

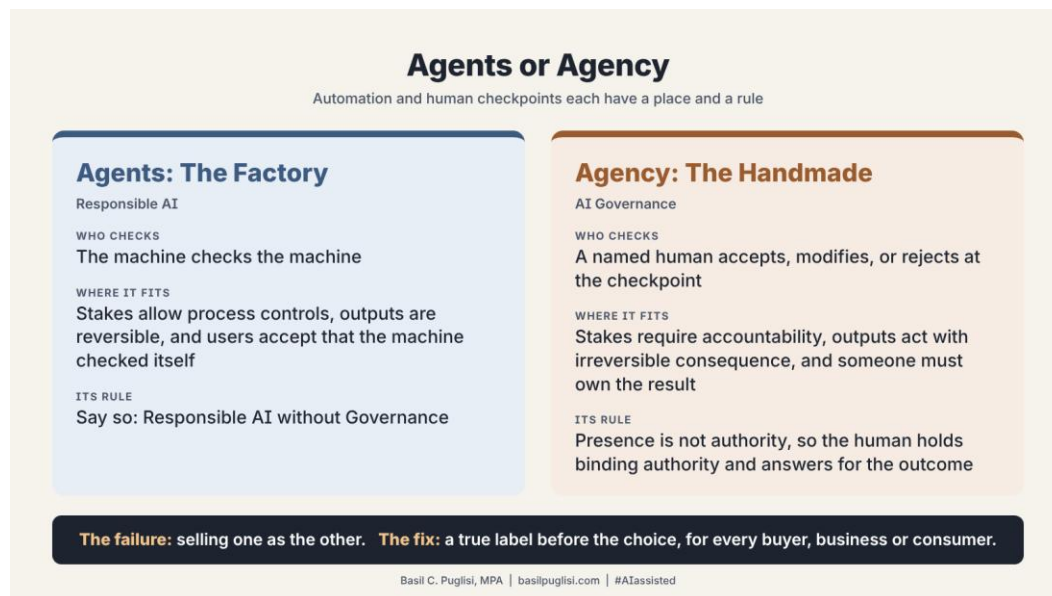
Agents or Agency: Transparency Is Success

Automation and human checkpoints each have a place and a rule. The principle that holds both together is that no one gets fooled, because factory or handmade is a choice every customer makes, business or consumer.

Every organization that puts AI to work makes the same choice, whether it names the choice or not. It can let agents carry the work, with machines checking machines from the first step to the final output. It can also keep human agency in the work, with a named person holding authority at checkpoints and answering for what comes out. The first is automation, which belongs to Responsible AI, and the second is human oversight and accountability, which is [AI Governance](#).

Those labels follow the [operational definitions](#) the author set out on January 1, 2026, and has built on since. "Ethical AI is the articulation of values." "Responsible AI translates values into machine behavior." "AI Governance assigns decision rights, escalation paths, override authority, and consequence ownership" (Puglisi, 2026a). The field has no settled definition of any of the three, which is why the author proposed his own. Readers working inside other frameworks will find the same words used more broadly there.

Neither choice is wrong. The failure comes when one is sold as the other, and the remedy for that failure is the subject of this article: telling the buyer, plainly and before the purchase, which one they are getting.



Agents or agency: the factory and the handmade product, each with a place and a rule.

Agents: The Factory

Responsible AI is the factory. An agent runs the steps, other software checks the work, and the system validates itself against its own parameters. The author summed up that condition in January 2026 in five words: "The machine checks the machine" (Puglisi, 2026b).

The factory has a real place. The same letter called factory quality "the appropriate standard when stakes allow process controls to suffice, when outputs are reversible, when users accept that the machine checked itself" (Puglisi, 2026b). Days later the author named where it belongs: "Consumer chatbots processing millions of interactions, recommendation engines serving billions of suggestions, content moderation at platform scale: these operate appropriately under Responsible AI with honest disclosure" (Puglisi, 2026c). He was equally plain that "Choosing Responsible AI over AI Governance is a legitimate organizational decision based on risk tolerance and consequence severity" (Puglisi, 2026c).

The factory also carries a rule. "Systems operating at speeds or scales that preclude human checkpoint authority should carry explicit disclosure: Responsible AI without Governance. This is not a limitation to hide. It is a category to name" (Puglisi, 2026b). A companion rule closes the gap from the other side: "Responsible AI must never be called AI Governance without direct human oversight" (Puglisi, 2026c).

Factory work still has an owner. After a customer relied on its website chatbot, Air Canada took the position, as the tribunal described it, that the chatbot was "a separate legal entity that is responsible for its own actions." The [British Columbia Civil Resolution Tribunal](#) called that "a remarkable submission" and wrote, "It should be obvious to Air Canada that it is responsible for all the information on its website" (*Moffatt v. Air Canada*, 2024). The factory can run without a hand on each output, but it cannot run without someone who answers for it.

Agency: The Handmade

AI Governance is the handmade product. [Checkpoint-Based Governance](#) puts a named human at defined points in the work with binding authority over what happens next, and that person accepts, modifies, or rejects what the machine produced (Puglisi, 2025). The human answers for the result through at least one of four channels: moral, professional, civil, and criminal (Puglisi, 2026e). [The January 28 letter](#) put the difference in one line, "The handmade product has a human hand on the output" (Puglisi, 2026b). The analysis that followed three days later stated the consumer's promise plainly: "someone inspected this specific item before it reached you" (Puglisi, 2026c).

Handmade does not mean a person reads every output. The same letter said so directly: "Governance does not require reviewing every output. It requires defining thresholds where human authority activates before irreversible consequence" (Puglisi, 2026b). Where outputs arrive faster than any person can read them, [the June article](#) moved the checkpoint upstream, so

that "the named human signs the policy that authorizes the class of outputs" (Puglisi, 2026e). The hand is on the decision that matters, wherever that decision sits.

The handmade product has its own place, "when stakes require accountability, when outputs act with irreversible consequence, when someone must own what the system produces" (Puglisi, 2026b). Its rule is that presence is not authority, so a person who only watches does not turn the factory into a workshop. The June article drew that line directly: "Any reduction in substantive human engagement at a checkpoint converts AI Governance back to Responsible AI, regardless of whether a human is physically present" (Puglisi, 2026e). [The January 31 analysis](#) said it in fewer words: "Factory with quality control remains factory with quality control" (Puglisi, 2026c).

A pending lawsuit raises the question of whether the people a policy names as deciding hold the authority it promises. UnitedHealthcare's [policy documents](#) told members that "UnitedHealthcare's Clinical Services Staff and Physicians make decisions on the health care services you receive" (*Estate of Lokken v. UnitedHealth Group, Inc.*, 2025). [The plaintiffs allege](#) that "Employees who deviate from the nH Predict AI Model projections are disciplined and terminated" (*Estate of Lokken v. UnitedHealth Group, Inc.*, 2024). UnitedHealth [denies the allegations](#), and the case is pending (*Estate of Lokken v. UnitedHealth Group, Inc.*, 2026). If the allegations were proven, the staff named as deciding would have been present without the authority the policy promised.

A name on a form isn't enough either. [Elish \(2019\)](#) described the moral crumple zone, where responsibility lands on a human who had limited control over an automated system. The rule asks for binding authority because a person placed at the checkpoint only to absorb blame is still part of the factory.

The handmade product can carry its own mark. Manufacturing traditionally had a system for this in quality assurance: the name or serial number of the employee responsible for the inspection. The 2015 edition of [ISO 9001](#), the international quality management standard, required that the record of a product's release trace back to the person or persons who authorized it (International Organization for Standardization, 2015). AI Governance can carry the same mark. A line such as "Inspected by #S17G89F" shows the customer that if something goes wrong, someone has to answer for it, and this is who. The mark faces the same test the June article set for any signature, because "a name on a form without authority, competence, evidence access, and a demonstrable record of override does not pass" (Puglisi, 2026e).

The Principle: Do Not Fool People

Both models work, and both fail the same way when the label lies. The letter of January 28 opened its comparison with a caution: "This is not a moral judgment. It is a quality distinction" (Puglisi, 2026b). It then named the failure point in four sentences: "Neither is universally correct. Organizations choose factory or handmade based on context and acceptable risk. The

problem emerges when factory quality is labeled as governance quality. Users, enterprises, and policymakers deserve to know which they are trusting" (Puglisi, 2026b).

The market already has a name for one form of that lie. [Gartner](#) used "agent washing" in June 2025 to describe "the rebranding of existing products, such as AI assistants, robotic process automation (RPA) and chatbots," without the capability the new label promised. It also estimated that "only about 130 of the thousands of agentic AI vendors are real" (Gartner, 2025). The June article carried the term one step further, to "the practice of selling autonomy as if it were governance," and argued that "selling factory mode as governance is the mislabeling that fuels the backlash" (Puglisi, 2026e). [The February article](#) traced the same move through marketing language, where a human in the loop can be "technically present but operationally decorative." The real test, it argued, is "the difference between a human who participates and a human who decides" (Puglisi, 2026d).

The plainest case sold the factory as a professional. DoNotPay promoted its service as "the world's first robot lawyer." The [Federal Trade Commission](#) alleged that the company "did not conduct testing to determine whether its AI chatbot's output was equal to the level of a human lawyer" (Federal Trade Commission, 2024). [The final order](#), issued in January 2025, required \$193,000 in monetary relief and notice to people who subscribed from 2021 through 2023. It also barred claims that the service "operates like a human lawyer" unless the company has competent and reliable evidence behind them (Federal Trade Commission, 2025).

Transparency, then, is not a courtesy added to either model. It is the condition that lets both exist side by side, because a buyer can weigh the speed of the factory against the accountability of the workshop only when each one says which it is.

The Cost of Saying So

Transparency carries a price, and the record should state it. Across 13 experiments, [Schilke and Reimann \(2025\)](#) documented that people who disclosed their own use of AI were trusted less than people who did not. The penalty held across the ways the disclosure was framed and whether disclosure was voluntary or required. It was smaller, though not eliminated, among evaluators with favorable views of technology and high confidence in AI accuracy. The study tested people and organizations disclosing their own AI use, not product labels, and research on labels points the same way.

In two preregistered experiments with 4,976 people in the United States and the United Kingdom, [Altay and Gilardi \(2024\)](#) found that an "AI-generated" label lowered the perceived accuracy of headlines whether a person or a machine had written them. They traced the drop to "expectations that headlines labeled as AI-generated have been entirely written by AI with no human supervision." The authors report that the labels "only induced skepticism when participants were given no AI definitions" or were told that AI chose the topic and wrote the whole article. A label defined as a journalist choosing the topic, even with AI writing the first

draft, did not induce it. Product descriptions that named artificial intelligence lowered purchase intentions through lost emotional trust, most of all for high-risk products ([Cicek, Gursay, & Lu, 2024](#)). In medicine, naming the human did not remove the penalty: identical advice labeled as coming from a physician working with AI was rated less reliable and less empathetic than the same advice credited to a physician alone ([Reis, Reis, & Kunde, 2024](#)). None of these studies tested a label that names who answers for the output, so how buyers would read a factory or handmade label remains an open question.

Comfort follows the human role, and people cannot see that role today. Across six countries, [the Reuters Institute](#) reported that just 12 percent of people were comfortable with news made entirely by AI. The figure rose to 21 percent with a "human in the loop" and to 43 percent "when a human leads with some AI help." Only 33 percent on average think journalists "always" or "often" check AI outputs (Simon, Nielsen, & Fletcher, 2025).

Schilke and Reimann recorded one outcome worse than disclosing: having the AI use exposed by someone else. The European rules described below make that outcome more likely, because the Code of Practice that accompanies them asks providers to mark AI output in machine-readable form and to offer a way to detect it (Paul, Weiss, 2026). Some hidden AI use grows easier to find, and the research shows that exposure is the costlier way for the truth to come out.

Two objections stand beside the trust penalty, and neither one goes away because the label is honest. Disclosure research outside AI records a further risk. Advisers who disclosed a conflict of interest gave more biased advice, and the people they advised felt more pressure to follow it ([Loewenstein, Sah, & Cain, 2012](#); [Sah, Loewenstein, & Cain, 2013](#)). Knowing that a system is a factory is also not the same as being able to choose another, since a person on the receiving end may have no other provider to turn to. The penalty and the objections are all real, and this article does not claim to have answered any of them.

The Buyer's Choice, Business or Consumer

Factory or handmade is the question every customer answers, whether the customer is a company buying an AI system or a person using one.

For a business buyer, the June article framed it as "a procurement question that any officer can answer," reduced to plain terms: "who can stop it, and who answers for it?" (Puglisi, 2026e). A vendor that can name the person, the checkpoint, and the authority that person holds is selling handmade. A vendor that cannot is selling factory, which may be exactly what the buyer needs, as long as the contract and the sales pitch both say so.

Adoption is running ahead of that question. In September 2026, [EY reported](#) that 91 percent of 202 senior AI decision-makers at large US public companies said their organizations use agentic AI, and that 49 percent of those users had not updated their governance for it (EY, 2026).

The label also has to reach past the buyer. When a bank buys a factory system that decides on customer accounts, the account holders never chose it, so the honest label belongs in front of them too. A bank that tells its account holders that machines decide, and how to reach a person, has put the risk in the open. A bank that records it only in its procurement file has not.

The Financial Stability Board's October 2026 [sound practices](#), which it calls non-prescriptive, describe a credit union doing that. AI handled data entry and the initial risk assessment, and human loan officers made the final approval. Applicants were told that AI was used in the initial stages, and members were made aware of their right to an explanation and an appeal. The same report does not call for disclosure to end-users in every case (Financial Stability Board, 2026).

The people a system decides about hold views of their own. In a [Pew Research Center survey](#) fielded in December 2022, 71 percent of Americans opposed AI making a final hiring decision, and 7 percent favored it (Pew Research Center, 2023). Job applicants never chose the screening software either. In *Mobley v. Workday*, [a federal court in May 2025](#) preliminarily let applicants aged 40 and over who allege age discrimination proceed together as a collective. The court noted that "The unified policy that Mobley has identified is Workday's AI recommendation system." A class certification hearing is set for March 9, 2027, and no court has ruled on the merits (*Mobley v. Workday, Inc.*, 2025, 2026).

A government system leaves the people it decides about no other provider at all. Australia's [Robodebt scheme](#) raised welfare debts by averaging income data automatically, with the aim "to reduce (and preferably obviate) human involvement" at the Department of Human Services (Royal Commission into the Robodebt Scheme, 2023, p. xxv). The Royal Commission called it "a crude and cruel mechanism, neither fair nor legal" (Royal Commission into the Robodebt Scheme, 2023, pp. xxix-xxx).

The Royal Commission's [Recommendation 17.1](#), which the government accepted, asks for a clear path to review and a plain-language explanation on departmental websites wherever automated decision-making is used. It also asks that business rules and algorithms be open to independent expert scrutiny (Department of the Prime Minister and Cabinet, 2026). In June 2026, [the Federal Court approved](#) a settlement of up to A\$548.5 million, on top of the roughly A\$112 million 2021 settlement (*Prygodicz v Commonwealth of Australia*, 2026).

For a consumer, the law now supplies part of the label. The transparency obligations in Article 50 of the EU AI Act apply from August 2, 2026. Generative systems already on the market before that date have until December 2, 2026, to mark their output under Article 50(2). The voluntary Code of Practice also asks its signatories to make watermark detection interoperable by February 2, 2027. The European Commission [finalized its guidelines](#) on these obligations on July 20, 2026 (European Commission, 2026a, 2026b; Paul, Weiss, 2026).

A [client memo from Paul, Weiss](#) summarizing the guidelines and the companion Code of Practice reports that AI systems interacting with people must tell them they are dealing with AI. Agents must disclose both their artificial nature and the person on whose behalf they act, and a

generic statement or a notice buried in terms and conditions does not meet the duty. AI-generated text published to inform the public on matters of public interest must be labeled unless it has gone through human review or editorial control. The exemption also requires that a natural or legal person hold editorial responsibility for it. Spell-checking or a cursory sign-off does not qualify, and the Code asks for documented workflows rather than assertions (Paul, Weiss, 2026).

That last rule reaches only text on matters of public interest, not every AI product, and it runs in a direction worth noticing. The law labels the text that no responsible person reviewed and leaves the reviewed text unmarked. The reader learns when a machine spoke alone and learns nothing extra when a responsible person stood behind the words. A rubber stamp still does not earn the exemption.

The United States reaches the same line by other routes. In March 2024, the [Securities and Exchange Commission](#) charged two investment advisers, Delphia and Global Predictions, with false and misleading claims about their use of AI. The firms paid \$400,000 in combined penalties, and the agency's chair called the practice AI washing (U.S. Securities and Exchange Commission, 2024). Colorado's [SB 26-189](#), signed May 14, 2026, and effective January 1, 2027, requires deployers to give consumers clear notice before automated decision technology is used to materially influence a consequential decision. Deployers must also explain an adverse outcome within 30 days and, on request, offer meaningful human review and reconsideration where commercially reasonable (Colorado General Assembly, 2026; [Dupree & Bates](#), 2026). Under the law, meaningful human review must come from an individual the employer designates who has authority to approve, modify, or override the decision. That reviewer must also be trained for the review and must not simply defer to the system's output ([Bryan Cave Leighton Paisner](#), 2026).

New York's courts took the other side of the same line. [Part 161](#), effective June 1, 2026, does not require attorneys or parties to disclose AI use in court papers, because the duties that attach to every signed paper already apply "regardless of whether AI tools were used." The model rule the courts may adopt goes one step further: "By signing such paper, an attorney or party certifies that such a review has been conducted" (New York State Unified Court System, 2026; Russo Law, 2026). Colorado puts the factory label in front of the people the system decides about, and New York lets the handmade product carry a signature in place of a label. Both point the people who rely on the output toward the same question: who answers for it.

Success Is a True Label

Agents or agency is a choice, not a contest. Agents give speed and scale, agency gives a person who answers, and each has a place where the other would be the wrong tool. Success isn't picking a side. It is a label true enough that every buyer, business or consumer, and every person the output acts on, knows which one is in play before it matters.

Most organizations surveyed say they keep a hand on the decisions that matter most, and most are not fully confident they could prove it. [Grant Thornton's 2026 survey](#) of 950 business leaders reported that only 5 percent allow agents to execute high-stakes decisions without human review (Grant Thornton, 2026b). The same survey reported that 78 percent "[lack full confidence](#) that their organization could pass an independent AI governance audit within 90 days" (Grant Thornton, 2026a).

Success can be counted as the share of labels that match a checkpoint record someone can independently check. A label that reads clearly but claims a checkpoint no one kept counts as a failure. So does an inspection mark that leads to no record of a person with the authority to act.

References

- Altay, S., & Gilardi, F. (2024). People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. *PNAS Nexus*, 3(10), pgae403. <https://doi.org/10.1093/pnasnexus/pgae403>
- Bryan Cave Leighton Paisner. (2026, September 14). *Colorado's updated AI law*. <https://www.bclplaw.com/en-US/events-insights-news/colorados-updated-ai-law.html>
- Cicek, M., Gursay, D., & Lu, L. (2024). Adverse impacts of revealing the presence of "Artificial Intelligence (AI)" technology in product and service descriptions on purchase intentions: The mediating role of emotional trust and the moderating role of perceived risk. *Journal of Hospitality Marketing & Management*, 34(1), 1-23. <https://doi.org/10.1080/19368623.2024.2368040>
- Colorado General Assembly. (2026). *SB26-189: Automated decision-making technology*. <https://leg.colorado.gov/bills/sb26-189>
- Department of the Prime Minister and Cabinet. (2026, March). *Robodebt Royal Commission implementation update*. Australian Government. <https://www.pmc.gov.au/sites/default/files/resource/download/robodebt-royal-commission-implementation-update-march-2026.pdf>
- Dupree, L. P., & Bates, L. N. (2026, May 26). *Colorado replaces landmark AI Act: An overview of the new SB 26-189 framework*. Finnegan. <https://www.finnegan.com/en/insights/articles/colorado-replaces-landmark-ai-act-an-overview-of-the-new-sb-26-189-framework.html>
- Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society*, 5, 40-60. <https://doi.org/10.17351/ests2019.260>
- Estate of Lokken v. UnitedHealth Group, Inc.*, No. 0:23-cv-03514 (D. Minn. filed Apr. 5, 2024) (first amended class action complaint). https://litigationtracker.law.georgetown.edu/wp-content/uploads/2023/11/Lokken_2024.04.05_AMENDED-COMPLAINT.pdf
- Estate of Lokken v. UnitedHealth Group, Inc.*, No. 0:23-cv-03514 (D. Minn. Feb. 13, 2025) (memorandum opinion and order). https://litigationtracker.law.georgetown.edu/wp-content/uploads/2023/11/Lokken_2025.02.13_MEMORANDUM-OPINION-AND-ORDER.pdf
- Estate of Lokken v. UnitedHealth Group, Inc.*, No. 0:23-cv-03514 (D. Minn. filed Sept. 4, 2026) (second amended answer). https://litigationtracker.law.georgetown.edu/wp-content/uploads/2023/11/Estate-of-Gene-B.-Lokken-The-et-al-v.-UnitedHealth-Group-Inc.-et-al_2026.09.04_SECOND-AMENDED-ANSWER-TO-AMENDED-COMPLAINT.pdf
- European Commission. (2026a, June 10). *Code of Practice on transparency of AI-generated content* [Frequently asked questions]. <https://digital-strategy.ec.europa.eu/en/faqs/code-practice-transparency-ai-generated-content>
- European Commission. (2026b, July 20). *Guidelines on transparency obligations for providers and deployers of AI systems*. <https://digital-strategy.ec.europa.eu/en/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems>
- EY. (2026, September 15). *EY survey finds that autonomous AI implementation outpaces oversight, yielding an AI governance gap* [Press release]. https://www.ey.com/en_us/newsroom/2026/09/ey-survey-finds-that-autonomous-ai-implementation-outpaces-oversight-yielding-an-ai-governance-gap
- Federal Trade Commission. (2024, September 25). *FTC announces crackdown on deceptive AI claims and schemes* [Press release]. <https://www.ftc.gov/news-events/news/press-releases/2024/09/ftc-announces-crackdown-deceptive-ai-claims-schemes>

- Federal Trade Commission. (2025, January 14). *In the matter of DoNotPay, Inc.: Decision and order* (Docket No. C-4812).
https://www.ftc.gov/system/files/ftc_gov/pdf/2323042_donotpay_decision_and_order_o.pdf
- Financial Stability Board. (2026, October 15). *Sound practices for responsible adoption of artificial intelligence (AI): Final report*. <https://www.fsb.org/uploads/P151026-1.pdf>
- Gartner. (2025, June 25). *Gartner predicts over 40% of agentic AI projects will be canceled by end of 2027* [Press release]. <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>
- Grant Thornton. (2026a, April 13). *Grant Thornton survey: A widening "AI proof gap" is emerging, but well-governed AI is showing results* [Press release].
<https://www.grantthornton.com/insights/press-releases/2026/april/grant-thornton-survey-on-ai-proof-gap>
- Grant Thornton. (2026b). *2026 AI impact survey report*.
<https://www.grantthornton.com/services/advisory-services/artificial-intelligence/2026-ai-impact-survey>
- International Organization for Standardization. (2015). *Quality management systems: Requirements* (ISO Standard No. 9001:2015), Clause 8.6. <https://www.iso.org/standard/62085.html>
- Loewenstein, G., Sah, S., & Cain, D. M. (2012). The unintended consequences of conflict of interest disclosure. *JAMA*, 307(7), 669-670.
<https://www.cmu.edu/dietrich/sds/docs/loewenstein/UnintendedConsq.pdf>
- Mobley v. Workday, Inc.*, No. 3:23-cv-00770-RFL (N.D. Cal. May 16, 2025) (order granting preliminary collective certification). https://www.govinfo.gov/app/details/USCOURTS-cand-3_23-cv-00770
- Mobley v. Workday, Inc.*, No. 3:23-cv-00770-RFL (N.D. Cal. July 13, 2026) (order granting motion to change time). <https://caselaw.findlaw.com/court/us-dis-crt-n-d-cal/239454.html>
- Moffatt v. Air Canada*, 2024 BCCRT 149.
<https://decisions.civilresolutionbc.ca/crt/crtd/en/525448/1/document.do>
- New York State Unified Court System. (2026). *Part 161: Use of artificial intelligence technology*.
<https://www.nycourts.gov/rules/part-161-use-artificial-intelligence-technology>
- Paul, Weiss, Rifkind, Wharton & Garrison LLP. (2026, August 4). *EU finalises transparency rules for AI-generated content* [Client memorandum].
https://www.paulweiss.com/media/204bwi3d/eu_finalises_transparency_rules_for_ai_generated_content.pdf
- Pew Research Center. (2023, April 20). *Americans' views on use of AI in hiring*. In *AI in hiring and evaluating workers: What Americans think*.
<https://www.pewresearch.org/internet/2023/04/20/americans-views-on-use-of-ai-in-hiring/>
- Prygodicz v Commonwealth of Australia (Robodebt settlement)* [2026] FCA 1029.
<https://www.judgments.fedcourt.gov.au/judgments/Judgments/fca/single/2026/2026fca1029>
- Puglisi, B. C. (2025, September 23). *Checkpoint-Based Governance: An implementation framework for accountable human-AI collaboration (v2 drafting)*. basilpuglisi.com.
<https://basilpuglisi.com/checkpoint-based-governance-an-implementation-framework-for-accountable-human-ai-collaboration-v2-drafting/>
- Latest update: Puglisi, B. C. (2026, September 11). *Checkpoint-Based Governance: Presence is not authority*. basilpuglisi.com. <https://basilpuglisi.com/cbg/>

Sources in this work: Bradley (2025); Congruity 360 (2025); European Parliament and Council (2024); ISACA (2025); International Telecommunication Union (2025); Lu et al. (2019); Lumenalta (2025); McKinsey & Company (2025); National Institute of Standards and Technology (2023); OECD (2025); Parasuraman and Manzey (2010); Precisely (2025); Splunk (2025); Stanford Institute for Human-Centered Artificial Intelligence (2024); Strobes Security (2025); Superblocks (2025); Visa (2025).

Puglisi, B. C. (2026a, January 1). *What we failed to define is how we fail*. basilpuglisi.com.

<https://basilpuglisi.com/what-we-failed-to-define-is-how-we-fail/>

- Follow-up: Puglisi, B. C. (2026, March 14). *AI Governance has no formal definition. Here is one*. basilpuglisi.com. <https://basilpuglisi.com/ai-governance-has-no-formal-definition-here-is-one/>
- Latest update: Puglisi, B. C. (2026, September 15). *Artificial intelligence: Technology, capability, governance, and human accountability*. basilpuglisi.com. <https://basilpuglisi.com/ai-artificial-intelligence/>

Sources in these works: Bai et al. (2022); Batool, Zowghi, and Bano (2025); Center for AI Safety (2023); CGTN (2024); Companies committed to responsible AI (2021); Council of Europe (2026); Dietterich (2000); European Commission (2019); European Union (2024); Floridi and Cowan (2025); Goffi (2022); Harper (n.d.); Henrich, Heine, and Norenzayan (2010); IBM (2025); International Organization for Standardization (2021, 2022, 2023); Irving, Christiano, and Amodei (2018); Jobin, Ienca, and Vayena (2019); Maas (2025); Merriam-Webster (n.d.); Microsoft (2022); MIT Sloan School of Management (2023); National Institute of Standards and Technology (2023, 2024); OECD (2019); OpenAI (2023); Oxford English Dictionary (n.d.); PBS (2023); Stix (2022); Tao, Viberg, Baker, and Kizilcec (2024); UNESCO (2021); Waldrop (1987).

Puglisi, B. C. (2026b, January 28). *The adolescence of governance*. basilpuglisi.com.

<https://basilpuglisi.com/the-adolescence-of-governance/>

Sources in this work: Amodei (2026); Anthropic (2026); Anthropic, Responsible Scaling Policy.

Puglisi, B. C. (2026c, January 31). *The missing governor: Anthropic's constitution and essay acknowledge what they cannot provide*. basilpuglisi.com. <https://basilpuglisi.com/the-missing-governor-anthropics-constitution-and-essay-acknowledge-what-they-cannot-provide/>

Sources in this work: A-LIGN; Amodei (2026); Anthropic (2026); Atari et al. (2023); ATD Fourth World; BigID; Center for Health Care Strategies; Emergo by UL; European Union (2024); eyreACT; Fricker (2007); Greenlight Guru; Henrich, Heine, and Norenzayan (2010); IAPP; International Organization for Standardization (2023); Kellogg Insight; KPMG; McKinsey & Company (2025); Mind Australia (2024); National Council for Mental Wellbeing; National Institute of Standards and Technology; National Mental Health Consumer and Carer Forum; OECD; Phillips (2014); RAND Corporation (2024); Samuel (2026); The AI Praxis; Trott; U.S. Food and Drug Administration; WHO Europe (2025).

Puglisi, B. C. (2026d, February 5). *The great AI language collapse: Why marketing is killing accountability*. basilpuglisi.com. <https://basilpuglisi.com/the-great-ai-language-collapse-why-marketing-is-killing-accountability/>

Sources in this work: Anthropic (2023); Bai et al. (2022); CEN-CENELEC (2025); European Parliament and Council (2024); Fink (2025); IEEE; International Association of Privacy Professionals (2022); International Organization for Standardization (2021, 2022, 2023, 2024); Jobin, Ienca, and Vayena (2019); Maas (2025); National Institute of Standards and Technology (2023, 2024); OECD; Orozco y Villa and Menendez (2025); PMC (2021); Stix (2022); UNESCO.

Puglisi, B. C. (2026e, June 4). *Why agentic AI was always going to fail*. basilpuglisi.com.

<https://basilpuglisi.com/why-agentic-ai-was-always-going-to-fail/>

Sources in this work: Anthropic Institute (2026); Blum (2010); Carnegie Mellon University (2025); Cloudflare and Prince (2026); Colorado General Assembly (2024, 2026); Connecticut General Assembly (2026); Council of Europe (2024); EDPB and WP29 (2018); Economist and YouGov (2026); Electronic Privacy Information Center (1999); European Commission, Council, and Parliament (2026); European Union (2016, 2024); EY (2025); Financial Crisis Inquiry Commission (2011); Friis and Riley (2025); Gartner (2025); Google (2025); Grant Thornton (2026); Hao (2025); Harris (2025); Human Statement and Future of Life Institute (2026); IBM (2026); IEEE Spectrum (2026); ISO/IEC (2023); ITIF and Morning Consult (2026); Microsoft; MIT NANDA (2025); National Institute of Standards and Technology (2023); OECD (2019); OpenClaw (2026); PCAOB (2024); Pew Research Center (2023); RAND (2024); S&P Global Market Intelligence (2025); Saeri et al. (2026); Salesforce (2025); Sarbanes-Oxley Act; The Moltbook Illusion (2026); UK ICO; United Kingdom (2025); U.S. Food and Drug Administration; xAI v. Weiser (2026).

Reis, M., Reis, F., & Kunde, W. (2024). Influence of believed AI involvement on the perception of digital medical advice. *Nature Medicine*, 30(11), 3098-3100. <https://doi.org/10.1038/s41591-024-03180-7>

Royal Commission into the Robodebt Scheme. (2023). *Report* (Vol. 1). Commonwealth of Australia. https://www.monash.edu/_data/assets/pdf_file/0007/3365503/report_of-the-royal-commission-into-the-robodebt-scheme.pdf

Russo Law LLC. (2026, June 18). *New York's Part 161: What attorneys need to know about the new AI rule*. Mondaq. <https://mondaq.com/unitedstates/court-procedure/1801812/new-yorks-part-161-what-attorneys-need-to-know-about-the-new-ai-rule>

Sah, S., Loewenstein, G., & Cain, D. M. (2013). The burden of disclosure: Increased compliance with distrusted advice. *Journal of Personality and Social Psychology*, 104(2), 289-304. <https://doi.org/10.1037/a0030527>

Schilke, O., & Reimann, M. (2025). The transparency dilemma: How AI disclosure erodes trust. *Organizational Behavior and Human Decision Processes*, 188. <https://doi.org/10.1016/j.obhdp.2025.104405>

Simon, F. M., Nielsen, R. K., & Fletcher, R. (2025, October 7). *Generative AI and news report 2025: How people think about AI's role in journalism and society*. Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/generative-ai-and-news-report-2025-how-people-think-about-ais-role-journalism-and-society>

U.S. Securities and Exchange Commission. (2024, March 18). *SEC charges two investment advisers with making false and misleading statements about their use of artificial intelligence* [Press release 2024-36]. <https://www.sec.gov/newsroom/press-releases/2024-36>

Disclaimer

I am not a lawyer, and this article does not provide legal advice. This is thought research and governance analysis based on public sources, cited materials, and human-AI review. It is intended to help executives, practitioners, insurers, and governance teams think more clearly about AI risk, liability exposure, and documentation practices. Readers should not rely on this article as a legal opinion, compliance determination, or substitute for qualified counsel. Any organization facing a legal, regulatory, contractual, or insurance question should consult its own attorney, broker, or professional adviser before acting.

The Other AI: Audio Briefings on Augmented Intelligence and AI Governance
[Spotify](#) | [Apple Podcasts](#) | [Amazon Music](#) | [YouTube Playlist](#)

Basil C. Puglisi, MPA

A Human-AI Collaboration

#AIassisted using the HAIA Ecosystem | CC BY-NC-SA 4.0

Free for personal, educational, and noncommercial research use with attribution. Commercial exploitation, paid productization, and enterprise commercialization require separate permission and licensing.