

Civil Society and Agility in AI Governance

Ganchimeg Namsrai

Johannes M. Bauer

Ruth Shillair

Department of Media and Information and Quello Center

Michigan State University

DRAFT FOR PURPOSES OF DISCUSSION

August 1, 2026

Prepared for presentation at the
54th Research Conference on Communication, Information, and Internet Policy
Washington, D.C. September 25-26, 2026

Abstract

Agile and adaptive methods of regulation and civil society participation are increasingly identified as institutional mechanisms that may assist in addressing the governance challenges associated with artificial intelligence (AI). These considerations are frequently addressed at a theoretical level and often examined independently. This paper integrates these perspectives to explore how agility, adaptivity, and civil society participation are beginning to be organized and how they may contribute to AI governance.

Keywords

Artificial Intelligence (AI), Civil Society, Governance, Agility, Adaptivity

JEL Codes

L31, L86, L88, L96, L98

INTRODUCTION^{1, 2}

Artificial intelligence (AI) governance can be defined as the set of technical, institutional, and social arrangements intended to align its outcomes with broader societal goals. This implies measures that improve the working of the components of the AI system, from physical infrastructure to applications and uses, measures to develop desirable uses and applications (positive alignment), and measures to prevent undesirable outcomes (negative alignment) (see, e.g., Christian, 2020; Laukkonen et al., 2026; Reuel & Undheim, 2024). Because there are numerous locations across the vast, interrelated AI value system that could serve as possible leverage points to achieve envisioned outcomes, and an equally differentiated set of governance instruments, operationalizing the most effective way of achieving desired governance outcomes is challenging.

Traditional concepts of externalities, public goods, and competitive choke points provide guidance to identify some cases in which governance can improve overall outcomes. However, they do not work well under conditions of rapid technological change and uncertainty (see, e.g., Bauer & Stocker, 2026). Risk-based approaches to governance fail at that stage of technological development, because risks and opportunities are not yet sufficiently known. Reliance on the precautionary principle will often result in overly stringent regulations. In contrast, a fully unregulated system may develop in irreversible, undesirable directions. This dilemma, that both too little and too much governance may result in lower rates or unwanted directions of innovation, is a key insight from recent innovation research. Agile governance, a range of institutional arrangements and instruments that can adapt more flexibly and faster than traditional forms of government policy, are increasingly embraced as a possible response to these challenges.

Agility emphasizes time to implement a measure, whether it is a law, a solution to a problem, or the enforcement of a rule or principle. Adaptivity is a related concept that

¹ The authors have no competing interests to declare. Research reported in this paper was conducted as part of our employment at Michigan State University.

² The authors thank Sandra Braman and Michael Stern for valuable comments on earlier iterations of this research project.

emphasizes the ability to respond to changing circumstances and new, possibly unexpected developments (Janssen, 2025; Janssen & van der Voort, 2016). It requires a different set of capabilities, such as the ability to interpret and make sense of shifting observations. Moreover, it requires mechanisms that allow changes in routines and institutional arrangements, and provisions that strengthen resilience (i.e., the ability to continuously adapt and to bounce back after failures) (Reuel & Undheim, 2024). Civil society organizations (CSOs) contribute to these arrangements.

Civil society can contribute in numerous ways to the development of a dynamic learning system that adapts as experience becomes available. Civil society organizations have been successful in organizing resistance to unwanted developments, contributing to court, regulatory and legal discussions that led to the establishment of guardrails, and to find dynamic responses to problems. They are also contributing to formulating positive visions for AI. Often, these contributions evolve as a process of local experimentation that advances by trial-and-error. Proposals for adaptive governance include “soft law,” regulatory sandboxes, outcome-based regulation, or data-driven governance, but the empirical evidence of their effectiveness is limited. All these approaches are informed by evidence and rely on informational feedback loops that enable learning about intended and unintended unwanted consequences, implicated values, and on the workability of governance interventions.

We explore proposed agile and adaptive governance as well as the contributions of civil society based on a narrative review of research literature and policy documents. Civil society organizations, in our paper is broadly understood as “an institutional realm of private associations, voluntarism, and active citizens” (Torfing, 2020) that exists independently of government and private corporate can make important contributions to such agile governance methods. Among other contributions, civil society organizations bring knowledge and values to governance discussions that traditional government and corporate actors may not consider. Civil society actors may represent groups that are harmed by AI; they may advocate for vulnerable and marginalized groups; may contribute to enhanced accountability; and ensure that AI development remains aligned with human rights and

other social values. Their involvement may also mitigate concerns about the capture of agile regulatory arrangements by corporate and government players.

The contributions of the paper include a critical discussion of the challenges, limitations, and potential of agile and adaptive AI governance solutions and an initial examination of civil society's contribution to AI governance. We explore these issues in seven steps. The following section discusses our motivating research questions and details of the methodologies that we used to explore them. Section three recaps analytically the unique conditions of AI governance and sketches how civil society can contribute to overcome them. Section four discusses how the concept of agility has been translated into specific institutional designs showing a long arc of precedents to the current discussion. Section five reports on our empirical research on how civil society participation is organized. It also discusses specific cases of civil society activity that illustrates the arguments made in sections three and four. Section six discusses the findings and section seven offers concluding reflections and an outlook.

2. RESEARCH QUESTIONS AND METHODOLOGY

There are widespread calls for AI governance and emerging methods often embrace agility and adaptive methods (Batool et al., 2025). At the same time civil society has the potential to offer valuable insights to this process to protect human flourishing. In this paper we explore (1) how agility and civil society participation are beginning to be organized and (2) how they may contribute to AI governance. This paper examines the potential role and intersection of agility and civil society participation in AI governance by addressing the following research questions: How are agile governance arrangements and civil society participation being organized and institutionalized in AI governance (RQ1)? How could agility and civil society participation contribute to AI governance (RQ2)?

2.1 Focus Areas of the Study

As the growth and diffusion of AI is a topic that impacts all levels of society, we start by defining the key concepts used in the study. We focus on the potential contributions of civil

society to improving the speed of governance responses (agility) and its ability to adjust to changing circumstances (adaptivity). Done well and effectively, both dimensions could help overcome the current challenges of AI governance, such as the pacing problem and the need to respond to unexpected developments.

AI governance: AI governance to refer to the vast array of social, economic, legal, and technical mechanisms that are in play to coordinate, constrain and enable the activities of players active in the AI ecosystem(s).

Agility: A range of institutional arrangements and instruments that can adapt more flexibly and faster than traditional forms of government policy. Agile systems often focus on quickly changing to meet stakeholder needs and changing the scope and conditions for governance. Agile governance focuses on seeing needs and responding quickly (Janssen & van der Voort, 2020).

Adaptivity: The ability to respond to changing circumstances and new, possibly unexpected developments. It requires a different set of capabilities, such as the ability to interpret and make sense of observations. Moreover, it requires mechanisms that allow changes in routines and institutional arrangements, and provisions that strengthen resilience (i.e., the ability to continuously adapt and to bounce back after failures) (Reuel & Undheim, 2024).

Civil society: Civil society organizations, in our paper broadly understood as “an institutional realm of private associations, voluntarism, and active citizens” (Torfing, 2020) that exists independently of government and private corporate can make important contributions to such agile governance methods.

2.2 Research Methods

To examine how civil society participation in AI governance is conceptualized, institutionalized, and implemented in practice, this study employs three complementary research methods. These include a structured review of academic and policy literature, an analysis of U.S. state legislation related to AI governance bodies, and a narrative analysis based on selected case vignettes of civil society organizations.

2.2.1 Structured Review of the Research Literature

Our discussion of the research literature builds on the work by Namsrai (2026). The literature selection process involved systematic identification, screening, and eligibility assessment to ensure transparency and analytical rigor, inspired by principles developed in the PRISMA framework (Page et al., 2021). Relevant academic publications were identified through major scholarly databases: Web of Science, Scopus, HeinOnline, Google Scholar as well as complementary web searches. Searches were conducted using multiple variations of terms such as “civil society roles in AI governance,” “civil society and AI,” and “multistakeholder governance.” In addition, research on the governance of dynamic adaptive systems was reviewed.

For the academic publications, after removing duplicate records, abstracts were screened to assess relevance, and sources that did not address AI governance or did not involve civil society participation were excluded. The full texts of the remaining records were then reviewed to determine whether they were suitable for inclusion, based on their relevance to the research topic and their ability to inform the research questions. Given the rapidly expanding body of literature on AI, the analysis covers only documents published between 2022 and February 2026 to maintain analytical focus and coherence in line with the scope of the study. However, key foundational theoretical works published prior to this period are included where relevant.

In addition to academic studies, the review includes policy documents and reports from a selected international and civil society organizations actively engaged in AI governance, as well as aspects of the U.S. policy framework. These sources were identified through targeted web searches and institutional repositories. In total, approximately 100 academic articles and 30 policy and regulatory documents were initially identified, of which about 60 articles and 12 documents were selected for in-depth analysis based on relevance and analytical fit.

2.2.2 Analysis of U.S. State Legislation

U.S. federal AI policy currently relies predominantly on Executive Orders by the President and voluntary agency frameworks, such as the risk management framework developed by the National Institutes of Standards and Technology (NIST). Neither of these documents has specific provisions guiding civil society participation. We therefore examined state level activities using the National Conference of State Legislatures (NCSL) Artificial Intelligence Legislation Database as a primary source of the analysis. The database compiles AI-related bills introduced across the 50 U.S. states, the District of Columbia, and the U.S. territories. The analysis covered bills included in the database from 2019 through the latest update on June 1, 2026.

In the first stage, approximately 200 bills proposing the establishment of new AI governance bodies or amendments to the composition or responsibilities of existing governance bodies were identified by reviewing the bill summaries provided in the NCSL database. The analysis included proposals establishing or modifying task forces, advisory groups, commissions, committees, boards, working groups, partnerships, councils, divisions, and similar governance structures.

In the second stage, the 39 bills that had been enacted or adopted were identified. The full text of each enacted bill was then reviewed to determine whether the governance body had a multistakeholder composition. This process identified 18 multistakeholder governance bodies, of which eight explicitly included civil society representatives in their institutional structure. These eight governance bodies constituted the empirical cases for further analysis.

Civil society participation was coded only when legislation explicitly designated civil society organizations as members of the governance body, reserved seats for them, or otherwise institutionalized their formal participation. General provisions requiring public consultation or opportunities for public input were not considered institutionalized civil society participation.

For each of the eight governance bodies, the analysis examined: (1) the AI governance issues or policy areas within its mandate; (2) its purpose and primary functions; (3) its duration or term of operation; (4) the total number of members and the number representing civil society; and (5) the types of civil society organizations represented. Based on these characteristics, the study analyzed the institutional roles assigned to civil society within each governance body.

To identify the civil society organizations represented, the analysis relied not only on the legislative texts but also on publicly available materials, including official websites, membership lists, reports, and other publicly accessible documents associated with each governance body.

2.2.3 Complementary Narrative Analysis

To complement the conceptual analysis and review of legislative proposals, we employed a narrative approach based on case summaries (vignettes) to examine selected CSOs engaged in AI governance. An initial set of organizations was identified through an open search using keywords such as “civil society organizations working on artificial intelligence.” From the search results, a list of 31 CSOs was compiled, and reviewed using publicly available information, including official websites and social media platforms. This set was selected to provide a broad overview of civil society engagement in AI governance and to help contextualize the field. To identify dominant patterns, information on each organization’s AI-related engagement, core mandate, target issues, primary activities, and sources of funding was collected from their official websites and analyzed using a qualitative, descriptive approach with simple frequency-based comparisons across cases.

From this initial set, five CSOs were selected for further analysis and presented as case vignettes. The selection was guided by the criterion of active involvement in AI governance. In this study, “active involvement” is defined as sustained and visible engagement in AI-related activities, including activities such as policy advocacy, research, public engagement, and capacity building. The analysis follows a qualitative, thematic approach to identify recurring patterns in how CSOs contribute to AI governance. Rather

than aiming for statistical generalization, this narrative approach enables a deeper understanding of the roles of civil society actors.

To identify examples of policy and institutional change resulting from the contributions of civil society organizations to AI governance, a targeted search was conducted using publicly available online sources. The search employed keywords such as "civil society success story" and "civil society advocacy." The identified cases were then verified against official information published by the respective organizations and other relevant sources before being included in the analysis.

3. CONCEPTUALIZING AI GOVERNANCE

Numerous research papers, policy analyses, and position papers describe principles and frameworks for the governance of artificial intelligence (see, e.g., Bullock et al., 2024; Dubber et al., 2020; Hagendorff, 2020; Schiff et al., 2020; Schmitt, 2022). In addition, there is growing interest in understanding the repercussions of the socio-technical characteristics of AI for the possibility of effective governance (see, e.g., Potts, 2025). Here we aim at a rigorous theoretical articulation of the role of AI governance from an analytical perspective, a characterization of the conditions under which AI governance must succeed, and the challenges that it must overcome. We use the term AI governance to refer to the vast array of social, economic, legal, and technical mechanisms that are in play to coordinate, constrain and enable the activities of players active in the AI ecosystem(s).

Typically, the intended or tacit goal of these forms of coordination is to increase the likelihood of desirable outcomes and reduce the likelihood of undesirable ones from applications and uses of AI. However, this is not always the case, as individual players or groups may have incentives to bias governance arrangements in favor of their own, strategic goals. How well the overall governance arrangement can mediate conflicts and tradeoffs has direct consequences for the overall outcomes. Agile and adaptive forms of governance are subsets of these arrangements.

Several recent studies explain that the complexity, decentralization, and fast-paced development of AI systems greatly increase uncertainty and limit the effectiveness of conventional, centralized governance. Kolt et al. (2025) argue that AI governance must be adaptive, evidence-informed, feedback-driven, and institutionally flexible because risks emerge across interconnected actors and sociotechnical contexts. Potts (2025) similarly conceptualizes AI as a distributed and partially opaque “hyperobject” that cannot be effectively governed through traditional top-down regulation. He calls instead for decentralized, community-driven forms of stewardship and governance embedded in algorithms. Sokhansanj (2025) highlights how the growth of locally deployed and open-source AI further weakens centralized oversight and therefore requires multilayered, polycentric, and participatory governance mechanisms. Schrepel (2025) adds that attempts to design “future-proof” regulation are unrealistic. Rather, regulatory review and adaptation must evolve. Together, these studies suggest that AI governance should be understood as an ongoing, distributed, and adaptive process rather than a fixed regulatory framework.

3.1 AI Governance as a Hard Institutional Design Problem

Governance can be viewed as a social coordination technology, recipes that are in place to resolve specific problems and achieve desirable outcomes (see, e.g., Baldwin, 2024; Enroth, 2014; Fairfield, 2021). This perspective sharpens the view that effective governance requires that a specific measure or a set of measures be designed so that they have predictable effects on outcomes. Research in other areas shows that this requires meeting several constraints created by the socio-technical system to be governed (see, e.g., Bauer, 2022). Sustainable governance must (1) be compatible with the socio-technical working of the system; (2) instruments capable of shaping system outcomes in desirable ways must exist, and (3) they must be politically feasible. These constraints are extraordinarily difficult to meet simultaneously. It is unlikely that solutions can be constructed in one sweep, such as by major legislation or anticipatory sector regulation. It is more realistic to expect that they will evolve over time in a piecemeal process of societal learning.

AI creates challenging conditions in all three areas. The AI value system requires services from providers on several interdependent layers: physical infrastructure (e.g., networks, energy, water), hardware and software resources (e.g., compute, data centers), cloud services (evolving earlier concepts such as software-as-a-service—SaaS), data and data sources, machine learning capabilities (e.g., foundation models, specialized models), and applications and uses. Some players operate in or more of these layers, but others, such as hyperscalers, are vertically integrated across all of them. Players on these layers relate in a complex web of competitive, co-opetitive, complementary, and cooperative business relations. This system develops in a fast-paced co-evolution driven by players in different locations, with innovation and new capabilities in one area affecting innovation opportunities in others. Key developments happen under conditions of deep uncertainty, which implies that no causal model of the potential positive and negative consequences is known.

As Clark and Claffy (2015) have shown for the case of the Internet, fast-paced co-evolution and diversity of players make it challenging to find a locus where governance could intervene in a way that achieves desired outcomes. Deep uncertainty and malleability of AI add further complications. The theory and practice of governance has identified three broad groups of rationales for policy interventions into a decentralized socio-technical system: a need to provide a legal and institutional framework for players (a constitution or “order” of markets, such as the protection of privacy), the existence of flaws in the ability of the system to coordinate well (market failures such as market power, externalities, public goods), and measures to correct undesirable outcomes (e.g., unacceptable biases, job losses, and inequalities). Designing such policies well requires a workable model of how the system works, and a good understanding of how any governance measure affects outcomes, both of which are often lacking in AI.

The third condition is that policymakers find sufficient agreement on the norms, values, and instruments that should be deployed. Conditions of incomplete knowledge permit the co-existence of widely different policy visions because information that would help select better models is lacking or is not yet available. Together with differing political

norms and values about justice and equity, this complicates finding sufficient agreement on how a system should be governed. Even where such agreement can be found, such as in the European approach to AI, doubts about the effectiveness remain that cannot be resolved by empirical evidence. This inability to find and implement feasible, sustainable governance will become visible in many forms. One possible outcome is incremental measures that result in continuous muddling through with or without achieving desired outcomes (Lindblom, 1959). The emergence of cycles of interventionist, libertarian, and other policy approaches that do not converge on a stable model is a second possibility. A third outcome is the absence of coherent policy, perhaps combined with ad hoc interventions.

3.2 AI Governance as Real-World Experiments

All this suggests that AI governance will necessarily develop as a process of experimentation and adjustment. Places with good governance will likely differ from those with poor governance by how well insights are generated from the emerging experience and how these insights are translated into adaptations and improvements of the governance approach. This will likely require reliance and orchestration of a wide range of governance options, from measures addressed to individuals to agreements at a global level. Like the layers in the socio-technical infrastructure, the institutional system in which governance must be anchored is nested and layered. At the lowest level are norms and incentives that guide decisions by individuals, such as AI developers, application developers, and users. Algorithms constitute a second layer, and organizational governance as a third layer. Voluntary sectoral rules (e.g., standards, codes of conduct); local, state, and national government policies; and international and global agreements complement these layers.

Governance measures at these different layers evolve at different time scales, from continuous adjustments to very long periods that can span decades, centuries, or even millennia as in the case of some cultural norms (see, e.g., Janssen & van der Voort, 2016; Williamson, 2000). In part, this is by design, in part by default. For example, constitutional law is designed as a framework envisioned to be stable over long periods of time, whereas statutes, regulations, and administrative orders are faster paced. Individual norms and organizational governance may evolve even more quickly. Rarely is there a formal,

deliberate process of meta-governance in place to assign tasks to these levels. Some places have adopted principles that guide aspects of this process. For example, the U.S. constitution assigns authority to the different levels of government but is not built on an overarching principle to guide the decisions. Other places, such as the European Union, are built on a principle of subsidiarity that implies assigning tasks to the lowest level capable of addressing an issue effectively.

Even though these constellations are not deliberately constructed but evolve, at any point in time, a place can be characterized by a constellation of measures across this institutional stack (Khan et al., 2026). If the measures are deliberately designed as part of a system, we will refer to it as *a governance system*. If this is not the case and they are the outcome of uncoordinated developments, we will refer to it as *a governance arrangement*. Findings in the theory of dynamic systems and in the theory of complex adaptive systems suggest that multiple, relatively stable constellations exist. This would explain why different countries or regions develop different approaches that coexist over long periods without convergence to one system of best practices. However, one would expect that these governance arrangements or systems are associated with characteristic patterns of outcomes that correspond to them.

4. AGILITY AND ADAPTIVITY AS INSTITUTIONAL RESPONSES

Demands for more agile public policy and the need for more agile business strategies recurred throughout the twentieth century. However, recent calls for agility and adaptivity are phrased differently, and must be seen as a response to new forces of transformation. Whereas historically the aim of agility and adaptivity was to develop appropriate *policies* faster, the governance debate has shifted some of the emphasis to the solution of *problems* (Enroth, 2014). One reason is that the pacing problem has become more pressing as technological change is accelerating in many areas while workable institutional responses are more difficult to develop. However, three other factors are in play: the changing cross-sectoral pattern of technological change in which biotechnology, nanotechnology, cognitive

technologies, and information technology are increasingly intertwined and convergent; the global nature of change; and challenges to enforce compliance with established rules (Bauer, 2022, p. 32). Agility, adaptivity, and civil society involvement are viewed as promising solutions to help address the pacing problem, facilitate the exploration of innovation opportunities, and improve alignment with important societal values to strengthen compliance, accountability, and transparency.

4.1 Responses to the Pacing Problem

Historical demands for more agile and adaptive public policy recurred throughout the twentieth century. They were often a response to real and perceived mismatches between the pace at which technology developed, market players acted, and the slower responses of formal government policy. These concerns contributed to more flexible government policies in information and communication industries (e.g., regulatory reform, deregulation), the emergence of new institutional forms of governance (e.g., stronger reliance on non-governmental organizations in Internet governance), and new business practices. They inspired efforts in management to adopt more agile competitive strategies since the 1980s (Nagel, 1992), and are recently considered in public administration, often associated with a shift from rigid, top-down bureaucratic control to bottom-up, iterative and citizen-centric approaches (Mergel et al., 2018).

These proposals are strongly inspired by the adoption of the agile principles in software development in the 1990s (Beck et al., 2001). In that context, changing parameters for a project and the need to respond to the concerns of the immediate stakeholders undermined the historical top-down waterfall approach. The core principles of agile project governance include iterative processes that incrementally reach management goals. This includes responsiveness to stakeholders and adopting parameters of a project to fit dynamic situations. The concept of agile governance has gone beyond projects and software development to governance at all levels as it allows ways to give governing bodies a chance to listen to stakeholders and be responsive to changing needs. However, agile systems have problems scaling if there are multiple stakeholders, especially if they have conflicting goals (Nuottila et al., 2022).

The desire to increase agility and adaptivity has contributed to the migration of governance tasks to non-government players and businesses. There, agile and adaptive forms of governance fulfil several interrelated and sometimes conflicting roles. They facilitate coordination between firms that provide complementary services, interaction with customers, and provide reprieve from other forms of government regulation. For example, platforms that develop and support AI models have vested interests in at least a bare minimum of governance, especially guidelines that support public perception of high quality and reliability. These industry standards usually are not legislated but seen as governance baselines and are accepted as default standards.

NIST standards usually go through a period of open comments and review processes by domain experts (e.g., NIST Cybersecurity Framework, 2024) and seen as industry standard, thus widely adopted (IBM, 2021). From NIST technological guidance, concepts such as interoperability and portability allow innovation and growth across platforms. The NIST AI guidelines include items that are seen as “best practices” and therefore often become industry standards (NIST, 2026). Thus, when we look at the most basic level of technological governance, there are established methods and protocols that have proven flexible and yet robust for decades. Platforms themselves, in addition to industry standards often promote their additional self-directed levels of governance to address common stakeholder concerns (e.g., Anthropic, 2026; Google, 2026).

4.2 Exploring New Innovation Opportunities

Some states are using AI governance in controlled “sandbox” governance models. It allows AI to be used for high-risk, but high-reward tasks that are repetitive. This allows for “adaptive governance through structured experimentation” (Gasser et al., 2025). For example, Utah (Utah SB 149 (2024), the AI Policy Act)³ makes it legal to use AI for routine refills for chronic conditions. The backbone of this policy is an agreement with the AI Learning Laboratory Program to cover renewals of prescriptions that legally needed a licensed medical practitioner to renew (Metnick et al., 2026). Even though at face value this

³ <https://le.utah.gov/~2024/bills/static/SB0149.html>

allows those who can't easily get to a doctor (e.g., elderly, rural, lower income) the ability to better control chronic conditions through access to necessary medications. From the medical professional's point of view, this loss of yearly contact point reduces the important interaction of a medical expert checking if the prescription is still appropriate given possible changes in the patient's needs.

For policymakers, perhaps more concerning is that something that was previously regulated by the FDA and medical experts is categorically switched to the control of the Department of Commerce in the state of Utah (see Utah SB 149 (2024), the AI Policy Act)⁴. In this sandbox example, AI is being used in a way for "efficiency" and providing sandbox isolation of the application, but the commerce department is doing the monitoring, not the medical board, nor the FDA who have more expertise. This is an example where although AI is being used in a "governed" way, the primary stakeholders were set aside, and the test application is moved forward.

The sandbox model offers a way for market and non-market players to explore innovation opportunities combined with a "prototyping" of governance. It allows examines strengths and weaknesses of the governance model before fully implementing a policy model on a larger scale (Gasser et al., 2025). Doing this well requires a rigorous framework for monitoring and evaluation, appropriate safeguards, and robust principles for translating the experience into revised governance. Many of the existing sandbox models do not meet these criteria. Often, monitoring and evaluation are insufficient so that no lasting learning effects are generated. In other cases, guardrails established around sandboxes have proven ineffective, as the "escape" of rogue Open AI and Anthropic agents during recent tests shows. Unbeknownst to the engineers running the tests, the agents apparently went rogue and hacked into other systems.⁵ Furthermore, at this early stage of experimentation with

⁴ Ibid.

⁵ See Open AI find evidence that other agents escaped as it widens hacking probe, Reuters, July 31, 2026, <https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.reuters.com/business/openai-finds-evidence-other-ai-agents-escaped-containment-it-widens-hacking-2026-07-31/>; Anthropic's Claude AI escapes to hack into three organisations, BBC, July 31, 2026, <https://www.bbc.com/news/articles/cz7dl7w8y7po>.

sandboxes, robust principles for translating the experience into adaptations of governance have not yet been developed.

4.3 Improving Alignment

Going beyond the window of technological design typically adds additional complexities that can only be sketched briefly here. As previously discussed, the combination of the fast pace of innovation and the decentralized socio-technical systems complicates finding a policy approach that the various stakeholders find agreeable. Typically, there will be real or perceived trade-offs between goals so that stakeholders with diverging priorities will advocate for different policy responses. Alternative responses affect players differently, creating varying patterns of costs and benefits. Both may complicate finding workable common ground. For example, security and innovation may be debated as conflicting rather than complementary goals.

Even where there is broad agreement on goals, stakeholders may bring different moral and normative imaginations to discussions of specific policies. Frameworks such as utilitarianism, rules-based approaches to evaluating social outcomes, or frameworks rooted in virtues ethics may result in different recommendations. Reconciling them would require foundational discussions that are happening, but not yet on a broad scale. Consequently, ongoing debates often oscillate between positions, such as the roller coaster of executive orders of the two most recent U.S. administrations. Some stakeholders and civil society advocates call for governance that is flexible, adaptive, and agile (e.g. UNESCO, n.d.; United Nations, 2024b; WEF, 2020)⁶. Others call for a moratorium on the development of superhuman intelligence capabilities. And yet others advocate for unregulated market forces.

As AI is foundational to a wide variety of critical applications, the number of those impacted by these systems scales rapidly. Thus, the proliferation of stakeholders who have a valid interest in seeing governance that protects their interests may result in political gridlock rather than workable solutions. Alternatively, it may result in some actors pursuing

⁶ The Future Society. <https://thefuturesociety.org/>

their narrow agendas that may not be aligned with broader public interest goals. Finally, even though agile and adoptive methods are seen as crucial for the success of AI governance (OECD, 2021, 2024), little is known about their efficacy (Bauer, 2023).

5. REIMAGINING THE ROLE OF CIVIL SOCIETY

As briefly discussed in the previous section, the shift from traditional government policy to governance as problem solving has created additional areas for contributions by civil society organizations. Here we discuss how civil society participation is reimagined at three layers of the governance system: selected international organizations, in the U.S. legal system, and on the ground in specific civil society engagements.

5.1 International Organizations

Several international organizations have embraced the participation of civil society in AI governance. Various initiatives and policy discussions promote multistakeholder collaboration. This subsection reviews the approaches of the United Nations (UN), the United Nations Educational, Scientific and Cultural Organization (UNESCO), the International Telecommunication Union (ITU), the Organisation for Economic Co-operation and Development (OECD), and the World Economic Forum (WEF).

United Nations (UN)

The UN recognizes the need for globally coordinated AI governance amid increasing fragmentation (United Nations, 2025). In 2023, it established the High-Level Advisory Body on AI (HLAB), which produced the *Governing AI for Humanity* report (2024), identifying gaps in representation, coordination, and implementation and calling for human rights–based, multistakeholder approaches. Concurrently, the Global Digital Compact adopted by the UN in 2024 constitutes a non-binding global policy framework for digital governance, including AI, grounded in human rights and multistakeholder principles (United Nations, 2024a). Building on this and the HLAB recommendations, the UN General Assembly created two mechanisms in 2025: the Global Dialogue on AI Governance and the Independent

International Scientific Panel on AI. The Dialogue serves as a forum for multistakeholder discussion, while the Panel provides independent, evidence-based assessments of AI risks and impacts.

This policy framework moved forward relatively quickly. In February 2026, the UN General Assembly appointed 40 members to the Independent International Scientific Panel on AI. Most members were from academia (70%), followed by the private sector (20%) and government (5%), while two members (5%) represented civil society. A similar pattern was observed in the HLAB, where civil society representation was also minimal, with two members (Namsrai, 2026). In July 2026, the Panel released its preliminary report, which assessed recent developments in AI capabilities, applications, risks, and societal impacts. It also identified gaps in existing evaluation and governance mechanisms, particularly in relation to rapidly advancing and increasingly agentic AI systems (United Nations, 2026). The report informed the inaugural session of the Global Dialogue on AI Governance, held in Geneva on 6–7 July 2026. The Dialogue was organized around four thematic clusters: AI opportunities and impacts, bridging AI divides, ensuring safe and trustworthy AI, and protecting and promoting human rights. Across these thematic areas, discussions also addressed international cooperation, capacity-building, and inclusive approaches to AI governance⁷.

The Dialogue brought together Member States, UN entities, intergovernmental organizations, the private sector, academia, civil society, the technical community, and other stakeholders, although publicly available information does not provide a breakdown of participants by stakeholder group. Stakeholders were also invited to provide written submissions ahead of the Dialogue, resulting in 1,427 contributions: 391 from civil society, 376 from the private sector, 357 from academia, 144 from international organizations, 119 from the technical community, 26 from governments, and 14 from other stakeholders⁸. These statistics indicate that civil society actively contributed to the written consultation process. However, further research is needed to assess how stakeholder submissions

⁷ <https://www.un.org/global-dialogue-ai-governance/en/programme>

⁸ <https://www.un.org/global-dialogue-ai-governance/en/insights-written-submissions>

shaped the design and agenda of the Dialogue and the extent and nature of civil society participation during the event itself. A civil society readout report submitted after the Dialogue is discussed in the following 5.3.3 subsection.

UNESCO

UNESCO is expanding civil society participation in AI governance through the Global Civil Society Organizations and Academic Network on AI Ethics and Policy (<https://www.unesco.org/ethics-ai/en/civil-society-organizations>), launched in 2025. Building on the 2021 Recommendation on the Ethics of AI, the initiative promotes collaboration, knowledge exchange, and human rights-based AI development and participation in governance (UNESCO, 2025). As of February 2026, the Network's public directory lists over 100 member organizations from 43 countries, with around 88% representing CSOs (Namsrai, 2026). It operates through thematic subgroups and a multi-stakeholder coordination structure, organizing policy dialogues and capacity-building activities (UNESCO, 2025). Compared to other initiatives, UNESCO's approach is more formally institutionalized through this network model. However, as it remains in an early stage, its impact is not yet clear (Namsrai, 2026).

ITU

The ITU's *AI for Good* platform (<https://aiforgood.itu.int/>), launched in 2017, aims to advance AI for the UN Sustainable Development Goals by bringing together governments, industry, researchers, and civil society. Its 2025 AI Governance Report highlights multistakeholder governance and emphasizes civil society's role in addressing risks, protecting marginalized groups, and promoting transparency (International Telecommunication Union, 2025a). The ITU also established an Academic Advisory Board in 2026 to provide research-based input (International Telecommunication Union, 2025b). Civil society engagement can be observed primarily through its annual AI for Good Global Summit and its associated working groups. These forums provide opportunities for stakeholders to express their perspectives on the societal impacts, risks, and ethical

implications of AI, participate in discussions on AI governance and standards, and present their own initiatives and experiences (Namsrai, 2026).

OECD

The OECD's AI Policy Observatory (<https://oecd.ai/en/>), launched in 2020, monitors national AI policies, supports the OECD AI Principles (the first intergovernmental AI standard), and connects governments, experts, and CSOs within a multistakeholder framework (Organisation for Economic Co-operation and Development, 2026). OECD.AI offers CSOs input-oriented participation through expert discussions, consultations, and Global Partnership on AI activities, enabling them to use OECD policy tools and contribute analysis. However, engagement is selective and expert-driven, and the Observatory is neither a decision-making body nor an open platform (Namsrai, 2026).

WEF

The WEF (<https://www.weforum.org>) exemplifies a private-led multistakeholder model through its AI Governance Alliance (AIGA), launched in 2023, which brings together governments, corporations, academia, and CSOs to develop AI policy and governance frameworks. Its 2024 Briefing Paper Series positions CSOs as essential partners in responsible AI, highlighting their role in representing underrepresented voices, particularly from the Global South (World Economic Forum, 2024). Their 2024 report *Governance in the Age of Generative AI: A 360° Approach* argues that generative AI risks require collaboration among industry, government, civil society, and academia, recognizing CSOs for their expertise on impacts across communities (World Economic Forum & Accenture, 2024). However, the model remains corporate-dominated, with civil society participation largely invite-only and lacking institutionalized representation.

Taken together, these initiatives highlight the importance of multistakeholder participation and civil society in AI governance. While bodies like the UN and UNESCO have established institutional arrangements to support such participation, civil society's role is generally confined to consultation, including contributing expertise, participating in policy discussions, and raising concerns about the risks and impacts of AI (Namsrai, 2026).

5.2 U.S. Legislation

According to the National Conference of State Legislatures (NCSL), 3,438 AI-related bills were introduced across U.S. states between 2019 and June 1, 2026. Of these, around 200 bills proposed the establishment of AI governance structures that include task forces, advisory groups, commissions, committees, boards, working groups, partnerships, and councils. Of the approximately 200 bills proposing AI governance bodies (Table 1), 39 were enacted or adopted. Among these, 18 established governance bodies with a multistakeholder composition (Table 2), of which eight explicitly included civil society representatives as members. These governance bodies consisted of three task forces, two working groups, two commissions, and one council (Table 3).

Table 1. AI-related bills and proposed AI governance bodies by year, 2019-2026

Year	Proposed N of Bills	Bills Proposing AI Governance Bodies
2019	75	17
2020	61	10
2021	84	10
2022	67	11
2023	135	14
2024	480	52
2025	1143	57
2026	1393	29
Total	3438	200

Table 2. Multistakeholder AI governance bodies by body type and year

Body type	Total	2019	2020	2021	2022	2023	2024	2025	2026
Task Force	7	0	0	1	1	0	3	1	1
Commission	3	0	1	0	0	0	1	0	1
Working Group	3	0	0	0	0	0	1	2	0
Council	2	0	0	1	0	0	1	0	0
Advisory Council	2	0	0	0	1	0	1	0	0
Advisory Board	1	0	0	0	0	0	0	1	0
Total	18	0	1	2	2	0	7	4	2

Table 3. Enacted/adopted multistakeholder AI governance bodies explicitly including civil society representatives (A more detailed table is provided in Appendix A.)

	Year founded / Jurisdiction	Bill Number (Title) / Established Body / Purpose	N of Members / CS Representatives
1	2020 / Massachusetts	-HB 5250 (Future of Work Commission) -Special Commission -Study AI's impacts on the future of work	17 / 6
2	2022 / Vermont	-H.410 (Use and Oversight of Artificial Intelligence in State Government) -Advisory Council -Advise oversight of government AI systems	10 / 2
3	2024 / Colorado	-HB 24-1468 (Artificial Intelligence Act) (Reconstituted and expanded from the 2022 Task Force for the Consideration of Facial Recognition Services) -Task Force -Develop AI policy recommendations	26 / 6
4	2024 / Washington	-ESSB 5838 (Artificial Intelligence Task Force) -Task Force -Assess AI trends and recommend legislation	19 / 4
5	2025 / Hawaii	-SB 742 (Data Sharing) -Working Group -Improve data sharing and AI guidelines	18 / 2 (1 of them found from the source)
6	2025 / Maryland	-HB 956 (AI Implementation Workgroup) - Workgroup -Monitor AI implementation and recommend regulation	33 / 8 (2 of them appointed)
7	2026 / Pennsylvania	-SR 244 (Child Protection in the Digital Age) -Task Force -Address AI-related child protection risks	17 / 2
8	2026 / Alabama	-HJR 51 (AI and Children's Internet Safety) -Study Commission -Study AI and children's online safety	14 / 1

The policy scope of these eight governance bodies primarily focuses on AI governance, regulation, ethics, human rights, and consumer protection. Some also address sector-specific AI policy issues, including the future of work and automation (Massachusetts), data governance (Hawaii), and children's online safety (Pennsylvania and Alabama).

These governance bodies examine the opportunities and risks associated with AI, convene diverse stakeholders for deliberation, synthesize scientific evidence and

stakeholder perspectives, develop AI policies and legislative recommendations, and submit these recommendations to legislative and executive authorities. To fulfill these responsibilities, they organize meetings and hearings and produce reports and policy recommendations.

Most of these governance bodies were designed to operate on a temporary basis. For example, Alabama's Study Commission was established for a six-month period, while Massachusetts' Special Commission on the Future of Work operated from 2021 to 2022 and completed its mandate upon issuing its final report. Most of the remaining bodies are scheduled to operate for approximately two to five years before being dissolved. At present, the Vermont AI Advisory Council, the Washington Artificial Intelligence Task Force, the Hawaii Working Group, the Maryland Workgroup, and the Pennsylvania Task Force remain active. By contrast, Colorado's Artificial Intelligence Impact Task Force is currently undergoing sunset review, while the Pennsylvania Task Force has no specified expiration date.

The level of civil society representation varies considerably across states and governance bodies. Based on statutory membership, civil society representatives account for approximately 7% to 35% of total membership. For example, six of the seventeen members of Massachusetts' Special Commission represented labor organizations, workforce development organizations, and worker advocacy groups, accounting for more than one-third of the Commission's membership. Similarly, six of the twenty-six members of the Colorado's Task Force represented organizations focused on labor rights, civil liberties, disability rights, consumer protection, privacy, and digital rights. However, in some cases, not all seats allocated to civil society organizations have been filled, resulting in lower actual representation than envisioned by statute. For example, although Maryland's Workgroup statutorily reserved eight of its thirty-three seats (approximately 24%) for civil society organizations, as of the 2026 report only two nonprofit representatives had been appointed, representing approximately 6% of the Workgroup's total membership.

The types of civil society represented are diverse rather than concentrated in a single sector. Labor organizations are the most common, with American Federation of Labor and

Congress of Industrial Organizations (AFL-CIO) <https://aflcio.org/> and affiliated organizations represented in Massachusetts, Colorado, and Washington. Civil liberties, privacy, and technology policy organizations, including the American Civil Liberties Union (ACLU) <https://www.aclu.org/>, the Center for Democracy & Technology <https://cdt.org/>, and the Future of Privacy Forum <https://fpf.org/>, contribute expertise on AI accountability, algorithmic discrimination, privacy, and digital governance.

Other represented organizations include the American Association of People with Disabilities <https://www.aapd.com/> (Colorado), which advocates for disability rights; House of Ruth Maryland <https://hruth.org/>, which provides services to survivors of domestic violence; and the Public Justice Center <https://www.publicjustice.org/en/> (Maryland), which engages in public interest legal advocacy. Alabama's enabling resolution also reserves one seat for an appointee of Voices for Alabama's Children <https://alavoices.org/>, a statewide nonprofit organization that advocates for children's well-being and public policy affecting children.

These examples suggest that legislators are seeking to involve organizations with relevant expertise and experience in specific AI policy areas in the policymaking process. First, civil society organizations contribute knowledge and expertise in areas such as labor, civil and human rights, consumer protection, privacy, ethics, disability rights, child protection, and data governance to policy deliberations. Second, they provide a channel through which the interests and perspectives of groups that may be disproportionately affected by AI, including workers, consumers, children, and people with disabilities, can be represented. Third, as formal members of these governance bodies, they participate in meetings, deliberations, research, report writing, and the development of policy recommendations.

However, these governance bodies are primarily advisory and research oriented. Although civil society representatives are formally included in their membership, further research is needed to examine the extent of their participation in the policymaking process and to assess their specific contributions to final reports, policy recommendations, and other governance outcomes. In addition to these eight cases, other multistakeholder

governance bodies include representatives from the private sector and academia but do not explicitly include civil society. In some instances, civil society could be interpreted as being indirectly represented through broader terms such as the public or experts. However, such general references do not guarantee meaningful civil society participation (Namsrai, 2026).

In addition to the cases discussed above, two pending New Jersey bills (SB 4253 in 2025 and SB 1668 in 2026) are noteworthy because they envision a broader role for nonprofit organizations. Unlike the governance bodies described above, these bills propose that nonprofit organizations serve not only as advisors but also as the primary implementation partners of an Artificial Intelligence Innovation Partnership to be established under the New Jersey Commission on Science, Innovation and Technology. The proposed Partnership would support AI innovation through one or more independent nonprofit organizations. These organizations would be responsible for developing regional AI innovation ecosystems, coordinating collaboration among government, businesses, universities, research institutions, and other stakeholders, supporting AI startups, administering grant and incubation programs, and helping attract investment. However, they remain under legislative consideration and have not been enacted.

Overall, these findings suggest that U.S. states are beginning to institutionalize civil society as formal participants in AI governance structures. This creates opportunities for civil society to participate more systematically in policymaking, engage more closely in decision-making processes, and facilitate more effective communication and information exchange (Namsrai, 2026). There are also early indications of a shift toward more substantive implementation roles. The limited representation of civil society organizations within AI governance bodies does not imply that they are unable to contribute to AI governance through their own activities and modes of engagement. The following subsections present examples of how civil society organizations operate and the potential contributions they make to AI governance.

5.3 Cases of Civil Society Engagement in the U.S.

This subsection is organized into three parts. First, it provides an overview of civil society organizations engaged in AI, including when they began their AI-related activities, their core mandates, target issues, primary activities, and sources of funding. Second, it presents in-depth vignette analyses of five U.S.-based civil society organizations to examine their activities, governance strategies, and potential contributions to AI governance. Third, it discusses selected examples of policy and institutional changes that resulted from the activities and advocacy of civil society organizations.

5.3.1 Overview of Civil Society Organizations Engaged in AI Governance

First, civil society engagement in AI is relatively recent, emerging mainly from the mid-2010s and accelerating in the 2020s. Some organizations were established with a direct focus on AI (AI-specific), while others have expanded their broader mandates, such as digital rights and human rights, to incorporate AI-related issues. At the same time, new areas such as AI safety, geopolitics, and AI literacy are becoming more prominent, reflecting the ongoing expansion and diversification of the field.

In terms of substantive focus, the most prominent issue areas include human rights and freedoms, algorithmic bias and discrimination, surveillance and privacy, AI governance and policy, and trustworthy and responsible AI.

The activities of CSOs are multi-dimensional. They engage in policy advocacy, conduct research, disseminate information to the public, monitor developments, and in some cases participate in litigation. This demonstrates that the role of civil society is not confined to a single function but instead expands in line with the complex and multifaceted nature of AI governance.

In terms of funding, most organizations rely on external sources such as grants, philanthropic foundations, and individual donations. Philanthropic and international funding plays a dominant role, while in some cases support from technology companies is also observed. (See Appendix B for a detailed table of the organizations included in the analysis.)

5.3.2 Illustrative Vignettes of Civil Society Engagement⁹

For the case vignettes, the activities of the Center for AI and Digital Policy (CAIDP), the Partnership on AI (PAI), the Center for Democracy & Technology (CDT), the Algorithmic Justice League (AJL), and Oakland Privacy are examined as cases of civil society engagement in AI governance.

Center for AI and Digital Policy

CAIDP (<https://www.caidp.org/>) is an independent nonprofit founded in 2020 in Washington, DC, U.S., focused on aligning AI policy with human rights, democratic governance, transparency, and accountability. It operates at both global and national levels through wide range of activities; only selected examples are highlighted below.

One of CAIDP's core activities is submitting formal comments and policy recommendations. For example, it has contributed to the UN General Assembly, where it emphasized principles of human rights, transparency, and inclusion, as well as to OECD consultations on AI risk thresholds and U.S. AI policy processes, demonstrating active engagement in these areas (Center for AI and Digital Policy, 2025c; 2025a).

Another key activity is its production of comparative research. Since 2021, it publishes the annual *Artificial Intelligence & Democratic Values Index*, evaluating national AI policies across criteria such as human rights, transparency, and accountability, enabling cross-national comparison and greater transparency (Center for AI and Digital Policy, 2025a).

In addition, CAIDP runs global capacity-building programs, including the AI Policy Clinic, which trains policymakers, researchers, and CSOs through practical work on real policy documents. Together with CAIDP Fellows programs, these initiatives help build an international network of AI policy actors (Center for AI and Digital Policy, n.d.).

Beyond these highlighted areas, CAIDP participates in international networks such as UNESCO's AI Ethics Network, supports the implementation of global AI principles,

⁹ This subsection and the next build on research reported in Namsrai (2026).

facilitates information-sharing through platforms like its weekly updates, and contributes to broader advocacy, research, and multistakeholder engagement in AI governance.

Partnership on AI

The Partnership on AI (<https://partnershiponai.org/>), takes a different approach. As a nonprofit, multistakeholder organization founded in 2016 in the U.S., it brings together technology companies, civil society organizations, academia, labor representatives, and policymakers to promote responsible AI development and governance (Partnership on AI, 2026a). Unlike advocacy-oriented organizations, PAI operates as a cross-sector platform focused on collaborative and implementation-oriented approaches.

Two key aspects characterize its work. First, PAI relies on multistakeholder Steering Committees that identify priorities and co-develop frameworks, guidelines, and reports across thematic areas, while supporting their dissemination and implementation (Partnership on AI, 2026c). Second, it develops practical governance tools, such as the Guidance for Inclusive AI, which outlines steps for engaging stakeholders from early stages, integrating feedback into decision-making, and establishing accountability, reflecting its view of AI development as a collaborative social process (Partnership on AI, 2026b).

Center for Democracy & Technology

The Center for Democracy & Technology (<https://cdt.org/>), founded in 1994 in Washington, DC, is a nonprofit organization focused on protecting civil rights and liberties in the digital environment, with increasing emphasis on AI issues such as algorithmic accountability, automated decision-making, privacy, and civil rights. CDT engages in U.S. federal, state, and EU policy discussions by providing formal input, technical expertise, and training on AI-related risks (Center for Democracy and Technology, 2024).

Its AI Governance Lab conducts research on algorithmic fairness, transparency, evaluation, and auditing, developing methods, standards, and best practices for policymakers, companies, and civil society (Center for Democracy and Technology, 2026b). In addition, CDT maintains institutional support structures, including an Advisory Council

and a Collaborating Attorney Network, which provide strategic guidance and legal support for policy advocacy and analysis (Center for Democracy and Technology, 2026a; 2026c).

Algorithmic Justice League

Second CSO is the Algorithmic Justice League (<https://www.ajl.org/>), founded in 2016 in the United States, works to identify and mitigate discrimination and harms in AI systems. It combines art, research, and advocacy to make the social impacts of AI visible and to promote more equitable and accountable AI. For example, through projects such as Gender Shades, AJL has exposed racial and gender bias in facial recognition systems and communicates these findings to the public and policymakers through videos and documentary media (Algorithmic Justice League, n.d.).

It also participates in policy debates on regulating and restricting facial recognition technologies. In addition, through its AI Harms Knowledge platform, AJL encourages individuals to document AI-related harms, demand transparency and accountability, and engage in policy discussions, thereby strengthening public participation in AI governance (Algorithmic Justice League, 2025).

Oakland Privacy

A third example is Oakland Privacy (<https://oaklandprivacy.org/>), a grassroots civil society coalition based in Oakland, California, which promotes transparency and accountability in government surveillance. Established in 2013 during debates over the Domain Awareness Center, it operates through citizen activism and community engagement.

The organization informs the public about surveillance technologies such as cameras, license plate readers, and facial recognition, while participating in local policymaking. It contributed to Oakland's Surveillance Technology Ordinance and helped establish the Oakland Privacy Advisory Commission, enabling participatory oversight (Oakland Privacy, 2020). More recently, it has engaged in debates on data-driven surveillance, including opposing Flock Safety cameras, raising concerns about algorithmic

rent pricing, and criticizing California SB 690. Through public records investigations, it has exposed data access issues and called for stronger oversight (Oakland Privacy, 2026).

In sum, these organizations illustrate distinct but complementary modes of civil society engagement in global AI governance. CAIDP emphasizes advocacy through normative frameworks, policy submissions, and accountability grounded in human rights. PAI adopts a multistakeholder approach, focusing on collaboration, practical tools, and implementation. CDT represents a policy-oriented model grounded in regulatory processes and expert analysis, AJL a public-facing model combining research and advocacy, and Oakland Privacy a community-driven model focused on participatory oversight.

Moreover, several CSOs focus on AI literacy and education, including AI and You (<https://aiandyou.org/>), which emphasizes women's participation, and the Linden AI Institute (<https://www.lindenaiinstitute.org/>), which promotes access for underrepresented groups. While they share a goal of improving public understanding, they differ in approach and audiences and often collaborate with and receive support from technology companies.

Taken together, the cases suggest that civil society participation is primarily expressed through conventional, external modes of engagement, while also expanding and becoming increasingly diverse across levels and orientations. Rather than a single type of actor, civil society constitutes a heterogeneous ecosystem of advocacy groups, multistakeholder platforms, research-driven watchdogs, grassroots movements, and education-focused initiatives.

They also indicate that civil society actors do not possess uniform levels of knowledge and technical expertise, with capacities varying significantly across organizations. Even among the relatively visible CSOs examined here, this variation is substantial and likely more pronounced beyond the sample.

5.3.3 Examples of Contributions of Civil Society Organizations to Policy and Institutional Change

The following examples illustrate how civil society organizations express their positions through the activities and policies of international organizations and state level authorities.

Before the UN Global Dialogue on AI Governance, and again following its conclusion, members of civil society organizations from five coalitions, namely the Global Digital Rights Coalition, the Global Digital Justice Forum, the Global South Alliance, MAP-AI, and the civil society members of the Gender in Digital Coalition, submitted written inputs¹⁰¹¹¹² and a Civil Society Read Out Report¹³, respectively. These documents addressed key issues including meaningful human rights-based AI governance, accountability and human oversight, capacity-building, gender equality, environmental impacts, and data governance.

In addition, they presented a range of recommendations on how to strengthen civil society participation in the Dialogue. These included ensuring meaningful and continuous civil society engagement throughout the process, increasing the representation of civil society organizations in plenary sessions, providing financial and logistical support for underrepresented groups, improving transparency regarding how stakeholder inputs are reflected in outcomes, and creating more inclusive and participatory consultation mechanisms. This indicates that civil society sought not only to contribute substantively to AI governance discussions but also to strengthen their representation and participation within the Dialogue process.

Beyond efforts to strengthen civil society participation in international AI governance dialogues, civil society organizations have sought to expand their representation in state-

¹⁰ <https://www.gp-digital.org/wp-content/uploads/2026/05/Global-Digital-Rights-Coalition-submission-to-the-AI-Global-Dialogue.pdf>

¹¹ <https://www.derechosdigitales.org/wp-content/uploads/2026/05/Global-Dialogue-on-AI-Governance-%E2%80%93-Written-submission-by-the-Global-Digital-Justice-Forum.docx.pdf>

¹² <https://globalnetworkinitiative.org/wp-content/uploads/final-15-may-ccg-gni-submission-to-ungd-consultation-909.pdf>

¹³ <https://www.apc.org/sites/default/files/civil-society-read-out-report-ai-dialogue-report.pdf>

level AI governance bodies. For example, in response to Maryland House Bill 956, a coalition of civil society organizations submitted a joint letter¹⁴ to the Senate Finance Committee requesting amendments to the composition of the Workgroup on Artificial Intelligence Implementation. The coalition argued that the original proposal was dominated by industry representatives and lacked adequate civil society and academic representation, creating a risk that the workgroup would prioritize corporate interests over the public interest. It therefore recommended adding representatives from nonprofit organizations focused on data privacy, civil rights, and consumer protection, as well as labor representatives, academic experts, and independent technologists.

The enacted version of H.B. 956¹⁵ incorporated several provisions closely aligned with these recommendations by reducing the number of industry representatives and adding nonprofit representatives focused on civil rights, consumer protection, and data privacy, along with labor-affiliated representatives, academics, and cybersecurity specialists. Although not all recommendations were adopted, the resulting composition was more balanced and suggests that civil society advocacy may have contributed, at least in part, to shaping the institutional design of the workgroup.

In addition, civil society organizations have played an important role in identifying AI-related harms and promoting institutional safeguards that limit the use of AI in specific contexts, strengthen accountability and transparency, and reduce erroneous or discriminatory applications. Several cases illustrate these efforts.

First, in Oakland, the Electronic Frontier Foundation (EFF) and local civil rights and privacy advocates supported efforts to prohibit government use of facial recognition technology during discussions on the city's surveillance policy. In 2019, the Oakland City

¹⁴ <https://epic.org/documents/maryland-h-b-956-workgroup-on-ai-implementation/#f1e8a3ef-6366-4059-8b41-21d93d674f7e>

¹⁵

https://sn.lexisnexis.com/opentext/eyJ0eXBlljoiYmlsbCIsImkljoiTUQyMDI1MDAwSDk1NiJ9.O6ASqsN492qr5xfrmTLv6tJZamvrr_4exaBDztZGFJU0/text

Council unanimously adopted an ordinance banning government use of facial recognition technology¹⁶.

Second, Americans for Responsible Innovation (ARI) and a coalition of more than 20 organizations advocating for child safety and responsible technology engaged in advocacy activities, including coalition building, legal analysis, public mobilization, and legislative outreach. Subsequently, the TAKE IT DOWN Act was enacted in 2025, establishing federal requirements for addressing the distribution of non-consensual intimate images, including AI-generated deepfakes¹⁷.

Third, ARI and other advocacy organizations campaigned against a proposed 10-year federal moratorium on state AI regulation through policy research, public campaigns, and legislative engagement. The provision was ultimately removed from the reconciliation bill by the U.S. Senate¹⁸.

Fourth, following litigation brought by the ACLU and co-counsel on behalf of Robert Williams, the Detroit Police Department adopted a revised facial recognition policy that prohibits relying solely on facial recognition results for arrests, requires independent corroborating evidence, and mandates officer training on the technology's limitations and risks¹⁹.

Finally, litigation brought by the ACLU of Illinois and other organizations under the Illinois Biometric Information Privacy Act (BIPA) resulted in a settlement requiring Clearview AI to significantly restrict access to its facial recognition database by most private entities nationwide and to comply with additional restrictions in Illinois²⁰.

These changes might not be attributed solely to the activities of civil society organizations, as other factors likely also contributed to these outcomes. However, the

¹⁶ <https://www.eff.org/deeplinks/2019/07/victory-oakland-city-council-votes-ban-government-use-face-surveillance>

¹⁷ <https://ari.us/success-stories/>

¹⁸ <https://ari.us/success-stories/>

¹⁹ <https://www.aclu.org/press-releases/civil-rights-advocates-achieve-the-nations-strongest-police-department-policy-on-facial-recognition-technology>

²⁰ <https://www.aclu.org/press-releases/big-win-settlement-sures-clearview-ai-complies-with-groundbreaking-illinois>

cases illustrate the potential of civil society organizations to make contributions to policy and institutional change in AI governance.

6. DISCUSSION

Our narrative review of the research literature and of policy documents has revealed conceptual arguments in support of the role of civil society actors in AI governance. It has also generated evidence of their active role in the development of governance responses and identified efforts to institutionalize civil society participation. We begin our discussion with reflections on the contributions of civil society organizations to address the hard challenges of AI governance and then continue with a discussion of the observations from our narrative review of the literature and of case vignettes.

6.1 Civil Society Organizations in the Governance System

Civil society organizations active in AI governance are heterogeneous in age, size, funding, scope of activities, and capabilities. They range from older civil society organizations, such as the Electronic Frontier Foundation (EFF), to new groups founded in response to the recent rapid growth of and interest in AI, such as the Center for AI and Digital Policy (CAIDP) and the Global Center on AI Governance in Cape Town, South Africa. Some are active locally while others operate nationally and globally. Some pursue a broad range of activities from advocacy to testimony in public proceedings, whereas others pursue more specific roles, such as representation in a government task force. Despite this diversity, there are common elements in how they can contribute to AI governance.

Table 4 builds on the insight that the layers of the institutional stack normally operate at varying time scales that range from real time decisions at the operational level to centuries for culturally embedded norms. This has been theorized extensively in research on institutions (see, e.g., Williamson, 2000), technological transitions (see, e.g., Geels, 2002), and on socio-technical ecosystems (see, e.g., Janssen & van der Voort, 2016). Civil society organizations can contribute across these scales, but they do so in different ways and with varying effectiveness.

Given their heterogeneity, the strengths and weaknesses of specific organizations will influence where they can contribute most effectively. Faster-paced decision processes at the lower end of the stack will likely be more challenging, but monitoring, auditing, and consultation offer opportunities to contribute. At the higher end of formal governmental policies and international agreements, decisions are influenced by many other factors, and active participation by civil society organizations will be facilitated if they are provided with a formal role. They probably can have the strongest role in networked, multi-stakeholder governance arrangements.

Table 4: Time scales of governance and potential roles of civil society organizations

Institutional mode	Time scale	Possible contributions
Embedded cultural norms	Decades/centuries	Advocacy
International agreements	Years/decades	Advocacy
National constitutional law	Years/decades	Testimony, advocacy
National statutory law	Years	Testimony, advocacy
National case law	Years	Testimony, advocacy
National regulation	Years	Testimony, advocacy
Regulatory sandboxes	Month/years	Monitoring, evaluation
Network governance	Month/years	Participation, consultation
Organizational governance	Weeks/months	Consultation, auditing, monitoring
Algorithmic governance	Days, real time	Auditing, monitoring
Operational decisions	Real time	Monitoring
Individual decisions	Real time	Ethical guidelines

Source: Own research. Inspired by Williamson (2000; Janssen & van der Voort (2016).

6.2 Strengths and Weaknesses of Civil Society

Examining the potential role of civil society organizations toward finding sustainable solutions offers additional insights (see Table 2). In each area, civil society organizations have potential strengths and weaknesses. For example, as white hat and black hat hackers show, these groups often have deep insights into the functioning of technology. However, in the case of AI these may be more limited in areas that are considered proprietary or limited due to national security concerns. Civil society organizations can bring expertise and insight

into the design of specific policy instruments, but they typically have limitations in implementing them.

Table 5: Civil society organizations and alignment

Area	Strengths of civil society	Weaknesses of civil society
Working of socio-technical system	• Access to local tacit and explicit knowledge of AI uses and their impacts • Deep understanding of local and specific aspects of AI and its impacts	• Limited insights into proprietary aspects of AI development • Resource constraints may limit ability to research outcomes, impacts of AI
Development of workable instruments	• Experience of (local) effects) of AI • Practical knowledge of how AI governance affects (local) outcomes • Familiarity with community concerns	• Powers to implement proposals limited to voluntary actions or contingent on other actors • Development of comprehensive views
Achievement of political feasibility	• Bring community values to the discussion • May contribute to development of shared mental models	• May further fragment discussion, inadvertently increasing the agency of powerful actors

Source: Own research. See also Bauer (2022).

They probably have to contribute most to the development of feasible policy solutions by articulating values that are shared by stakeholders and sharing their tacit knowledge of uses and effects of AI, including resistance to unacceptable ones. At the same time, there is a risk that their presence leads to further fragmentation of the polity with the inadvertent effect that resourceful players may gain additional leeway to pursue their goals. Thus, much will depend on the institutional arrangements in which they participate, their specific capabilities, and their willingness to contribute to the political process.

7. CONCLUSION

Civil society organizations span a wide range on the social and political spectrum and operate with varying effectiveness. As our case vignettes show, there are important

instances in which they succeeded in representing user interest and changing the course of policy. Studies of the role of civil society organizations in other contexts, such as Internet governance, suggest that they are granted a level of legitimacy despite limited formal democratic controls and oversight. Whether this will be replicated in AI governance remains to be seen.

There are also many areas of ambiguity which civil society organizations involved in AI governance must navigate. Some of them rely heavily on funding from big tech companies. Others diversify their resources more widely to retain independence, but they may be tacitly beholden to industry positions. Even if they can successfully avoid these potential tensions, the efforts of institutionalizing their participation may inadvertently lead to an embedding in established networks of state and corporate power. Access to information and data is another handicap that many CSOs face, even though there are successful examples of white hat hacker communities that uncovered biases and flaws in AI models and applications as well as numerous other successful initiatives as discussed earlier.

Despite these challenges, our work revealed many areas in which they can contribute meaningfully and effectively, especially if part of broader networks of knowledge exchange. Public policy and the private sector could strengthen these positive impacts in several ways. Legal frameworks that improve access to data could be an important step. Non-conditional funding to CSOs from non-private sources could create an important independent voice that can help navigate the many challenges of governing fast paced technology and its implications.

AI Statement

During the preparation of this work, the authors used ChatGPT, Claude, Copilot, and Gemini for preliminary background research. The authors reviewed and verified the output as needed. They assume full responsibility for the content of the published article. AI tools were not used for writing purposes.

REFERENCES

- Algorithmic Justice League. (n.d.). *Projects*. Retrieved March 29, 2026, from <https://www.ajl.org/library/projects>
- Algorithmic Justice League. (2025). *Harms Reporting*. <https://www.ajl.org/harms>
- Anthropic. (2026). *AI policy*. <https://www.anthropic.com/policy>
- Baldwin, C. Y. (2024). *Design rules, volume 2: How technology shapes organizations*. MIT Press. <https://doi.org/https://doi-org.proxy2.cl.msu.edu/10.7551/mitpress/7374.001.0001>
- Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. *AI and Ethics*, 5(3), 3265–3279.
- Bauer, J. M. (2022). Toward new guardrails for the information society. *Telecommunications Policy*, 46(5), 102350. <https://doi.org/10.1016/j.telpol.2022.102350>
- Bauer, J. M. (2023). Agile regulation: Experiments in flexible governance. In F. Birke, G. Brunekreeft, V. Stocker, & H.-J. Weiß (Eds.), *Perspektiven der Netzökonomie – Beiträge zu Ehren von Günter Knieps* (pp. 31–49). Nomos.
- Bauer, J. M., & Stocker, V. (2026, July 28). *The Interplay of Regulation and Innovation at the Technological Frontier*. 25th Biennial Conference of the International Telecommunications Society (ITS), Tokyo, Japan.
- Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., & Thomas, D. (2001). *Manifesto for Agile Software Development*. Manifesto for Agile Software Development. <https://agilemanifesto.org/>
- Bullock, J. B., Chen, Y.-C., Himmelreich, J., Hudson, V. M., Korinek, A., Young, M. M., & Zhang, B. (2024). *The Oxford handbook of AI governance*. Oxford University Press.
- Center for AI and Digital Policy. (n.d.). *Global Academic Network*. Center for AI and Digital Policy. Retrieved March 29, 2026, from <https://www.caidp.org/global-academic-network/>

- Center for AI and Digital Policy. (2025a). *Artificial Intelligence and Democratic Values 2025*.
<https://www.caidp.org/reports/caidp-index-2025/>
- Center for AI and Digital Policy. (2025b). *CAIDP Statements*. Center for AI and Digital Policy. <https://www.caidp.org/statements/>
- Center for Democracy and Technology. (2024). *2024 Annual Report*. <https://cdt.org/2024-annual-report/>
- Center for Democracy and Technology. (2026a). *CDT Advisory Council*.
<https://cdt.org/about/advisory-council/>
- Center for Democracy and Technology. (2026b). *CDT AI Governance Lab*.
<https://cdt.org/cdt-ai-governance-lab/>
- Center for Democracy and Technology. (2026c). *Collaborating Attorney Network*.
<https://cdt.org/about/collaborating-attorney-network/>
- Christian, B. (2020). *The alignment problem: Machine learning and human values*. W. W. Norton & Company.
- Clark, D. D., & claffy, k c. (2015). Anchoring policy development around stable points: An approach to regulating the co-evolving ICT ecosystem. *Telecommunications Policy*, 39(10), 848–860. <https://doi.org/https://doi.org/10.1016/j.telpol.2015.07.003>
- Dubber, M. D., Pasquale, F., & Das, S. (2020). *The Oxford handbook of ethics of AI*. Oxford University Press.
- Enroth, H. (2014). Governance: The art of governing after governmentality. *European Journal of Social Theory*, 17(1), 60–76.
<https://doi.org/https://doi.org/10.1177/1368431013491818>
- Fairfield, J. A. T. (2021). *Runaway technology: Can law keep up?* Cambridge University Press.
- Gasser, U., Siewert, M., Español, A. G., & Molnar, A. (2025). *AI Sandboxes – A Key Tool for Adaptive AI Governance*. Tum Think Tank | Global Network of Internet and Society Centers. https://tumthinktank.de/wp-content/uploads/PolicyBrief_Final.pdf

- Geels, F. W. (2002). Technological Transitions as Evolutionary Reconfiguration Processes: A Multi-level Perspective and a Case Study. *Research Policy*, 31, 1257–1274.
[https://doi.org/https://doi.org/10.1016/S0048-7333\(02\)00062-8](https://doi.org/https://doi.org/10.1016/S0048-7333(02)00062-8)
- Google. (2026). *AI Principles*. Google AI. <https://ai.google/principles/>
- Hagendorff, T. (2020). The ethics of AI: An evaluation of guidelines. *Minds and Machines*, 30, 99–120. <https://doi.org/https://doi.org/10.1007/s11023-020-09517-8>
- IBM. (2021). *What is the NIST Cybersecurity Framework?* | IBM.
<https://www.ibm.com/think/topics/nist>
- International Telecommunication Union. (2025a). *The Annual AI Governance Report 2025: Steering the Future of AI* (ISBN 978-92-61-41441-2 (electronic)).
<https://www.itu.int/epublications/en/publication/the-annual-ai-governance-report-2025-steering-the-future-of-ai>
- International Telecommunication Union. (2025b, September 5). New ITU academic collaboration on technology trends. *ITU*. <https://www.itu.int/hub/2025/09/new-itu-academic-collaboration-on-technology-trends/>
- Janssen, M. (2025). Responsible governance of generative AI: Conceptualizing GenAI as complex adaptive systems. *Policy and Society*, 44(1), 38–51.
<https://doi.org/10.1093/polsoc/puae040>
- Janssen, M., & van der Voort, H. (2016). Adaptive governance: Towards a stable, accountable and responsive government. *Government Information Quarterly*, 33(1), 1–5. <https://doi.org/10.1016/j.giq.2016.02.003>
- Janssen, M., & van der Voort, H. (2020). Agile and adaptive governance in crisis response: Lessons from the COVID-19 pandemic. *International Journal of Information Management, Impact of COVID-19 Pandemic on Information Management Research and Practice: Editorial Perspectives*, 55, 102180.
<https://doi.org/10.1016/j.ijinfomgt.2020.102180>
- Khan, M. S., Poyer, A., & Khan, M. J. (2026). *The Institutional Stack: Realigning US AI Governance* [Unpublished manuscript].

- Kolt, N., Shur-Ofry, M., & Cohen, R. (2025). Lessons from complex systems science for AI governance. *Patterns*, 6(8), 101341. <https://doi.org/10.1016/j.patter.2025.101341>
- Laukkonen, R., Krier, S., Bakalar, C., Chandaria, S., Kringelbach, M., Elwood, A., Ford, D., Rosas, F., Bohacek, M., Franklin, M., Tomašev, N., Chan, S., Rieser, V., Patel, R., Levin, M., & Rao, A. (2026). *Positive Alignment: Artificial Intelligence for Human Flourishing* (arXiv:2605.10310). arXiv. <https://doi.org/10.48550/arXiv.2605.10310>
- Lindblom, C. E. (1959). The science of “muddling through.” *Public Administration Review*, 19(2), 78–88.
- Mergel, I., Gong, Y., & Bertot, J. (2018). Agile government: Systematic literature review and future research. *Government Information Quarterly*, 35(2), 291–298. <https://doi.org/10.1016/j.giq.2018.04.003>
- Metnick, C., Becton, E., & Fogelhut, J. (2026). *Testing the Boundaries of Artificial Intelligence in Care Delivery: Utah’s Prescription Renewal Pilot Program*. Shepherd. <https://www.sheppard.com/insights/blogs/testing-the-boundaries-of-artificial-intelligence-in-care-delivery-utahs-prescription-renewal-pilot-program>
- Nagel, R. N. (1992). *21st Century Manufacturing Enterprise Strategy Report (AMEF N0001-92)*.
- Namsrai, G. (2026). *Roles of Civil Society in AI Governance* [Michigan State University]. <https://doi.org/https://doi.org/doi:10.25335/an3a-mw04>
- NIST. (2024). *The NIST Cybersecurity Framework (CSF) 2.0* (NIST CSWP 29; p. NIST CSWP 29). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.CSWP.29>
- NIST. (2026). *NIST AI Resource Center: Govern*. National Institute of Standards and Technology: U.S. Department of Commerce. <https://airc.nist.gov/airmf-resources/playbook/govern/>
- Nuottila, J., Aaltonen, K., & Kujala, J. (2022). Challenges of adopting agile methods in a public organization. *International Journal of Information Systems and Project Management*, 4(3), 65–85. <https://doi.org/10.12821/ijispm040304>

- Oakland Privacy. (2020). *Oakland Privacy Advisory Commission Documents*.
<https://oaklandprivacy.org/oak-advisory-commission/>
- Oakland Privacy. (2026). *Oakland Privacy*. <https://oaklandprivacy.org/>
- OECD. (2021). *Practical guidance on agile regulatory governance to harness innovation*. 20 July 2021 [GOV/RPC(2021)9/REV2] [[GOV/RPC(2021)14/FINAL].
- OECD. (2024). *Framework for anticipatory governance of emerging technologies*. Science, Technology and Industry Policy Papers No. 165. Organisation for Economic Co-operation and Development.
- Organisation for Economic Co-operation and Development. (2026). *The OECD Artificial Intelligence Policy Observatory*. <https://oecd.ai/en/>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71.
<https://doi.org/10.1136/bmj.n71>
- Partnership on AI. (2026a). Annual Report 2025. *Partnership on AI*.
<https://partnershiponai.org/annual-report-2025/>
- Partnership on AI. (2026b). Guidance for Inclusive AI: Practicing Participatory Engagement. *Partnership on AI*. <https://partnershiponai.org/guidance-for-inclusive-ai/>
- Partnership on AI. (2026c). *Our Work*. Partnership on AI. <https://partnershiponai.org/work/>
- Potts, J. (2025). Governing hyperobjects: The new economics of ai regulation: The law & technology & economics of AI. *Network Law Review* (Ed. Adrian Kuenzler, Thibault Schrepel & Volker Stocker), 85(Summer).
<https://heinonline.org/HOL/P?h=hein.journals/ntwklwrw2025&i=295>
- Reuel, A., & Undheim, T. A. (2024). *Generative AI Needs Adaptive Governance* (arXiv:2406.04554). arXiv. <https://doi.org/10.48550/arXiv.2406.04554>
- Schiff, D., Biddle, J., Borenstein, J., & Laas, K. (2020). *What's next for ai ethics, policy, and governance? A global overview*. 153–158. <https://doi.org/10.1145/3375627.3375804>

- Schmitt, L. (2022). Mapping global AI governance: A nascent regime in a fragmented landscape. *AI and Ethics*, 2(2), 303–314. <https://doi.org/10.1007/s43681-021-00083-y>
- Schrepel, T. (2025). The Future-Proof Fantasy of AI Regulation. *Network Law Review*, (The Law&Technology&Economics of AI).
- Sokhansanj, B. A. (2025). Local AI Governance: Addressing Model Safety and Policy Challenges Posed by Decentralized AI. *AI*, 6(7), 159. <https://doi.org/10.3390/ai6070159>
- Torring, J. (2020). Governance through civil society. In *Oxford Research Encyclopedia of Politics*.
- UNESCO. (n.d.). *Empowering Citizens and CSOs in AI governance | Global AI Ethics and Governance Observatory*. Retrieved July 31, 2026, from <https://www.unesco.org/ethics-ai/en/node/335>
- UNESCO. (2025). *Global Civil Society Organizations and Academic Network on AI Ethics and Policy*. <https://www.unesco.org/ethics-ai/en/civil-society-organizations>
- United Nations. (2024a). *Global Digital Compact*. <https://www.un.org/global-digital-compact/en>
- United Nations. (2024b). *Governing AI for humanity: Final report*. United Nations.
- United Nations. (2025). *Artificial Intelligence (AI)*. <https://www.un.org/en/global-issues/artificial-intelligence>
- United Nations. (2026). *Preliminary Report of the Independent International Scientific Panel on AI: Evidence-Based Assessment of Opportunities, Risks and Impacts of Artificial Intelligence* [Preliminary Report]. United Nations. https://www.un.org/independent-international-scientific-panel-ai/sites/default/files/2026-07/en_Preliminary%20Report_.pdf
- WEF. (2020). *Agile regulation for the Fourth Industrial Revolution: A toolkit for regulators*.
- Williamson, O. E. (2000). The New Institutional Economics: Taking Stock, Looking Ahead. *Journal of Economic Literature*, 38(3), 595–613. <https://doi.org/10.1257/jel.38.3.595>

World Economic Forum. (2024). *AI Governance Alliance: Briefing Paper Series*.

<https://www.weforum.org/publications/ai-governance-alliance-briefing-paper-series/>

World Economic Forum & Accenture. (2024). *Governance in the Age of Generative AI: A 360° Approach for Resilient Policy and Regulation* (AI Governance Alliance Briefing Paper Series). <https://www.weforum.org/publications/governance-in-the-age-of-generative-ai/>

APPENDIX A. AI GOVERNANCE BODIES INCLUDED CIVIL SOCIETY REPRESENTATIVES IN U.S STATES. Source: Own research.

	Year founded / Jurisdiction	Bill Number (Title) / Established Body / Purpose	Duration of Body	N of Members / Civil Society Representatives	Civil Society Representatives
1	2020 / Massachusetts	-HB 5250 (Future of Work Commission) -Special Commission -Study AI's impacts on the future of work	2021-2022	17 / 6	Members of the labor community with experience in future work issues: Massachusetts AFL-CIO (Steven Tolman) Members of the labor community with experience in future of work issues: North Atlantic States Regional Council of Carpenters (Joanne Goldstein) Expertise in fair labor and workers' rights: Fair Employment Project (Lisa J. Bernt) Experience in workforce training and education: Just-A-Start Corporation (Miriam Ortiz) Members of the labor community with experience in future of work issues: MassCOSH (Al Vega) Experience in workforce training and education / workforce development: Massachusetts Workforce Association (Tonja Mettlach)
2	2022 / Vermont	-H.410 (Use and Oversight of Artificial Intelligence in State Government) -Advisory Council -Advise oversight of government AI systems	2022-June 30, 2027 (scheduled repeal on June 30, 2027)	10 / 2	American Civil Liberties Union of Vermont (Joe Near), Ethics and human rights expert: Norwich University Applied Research Institutes (Phil Susmann)

3	2024 / Colorado	-HB 24-1468 (Artificial Intelligence Act) (Reconstituted and expanded from the 2022 Task Force for the Consideration of Facial Recognition Services) -Task Force -Develop AI policy recommendations	2024– September 1, 2027; currently undergoing sunset review.	26 / 6	A coalition of statewide labor organizations actively involved in representing the workforce impacted by artificial intelligence systems or automated decision systems: Colorado AFL-CIO (Kjersten Forseth) A statewide civil liberties organization: ACLU of Colorado (Anaya Robinson) An organization that advocates on behalf of individuals with disabilities: American Association of People with Disabilities (Henry Claypool) A national nonprofit, nonpartisan organization that focuses on technology, policy, and civil rights: Center for Democracy & Technology (Matt Scherer) A nonprofit organization that is focused on privacy and that has a membership consisting of developers and deployers: Future of Privacy Forum (Tatiana Rice) An organization that advocates for consumers and the First Amendment: Consumer Reports (Grace Gedye)
	Year founded / Jurisdiction	Bill Number (Title) / Established Body / Purpose	Duration of Body	N of Members / Civil Society Representatives	Civil Society Representatives
4	2024 / Washington	-ESSSB 5838 (Artificial Intelligence Task Force) -Task Force -Assess AI trends and recommend legislation	2024–2027 (expires June 30, 2027)	19 / 4	Three representatives from community advocate organizations representing communities disproportionately vulnerable to being harmed by algorithmic bias: buildJUSTLY (Vicky Tamaru); WA Build Back Black Alliance (Paula Sardinas); ACLU-Washington (Dr. Tee Sannon) One representative of statewide labor organizations: Washington State Labor Council, AFL-CIO (Cherika Carter)
5	2025 / Hawaii	-SB 742 (Data Sharing) -Working Group -Improve data sharing and AI guidelines	July 1, 2025– June 30, 2027	18 / 2 (1 of them found from the source)	Members representing nonprofit organizations having experience in data sharing and governance: HawaiiKidsCAN (David Sun-Miyashiro)

6	2025 / Maryland	-HB 956 (AI Implementation Workgroup) - Workgroup -Monitor AI implementation and recommend regulation	July 1, 2025– June 30, 2029	33 / 8 (2 of them appointed)	Under the statute, 8 seats are designated for civil society representatives from organizations focused on civil rights and liberties, consumer protection, data privacy, and labor; however, according to the 2026 report, only 2 nonprofit civil society representatives had been appointed: Ivy A. Finkenstadt of House of Ruth Maryland and Ashley J. Woolard of the Public Justice Center.
7	2026 / Pennsylvania	-SR 244 (Child Protection in the Digital Age) -Task Force -Address AI-related child protection risks	Recurring body with no specified final expiration date.	17 / 2	Director (or designee) of a rape crisis center - 1 Representative of a children's advocacy center - 1 Not yet publicly available
8	2026 / Alabama	-HJR 51 (AI and Children's Internet Safety) -Study Commission -Study AI and children's online safety	May 1, 2026– November 1, 2026	14 / 1	One member appointed by Voices for Alabama's Children - Not yet publicly available

APPENDIX B. EXAMPLES OF CIVIL SOCIETY ORGANIZATIONS WORKING ON ARTIFICIAL INTELLIGENCE. Source: Namsrai (2026).

	Name of Organization	Website	Geographic scope	Year founded / AI engagement since	Core Mandate (Expanded or AI specific)	Target issues/area	Primary activities	Funding
1	Center for AI and Digital Policy	https://www.caidp.org/	Global	2020	AI-specific	AI governance, democratic values, human rights, rule of law, assessment of national AI strategies and regulations	AI and Democratic Values Index; policy research; advising governments and international organizations; AI policy training/clinics; advocacy around AI Bill of Rights and global AI norms	Funded by philanthropic foundations and individual donations
2	The Future Society	https://thefuturesociety.org/	Global	2014	AI-specific	AI safety, security, geopolitics; trustworthy AI; risk management	Policy research and recommendations; convening roundtables and summits (e.g., US-China AI dialogue, Dubai Global AI Governance Roundtable); advising governments (EU AI Office, France AI Summit); standards development; expert consultations	Grants, donations
3	Institute for AI Policy and Strategy	https://www.iaps.ai/	US	2020	AI-specific	AI frontier security; compute policy; international strategy (US-China competition, global coordination)	Policy research/papers; policy talent development (AI Policy Fellowship); stakeholder engagement (Congress, executive, industry, civil society); expert convening	Sponsored by Rethink Priorities, a US-based 501(c)(3); Philanthropic grants

4	Partnership on AI	https://partnershiponai.org/	Global	2016	AI-specific	Responsible AI practices; safety, accountability, transparency; AI incidents; shared prosperity/economy; media integrity; justice/equity	Research/reports (e.g., AI Incident Database, Safety Critical AI); guidelines/standards (e.g., Safe Foundation Model Deployment); convening forums/steering committees; policy advocacy	Philanthropic foundations; corporate contributions
5	American Civil Liberties Union	https://www.aclu.org/	US	1920/2018	Expanded: core civil liberties/human rights defense; AI work is part of broader privacy, racial justice, criminal justice	AI surveillance; algorithmic discrimination/bias in criminal justice, employment, housing, credit; civil rights protections against automated decision-making; government	Litigation (e.g., ACLU v. Clearview AI); policy advocacy (moratoria on government facial recognition); legislative testimony; coalition campaigns; public education/reports	Membership dues; donations; foundation grants; fundraising campaigns
6	Center for Civil Rights and Technology	https://civilrights.org/	US	1950/2023	Expanded: coalition of 240+ civil rights organizations defending civil/human rights;	AI bias/discrimination; data privacy and algorithmic surveillance; industry accountability for tech harms; broadband equity; civil rights in voting, employment, housing, healthcare	Policy advocacy/letters; convenings (e.g., Regulatory Code: AI, Civil Rights, and Democracy); research/reports (e.g., Innovation Framework: Civil Rights Approach to AI); coalition building; public education	Membership dues; major donors/foundations; fundraising
7	Algorithmic Justice League	https://www.ajl.org/	US	2016	AI-specific	Algorithmic bias; facial recognition and biometric surveillance; automated discrimination; civil	Audits and research on AI systems; policy and legislative advocacy; expert testimony; public campaigning and education (including via film and media)	Grants and philanthropy via nonprofit sponsors; individual donations

						and human rights impacts of AI		
8	Electronic Privacy Information Center	https://epic.org/	US	1994/2018	Expanded: defending privacy rights; AI work within surveillance/privacy portfolio	AI policy/human rights; facial recognition/biometrics; government AI use; algorithmic transparency/accountability; data minimization	Policy advocacy/comments; litigation; research/reports; coalition work; public education	Membership donations; foundation grants; individual contributions
9	AI Now Institute	https://ainowinstitute.org/	US	2017	AI-specific	AI accountability; public sector AI governance; bias/discrimination; power concentration in AI; national security and geopolitics in AI; labor/work impacts	Policy research/reports (e.g., annual AI reports, AI Lexicon); advocacy for public interest AI; narrative development; coalition partnerships; public commentary	Philanthropic foundations; no corporate/tech industry funding since 2022
10	AI4ALL	https://ai4all.org/	US	2017	AI-specific	AI diversity/inclusion; underrepresented groups (women, BIPOC, rural, low-income); responsible AI education; AI literacy/ethics	Summer programs/academies for high school/undergrads; AI4ALL Ignite accelerator; Open Learning curriculum for teachers; storytelling initiatives	Philanthropic grants (Google.org \$2M, Patrick J. McGovern \$300K, Pivotal Ventures, Salesforce, GM)

11	Civic AI Security Program	https://civai.org/	US	2024	AI-specific	Public awareness of AI risks; AI cybersecurity threats (phishing, deepfakes, voice cloning, vulnerabilities); ethical/social/political implications of AI	Educating public/decision-makers about AI capabilities and dangers through interactive software demonstrations (not research papers); Live demos (Deepfake Sandbox, AI Values explorer, cybersecurity threats); briefings for decision-makers/civil society; scalable public educational tools	Grant-based (specific donors not public)
12	Data & Society	https://datasociety.net/	US + global	2014	Expanded mandate (technology and society research institute with strong focus on AI and algorithmic governance)	Civil liberties and civil rights impacts of data-driven decision-making; data governance; algorithmic accountability; platform governance;	Independent research; policy reports; convenings and workshops; expert advisory roles; public scholarship; fellowships = knowledge production and policy advisory	Philanthropic foundations; research grants; donations
13	Human Rights Watch	https://www.hrw.org/	Global	1978/2020s	Expanded: investigates/protects human rights worldwide; AI work within technology & rights portfolio	AI-enabled surveillance; facial recognition; predictive policing; autonomous weapons systems; migration technologies; discrimination and rights violations linked to automated systems	Investigative reports; policy advocacy; policy recommendations; media engagement; joint statements; UN+ engagement; litigation support	Individual donations; foundation grants; no government/corporate funding
14	Electronic Frontier Foundation	https://www.eff.org/	US	1990/2020s	Expanded: digital civil liberties (privacy, free speech, surveillance); AI work within	Facial recognition; biometric surveillance; predictive policing; algorithmic transparency; automated decision-making in public	Strategic litigation; policy advocacy; legislative testimony; public campaigns; investigative research; technical analysis	Individual donations; membership contributions; philanthropic foundations; grants

					surveillance/privacy portfolio	services; privacy and free expression		
15	World Privacy Forum	https://worldprivacyforum.org/	Global	2003/2020s	Expanded: privacy and data protection research center; AI work via Trustworthy AI governance project	AI governance frameworks; trustworthy AI implementation; facial recognition/privacy; data protection laws and AI intersection	Policy research/reports (e.g., Global AI Governance Index); submissions to governments and international bodies (e.g., India AI Summit 2026); UNESCO/GPAI engagement; privacy impact assessments; testimony; expert consultation; advocacy	Nonprofit; foundation grants and donations (specifics not public)
16	WITNESS	https://www.witness.org/	Global	1992/2020s	Expanded: human rights and media justice organization	AI-generated content/deepfakes; audiovisual evidence integrity; video verification; surveillance vs. human rights documentation	Training activists/journalists on safe/ethical video use; AI Impact Summit advocacy; tools for video verification; policy on AI and human rights evidence	Nonprofit; philanthropic foundations, grants, donations
17	DataEthics 4All	https://dataethics4all.org/	US	2020	AI-specific	AI ethics; responsible AI; data literacy; algorithmic accountability; workforce development; inclusion in AI	Education and training programs; curriculum development; youth engagement; partnerships with universities and industry; awareness campaigns; community-building	Grants; partnerships; sponsorships; nonprofit funding
18	Youth for Privacy	https://www.youthforprivacy.org/	US	2022	Expanded: youth digital rights and privacy	Youth digital privacy; data protection in AI systems; cybersecurity; ethical AI; children's online safety	Workshops/bootcamps (e.g., Data Values Boot Camp); UN engagement (Open-ended Working Group on ICTs, #Leaders4Tomorrow); fellowships; advocacy campaigns; public education	Fiscally sponsored by Aspiration

19	Oakland Privacy	https://oaklandprivacy.org/	US	2013	Expanded: Originally focused on privacy, surveillance oversight, and civil liberties at the municipal level.	Surveillance technologies; AI-based policing tools; Community oversight and impact-based review of surveillance technologies, including surveillance impact reports and use-policy requirements	Advocating municipal surveillance regulation (e.g., SERO); Monitoring government compliance (SIRs, SUPs); Filing public records requests on surveillance and AI; Promoting transparency and participating in local oversight	Primarily foundation grants and donations (exact funding streams not always publicly itemized in detail)
20	Surveillance Technology Oversight Project	https://www.stopspying.org/	US	2019	Expanded: Civil rights organization focused on challenging government surveillance and advocating for privacy protections.	NYPD and local government surveillance technologies; AI-based policing tools, including facial recognition and other biometric-based identification systems; Discriminatory impact of surveillance on communities of color, immigrants, and protesters.	Litigation to obtain NYPD surveillance records (including facial recognition, drones, and license-plate readers); Advocacy for the POST Act and stronger civilian oversight of surveillance; Research and reports on AI-enabled surveillance abuses; Public campaigns (e.g., #BanTheScan, #PowerDownSurveillance) to limit or ban high-risk AI surveillance.	Hosted by the Urban Justice Center and funded through grants, individual donations, and pro-bono legal partnerships.
21	All Tech Is Human	https://alltechishuman.org/	US	2017	Expanded: Dedicated to building the responsible technology ecosystem. AI is addressed within a broader “Responsible Tech” framework	Responsible and ethical AI (bias, fairness, accountability, transparency); Public awareness and discourse around AI’s role in society (surveillance, labor, democracy, education); Norm-setting and cross-sector collaboration on AI governance and practice.	Hosting conferences and events on AI ethics and responsible technology; Publishing ecosystem mappings to connect responsible tech practitioners; Organizing public panels and workshops on AI risks and governance.	Sponsorships, partnerships, philanthropy, and event-based funding (corporate participation in responsible tech ecosystem)

22	Communities United for Police Reform	https://www.changethenypd.org/	US	2012	Expanded: increasingly engages with AI-enabled surveillance and data-driven policing as part of its broader accountability agenda.	Discriminatory and abusive policing practices; NYPD surveillance and AI-enabled data-driven policing; Civilian oversight and transparency in policing	Coalition-based advocacy for policing reform; Know-your-rights education and grassroots mobilization; Engagement with oversight and surveillance transparency reforms	Philanthropic foundations, nonprofit grants, and member organization support
23	Center for Democracy & Technology (CDT)	https://cdt.org/	US + global	1994/2010s	Expanded: focused on digital rights, free expression, privacy, and security; increasingly AI-specific	AI fairness, risk, assurance, and safety; AI in public-sector decision-making, open-source models, and personalization; Privacy, surveillance, and democratic-process-related uses of AI (e.g., elections, government procurement of AI).	Policy advocacy (federal, state, and EU-level, including AI-regulation processes and AI-safety frameworks); Research and reports through the AI Governance Lab (e.g., AI assurance, AI-safety training, AI-personalization, AI-documentation); Technical-policy guidance and recommendations for governments and companies using AI in public-sector services.	Foundation grants, philanthropic funding, and corporate support

24	AI LA	https://www.joinai.la/	US	2016	AI-specific	AI education and literacy for professionals, students, and the public; Responsible AI, ethics, and the societal impact of AI (via the Responsible AI Reading Group and related programs); Inclusivity and diversity in AI / STEAM, including programs for underrepresented groups.	Hosting in-person and online AI-focused events (panels, workshops, hackathons, summits, reading groups); Running AI education and talent programs such as IDEAS™ to train underrepresented groups and link them with industry; Producing newsletters, online communities (Discord), and membership tiers to support AI-literate professionals and the LA tech ecosystem.	Supported through sponsorships, partnerships, event funding, and donations
25	Access Now	https://www.accessnow.org/	Global	2009/2018s	Expanded: digital rights and human rights in the digital age	Digital rights; AI governance; internet shutdowns; surveillance; data protection; freedom of expression; cybersecurity; civic space	Policy advocacy; technical support (Digital Security Helpline); research and reporting; global convening (RightsCon); grantmaking to grassroots organizations; coalition-building (e.g., #KeepItOn)	Foundations and development agencies; also supported by companies, individuals, courts, and civil society organizations
26	Global Center on AI Governance	https://www.globalcenter.ai/	Global (Africa-based)	2021	AI-specific	AI governance; responsible and equitable AI; global inequality; AI for development; human rights and inclusion; Global South representation	Policy research; global comparative analysis (e.g., Global Index on Responsible AI); advisory to governments and multilateral actors; capacity building and training; convening networks and partnerships; advocacy for inclusive AI governance	Funded by international development agencies and philanthropic partners (e.g., IDRC, USAID, FCDO); project-based grants

27	Al Sur	https://www.alsur.la/	Regional (Latin America)	2018 / 2020s	Expanded: digital rights and human rights in the digital environment	Digital rights; AI governance; surveillance; data protection; freedom of expression; cybersecurity; internet governance; inequality in Global South contexts	Regional policy research and reports; advocacy in international forums; coalition-building across CSOs and academia; capacity building and knowledge-sharing; participation in global governance processes; convening and joint submissions	Primarily funded through project-based grants, international donors, and partner organizations (not fully publicly specified)
28	Algorithm Watch	https://algorithmwatch.org/en/	Regional (Europe)	2017	AI-specific	Algorithmic decision-making; AI governance; transparency and accountability; fundamental rights; discrimination and bias; platform governance; power asymmetries	Research and analysis of algorithmic systems; policy advocacy; public communication and education; investigative research; development of policy recommendations; convening expert networks; campaigns on algorithmic accountability	Funded primarily by philanthropic foundations and individual donations
29	Digital Rights Foundation	https://digitalrightsfoundation.pk/	National /Pakistan	2012/ 2020s	Expanded: digital rights, human rights, and digital governance	Digital rights; online harassment (especially gender-based); data privacy; surveillance; freedom of expression; digital security; AI governance (emerging); civic space	Policy advocacy; research and reports; digital security trainings; cyber harassment helpline (support services); public awareness campaigns; coalition engagement	International grants, development agencies, and philanthropic support; project-based funding

30	AI and You	https://aiandyou.org/	US	2019	AI-specific	AI literacy; public understanding of AI; inclusion of marginalized communities; workforce and education impacts of AI; digital inequality	AI literacy training and educational content; public awareness campaigns; community events and summits; policy education and engagement; partnerships with civil society, government, and industry	Primarily funded through philanthropic grants, foundations, and nonprofit partners
31	Linden AI Institute	https://www.lindenaiinstitute.org/	US	2025	AI-specific	AI literacy; digital inclusion; education and workforce readiness; ethical AI awareness; reducing digital inequality	AI literacy workshops (K–12 educators, parents, communities); curriculum and resource development; community outreach programs; public awareness and training on ethical AI; partnerships with schools and civic organizations	Not publicly specified; likely through nonprofit partnerships, community programs, and philanthropic funding