AI service is authorised to do. The AI Receipt provides evidence of what actually happened. A test purchase then asks whether reality matched the promise.

But that cannot be the end of the process.

Evidence must travel back inside the organisation. It may come from complaints, failed escalations, incidents and unexpected behaviour, but a mature system does not wait for failure. Routine monitoring and test purchases can expose drift or broken processes before a customer does, while consumer bodies, journalists, regulators and other external scrutiny may provide evidence the organisation did not generate itself.

The loop therefore needs both triggers and ownership. Re-induction should occur proportionately when material change or evidence calls the existing authorisation into question: after significant model or system updates, changes in data or integrations, observed drift, scheduled monitoring, material changes to the use case, or a repeat check following remediation. A named person or organisational role must remain responsible for ensuring that evidence is reviewed and the loop turns.

INDUCT → AUTHORISE → OPERATE → EVIDENCE → REVIEW → RE-INDUCT

That is why responsible AI cannot be treated as a one-time implementation project. AI systems change. Businesses change. Customers change. Regulation changes. The conditions under which an AI system was originally authorised will not remain static indefinitely.

The practical responsibility of leadership is therefore not to predict every future AI development. It is to create an organisational system capable of responding to change.

Stage One establishes what AI is permitted to do. Stage Two makes that permission visible and testable from the customer's side. Together they create something more useful than either an internal governance framework or an external transparency statement on its own: a continuous chain between organisational intent, delegated authority, customer experience and evidence.

That chain matters because trust in AI will not ultimately be created by telling customers that an organisation uses advanced technology. It will be created by being able to answer much simpler questions: Why is AI here? What is it allowed to do? Where does its authority end? Who remains responsible? And can the organisation prove that what it promised is what actually happened?

The organisations able to answer those questions will be better placed not simply to adopt AI, but to remain accountable for it as its role develops.

And when the answers change, the loop turns again.

Contact

Sergei Ponomarev, PhD, Founder, AI Business

Building a consumer quality assessment system for AI agents for business.

aibusiness.vc | info@aibusiness.vc | linkedin.com/in/sergei-ponomarev/

Amanda Cunningham, Founder, Savvy Pixel®

Digital and AI capability advisor supporting businesses to turn organisational knowledge into structured, accountable digital capability.

savvypixel.co.uk | amanda@savvypixel.co.uk | linkedin.com/in/amandasavvypixel

References

[1] Department for Science, Innovation and Technology (2026), Guidance for using the AI Management Essentials tool, updated 6 February 2026.

[2] Department for Science, Innovation and Technology (2026), AI Management Essentials tool (accessible), updated 6 February 2026.

[3] Information Commissioner's Office, Guidance on AI and data protection: Accountability and governance implications of AI.

[4] Information Commissioner's Office, Guidance on AI and data protection: About this guidance. Current ICO AI guidance should be checked against the latest position following changes to UK data-protection law.

[5] National Institute of Standards and Technology, AI Risk Management Framework 1.0, AI RMF Core. NIST notes that AI RMF 1.0 is being updated.

[6] Regulation (EU) 2024/1689 on artificial intelligence, Article 50: transparency obligations for systems interacting with people. Applicable from 2 August 2026.

[7] Ibid., Article 86: the right to an explanation of an individual decision made by a high-risk system.

[8] Quebec, Law 25 (Loi 25): notification of a decision based exclusively on automated processing, and the rights of the person concerned.

[9] Directive 2011/83/EU on consumer rights: pre-contractual information requirements.