

Theme 6: AI Governance

Stories of AI implementations having all the restraint of the wild west aside, there are substantive efforts at governing AI (or more accurately, governing the *use* of AI) at multiple scopes, from the single institutional level to the international. In this theme, we will survey some attempts at AI governance as well as some metastudies which attempt to explore one aspect of AI governance over a broad area.

With deference to Stan Lee, “*with great power there must also come — great responsibility*”¹. The City of Albuquerque, New Mexico’s AI Working Group developed an AI policy document (2025) that seeks to “improve lives, enhance public services, preserve public trust, and is rooted in our shared values of access, advancement, and accountability” while positioning Albuquerque to be a “national leader in the responsible and sustainable use of AI (2025)”

- ABQ AI WG. (2025, August 29). *Artificial Intelligence Policy 2025*. Albuquerque Technology & Innovation Dept. <https://cabq.legistar.com/gateway.aspx?M=F&ID=34f61027-6d9c-48a0-a56a-60340f3f359e.pdf>

A quandary that each publicly traded corporation must wrestle with is “whose benefit supersedes whose?” In a 2004 interview², cofounder and then-chairman of Southwest Airlines Herb Kelleher maintained that placing employees first led to satisfied customers which ultimately drove shareholder value. Likewise, a 2023 paper by Tay et al. echoed and expanded this claim³ by pushing back against performance-based valuation in favor of what they called “optimal functioning” in which worker, organization, customer, and societal/environmental well-being are all equally regarded. We start our examination of corporate AI governance with a paper by Cordeiro et al. which similarly advances the claim that organizations that integrate AI into their operations will find themselves poorly served by the so-called “shareholder primacy” which has been at the center of corporate operations from the 1970s through the early 2000s. The use of AI, they say, will move organizations away from “shareholder primacy” to a “stakeholder-centric” model (2026).

¹ page 14, panel 8 of Lee, S., & Ditko, S. (1962, September). [Amazing Fantasy \(Comic No. 15; Vol. 1\)](#). Marvel Comics. (Original work published Atlas Magazines)

² Lucier, C. (2004, June 1). *Herb Kelleher: The Thought Leader Interview* [Business Strategy Online Magazine]. Price Waterhouse Coopers Limited. Strategy+business. <https://www.strategy-business.com/article/04212>

³ Tay, L., Batz-Barbarich, C., Yang, L.-Q., & Wiese, C. W. (2023). Well-being: The Ultimate Criterion for Organizational Sciences. *Journal of Business & Psychology*, 38(6), 1141–1157. Business Source Complete (173517292). <https://doi.org/10.1007/s10869-023-09908-5>

A common response to concerns about corporate AI governance is for tech companies to publish a variety of documents inventorying the ways in which they are responsible, transparent, effective, reliable, and inclusive: and how customers who use their AI products can also manifest those traits. A critical reader might ask themselves: “Are these claims proactive or reactive?” Microsoft’s Responsible AI Principles and Approach site (2026) is a good example of such content. Should we conclude anything from the fact that the transparency report has been updated in both 2024 and 2025, while the Responsible AI Standard hasn’t been updated since 2022? This site is a gateway to a variety of materials ranging from those intended to look like industry standards, to quasi-marketing case studies, and training. What does the choice of topics, and the choice of topics omitted say about Microsoft’s goals for this site, and their view of their AI products’ strengths and weaknesses?

- Cordeiro, C. M., Adomaitis, L., & Huang, L. (2026). The AI-policy-governance nexus: How regulation and AI shift corporate governance toward stakeholders. *Technology in Society*, 84, 103117. <https://doi.org/10.1016/j.techsoc.2025.103117>
- Microsoft. (2026). *Responsible AI Principles and Approach*. Microsoft. Microsoft AI. <https://www.microsoft.com/en-us/ai/principles-and-approach>

Of course, we see AI being implemented by more types of institutions than just companies. Jobin et al. conducted a scoping review (2019) using 84 instances of AI ethics documents in grey literature from private companies, government agencies, academic and research organizations, NGOs, intergovernmental/supernational organizations, non-profits, professional and scientific societies, and so forth. Common themes emerged from nearly all of the documents: transparency, justice and fairness, non-maleficence, responsibility, and privacy. On the surface, this would seem to indicate commonality, but the operational details of these themes varied wildly between documents. The authors, at least in 2019, found their search methods yielded documents in the global north more easily than elsewhere. They also found that documents tended to focus on harm prevention rather than promotion of the benefits of AI, almost as if the goal of the entire ethics exercise is to manage exposure rather than advance more effective ways of working.

- Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>

As we’ve seen a few times already, AI could be regarded as a confounding variable. As the Financial Services Sector Coordinating Council observes in “Published Documents: Financial Sector Artificial Intelligence Executive Oversight Group Deliverables (2026)”, AI tends to be an accelerant. For some actors, it provides an enhanced ability to compromise financial systems and engage in a variety of cyberattacks. For others, it provides a set of tools for enhanced transparency and development of innovative products. At nine intervals of various lengths since November 2023, the FSSCC has issued a set of reports regarding

identified gaps in the financial sector's use of and preparedness for AI. The cited reference is this year's master document which describes and links to each of the subordinate gap documents. This year, they address a lexicon; a risk management framework; identity and authentication; explainability; data quality and alignment with state, federal, and international governance; and enhanced fraud.

- FSSCC. (2026). Published Documents: Financial Sector Artificial Intelligence Executive Oversight Group Deliverables [Financial Industry Association]. *Financial Services Sector Coordinating Council*. <https://fsscc.org/aieog-ai-deliverables/>

The next two resources examine artificial intelligence from an international perspective.

Like the FSCC resource above, the UN's Artificial Intelligence Subject Briefing ([2025] 2026) is a launchpad for several in-depth documents relating to AI. Those twelve resources are "Resources" heading, about 73% of the way down the main page. The author(s) of this briefing see AI as helping the UN fulfill up to eighty percent of its sustainable development goals as well as strengthening the work of the entire UN system. These possibilities, however, come at the cost of "coordinated global governance", which includes pushing back against the "shadows" of AI: disinformation, human rights violations, voter and public opinion manipulation, and the undermining of trusted institutions. Aspirationally, the UN calls for all governments and stakeholders to work together to develop a common governance framework that supersedes any existing criteria and resolves existing gaps in AI governance.

- UN. (2026). *Artificial Intelligence Subject Briefing* [Subject briefing]. United Nations. United Nations. <https://www.un.org/en/global-issues/artificial-intelligence> (Original work published 2025)

The next resource with International scope is The National Judicial College's Massive Open Online Course (MOOC), "Artificial Intelligence and the Rule of Law. ([2015] 2026)" Developed with UNESCO, The Future Society, and the IEEE Standards Association, this six-module course's topics include adoption of AI across justice systems, online courts, algorithmic bias and safeguarding human rights, and AI ethics and governance relevant to judicial operators.

- NJC. (2026). *Artificial Intelligence and the Rule of Law* [UNESCO AI Course]. The National Judicial College. https://judges.org/ai_and_law/english/ (Original work published 2015)

The next three resources have a national focus: first on South Africa, then on the US, and finally on China.

A common theme we see in AI governance documents is the acknowledgement that using it is a risk that carries with it an opportunity to overcome historic technical and social debt

and act as a creative and labor force multiplier. In the “South Africa National Artificial Intelligence Policy Framework (2024)”, The Republic of South Africa’s Department of Communications & Digital Technologies sets the table for discussions that would result in AI regulations or perhaps an omnibus AI act. The framework places the plausible future for AI as one that balances the pull of the future, the push of the present, and the weight of the past.

- DCDT. (2024, August). *South Africa National Artificial Intelligence Policy Framework*. Department of Communications & Digital Technologies: Republic of South Africa. <https://techcentral.co.za/wp-content/uploads/2024/08/South-Africa-National-AI-Policy-Framework.pdf>

“The AI Bill of Rights”, originally named “Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People” was published in 2022 by the US White House Office of Science and Technology Policy as a collaborative effort by scholars, human rights groups, corporations and non-profits as a kind of advisory firewall between potential challenges presented by AI and democratic values. This document from IBM, “What is the AI Bill of Rights? (Krantz and Jonker [2022] 2026)” gives an overview of the tenets of the original document, the impact it has had on US agencies, and some overlapping US policies. The original framework focused on five core principles: methods to protect end-users from unsafe or ineffective AI; transmission of bias related to legally protected characteristics in AI systems, or algorithmic bias; data protection; transparency; and the ability to opt-out of AI entirely and engage with a human-centric alternative workflow.

- Krantz, T., & Jonker, A. (2026, May 24). *What is the AI Bill of Rights?* IBM. <https://www.ibm.com/think/topics/ai-bill-of-rights> (Original work published 2022)

No discussion of global AI governance would be complete without discussing China. In her analysis of Chinese AI governance, “Global AI Governance Law and Policy: China (2025)”, Barbara Li walks us through China’s plan to integrate AI into six economic sectors: “science and technology development, industrial utilization, consumer services, public welfare, governance and security, and international collaborations.” The plan is to achieve 70% AI penetration in those sectors by 2027, 90% by 2030, and arrive at a “fully AI-powered economy” by 2035.

- Li, B. (2025, November 12). *Global AI Governance Law and Policy: China*. International Association of Privacy Professionals (IAPP). <https://iapp.org/resources/article/global-ai-governance-china>

In their regional recommendations for AI and Africa, The African Observatory on Responsible Artificial Intelligence is another institution that is mindful of the bifurcated nature of AI, especially with regard to the need to both further deshackle Africa from the remnants of colonialism as well as avoid entering into technological colonization. In

“Responsible AI Governance in Africa: Prospects for Outcomes-Based Regulation (2024)”, authors Fola Adeleke and Funbi Akinwale strategize about ways to govern around the stochastic growth of AI in such a way that the governance itself doesn’t have to be continually updated. Instead of defining narrow-focus, institutional rules, they advocate “outcome-based, risk-weighted regulations.” Instead of scoping the regulations at the corporate or any institutional level, or even at a national level, the authors advocate scoping regulations and strategies at the regional/continental level, and around individual sectors. They are mindful of the fact that the African continent represents a wide divergence of levels of development and preparedness for AI governance. As of 2024, only ten of the fifty-some nations in Africa had national AI governance strategies.

To support governance as an adaptive learning-based process, the authors advocate for regulatory sandboxes in which industry and government can work together for each state and sector’s best solution.

- Adeleke, F., & Akinwale, F. (2024, March). *Responsible AI Governance in Africa: Prospects for Outcomes-Based Regulation*. African Observatory on Responsible Artificial Intelligence.

We finish this theme with Haie et al.’s governance comparison, “A Comparative Analysis of the EU, US and UK Approaches to AI Regulation (2024)”. A primary takeaway from this analysis is that while different entities may have similarly worded top-level goals for AI governance, going even one level down into operational details reveals substantial differences in perspective. Some areas, like a baseline definition of AI, are similar for the EU and the US, but aren’t addressed at all in UK documents reviewed by the authors. The US and UK adapt governance to different economic sectors, and the EU applies the same regulations horizontally across the board without regard for sector.

The three jurisdictions are compared on the basis of whether they’re horizontal or sectoral; how they define AI; amount of focus on risk, and on which AI model; how testing and monitoring take place; whether there’s a regulatory carve-out for open source AI; and how intervention and sectoral regulation should happen.

- Haie, A.-G., Cohen, T., Golodny, A., Avramidou, M., Goodwin, E., & Arethusa, V. (2024, April 30). *A Comparative Analysis of the EU, US and UK Approaches to AI Regulation*. STEPTECHTOE.
<https://www.steptoe.com/a/web/mdWt44Hg7mmGXDAcLsAr7n/comparative-table-on-approaches-to-ai-regulation-in-the-eu-us-and-uk-2.pdf>