



A Practical AI Roadmap for NGO Leadership Teams

AI use is now common in charities, but common use is not the same as a safe or useful operating model. This flagship for the NGO Pathways series sets out how leadership teams can reduce administrative drag without treating care, judgement or trust as automation problems.

David Saliba . 2026-07-25 . aimonger.com/whitepapers/ngo-practical-ai-roadmap-leadership/

The problem in one sentence

Start with where time is lost. Technology comes second. Trust comes first.

This paper is for the director who sees a team stay late to finish notes, the facility lead who cannot find the current version of a policy, and the programme manager who starts every funder report by searching inboxes. It is not a case for adding another dashboard to an already crowded week. It is a way to decide whether a small, governed change can protect more time for people without making the work less safe or less human.

The pressure is not imaginary, and it is not confined to one type of organisation. Social Employers' 2023 sector research reports a Dutch care-work pattern in which administration takes about 36 per cent of time. Representatives in Austria and Belgium described documentation growing more complex, with technology sometimes adding to everyday burden rather than relieving it. Eurofound's 2023 account of social services describes staff shortages, strained conditions and a digital divide, with about one in five EU social-care workers never using digital devices at work. These are sector patterns, not a diagnosis of every NGO. They do show why a tool that saves a few minutes but creates another login, another data entry step or another approval queue can make a hard week worse.

The human cost also deserves more care than a productivity slogan. OECD's 2023 Beyond Applause? found that 31 per cent of nurses and personal-care workers named workload and time pressure as their most important mental-health risk, compared with 19 per cent of all employees. The 2025 WHO Europe MeND survey, drawing on about 90,000 valid responses from doctors and nurses in 29 countries, found roughly one in three reporting depression or anxiety symptoms and around one in ten reporting passive suicidal thoughts. It is not an NGO census, and should not be presented as one. It is a warning light for every helping profession: when the working day is already emotionally demanding, administrative friction is not a harmless inconvenience.

Administrative burden, plainly: the time and mental effort taken by forms, duplicated facts, delayed notes, email chasing and report reconstruction around the real service.

A familiar Monday: a practitioner finishes a valuable session, then writes the note from memory, seeks a missing signature and answers three overlapping messages about the same person.

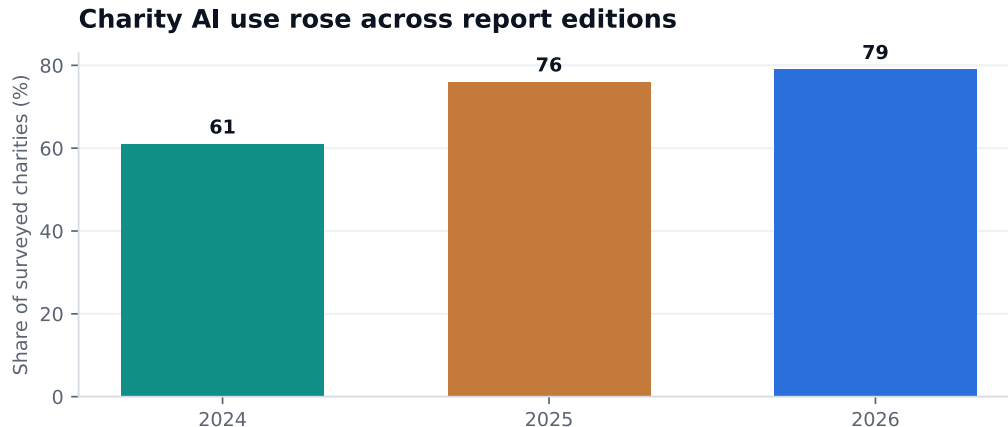
Hours returned to direct support means: time that staff can actually use with people, preparation or recovery because a needless administrative step has been removed or shortened. It is not an invitation to fill every reclaimed minute with a bigger caseload.

Adoption has accelerated. The operating model has not caught up.

The latest Charity Digital Skills Report is important because it stops this discussion being theoretical. Its 2026 edition, launched on 9 July 2026 and based on 807 responses, says 79 per cent of charities are using AI, rising to 92 per cent of large charities. That is a marked increase from 76 per cent in the 2025 edition and 61 per cent in 2024. The trend says that AI use is becoming ordinary. It does not say that the use is safe, strategic or evenly useful.

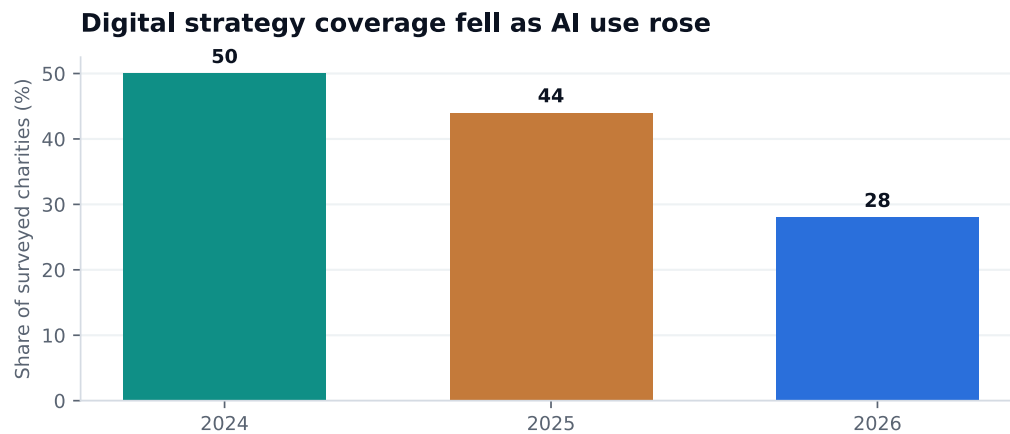
The same 2026 survey shows the contradiction leadership teams need to face. Eighty-one per cent reported digital progress, yet only 28 per cent had a digital strategy, down from 44 per cent in 2025 and 50 per cent in 2024. Sixty-three per cent use AI for administration and project management, precisely where paperwork pressure is felt. But 56 per cent named skills as their biggest AI barrier, 63 per cent named squeezed finances as their biggest barrier to digital progress, and 35 per cent said they do not trust AI tools. The last figure is more than double the 2025 result. Adoption is therefore not proof of confidence. It may equally describe staff trying unsanctioned tools because the official route is absent.

This is where directors should resist a simplistic reading of the numbers. A smaller organisation can be using AI to draft a meeting agenda while lacking a policy for client data, a training plan or someone who can answer a difficult question from a worried colleague. The 2026 report records that 33 per cent of boards have poor AI skills, 44 per cent identify staff training as their leading funding need, and 60 per cent say sector-wide AI training is essential. Only 17 per cent received dedicated digital funding. It is unreasonable to demand mature governance from a sector asked to modernise while funding the work only as an afterthought.



Charity Digital Skills Report editions 2024-2026 (survey). AIMonger redraw.

Figure 1. Reported charity AI use rose from 61% in the 2024 edition to 76% in 2025 and 79% in the 2026 edition. Read this as a survey trend, not a measure of safe or effective use. Source: Charity Digital Skills Report editions 2024-2026. AIMonger redraw.



Charity Digital Skills Report editions 2024-2026 (survey). AIMonger redraw.

Figure 2. The reported share of charities with a digital strategy fell from 50% in the 2024 edition to 44% in 2025 and 28% in 2026. The series does not establish cause, but it shows why AI adoption cannot stand in for a written operating plan. Source: Charity Digital Skills Report editions 2024-2026. AIMonger redraw. The funder position matters too. The 2026 report says 55 per cent of funders do not know whether applicants use AI, while 58 per cent say it is changing the applications they receive. This creates an awkward gap. Teams may be expected to show digital competence in bids while having little clarity on what funders consider acceptable, and while still carrying responsibility for privacy, accuracy and the final account of how money was spent. A director should not solve that ambiguity by forbidding everything or quietly permitting everything. The better response is a short, written rule for approved work, prohibited data and human review.

Where technology helped, and where it stalled

The strongest reason to begin with one workflow is that evidence for administrative technology is mixed. Ambient documentation, which listens to a consultation and drafts a note for a professional to review, is a useful test case. It is more mature than many NGO uses of generative AI and has been studied in real settings. It shows both what can be gained and why a tool alone does not change the shape of a day.

A 2026 multi-site US study led by Mass General Brigham and published in JAMA associated AI scribes with modest daily reductions of around 13 minutes in total electronic-record use and 16 minutes in documentation time, or approximately 3 per cent and 10 per cent. The gains were stronger for people using the technology in more than half of their visits. Yet after-hours record time did not differ significantly between groups. That result deserves attention. A drafting tool can improve a slice of the task without guaranteeing that evening work disappears.

The early randomised evidence is similarly qualified. A 2025 three-arm trial of 238 outpatient physicians found that one product reduced note-writing time by 9.5 per cent compared with usual care, while the other showed no significant benefit. A 2026 BMJ Digital Health commentary makes the right methodological point: favourable pilots, vendor claims and before-and-after comparisons are not enough. Organisations need to measure time, quality, error patterns, staff experience and what happens to the time supposedly saved. In other words, a polished draft is an intermediate outcome, not the mission outcome.

There is positive evidence too. A Great Ormond Street Hospital-led NHS study of the TORTUS ambient-scribing tool across nine London sites, covering more than 17,000 encounters, reported a 23.5 per cent rise in direct patient interaction time and an 8.2 per cent fall in appointment length. Its safety boundary was clear: the tool drafted notes and letters, while clinicians checked and edited them, and it did not make clinical decisions. This is the model worth borrowing, not because an NGO is a hospital, but because it names the human responsibility and measures a human result.

For NGO leadership, the lesson is practical. Technology may help with a first draft, locating a current policy, identifying a duplicated request or producing a structured handover. It stalls when documentation rules expand faster than drafting speed, when a separate system creates double entry, or when reclaimed minutes are silently replaced with another target. It also stalls when trust is treated as a communications problem. The 35 per cent trust concern in the Charity Digital Skills Report 2026 is evidence that staff need a credible boundary, training and the ability to challenge outputs, not a cheerful announcement that a tool is here to help.

An agentic layer, not a universal fix

Not every fix in this roadmap needs an AI system sitting inside it. Naming an owner for a workflow, writing the one-page data rule set out earlier, or retiring a template nobody trusts are acts of organisational discipline, not automation, and no AI system did any of that work. Dressing a decision to appoint a reviewer or delete a dead form as an “AI project” spends the small amount of trust and budget an NGO has for this kind of change on something a spreadsheet and a Friday afternoon could have delivered alone. A director who cannot tell the difference will eventually buy a tool to solve a problem that was never technical.

The picture changes once the lens widens beyond a single workflow. A note takes too long on a Tuesday, a handover drops a detail on a Wednesday, a funder report reconstructs the same attendance figures for the third time this quarter, and each looks like an isolated inconvenience from inside its own task. Every Pathway Brief in this series treats it that way on purpose, because a bounded fix with a named owner is what actually gets built in a week. But the pattern across notes, handovers, funder reports, onboarding, meetings and parent communication, where time concentrates, where the same interruption recurs, where a safeguarding concern sits quietly for days before anyone notices, is not visible from inside any single workflow. This is the ground an agentic layer, a small set of connected AI agents that watch several bounded workflows, reason about where the recurring bottleneck or risk actually sits, and prompt a named person before a deadline or safeguarding concern turns urgent, can occupy that a one-shot answer or a per-task chatbot cannot reach. AIMonger’s position is specific: an agentic layer belongs on top of workflows that already work, watching and flagging across the organisation, and it is not a platform sold to replace them.

Agentic layer, plainly: a small set of connected AI agents, each given a bounded view of specific workflows, that reason about where the pattern across those workflows sits and prompt a named person before a risk or deadline turns urgent, rather than waiting to be asked a question.

The counter-position has a name and a recognisable shape. Gartner’s analysis of 25 June 2025 predicts that over 40 per cent of agentic AI projects will be cancelled before the end of 2027, citing escalating cost, unclear business value and thin risk controls, and it coins the term “agent washing” for vendors who relabel an existing chatbot or a piece of robotic process automation as an autonomous agent, estimating that only around 130 of the thousands of vendors now claiming agentic capability are doing anything of the kind. Many of the use cases sold as agentic today, the same analysis states, do not need an agentic system at all, and Gartner’s 2026 review of generative AI projects puts the wider failure rate higher still: at least half were abandoned after proof of concept by the end of 2025, for the same reasons of cost, patchy data and unclear value. McKinsey’s 2025 State of AI survey shows why: 88 per cent of organisations report using AI somewhere in the business, yet only around 6 per cent attribute more than 5 per cent of EBIT impact to it, because most bolt AI onto a process they never redesigned around it. A vendor that pitches AI as the fix for everything, without first asking what already works here, an escalation route staff already trust, a practitioner’s judgement, a template nobody wants replaced, cannot make that advantage compound; each new module it sells resets the organisation to a generic workflow instead of building on the one already earning trust.

This is the doctrine already set out earlier in this paper, restated for the agentic layer rather than for a single tool: listen before buying, one workflow at a time, a human gate at the point of consequence. SAS’s July 2026 research on agentic AI return gives that instinct a number: organisations that use

automation to scale staff judgement rather than replace it are reported around 60 per cent more likely to double the return on an AI project, in what the researchers call “human in the lead”, a step beyond “human in the loop”. An agentic layer earns its place here on the same terms as everything else in this roadmap: added to what already works, watched by a named person, never sold as a replacement for it.

Listen before buying

The most useful first artefact is a map of the real working week, made with the people who carry it. Spend a week listening across roles. Ask where a note is retyped, where evidence becomes archaeology, what waits for one experienced person and which messages need a judgement call rather than a faster reply. Do not begin by asking which product has the most features. A product cannot tell you which delay matters most to a family, a funder or a practitioner.

This listening work turns vague frustration into a manageable operating question. A programme manager may discover that the hardest part of a funding report is not writing but reconciling three versions of attendance data. A facility lead may find that handovers fail because observations are buried in free text, not because staff lack commitment. A clinical or educational lead may learn that people are avoiding a shared system because it does not match the order in which care or support is actually delivered. Those findings can lead to templates, clearer ownership and cleaner source records before AI is involved.

Workflow, in plain English: the repeatable route by which a piece of work moves from a request to a checked outcome, including the people, records and decisions in between.

Worked example: “prepare the monthly funder update” starts with named evidence sources, passes through a draft with gaps marked, and ends only when a manager checks the figures and narrative against the source record.

Protect people and data before prompts

The protected path must be easier than the risky path. GDPR already requires care around personal data and special-category data. The EU AI Act adds duties that become especially relevant when a system affects access to services or is used in a consequential context. NIST AI 600-1 and UNESCO’s education guidance are useful here because neither asks a small organisation to become a laboratory. Both point towards ordinary disciplines: know which system is being used, know what data enters it, preserve provenance, test outputs, keep human oversight where harm could follow.

For a small NGO, this should become a one-page rule, not a legal library. It should say that identifiable client, child, health, disability, safeguarding and case material does not go into consumer AI accounts. It should name the approved route for any sensitive material, who can authorise an exception and what must be checked before an external message, plan or report is sent. It should also say that staff can stop and ask when a tool behaves oddly. A policy that staff cannot remember at 5.30 p.m. is not a control.

Three controls are more valuable than a grand launch:

1. **Classify the work before choosing the tool.** Public training material and a generic agenda do not carry the same risk as a client plan or safeguarding concern. Writing down a small number of categories gives staff a usable answer rather than asking them to make a legal judgement under pressure.
2. **Keep a human gate at the point of consequence.** A draft can save time, but a person remains accountable for a clinical note, educational plan, family message, funder report or referral. This is not ceremonial review: the reviewer needs source material, permission to reject the draft and enough time to do so.
3. **Record enough to investigate a problem.** An inventory of approved tools, purposes, owners and data categories is a modest administrative task that prevents a much larger one after an

incident. It also allows leadership to retire unused tools instead of letting subscriptions and informal practices multiply.

Choose one workflow, not a platform

Once the listening map and data boundary exist, choose one workflow where the organisation can see a result in weeks rather than a vision in a year. A good first candidate is frequent, bounded and reviewable. It has a recognisable beginning and end, existing source material, a named owner and a clear human check. Drafting action notes from approved meeting notes, retrieving the latest policy, preparing a funder-report skeleton or producing a structured handover may qualify. Automated decisions about eligibility, safeguarding, diagnosis or behaviour do not.

The choice should also account for who will use the result. A person writing family communication needs tone, accuracy and approval, not a generic “assistant”. A new starter needs a clear step sequence and a named person to ask, not a large folder of old examples. A manager preparing a report needs traceable evidence and gaps surfaced early, not fluent prose that conceals missing data. Different work has different review standards. One generic chatbot is rarely an operating model.

Skill recovery means: capturing safe methods, examples and watch-outs so the departure of an experienced colleague does not force the next person to reconstruct the job by rumour.

In practice: a short, reviewed playbook for the first parent call after an assessment, with anonymised examples and a named buddy, is more useful than an unread ninety-page induction file.

Measure the hours that return, then protect them

The measure is not how many prompts were sent or how impressive a demonstration looked. Track the time from request to checked completion, the number of duplicated asks, late notes, rework, and staff’s own account of whether the path became easier. Where possible, compare a small sample of the old and new path. If a handover takes less time but misses a safeguarding detail, it has failed. If a report drafts faster but creates an evening of fact-checking, the gain is smaller than it first appeared.

Leadership must then decide what the reclaimed capacity is for. The honest answer may be direct support, better preparation, same-day note closure, calmer supervision or an earlier finish. It should not automatically become another batch of work. This is the distinction the ambient-scribe evidence brings into focus: reducing note time does not automatically create a better day. The organisation has to protect the time it recovers.

The following phases keep the doctrine simple. The table supports the discussion; it is not a substitute for it.

Phase	Leadership task	Evidence of progress
0. Listen	Walk the week with staff before proposing a tool.	A pain map staff recognise as accurate.
1. Protect	Set data categories, approved routes and human gates.	A one-page rule people can use under pressure.
2. One workflow	Improve a bounded, frequent task with a named owner.	A checked result in the same working week.
3. Measure	Compare time, rework, quality and direct-support capacity.	A small set of measures the team believes.
4. Train	Practise the real path in short sessions and refresh it.	People can use it without relying on one enthusiast.
5. Widen	Add a second workflow only after the first is stable.	Fewer parallel tools and no abandoned pilot.

Train for judgement, not button pressing

Training is not an optional extra after procurement. The Charity Digital Skills Report 2026 records both the skills barrier and the funding gap. Short practice sessions should use realistic but safe examples: identify what must never be pasted into an unapproved tool, compare a weak draft with source material, revise tone for a family audience and decide when not to use AI at all. This builds discernment, which is more useful than a lesson on a changing interface.

Leaders also need training. When a third of boards are reported to have poor AI skills, governance cannot be delegated entirely to a junior digital champion. Directors and trustees do not need to become technical specialists. They do need to ask who owns the workflow, what data crosses a boundary, what is checked by a person, what evidence shows benefit, and whether staff feel able to challenge the system. A team that knows those questions can make a cautious first step without pretending certainty.

Widen only when the first path is ordinary

Scale is a tempting answer to scarce funding, but it is usually the wrong first answer. Add a second workflow when the first has a named owner, an understood data boundary, a repeatable review step and measures that have held for long enough to be credible. The signs of readiness are boring: fewer versions of the same document, fewer late notes, less rework and no need for the original project champion to rescue every case.

This is how the shorter NGO Pathway Briefs should be used. They apply the shared doctrine to a specific path such as notes, handovers, parent communication, funder reporting or onboarding. They are starting points for a local conversation, not a template to copy unchanged. The listening map decides which brief is relevant, and local policy decides whether the proposed route is acceptable.

Counter-position

Another leadership team may say that it is too small, too regulated or too tired for any AI work. That caution can be justified. Paper, a clearer template and a better shared folder may be the correct first intervention. A consumer chatbot is not an acceptable shortcut for sensitive client material, and an automated external message is not a substitute for professional judgement.

The counter-position becomes less convincing when it treats every change as equally risky. Doing nothing also has a cost: staff remain dependent on personal memory, personal inboxes and late-night reconstruction; unofficial tools become more attractive; and knowledge leaves with the people who know how the work is actually done. The proposed alternative is not a large transformation programme. It is a bounded improvement, with a human gate, evidence of benefit and permission to stop if the burden rises.

Decision criteria before funding anything

- 1. Can staff name the friction in their own words?** A workflow chosen from a listening map is more likely to solve a real delay than one chosen from a vendor demonstration. If the description is vague, continue listening rather than buying a broad platform.
- 2. Is the data boundary clear enough for a busy afternoon?** Staff must know what never enters a consumer tool, what can use an approved route and who can answer an uncertain case. A long policy that cannot be applied in the moment leaves the risky decision to the individual.
- 3. Does the output have a named human reviewer?** The reviewer must be able to compare it with source material and reject it without penalty. Where the consequence is clinical, educational, safeguarding-related or external, that accountability cannot be automated away.

4. **Will the team measure a human result within ninety days?** Record time to checked completion, rework, late notes, duplicated requests and direct-support capacity. Prompt volume and licence use can explain cost, but they do not show whether the service improved.
5. **Can a new starter use the path safely?** If the workflow only works for the two people who designed it, it has not reduced organisational fragility. Training, a short playbook and a named support person turn a clever workaround into a shared practice.
6. **Is there a reason to widen now?** Add the next workflow only when the first remains useful without constant rescue and does not increase burden elsewhere. A portfolio of half-used pilots is not digital progress.

FAQ

Who is this paper for?

It is for NGO directors, facility leads and programme managers, as well as the colleagues who carry the practical work around them. The examples travel across care, disability support, education and community services, but they must be adapted to local policy, professional duties and the people served.

Is this about replacing therapists, teachers or carers?

No. The paper is about reducing avoidable administrative drag and protecting time for direct support, preparation and recovery. Professional judgement, safeguarding decisions and responsibility for external outputs remain with people.

What should we do first if there is no budget for a large project?

Map one week of friction with staff. Choose a single bounded document, handover or evidence path, improve the source record and template first, then test a governed drafting or retrieval step only if it removes a real burden. The matching Pathway Brief can help frame that local test.

How can we stay safe with children's or clients' data?

Keep identifiable and special-category material inside approved systems and routes. Do not paste it into consumer AI services. Maintain a simple tool inventory, require human review for consequential outputs and seek specialist privacy or professional advice where the work calls for it.

Why not judge success by time saved alone?

Because a fast draft can create rework, errors or a new queue elsewhere. Measure quality, delay, duplication and whether staff gained usable direct-support capacity. The aim is a better working week and safer continuity, not a higher count of automated actions.

What should funders ask?

Ask what workflow is being improved, what data is protected, how staff are trained, who reviews outputs and what human outcome will be measured. The 2026 Charity Digital Skills Report shows many funders still do not know whether applicants use AI; clear questions can support responsible work rather than force hidden use.

References

1. Charity Digital Skills Report, Charity Digital Skills Report 2026 (launched 9 July 2026; 807 respondents). <https://charitydigitalskills.co.uk/report/>

2. Federation of European Social Employers, Evolving jobs, skills and training needs in the social services sector and the role of social partners in managing changes (2023). https://www.socialemployers.eu/wp-content/uploads/2023/09/report_Evolving_jobs_skills_and_training_needs_in_the_social_services_sector.pdf
3. Eurofound, Social services in Europe: Adapting to a new reality (2023). <https://www.eurofound.europa.eu/en/publications/2023/social-services-europe-adapting-new-reality>
4. OECD, Beyond Applause? Improving Working Conditions in Long-Term Care (2023). <https://doi.org/10.1787/27d33ab3-en>
5. WHO Regional Office for Europe, Mental Health of Nurses and Doctors survey in the European Union, Iceland and Norway (2025). <https://www.who.int/europe/publications/i/item/WHO-EURO-2025-12709-52483-81031>
6. Rotenstein, L. et al., “Changes in Clinician Time Expenditure and Visit Quantity With Adoption of Artificial Intelligence-Powered Scribes,” JAMA (2026), reported by Mass General Brigham. <https://www.massgeneralbrigham.org/en/about/newsroom/press-releases/ai-scribes-linked-to-modest-reductions-in-ehr-documentation-time>
7. Lukac, P. J. et al., “Ambient AI Scribes in Clinical Practice: A Randomized Trial,” NEJM AI (2025). <https://pmc.ncbi.nlm.nih.gov/articles/PMC12768499/>
8. BMJ Digital Health, “Does ambient voice technology in healthcare really save time? The need for rigorous evaluation to test the promise” (2026). <https://doi.org/10.1136/bmjdh-2026-000039>
9. Great Ormond Street Hospital, “GOSH-led trial of AI-scribe technology shows transformative benefits for patients and clinicians across London” (2025). <https://www.gosh.nhs.uk/news/researchgosh-led-trial-of-ai-scribe-technology-shows-transformative-benefits-for-patients-and-clinicians-across-london/>
10. NIST, Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST AI 600-1 (2024). <https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence>
11. Regulation (EU) 2024/1689, Artificial Intelligence Act. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
12. Regulation (EU) 2016/679, General Data Protection Regulation. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
13. UNESCO, Guidance for generative AI in education and research (2023). <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
14. CIPD, Employee induction (2025). <https://www.cipd.org/en/knowledge/guides/employee-induction/>
15. ILO, Decent work and the care economy, ILC.112/Report VI (2024). <https://www.ilo.org/media/534421/download>
16. European Commission, European Care Strategy, COM(2022) 440 final. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022DC0440>
17. Gartner, “Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled By End of 2027” (25 June 2025). <https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027>
18. Gartner, “Why Half of GenAI Projects Fail: Avoid These 5 Common Mistakes” (2026). <https://www.gartner.com/en/articles/genai-project-failure>

-
- 19.** McKinsey & Company, The State of AI: Global Survey (2025 edition). <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>
 - 20.** SAS, “For Agentic AI ROI, Invest in Human Judgment” (22 July 2026). <https://www.prnewswire.com/news-releases/sas-for-agentic-ai-roi-invest-in-human-judgment-302831330.html>
-

Published by AIMonger . <https://aimonger.com> . Malta . Europe