

# Governance Principles for AI in Purpose-Led Organisations

*A cross-framework synthesis informed by regulatory frameworks,  
sector practice, and emerging international thinking on AI governance*

---

**Organisation:** Jannaways

**Version:** 3.0

**Date:** April 2026

**Status:** Working document for internal use and client engagement

**Supersedes:** Version 2.0, February 2026

## **A note on how this document was produced**

This document was developed using multiple AI models as research, drafting, and structuring tools. The author's role was orchestration: identifying the right questions, selecting and verifying source frameworks, directing the synthesis, exercising judgement on every editorial decision, and taking responsibility for the result. AI contributed speed and breadth. The author contributed direction, discrimination, and accountability.

*This reflects the 'human in the lead' principle that runs through the document itself: the quality of the output depends on the quality of the human directing the process, not on the tools used to execute it.*

---

## Introduction

### Why governance matters now

Governance does not change when the tools change. What changes is the speed at which poor governance becomes visible. AI amplifies organisational reasoning at scale. That is its power, and its risk. If the thinking is sound, AI makes it faster and more consistent. If the thinking is flawed, AI makes it faster and more consistent too.

This is not primarily a technology question. It is a governance question. Organisations that cannot articulate how they think and decide cannot govern AI. The work of AI governance begins with making implicit reasoning explicit.

### Governance as confident action

These principles are not a brake on adoption. They are the foundation for confident action. The risk of ungoverned AI use is real. So is the risk of paralysis. Organisations that engage thoughtfully with AI governance are better placed to adopt AI with confidence, explain their decisions to regulators and funders, and maintain the trust of the people they serve.

As Alastair Denniston, Chair of the National Commission for the Regulation of AI in Healthcare, has noted: recognising the risks of staying still is as important as recognising the risks of moving forward. Governance enables movement. It does not prevent it.

### AI must serve communities, not just comply with rules

Rule compliance is necessary but not sufficient. The deeper governance question is whether AI serves the actual people and communities an organisation exists for. This requires more than technical compliance. It requires deliberate attention to whose needs are being met, whose voice is present in decisions about AI adoption, and whether the systems in use reflect the diversity and specificity of the people they affect.

The civic.ai 6-Pack of Care framework, developed at Oxford's Institute for Ethics in AI by Audrey Tang and Caroline Green, offers a useful principle here: AI governance that is designed only from the top down will fail the people it is meant to serve. Attentiveness, responsibility, competence, responsiveness, solidarity, and symbiosis must be designed in, not assumed.

### Rights and dignity as the foundation

These principles are grounded in a rights-aware understanding of governance. The responsible use of AI means that AI systems support and do not undermine, harm, or unfairly breach fundamental values including human rights, independence, choice and control, dignity, equality, and wellbeing. This definition, developed through the Oxford Project on responsible AI in social care, applies across all sectors. Purpose-led organisations in particular hold a responsibility to uphold these values in how they adopt and govern AI, not just in what they deliver.

This rights foundation is not confined to a single principle. It runs through accountability (who is protected and by whom), proportionality (the more significant the rights at stake, the stronger the governance required), transparency (people have a right to understand decisions that affect them), fairness (AI must not undermine equality or dignity), and contestability (people must have genuine routes to challenge). Readers should bring a rights lens to the whole document, not only to the sections where rights are explicitly named.

---

## The longer horizon

These principles are grounded in current regulatory frameworks and present-day governance practice. They are also designed to build institutional muscle for larger shifts ahead. The questions AI raises about accountability, transparency, and human judgement will not diminish as AI capabilities grow. Organisations that develop sound governance practice now are building the foundations they will need regardless of how the technology develops.

## Why a unified framework

This document synthesises governance principles from multiple authoritative frameworks into a single working reference for purpose-led organisations considering, adopting, or governing AI use. It is not a compliance checklist. It is a foundation for organisational conversation about how governance expectations apply when AI enters the picture.

Previous versions existed in separate editions for charities and for commercial SMEs. In practice, the governance principles are the same. The sector-specific difference is not in the principles but in the regulatory context, the source of legitimacy, and the language used. Where regulatory context differs meaningfully between sectors, this is noted in the text beneath each principle. The principle itself is universal.

## How to use this document

This is a reference document. Its purpose is to provide a comprehensive, sourced synthesis of governance principles that boards, senior leaders, and governance advisors can draw on when considering, adopting, or reviewing AI use. It is not a checklist to work through in sequence.

Each principle includes its source frameworks, a practical test or implication, and where relevant a failure mode or red flag. The principles are grouped into eight themes that create a natural narrative from accountability through to technology-specific discipline.

The question to bring to this document is not 'do we comply with all eighteen?' but 'where are we strong, where are we exposed, and what do we need to address first?' For practical implementation tools, see the companion document From Principles to Practice and the AI Governance Readiness Diagnostic.

## Where to start

If you are beginning AI governance work for the first time, start with Principles 1, 3, and 7. These establish human accountability, purpose alignment, and transparency as the non-negotiable foundations before anything else is addressed.

If you are reviewing existing AI governance practice, the most common gaps are found in Principles 5, 9, 14, and 15: proportionality (are controls matched to actual risk?), contestability (can people challenge decisions?), voice (have those affected had genuine input?), and currency (is governance live or has it drifted?).

If you are dealing with a specific governance concern, use this index: data and privacy (10, 11); safeguarding (12); fairness and bias (13); employment (14); procurement (10, note on procurement); board oversight (1, 2, 8, 17); AI-assisted decision-making (1, 6, 9, 18).

***The governing question running through this document: if AI scales how your organisation thinks and decides, is that thinking worth scaling?***

## Source frameworks

Framework	Core contribution
UK GDPR / Data Protection Act 2018 / Data (Use and Access) Act 2025	Accountability, proportionality, data stewardship, transparency, lawful basis, updated automated decision-making provisions
Charity Governance Code (2025)	Purpose, ethics and culture, board effectiveness, equity, digital governance, AI policy expectation
Good Business Charter / B Corp governance	Stakeholder accountability, responsible employment, environmental responsibility, multi-stakeholder value creation
UK Government AI Principles (2023/24)	Safety, transparency, fairness, accountability, contestability and redress
Nolan Principles of Public Life	Selflessness, integrity, objectivity, accountability, openness, honesty, leadership
ICO AI and Data Protection Guidance	DPIAs, fairness in AI, explainability, lawful processing, bias mitigation
IoD Corporate Governance Standards	Duty of care, competence, continuous judgement, oversight under pressure
Equality Act 2010	Non-discrimination including algorithmic bias, indirect discrimination through proxies
Employment law framework	UK GDPR Article 22, worker monitoring, consultation obligations, Employment Rights Bill
Safeguarding legislation	Care Act 2014, Working Together 2023, Charity Commission safeguarding guidance, duty of care to vulnerable people
civic.ai 6-Pack of Care (Tang & Green, Oxford 2025)	Attentiveness, responsibility, competence, responsiveness, solidarity, symbiosis as governance design principles; plurality and community accountability
Oxford Project— Responsible AI in Social Care definition (2025)	Rights-grounded definition of responsible AI: human rights, independence, dignity, choice, equality, wellbeing
Alliance on AI in Social Care — I/We Guidance (2025)	Co-production in the AI lifecycle; sustainable technology as governance domain; bias and accessibility
Denniston / CERSI AI — National Commission for Regulation of AI in Healthcare (2026)	Governance as enabler of confident action; risk of inaction; assurance as a system-wide function
University of Birmingham / CERSI-AI — AI Readiness Assurance Framework and AI Readiness Checklist v5 (February 2026, funded by The Health Foundation)	Deployment-specific readiness as a distinct governance level; five-category harm taxonomy (Workforce, Operational, Financial, External, Service Outcomes); problem formulation as pre-adoption discipline; technology-dependent behaviour as a named governance risk; AI information asset register as live monitoring tool (cersi-ai.org)

## Regulatory context: why this matters now

The UK is moving toward statutory AI regulation. The Data (Use and Access) Act 2025 received Royal Assent in June 2025, updating UK GDPR provisions for automated decision-making. A comprehensive AI Bill is expected in the second half of 2026, likely establishing an AI Authority and codifying the five UK Government AI principles into binding duties.

The 2025 update to the Charity Governance Code includes AI policy as suggested evidence of good governance for the first time. The ICO's AI and Data Protection Guidance is currently under review. The EU AI Act applies from August 2026, categorising most employment AI as high-risk with mandatory documentation requirements.

Expectations regarding equality and non-discrimination are increasing across all sectors. Regulators, funders and partners increasingly expect anticipatory fairness where AI influences decisions affecting service users, customers, or employees.

## A note on procurement

Many organisations will not build custom AI tools. They will subscribe to or procure AI-enabled services from third parties. The principles in this document apply equally to bought and built AI. Procurement due diligence for AI-enabled services should include: where data is processed and stored, what data is used for model training, what happens to data when the contract ends, what transparency is provided about how the AI operates, and whether the service meets the organisation's data protection, safeguarding and employment requirements. These are governance questions, not technical ones, and they belong at board or senior leadership level.

## The 5 Lenses Framework

The 18 principles in this document are organised using the 5 Lenses — Jannaways' proprietary framework seeing how the real work happens. The Lenses are not a checklist; they are five distinct lines of questioning, each designed to surface a different category of governance risk or decision.

Every principle in this document is tagged with its primary lens (and secondary lens where relevant). The tags are a navigation aid: they show which line of questioning each principle primarily serves, and help practitioners identify where gaps in their current governance approach lie.

**Purpose** Does this serve our mission? Tests whether AI use is genuinely aligned with the organisation's purpose and values, not just operationally convenient.

**People** Who is affected, and how? Examines the human impact — on service users, staff, and communities — including who may be excluded, harmed, or left without a voice.

**Power** Who decides, and who is accountable? Interrogates authority, accountability, and influence: who has the power to act, who is affected by that power, and whether the right people are in control.

**Process** Can we actually govern this? Assesses operational readiness: whether the organisation has the systems, capacity, and discipline to implement and sustain governance in practice.

**Practice** Will this change how people actually work? Focuses on day-to-day reality — whether governance commitments translate into changed behaviour, not just written policy.

In practice, most governance decisions touch more than one lens. The tags indicate emphasis, not exclusivity. The 5 Lenses work as a set.

# The Principles

## A. Accountability and authority

### 1. Human accountability is never delegated

5 Lenses: **Power**

Decisions with legal, ethical or material impact must always have a named human owner. AI tools can advise, surface risk, or structure thinking. They must never be the final authority. Those carrying ultimate governance responsibility retain that responsibility regardless of the tools used to reach their decisions.

**Sources:** UK GDPR (accountability principle); Charity Commission trustee guidance; UK Gov AI Principles (accountability and governance); IoD director duties; Good Business Charter Standard 9

**Governance test:** *If this goes wrong, who carries the consequence, and did they have the power to shape the decision?*

**AI implication:** *Every AI-assisted process needs a named human accountable for its outputs. Boards and leadership teams should maintain a register of AI tools in use and assign oversight responsibility for each. Contracting or subscribing to an external AI service does not transfer accountability for outcomes, legal duties, or reputational impact.*

For charities: Trustees retain ultimate responsibility for governance outcomes. For commercial organisations: Directors are personally liable for negligent governance under IoD standards and company law.

Human accountability also requires that those making decisions can see and question how the question itself was framed. AI that summarises, filters, or structures information before a human sees it is already shaping the decision. A board that approves an AI-assisted recommendation may be nominally accountable but is not genuinely so if the framing, the options presented, and the assumptions embedded in the analysis were set by the AI without scrutiny. Governance must include the framing, not just the final call.

**Framing test:** *Before approving any AI-assisted recommendation, ask: who framed the question, what options were excluded before this reached us, and what assumptions are embedded in how this was presented?*

### 2. Authority must match consequence

5 Lenses: **Power**

Those carrying risk must have the authority and information to act on it. This is a recurring failure mode in organisations of all kinds, creating silent workarounds and post-hoc justification. It is often misdiagnosed as a 'process' issue when it is actually a governance design problem.

**Sources:** Charity Governance Code (decision making); IoD governance standards; Nolan Principles (accountability); Good Business Charter Standard 9

**Failure mode:** *If a staff member is expected to make judgement calls about AI use but has no authority to set boundaries, governance has already failed.*

## B. Purpose and legitimacy

### 3. Use must align with organisational purpose

5 Lenses: **Purpose**

Data and AI use must clearly support the organisation's core purpose. Every AI adoption should be traceable to the mission, whether that is defined by charitable objects, articles of association, or a stakeholder governance commitment. This extends to the environmental and social impact of technology choices: AI infrastructure has energy costs, and organisations committed to responsible resource use should consider those costs as part of purpose alignment.

Purpose alignment also requires asking whose needs are actually being served. AI adopted for operational efficiency may not serve the people the organisation exists for. The question is not only 'does this serve our purpose?' but 'does this serve the specific communities we exist for, in ways they would recognise?'

**Sources:** *Charity Governance Code (organisational purpose); Charity Commission legitimacy tests; Nolan Principles (selflessness); Good Business Charter Standards 6 and 9; B Corp purpose alignment; public sector AI guidance; Alliance on AI in Social Care I/We Guidance (sustainable technology domain)*

**Red flag:** 'We can' without a clear 'why for us' and 'for whom specifically'.

**AI implication:** *Before deploying any AI tool, the organisation should articulate how it serves the people it exists for. If you cannot draw a clear line from the tool to your purpose and to the people it affects, it does not belong in your operations yet. Purpose alignment also includes the discipline of problem formulation: before evaluating any AI tool, the organisation should have clearly defined what specific problem it is trying to solve, for whom, the boundaries of appropriate use, and what the tool must never do. Adopting AI without this prior clarity creates a governance exposure that oversight cannot fully repair after the fact. Problem formulation is a pre-adoption discipline — it belongs before procurement, not inside it.*

### 4. Trust is a governance asset

5 Lenses: **Purpose** · also: **People**

Reputational impact is not secondary to legal compliance. Trust, whether from beneficiaries, customers, employees, funders, or the public, is foundational to organisational legitimacy. AI systems amplify reputational risk rapidly and at scale.

**Sources:** *Charity Governance Code (ethics and culture); UK GDPR (fairness); Nolan Principles (integrity, honesty); Good Business Charter Standard 5; B Corp stakeholder governance*

**Practical rule:** *If you would struggle to explain this decision publicly, governance is not done.*

**Rights awareness:** *Where AI-informed decisions influence access to essential services or support, organisations should remain alert to the potential engagement of rights relating to dignity, privacy and family life.*

## C. Proportionality and risk

### 5. Governance must be proportionate to risk

5 Lenses: **Process**

Controls should scale with sensitivity, impact and reversibility. This is a core GDPR principle and is embedded in UK risk management standards. It is often lost in AI debates, where organisations either over-govern low-risk uses or under-govern high-risk ones. Commercial and vendor risk falls within this principle. The depth of due diligence before adopting an AI tool — including vendor stability, contract terms, dependency risk, and the organisation's capacity to exit or switch — should reflect how exposed the organisation would be if the tool failed, changed terms, or was withdrawn.

**Sources:** *UK GDPR (proportionality); UK Gov AI Principles (safety, security and robustness); Charity Commission risk guidance (CC26); IoD risk management; B Corp stakeholder impact assessment*

**Failure mode:** *Heavy governance where risk is low, and light governance where risk is high.*

**AI implication:** *Using AI to draft a newsletter and using AI to triage safeguarding referrals are not the same governance question. Proportionality means matching the intensity of oversight to the potential for harm.*

Proportionate governance depends on separating risk into distinct harm pathways. Treating 'AI risk' as a single object obscures material differences in severity, reversibility and exposure.

Harm pathway analysis also clarifies which level of authority a decision requires. Changes to governing principles - the fundamental protections and commitments the organisation makes to people - affect what is normatively guaranteed and belong with governance authority: the board or trustees. Improvements to how those principles are implemented - making existing commitments more reliable or consistent - sit with technical authority. Operational adjustments sit with the system owner. The escalation trigger is which harm pathway is affected, not simply where in the organisation the change originates.

The University of Birmingham/CERSI-AI AI Readiness Assurance Framework provides a useful taxonomy for harm pathway analysis at the point of adopting a specific AI tool, mapping causes of harm across five categories: Workforce (staff capability, change management, and technology-dependent behaviour), Operational (deployment and process integration), Financial (vendor dependency, uncertain business case, liability, and exit cost), External (regulatory environment and community impact), and Service Outcomes (quality and safety). This taxonomy can be used to structure harm pathway assessment for any individual AI adoption decision.

### 6. Ambiguity must be governed, not ignored

5 Lenses: **Practice**

Most governance failures occur in grey zones, not rule breaches. GDPR explicitly addresses judgement calls. Boards routinely underestimate ambiguity. AI increases the frequency of edge cases because it operates across more decision points, faster, than human processes.

**Sources:** *UK GDPR (fairness, accountability); Charity Governance Code (decision making); ICO AI guidance; Good Business Charter Standard 9*

**Design principle:** *Define thresholds, not just permissions. A governance triage system (green/amber/red) can make expectations visible at the point of decision.*

## D. Transparency and explainability

### 7. Decisions must be explainable to the people affected

5 Lenses: **People** · also: **Power**

Not technically explainable, meaningfully explainable. Organisations must be open about where and how AI is being used. This applies to beneficiaries, customers, employees, funders, regulators, and staff. People affected by AI-informed decisions have the right to understand how those decisions were reached. Explainability must be proportionate to the impact.

**Sources:** *UK GDPR (transparency, fairness); UK Gov AI Principles (transparency and explainability); Charity Governance Code (openness); ICO Explaining Decisions Made with AI; Nolan Principles (openness); Good Business Charter Standard 5; Consumer Protection Regulations*

**Litmus test:** *Could you explain this to a service user, customer, employee, or trustee without defensiveness?*

For customer-facing organisations: People interacting with your organisation have a right to know when they are engaging with AI rather than a human. Intentionally obscuring AI use undermines trust and potentially violates consumer protection law.

### 8. Assumptions must be surfaced, not buried

5 Lenses: **Process** · also: **Practice**

Good governance exposes assumptions early. This reduces cognitive load at board level, improves the quality of challenge, and prevents retrospective rationalisation. Decisions to adopt, modify, or discontinue AI systems should be documented with clear reasoning, not left to institutional memory. This is a UK GDPR accountability requirement.

**Sources:** *Charity Governance Code (board effectiveness); IoD professional development; ICO accountability principle; UK GDPR (accountability principle)*

**Warning sign:** *If the board or leadership team is routinely surprised by what staff are doing with AI, assumptions have been buried rather than surfaced.*

**AI implication:** *This is where AI, if used well, can add most value: upstream of the decision, during preparation and sense-making, rather than in the decision itself. Undocumented decisions create governance gaps and leave the organisation vulnerable.*

Governance by design means surfacing assumptions early enough to shape implementation, not just review it. Trustee or board visibility brought into the design phase - before commitments are made, before systems are configured - is more effective than retrospective scrutiny. Questions raised after adoption are valuable, but the governance habit is to surface what matters before the decision is hard to reverse.

### 9. Contestability: clear routes to challenge AI-informed decisions

5 Lenses: **Power** · also: **People**

People affected by AI-informed decisions must have clear routes to challenge those decisions and receive a human review. This includes the right not to be subject to solely automated decision-making with significant effects. Organisations must be able to review and reverse AI-influenced decisions when challenged.

Contestability is not only an individual right. Where AI systems affect groups of people with shared experiences, communities should have meaningful routes to raise collective concerns and require systemic repair. This is particularly important in social care, housing, and other settings where AI affects people with limited capacity to challenge decisions individually.

**Sources:** UK Gov AI Principles (contestability and redress); UK GDPR (rights of data subjects, Art. 22); Data (Use and Access) Act 2025; Equality Act 2010; Good Business Charter Standard 5; B Corp stakeholder engagement; civic.ai 6-Pack Pack 4 (Responsiveness)

**Practical requirement:** Every AI-informed process that affects individuals needs a documented route for challenge. Staff need to know how to escalate concerns about AI outputs, and the people affected need to know they can question decisions.

**Red flag:** Making the challenge process onerous is functionally the same as having no challenge mechanism.

## E. Data stewardship and integrity

### 10. Data is stewarded, not owned

5 Lenses: **Process**

Organisations hold data in trust. This is explicit in GDPR and particularly acute for organisations working with vulnerable groups. Just because data exists does not mean it should be used. When staff paste personal information into a consumer AI tool, that is a data stewardship failure, not a minor process slip.

Organisations must have a lawful basis to process personal data for AI purposes and cannot repurpose data collected for one thing to train AI for something else without proper justification. Data shared with AI tools leaves the organisation's direct control in ways that traditional data sharing does not.

**Sources:** UK GDPR (all seven principles); Data Protection Act 2018; Data (Use and Access) Act 2025; ICO AI guidance; Charity Commission guidance; Good Business Charter Standard 10; B Corp supply chain governance

**Red flag:** The instinct to 'upload everything' into AI tools conflicts directly with data protection principles. Every piece of data shared with an AI tool should be justified.

**AI implication:** Organisations should establish clear policies about what categories of data can and cannot be used with different types of AI tool. Free or consumer AI tools may not meet the security standards required for personal data.

### 11. Data quality underpins decision quality

5 Lenses: **Process**

Poor data governance is a decision risk, not a technical flaw. Boards and leadership teams routinely underestimate this. AI systems magnify data weaknesses because they process at scale and with apparent confidence. Those carrying governance responsibility remain accountable for decision quality regardless of the tooling used.

**Sources:** UK GDPR (accuracy, data minimisation); Charity Governance Code (decision making); ICO AI guidance; IoD governance standards

**Board-level framing:** Bad data is a governance issue, not an IT issue.

## F. Fairness, rights, and duty of care

### 12. Safeguarding and duty of care decisions remain exclusively human

5 Lenses: **People** · also: **Power**

AI must never compromise safeguarding or the duty of care owed to vulnerable people. Decisions affecting vulnerable adults, children, employees in vulnerable situations, or customers reliant on essential services must not be delegated to or unduly influenced by AI systems without appropriate human oversight. AI-generated content must be reviewed before it reaches vulnerable people. Safeguarding concerns identified by AI must still follow established human reporting routes.

**Sources:** *Care Act 2014; Working Together to Safeguard Children 2023; Charity Commission safeguarding guidance; Charity Governance Code; Employment Rights Bill; Good Business Charter Standard 8; Health and Safety at Work Act 1974*

**Non-negotiable:** *AI may flag patterns or support triage, but the safeguarding or duty of care response must follow established procedures with trained professionals making the judgement calls. No AI output overrides a duty to report or act on a safeguarding concern.*

For charities and care providers: This includes explicit board-level approval for AI touching safeguarding pathways. For commercial organisations: Duty of care extends to employees and to customers where decisions affect access to essential services.

### 13. Fairness, non-discrimination, and plurality

5 Lenses: **People**

AI systems must not discriminate against individuals or produce unfair outcomes. AI can embed and scale existing biases in organisational data or reasoning. Fairness must be actively tested, not assumed. AI exhibits bias through proxies even when protected characteristics are not collected directly. The Equality Act applies to indirect discrimination.

Fairness also requires attention to plurality. AI tools designed for majority populations may actively disadvantage minority groups, people with specific communication needs, or communities with distinct cultural contexts, even when those tools pass generic bias tests. Organisations serving specific communities must test AI tools for fitness in their particular context, not only against general benchmarks.

This means actively involving people with lived experience in evaluating whether AI systems work for them, not just assessing outputs from the outside.

**Sources:** *UK Gov AI Principles (fairness); UK GDPR (fairness); Equality Act 2010; Charity Governance Code (equity, diversity and inclusion); ICO AI fairness guidance; EHRC 2025-2028 strategy; Good Business Charter Standard 2; civic.ai 6-Pack Pack 1 (Attentiveness) and Pack 6 (Symbiosis); Alliance on AI in Social Care I/We Guidance (bias and discrimination domain)*

**Key question:** *Who might be disadvantaged by this AI-informed decision, particularly where those affected include protected groups, vulnerable people, or employees? Does this tool work for the specific community we serve?*

**Red flag:** *'We do not collect protected characteristics, so bias is not relevant' is incorrect.*

---

## 14. People affected by AI must have a genuine voice in its adoption

---

5 Lenses: **People** · also: **Power**

When AI is introduced in ways that affect how people work, or how services are delivered to people who depend on them, those affected should have genuine input into decisions, not just be informed after the fact. For employees, this is both a legal obligation and a governance standard. For people who use services, this is an ethical requirement and increasingly a sector expectation.

Co-production is not the same as consultation. Consultation asks people for their views on decisions already forming. Co-production involves people with lived experience in shaping the questions, designing the systems, and evaluating whether they work - as a continuous discipline, not a one-off exercise. An organisation that asks service users whether they like the new AI tool after it has been chosen, configured, and deployed has consulted. An organisation that involves them in defining what the tool is for, what it must never do, and whether it is working for them has co-produced. The distinction matters because only the second approach catches the assumptions that would otherwise remain invisible.

**Sources:** *Good Business Charter Standard 8; B Corp worker governance; Employment Rights Bill; UK GDPR (Art. 22); TUC AI Bill proposals; Charity Governance Code (ethics and culture); Health and Safety at Work Act 1974; Alliance on AI in Social Care I/We Guidance (co-production domain); civic.ai 6-Pack Pack 1 (Attentiveness)*

**Practical test:** *When introducing AI that changes roles, workload, services, or how people experience support, did the people affected have genuine input? Or were they informed after the decision?*

**Red flag:** *Using AI for productivity monitoring or performance analytics without consultation creates legal risk under UK GDPR. The direction of travel in employment law is toward mandatory consultation for high-risk employment AI.*

---

## G. Oversight, learning, and adaptability

### 15. Governance must be live, not static

---

5 Lenses: **Practice**

Policies do not govern. Practice does. AI systems evolve faster than policies. Governance that was adequate six months ago may not be adequate now. Those responsible for governance must invest in their own AI literacy and review governance arrangements regularly.

AI deployments should also have defined scope and exit conditions. An AI system adopted for a specific purpose should have clear boundaries around what it is permitted to do, and a mechanism for reviewing whether it should continue, be modified, or be discontinued. Scope creep in AI systems, where a tool gradually takes on functions beyond its original mandate, is a governance failure that builds up quietly and is difficult to reverse. A related and undernamed risk is technology-dependent behaviour: where staff progressively delegate more judgement to AI systems than was originally intended, and where the professional capability to question AI outputs or work independently of them diminishes over time. This accumulates quietly and is rarely surfaced until something goes wrong. Governance should include periodic review of whether AI tools are being used within their original mandate and whether independent professional capability is being maintained alongside them.

**Sources:** *Charity Governance Code (board effectiveness); IoD professional development; ICO accountability principle; Nolan Principles (leadership); Good Business Charter Standard 9; civic.ai 6-Pack Pack 6 (Symbiosis — boundedness and sunset-readiness)*

**Requirement:** *Feedback loops that show how decisions are actually being made, not just how policies say they should be made. AI tools in use should be reviewed at defined intervals, with explicit decisions about whether they continue.*

**Red flag:** *Deferring entirely to technical teams or vendors without understanding governance implications is an abdication of leadership responsibility.*

### 16. Challenge must be safe and expected

---

5 Lenses: **Practice** · also: **Power**

Effective governance depends on people feeling able to surface concerns early. Psychological safety is a governance condition, not a cultural nice-to-have. Without it, escalation routes fail. No framework compensates for fear.

**Sources:** *Charity Governance Code (ethics and culture, board effectiveness); Nolan Principles (honesty, openness); IoD governance standards; Good Business Charter Standard 4*

**Reality check:** *If raising a concern about how AI is being used feels risky, governance is already broken.*

### 17. Governance must be visible in both directions

---

5 Lenses: **Power** · also: **Process**

Governance must flow in both directions. Staff need to see up into governance requirements so they know what is expected at the point of decision. Leaders need to see down into decision patterns so they can spot drift, ambiguity, and emerging risk. One without the other leaves gaps.

**Sources:** *Charity Governance Code (decision making, openness); ICO accountability; Good Business Charter Standard 9; BYOCAI methodology*

**Design principle:** *AI governance expectations should be visible and accessible to every person who uses AI, not locked in board papers or policy documents. Usage patterns and concerns should flow upward so leadership can see where the real questions are.*

---

*This is governance as a team sport: everyone operating from shared visible reasoning, not just the same documents.*

One further governance question belongs here: who has the authority to decide whether a proposed change is a principle-level decision requiring board visibility, or an implementation improvement that can be handled at a lower level? That classification is itself a governance decision. It should not be made unilaterally by technical or operational teams. Organisations should name in advance who holds that authority, and ensure it is genuinely independent of the teams proposing the change.

## H. Technology-specific discipline

### 18. Do not automate unresolved judgement

---

5 Lenses: **Purpose** · also: **Process**

Automation should follow clarity, not replace it. If the organisation cannot explain the judgement, it should not automate it. The role of technology is to reduce cognitive load, not decision ownership. AI is most valuable when it operates upstream of the decision point: during preparation and sense-making, helping surface assumptions, clarify framing, and hold risk consistently before meetings rather than inside them. The aim is not automation. It is clarity.

**Sources:** UK Gov AI Principles (accountability, transparency); UK GDPR (accountability, Art. 22); ICO AI guidance; IoD governance standards; BYOCAI methodology; Charity Governance Code (board effectiveness)

**Rule:** *If you cannot explain the judgement clearly to a person, you should not encode it into a system.*

**Failure mode:** *Automating a process that humans were already doing inconsistently does not solve the inconsistency. It scales it.*

*Designing AI to carry governance memory helps create the conditions in which thoughtful discussion can occur. Board members and trustees are offering attention, accountability and care. Senior leaders are carrying pace, complexity and consequence. Treating governance as something worth preparing for properly respects both roles.*

## The governing question

***If AI scales how your organisation thinks and decides, is that thinking worth scaling?***

Before an organisation can govern AI well, it needs to have articulated how it actually thinks and decides. That means making implicit reasoning explicit, testing whether it is consistent and current, and ensuring it is shared across the organisation.

The unifying thread across every framework, every principle, and every governance test in this document is the same: supporting better human judgement, earlier, with less friction.

## Document status and version history

Field	Detail
Document title	Governance Principles for AI in Purpose-Led Organisations
Version	3.0
Date	April 2026
Author	Rachel Jannaway, Jannaways
Status	Working document for internal use and client engagement
Supersedes	Version 2.0 (February 2026), which superseded separate charity and commercial editions v1.x
Key changes in v3.0	<p>New introduction: governance as confident action, rights as foundation, civic AI framing, longer horizon. Principles 3, 9, 13, 14, 15 updated to incorporate co-production, plurality, community contestability, boundedness/sunset. Source table updated with civic.ai, Oxford/Digital Care Hub, Alliance on AI in Social Care, and Denniston/CERSI. Theme F retitled to include rights.</p> <p>Principle 1 extended: governance must include scrutiny of how AI has framed the question, not just the final decision. Principle 5 extended: consequence-based escalation structure distinguishing principle changes (governance authority) from implementation improvements (technical authority).</p> <p>Principle 8 extended: governance by design, bringing trustee visibility into the design phase before commitments are made. How to use section updated: document purpose stated explicitly as reference document; governing question moved to introduction. Principle 14 co-production language sharpened to distinguish co-production from consultation. Principle 17 extended: authority to classify decisions as principle-level vs implementation must be named and independent. Review schedule updated with owner and triage criteria.</p>
Review schedule	Quarterly, or when source frameworks are updated. Review owner: Rachel Jannaway. Triage note: not every source framework update requires a version change. Changes to UK GDPR, the Charity Governance Code, or the AI Bill trigger a review. Updates to sector guidance or supplementary sources trigger a note for the next scheduled review.
Production method	Developed using multiple AI models as research and drafting tools. Author role: orchestration - directing the synthesis, verifying sources, exercising editorial judgement, taking responsibility for the output. AI contributed speed and breadth; the author contributed direction, discrimination, and accountability.
Sources verified	April 2026. All v2.0 sources plus: civic.ai 6-Pack of Care (Tang & Green, Oxford 2025); Oxford Project Responsible AI in Social Care definition (March 2025); Alliance on AI in Social Care I/We Guidance (2025); Denniston/CERSI National Commission for the Regulation of AI in Healthcare (2026); University of Birmingham/CERSI-AI: AI Readiness Assurance Framework and AI Readiness Checklist (v5, February 2026, funded by The Health Foundation) — a companion deployment-specific readiness tool for organisations preparing to adopt a specific AI technology. Available at <a href="https://cersi-ai.org">cersi-ai.org</a> .