Digital Society (2025) 4:32
https://doi.org/10.1007/s44206-025-00176-9

ORIGINAL PAPER

Perceptions of AI Ethics Policies Among Scientists and Engineers in Policy-Related Roles: An Exploratory Investigation

James Weichert[1] · **Qin Zhu**[2] · **Dayoung Kim**[2] · **Hoda Eldardiry**[1]

Received: 23 April 2024 / Accepted: 19 February 2025 / Published online: 24 April 2025

Abstract

The explosive growth of artificial intelligence (AI) over the past few years has focused attention on how diverse stakeholders regulate these technologies to ensure their safe and ethical use. Increasingly, governmental bodies, corporations, and nonprofit organizations are developing strategies and policies for AI governance. While existing literature on ethical AI has focused on the various principles and guidelines that have emerged as a result of these efforts, just how these principles are operationalized and translated to broader policy is still the subject of current research. Specifically, there is a gap in our understanding of how policy practitioners actively engage with, contextualize, or reflect on existing AI ethics policies in their daily professional activities. The perspectives of these policy experts towards AI regulation generally are not fully understood. To this end, this paper explores the perceptions of scientists and engineers in policy-related roles in the US public and nonprofit sectors towards AI ethics policy, both in the US and abroad. We interviewed 15 policy experts and found that although these experts were generally familiar with AI governance efforts within their domains, overall knowledge of guiding frameworks and critical regulatory policies was still limited. There was also a general perception among the experts we interviewed that the US lagged behind other comparable countries in regulating AI, a finding that supports the conclusion of existing literature. Lastly, we conducted a preliminary comparison between the AI ethics policies identified by the policy experts in our study and those emphasized in existing literature, identifying both commonalities and areas of divergence.

Keywords Artificial intelligence · AI ethics · AI governance · EU AI act · Blueprint for an AI bill of rights

✉ Hoda Eldardiry
hdardiry@vt.edu

[1] Department of Computer Science, Virginia Tech, Blacksburg, VA, USA

[2] Department of Engineering Education, Virginia Tech, Blacksburg, VA, USA

1 Introduction

In the global context, there has been an overwhelming amount of 'hype' surrounding the transformative potential of artificial intelligence (AI) across nearly every facet of human existence. Anticipating the significant opportunities that AI can bring to the global society, governments, corporations, and organizations have been actively developing ethical principles, frameworks, standards, and policies to govern the social and ethical implications of AI-enabled technologies (Jobin et al., 2019). Compared to other normative mechanisms—particularly ethical principles and frameworks—ethics *policies* wield broader impacts and are more practical to implement in guiding AI research and innovation towards socially desirable outcomes (Coeckelbergh, 2020; Floridi, 2023). Despite the emergence of AI policies across sectors, especially those aimed at ensuring that AI systems adhere to ethical norms and values (referred to as AI ethics policies), the implementation of these policies still faces certain challenges. Most notably, there is a lack of empirical evidence or experimentation to thoroughly assess the impact and effectiveness of AI ethics policies (Gianni et al., 2022). Furthermore, the formulation and implementation of AI ethics policies rarely incorporate the values and perspectives of diverse stakeholders, particularly through the inclusion of the public in democratic processes (Gianni et al., 2022). Moreover, researchers lack a systematic framework or typology to categorize these policies, particularly considering their complexity and diverse origins from various sectors and cultures (Corrêa et al., 2023). Most AI ethics policies—including discussions around these policies—were initiated by scholars across engineering, science, social studies and humanities disciplines (D. Schiff et al., 2020). Very few studies have focused on how AI policy practitioners, individuals who engage with these policies on a daily basis, perceive AI ethics policies. Assessing the effectiveness of AI ethics policies depends not only on their development being aligned with rigorous processes but also on their actual functionality, which can be partially measured by the experiences of policy practitioners who engage with these policies on a daily basis as part of their work. Therefore, we argue that understanding *how* policy practitioners actively engage with, contextualize, and deliberate on AI ethics policies is instrumental in assessing the effectiveness of these policies in governing AI research, innovation, and use. Such understanding can also aid in optimizing the implementation and operationalization of these policies.

To address this gap in the literature, this paper investigates the perspectives of a specific group of AI policy practitioners—individuals trained as scientists or engineers who have transitioned into policy-related AI roles—regarding their perceptions of AI ethics policies. Exploring the experiences of this particular group of people has unique values. Most importantly, given their interdisciplinary background, they have knowledge of both the technical foundation undergirding AI ethics policies and the practical and political processes through which these policies are formulated and implemented.

More specifically, this paper addresses the following two research questions: (1) What are the major categories of AI ethics policies, as identified by AI professionals who were trained as scientists and engineers but transitioned to policy-related roles?

(2) What are commonly shared perceptions of these experts toward the development and implementation of these policies?

In answering these questions, this paper provides the following key contributions to AI ethics policy literature:

1. We develop a simple yet comprehensive taxonomy for clustering AI ethics policy efforts according to originating organization, jurisdiction, and development stage.
2. We conduct qualitative interviews with experts in AI policy-related roles and analyze their commonly shared perceptions of AI ethics policies using our clustering taxonomy.
3. We compare the policy efforts identified through our interviews with experts, and compare these with the policies broadly discussed in the literature.
4. We evaluate the expert responses within the current context of AI ethics policy development and discuss the implications of our findings for future AI ethics policy.

2 Literature Review

This section reviews literature on the following topics: (1) the *motivations* and *processes* for developing AI ethics policies, including the involvement of key stakeholders; and (2) AI ethics policies in the United States, European Union, and China.

2.1 Motivations and Processes for AI Policymaking

The process of developing strategies for AI governance is tightly linked to the economic and political contexts of the strategy's originator. For a state, this entails navigating the intricacies of its political and administrative apparatus. Thus, there is no uniform roadmap for how policies transform from ideas to operationalizable regulations. Nevertheless, the work of Schiff et al. (2020) is helpful for understanding common factors that may motivate and influence the development of AI ethics principles, strategies, and policies. The authors review 88 ethics documents from across the public, private, and nongovernmental (NGO) sectors published between 2016 and 2019 to identify overlapping themes, furthering understanding about *who* produces AI ethics documents and *why* they are produced. These documents, as Schiff et al. note, originate predominantly from wealthy countries, chief among them the US and China. Yet amid competition between the US and China for AI dominance, smaller nations may look to their AI strategy to differentiate themselves from competing states for investment in AI innovation. Finally, Schiff et al. identify five motivations for AI ethics strategy/policy development beyond social responsibility: (1) creating *competitive advantage* by signaling openness to private investment in AI under a favorable regulatory regime; (2) internal *strategic planning*; (3) *strategic intervention* by corporations to preempt government regulation; (4) *signaling social responsibility* to improve one's reputation; and (5) *signaling leadership* on the international stage, so as to be seen as a major 'player' with respect to AI regulation.

Below, we examine in detail the policymaking and political landscapes surrounding AI regulation for three state-level actors: the US, the European Union (EU), and the People's Republic of China (hereafter, "China"). We focus on these two states and the EU because they represent the largest governmental actors concerning AI policymaking and regulation (Roberts et al., 2021). We stress, however, the importance of broadening the scope of AI policy analysis in future work, with a particular focus on states and organizations in the Global South. As Schiff et al. (2020) caution, the reinforcement of the current hegemony of wealthy Western states in AI ethics discourse risks exacerbating existing global inequalities, contradicting the goal of AI ethics to 'level the playing field.'

2.1.1 United States

From a policymaking perspective, the US stands out because of the diffusion of governance power in two key ways: (1) the separation of powers across the executive, legislative, and judicial branches of government, which plays an important role in federal policymaking (Rose-Ackerman, 2022); and (2) the devolution of policy-making power to the states (Krane, 1993). With respect to the latter, the prerogative of the federal government to regulate AI is clear given the technology's impact on interstate commerce. However, the federal government has yet to fully exercise that power, paving the way for states to set their own regulations while a "dominant set of ideas" about how to regulate AI remains missing from the national discourse (Parinandi et al., 2024). The lack of federal oversight has allowed states to act as "laboratories of democracy," giving researchers a sense of what AI regulations may soon be viable on a national scale. Parinandi et al. (2024) analyze state-level AI legislation from 2018 to 2022 and find partisan and bipartisan trends relating to the adoption of AI legislation. For example, the study found that Republican legislators are significantly less likely to vote for AI legislation centered around consumer protection compared to Democratic legislators, and that states are generally more likely to see AI regulation enacted if governed by Democrats. The likelihood of success for federal legislation may therefore depend heavily on whether that legislation is consumer-protection-focused or business-focused.

On the national level, both Congress and the executive under the president have the prerogative to formulate government 'policy' on how to govern the development and use of AI technology. 'Policy' here is loosely construed since it can encompass both executive branch policy preferences related to how laws are enforced and legislation from Congress. Ultimately, effective AI regulation must have both a legislative mandate from Congress and an executive branch willing to enforce the legislation.

That Congressional mandate is currently lacking, and clear policy preferences from the Biden administration are only beginning to take shape. As such, policy statements from the administration are beneficial in deciphering the direction of future legislative and regulatory efforts. With respect to standards-setting, the US should be expected to lag behind other comparable nations while federal standards agencies wait for a clear set of ethical principles and an industry consensus on standards to form. As the National Institute of Standards and Technology (NIST) notes in a plan prepared following President Trump's 2019 Executive Order 13859 titled "Main-

taining American Leadership in AI," current US law requires government standards bodies like NIST to rely on a consensus-based approach to their work (Tabassi et al., 2019). In other words, "standards flow from principles," which must be well-established in the private sector before being adopted in federal regulation (Tabassi et al., 2019). Despite robust public engagement requirements for federal rulemaking, it is unclear to what extent non-industry actors—e.g. consumers and the American public—can influence this regulatory process. The lack of action from Congress, then, means a limiting of the opportunities for representational democracy to influence AI policy.

Finally, it is worth acknowledging that the executive branch itself cannot be treated as a single unchanging entity in discussions surrounding AI policymaking (Hine & Floridi, 2022). The myriad agencies charged with implementing broad executive policy directives across diverse yet often overlapping domains create a complex fabric of federal regulation that is difficult for policy experts—let alone non-experts—to navigate. But perhaps more importantly, the federal bureaucracy is subject to shifts in direction more stark and more frequent than in comparable countries each time the occupancy of the White House changes hands. Hine and Floridi (2022) analyze AI policy differences across the Obama, Trump, and Biden administrations, and document the evolution of US AI policy from a more laissez-faire free market attitude under Obama and Trump to a focus on human rights and freedoms under Biden.

2.1.2 European Union

The European Union, a political and economic union of 27 member states, is another unique governmental organization with respect to AI governance, since EU laws and regulations coexist with 27 sets of national laws and regulations. Nevertheless, it appears that the political organs of the EU—especially the European Commission and the Parliament—are the driving forces behind AI policymaking, with the EU AI Act establishing a common approach to AI ethics policy across the Union (Roberts et al., 2021). This effort began in 2016 with a European Parliament report on governing autonomous robots and quickly merged into a Commission effort to coordinate a path forward regarding AI governance. In 2018, the Commission created a High-Level Expert Group on AI (HLEG), and in 2021 published the first draft of the EU AI Act, which was approved by the member states at the end of 2023 and passed by the Parliament in March 2024. In general, the EU's policy-making process on AI can be described as a heavily technocratic one, with the HLEG and particularly the European Commission playing key roles in shaping AI policy (Roberts et al., 2022). While the AI Act received revisions and required approval from the European Parliament and the governments of the member states, the Commission is the legislation's author and proponent. As with all EU law, the Commission has the exclusive right to initiate the legislative process; the Parliament cannot draft legislation itself (Ponzano et al., 2012). Roberts et al. (2021) note that this undemocratic approach may be cause for concern. It is also worth noting that the entire process of developing and passing the AI Act occurred after the 2019 EU Parliament elections and before the summer 2024 elections.

2.1.3 China

Processes in China for developing and implementing AI policies are heavily dependent on the guiding ideological hand of the Chinese Communist Party (CCP) under a governance model of "fragmented authoritarianism," whereby the national government establishes ambitious national goals which local governments strive to implement (Roberts et al., 2022). In particular, the 2017 "A New Generation AI Development Plan" (AIDP) created by the State Council acts as a "wish list" for the CCP, leaving local authorities to determine specific implementation details (Hine & Floridi, 2022). Local politicians are motivated by short term limits and the prospect of promotions within the Party to align their locality with the national goals and to do so in a short period (Roberts et al., 2022). However, these incentives may encourage local governments to inflate their economic targets concerning AI innovation or to fail to diversify the types of AI businesses they attract (Hine & Floridi, 2022). The expiration of the AIDP in 2020 heralded a slight shift in national policy, as AI-related policy goals were integrated into the larger science and technology section of the 2021 Five-Year Plan (Hine & Floridi, 2022). In this sense, AI is increasingly seen as one of many technologies that will spur economic growth over the coming years. Simultaneously, the AIDP anticipates that by 2025, China will transition from a more exploratory stance on AI to establishing concrete regulations before becoming the "world's primary AI innovation centre" by 2030 (Hine & Floridi, 2022).

2.2 Existing AI Policies and Legislation

Below, we identify and briefly elaborate on major existing guidance, policies, and legislation governing the use of artificial intelligence in the US, EU, and China. These policy efforts are, for the most part, the ones broadly identified in existing academic literature. In Sect. 5, we examine the extent to which the policy experts we interviewed also highlighted the importance of these policies.

2.2.1 United States

1. *Blueprint for an AI Bill of Rights*

The "Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People" (hereafter, "Blueprint") is a white paper released by the White House Office of Science and Technology Policy (OSTP) in October 2022. The Blueprint outlines major principles to "guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence," (Office of Science and Technology Policy, 2022). These principles are: *safe and effective systems, algorithmic discrimination protections, data privacy, notice and explanation*, and *human alternatives, consideration, and fallback*. As a non-binding "framework," the Blueprint is unenforceable. Nevertheless, it offers a first look at the Biden administration's outlook on AI regulation, and its influence is seen in Executive Order 14110 signed a year later. In their evaluation of the Blueprint, Hine and

Floridi (2023) remark on the focus of Blueprint on "automated systems" as opposed to AI specifically, which they argue is too broad.

2. *NIST AI Risk Management Framework*

The National Institute of Standards and Technology (NIST) released its "Artificial Intelligence Risk Management Framework" (RMF) in January 2023 following a year of public input and revisions (National Institute of Standards and Technology, 2023). The RMF is the result of the 2020 National Artificial Intelligence Initiative Act, which directed NIST to develop "voluntary standards for artificial intelligence systems," (National Institute of Standards and Technology, 2023). In this regard, the Framework is nonbinding, "rights-preserving, non-sector-specific, and use-case agnostic," (National Institute of Standards and Technology, 2023). The RMF organizes these standards into four functions—*Govern, Map, Measure*, and *Manage*—each with a list of (vague) recommendations.

3. *Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*

On October 30, 2023, President Biden signed Executive Order 14110: *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* ("E.O."). While all of our interviews with policy experts were conducted before October 2023, we include an overview of the E.O. here because of the Order's important emerging role in US AI governance. We also return to the E.O. later in this paper in discussing how the policy gaps identified by our interviews may be ameliorated by the introduction of the Order. The E.O. expands on both the Blueprint and RMF by directing specific actions to be taken by a wide range of federal agencies across eight policy areas: *safety and security, innovation and competition, worker support, consideration of AI bias and civil rights, consumer protection, privacy, federal use of AI*, and *international leadership* (Harris & Jaikaran, 2023). Some directed actions include the formation of a White House AI Council, directing the NSF Director to "support AI-related education and AI-related workforce development" and convening the Department of Justice and other agencies to discuss the "comprehensive use of their respective authorities and offices" to prevent AI-related discrimination (Exec. Order No. 14110, 2023).

2.2.2 European Union

1. *General Data Protection Regulation*

The European Union's General Data Protection Regulation (GDPR) came into effect in 2018 and because of its strong enforcement authorities has become "the most consequential regulatory development in information policy in a generation," (Hoofnagle et al., 2019). Most significant are the GDPR's restrictions against the sharing of personal data without valid consent (Roberts et al., 2022). The impact of this component of the GDPR is most commonly seen by users when being prompted to accept

or decline the use of cookies on websites compliant with the GDPR. Nevertheless, the Regulation does allow for some exceptions to the restriction on the sharing of sensitive data, particularly when such action is a "substantial public interest" such as public health (Hoofnagle et al., 2019).

2. *EU AI Act*

The EU Artificial Intelligence Act ("AI Act") was first proposed by the European Commission in April 2021, ratified by EU member states in December 2023, and passed by the European Parliament in March 2024 (Madiega, 2024). The legislation encompasses all "AI systems" used in the EU, a term which in the original Commission draft included common "statistical" approaches but has since been narrowed in scope (Council of the European Union, 2022; Madiega, 2024). A key feature of the Act is the categorization of AI practices into four risk levels, including *high risk* applications which are subject to stricter regulation, and *unacceptable risk* applications which are prohibited. The latter category encompasses subliminal manipulation, social scoring regimes and "real-time" biometric surveillance except in some law enforcement use cases (Veale & Zuiderveen Borgesius, 2021). The legislation also contains an obligation for developers of "General-purpose AI" models to disclose the data used to train their models and to implement a policy to mitigate copyright infringement (Madiega, 2024).

2.2.3 China

1. *AI Development Plan*

"A New Generation AI Development Plan" (AIDP) is China's earliest roadmap for AI policy (Hine & Floridi, 2022). The Plan was published by the State Council and establishes national goals for investment in AI-related technologies. The AIDP is primarily an economic document that aims to position AI as the "main driving force behind China's industrial upgrading and economic transformation," (Roberts et al., 2022). For example, the State Council set a target of the 'core' AI industry achieving a gross economic output of RMB 150 billion (USD 22.5 billion) by 2020 and RMB 400 billion (USD 60.3 billion) by 2025 (Ding, 2018). While the national government sets out these ambitions, the actualization of these targets is left to local governments. The drawback of this "fragmented authoritarianism" approach can be seen in the fact that a 2019 estimate valued the core industry at only RMB 57 billion (Hine & Floridi, 2022). Nevertheless, the AIDP does lay out some core principles for the Chinese government with respect to ethical AI development, particularly concerning "preserving social stability," (Roberts et al., 2022).

2. *Personal Information Protection Law*

The Personal Information Protection Law (PIPL) was passed by the Standing Committee of the National People's Congress (the Chinese national legislative body) in 2021 and mirrors in many respects the data privacy protections of the EU's GDPR, including requirements for obtaining informed consent (Roberts, 2022; Roberts et al.,

2022). However, there is a key exception to privacy protections when they "impede state organs' fulfillment of their statutory duties and responsibility," (Roberts, 2022). Roberts et al. (2022), in comparing the PIPL with the GDPR, caution that the introduction of invasive surveillance technologies, particularly among the Uyghur minority in Xinjiang, "contradicts a holistic right to privacy and displays the CCP's greater willingness to exploit AI technology in the name of national security than the EU," (Roberts et al., 2022).

3 Methodology

3.1 Data Collection

This paper contributes to a broader study investigating perspectives on AI ethics and policy among professionals initially trained as scientists or engineers, who have transitioned into policy-related roles. The study also explores their career trajectories, shedding light on the intersection of technical expertise and policy engagement within the AI field.

For the larger project, we conducted semi-structured interviews with 15 scientists and engineers (Zhu et al., 2023). We recruited participants based on two specific criteria: (1) possessing at least one STEM degree, and (2) actively engaging in AI policy-related work (broadly construed) as part of their daily responsibilities. We define engagement in AI policy-related work as participation in job tasks that require the consideration of technical AI systems in the context of legal, regulatory, or business requirements. Thus, purely technical development work without interaction with policy development or compliance does not qualify under this definition. First, we distributed a short survey to potential study participants to obtain their basic demographic and career information. The survey was distributed to the alumni of the *Christine Mirzayan Science and Technology Policy Fellowship* of the National Academies of Sciences, Engineering, and Medicine. Additional snowball sampling was also conducted. Except for the individuals who did not meet the recruitment criteria, we invited almost every individual who completed the survey to the interview. We conducted semi-structured interviews which lasted approximately 60 to 90 minutes each.

The interview protocol for the larger project consisted of two sections: 1) career pathways of AI professionals transitioning from scientists and engineers to policy practitioners; and 2) their perceptions of AI ethics policy. We asked questions related to career pathways because the overarching project also aims to establish a better understanding of the career pathways of AI professionals to become policy experts.

For this paper, we analyzed participant responses to a specific interview question in Part 2 of the protocol: "Are you aware of any ongoing policy efforts aimed at ensuring the ethical development and utilization of AI? If so, which groups are actively involved in these endeavors?"

3.2 Data Analysis

We audio-recorded the interviews and transcribed the audio files through an external transcription service approved by our university. After transcription, the data was de-identified, with each participant randomly assigned a pseudonym. After generating transcripts of each interview, we extracted participant responses to the specific interview question mentioned in Sect. 3.1 into a separate document. We started data analysis by conducting a priori coding aligned with the two research questions.

More specifically, in the initial round of coding, we systematically identified and highlighted any references to specific policy documents or regulatory efforts, general policy discussions mentioned by the participants, and any other notable or unique comments from the participants. In the second round of coding, we conducted thematic analysis (Braun & Clarke, 2021), leading to the development of a clustering taxonomy (described in Sect. 4.1) and a synthesis of participants' shared perceptions regarding AI ethics policies (described in Sect. 4.4). Additionally, all specific policy documents or efforts were combined to create Table 2. Finally, we consulted our coding notes, generated themes, and original interview transcripts in writing our analysis for Sects. 4 and 5.

3.3 Policy Identification

Due to the qualitative and open-ended nature of our research approach, not all references to specific policy documents or efforts were consistent across interviewees. Most document references had slight variations in name (e.g. "Blueprint for an AI Bill of Rights" or "the AI Bill of Rights") which were straightforward to reconcile. For these cases, we counted all references to the same document together and use the document's official title in this paper. We provide a complete hyperlinked list of these documents and the abbreviations we use in the appendix. In some cases, references to specific policy documents or efforts were not made using the title of the document (or a variation of the title). Rather, the interviewee described the policy effort in such a way as to uniquely identify it. In these cases, we used quotes and context from the interviewee to trace the reference to a named document. Once we trace a reference to a particular policy document or effort, we refer to the effort by the name of the document, if applicable.

For example, James' interview includes the following quote: "…we follow a variety of legislative proposals that have come on Capitol Hill. Just as an example—I forget the details, but it was in the news this week about regulation that would make it illegal for there to be algorithms directly used in the decision to launch nuclear weapons and thus kind of requiring people always be in the loop for that decision." An online search using the keywords "Congress AI nuclear weapons" reveals only one possible candidate: *S.1394: Block Nuclear Launch by Autonomous Artificial Intelligence Act of 2023* introduced by Senator Markey and Representatives Lieu, Beyer, and Buck. The content and date of publication of the legislation correspond with the context of James' interview answer, confirming that S. 1394 is the document to which James was referring.

The authors all agree on the identification of ambiguous policy references, and in all but one case the reference can only be to one document. The exception is Irene's response: "IBM has a framework, Microsoft has a really good framework and lots of those things, and some of them are made public." We traced the reference to a "Microsoft framework" to Microsoft's *Responsible AI Standard*, but an "IBM framework" on AI was more elusive. A search for "IBM AI Framework" does not yield a clear answer, but a search for "IBM AI principles" leads to IBM's *Principles for Trust and Transparency* page accessible from the IBM AI Ethics webpage. We believe that this is the closest match for Irene's reference, and have therefore included it (with a disclaimer) in Table 2.

4 Findings

4.1 Policy Documents Clustering

Our review of the interview responses resulted in the formulation of seven distinct clusters of AI policy. Below, we describe each cluster and its coding requirements. We believe that this clustering provides a simple yet comprehensive taxonomy for categorizing AI ethics policy efforts according to originating organization (e.g. public or private sector), jurisdiction (US, non-US national, and multinational or global), and development stage (i.e. general discussions or formulated policy). Figure 1 visualizes the seven clusters according to these distinctions.

1. ***US Policy Documents*** — This cluster includes any named reference to a specific policy document or regulation from the US executive branch. 'Policy document' here is broadly construed to include any document published by a government office or agency that outlines either nonbinding guidance or official government policy related to AI. To be included in this cluster, the document or regulatory effort must be (1) named by the interviewee, or described in enough detail to identify it; and (2) be traceable to a specific government

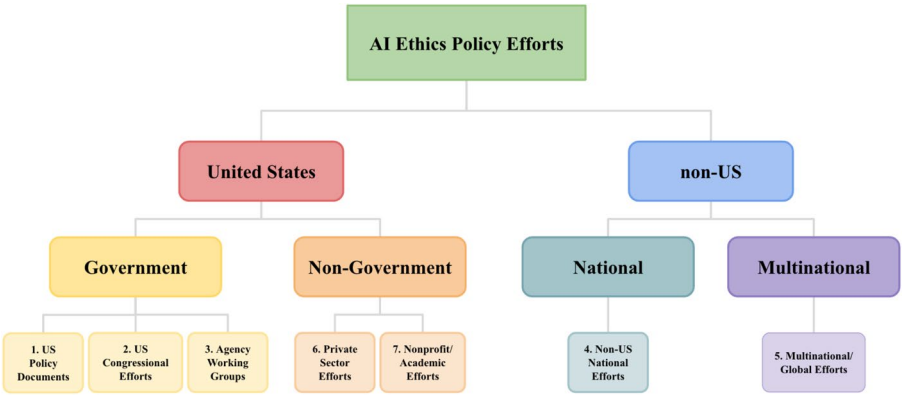

Fig. 1 Tree diagram of our clustering taxonomy according to jurisdiction and originating authority

entity that issued the document or regulation. Examples of mentioned US policy documents include the nonbinding *Blueprint for an AI Bill of Rights*, the USAID *Artificial Intelligence Action Plan*, and the NIST *AI Risk Management Framework*.

2. ***US Congressional Efforts*** — This includes any specific reference by an interviewee to an effort by any members of the US Congress to propose or pass legislation related to AI. The reference to the legislative effort must be (1) related to a specific proposal—either a draft bill or an outline for proposed actions to be taken by Congress—to govern an aspect of AI in the US, and (2) reasonably traceable to a member or group of members proposing the legislative action. Thus, statements such as "…there are lots of initiatives in Congress right now, at the White House, and all sorts of federal agencies…" (Daley) would not qualify for this cluster. The only interview to fall under this category was that of James, who mentioned "…regulation that would make it illegal for there to be algorithms directly used in the decision to launch nuclear weapons…". We traced this reference to S.1394, the *Block Nuclear Launch by Autonomous Artificial Intelligence Act of 2023* (see Sect. 3.3).
3. ***US Public Sector Agency Working Groups or General Conversations*** — References to general policy conversations or policy working groups taking place within US government agencies are included in this cluster, provided that (1) the reference does not refer to a particular policy document issued by the agency (cluster 1), and (2) the interviewee references a specific government agency or agencies where the policy conversations are taking place. Thus, a statement like "…there was a working group in kind of the AI space for a little while with a few of the agencies in D.C. that touched this area…" (Julia) would not qualify for this cluster as it cannot be traced to specific government agencies, but a statement such as "At State Department, there might be … 20 people, 25 people who are, on a day-to-day basis, involved in AI policy conversations," (Benjamin) would. By definition, this cluster includes no specific policy documents or efforts (see Table 2) since this group encompasses only informal and in-progress efforts by US government agencies.
4. ***Non-US National Guidance, Regulations and Legislative Efforts*** — Any reference to a legislative or regulatory effort by a country other than the US regarding AI falls under this cluster, which is the combination of clusters 1 and 2 for non-US countries. To be included in this category the reference must (1) relate to a specific country's legislation, regulation, or government plans to govern AI, and (2) be an action taken solely by a single government, not a multinational agreement or policy. References to Canada's *Artificial Intelligence and Data Act* and the UK's *AI Strategy* are included in this cluster, whereas references to the *EU AI Act* or the African Union's *Draft Strategy for AI* are included in cluster 5. Simply naming a country in reference to general AI regulation (e.g. "I think [in] most European countries, I think France, Canada, there are policies that are being developed in terms of what are the ethical uses for AI" [David]) does not qualify for this cluster.

   We combine both administrative policy and legislation into a single cluster for non-US countries for two reasons: (1) there were not enough references to policy

and legislation individually to justify two separate (and small) categories, and (2) we recognize that, outside of the US, there is perhaps less of a meaningful difference between administrative policy and legislative action. This is especially true for parliamentary democracies, where the executive government must command majority support from parliament, and for non-democratic countries such as China. We discuss the implications of the US separation of powers on AI regulation in detail in Sect. 5.

5. ***Multinational and Global Policy Efforts*** — This cluster includes specific references to multinational or global policy statements, strategies, treaties, or legislation, whether binding or nonbinding. 'Multinational' includes any group of two or more countries, including bilateral or multilateral statements, multinational governments such as the European Union, and regional associations such as the African Union or the Association of Southeastern Asian Nations (ASEAN). 'Global' refers generally to worldwide diplomatic organizations such as the Organization for Economic Cooperation and Development (OECD) or United Nations (UN). Examples of efforts included in this cluster are the US-EU *Joint Statement of the Trade and Technology Council*, the G7 *Statement on Hiroshima AI Process, EU AI Act*, the Council of Europe *Draft Framework Convention on AI*, and the UN *Global Digital Compact*.
6. ***Private Sector Groups and Policy Efforts*** — This cluster includes references to policy efforts or ethics frameworks originating from the private sector that are connected to a specific company or set of companies. Efforts related to nonprofit institutes or academic research are not included in this cluster (see cluster 7). An example statement that falls under this cluster is "IBM has a framework, Microsoft has a really good framework and lots of those things, and some of them are made public," (Irene). This statement acts as two distinct references, one to IBM and one to Microsoft.
7. ***Nonprofit Organizations, Institutions or Academia*** — References to non-governmental groups developing frameworks or policy recommendations not connected to a specific company fall under this cluster. Examples include references to research institutions connected to universities such as the AI Now Institute (NYU) and the Stanford Institute for Human-Centered AI, and think tanks such as the Future of Privacy Forum and the Special Competitive Studies Project. This cluster yielded no specific policy references since no publications by the referenced nonprofit organizations were mentioned.

4.2 Thematic Clusters

Table 1 shows the cluster labels assigned to each interviewee. Since the interviews were open-ended, 9 of 15 interviewees discussed more than one thematic cluster in their responses, while only four interviewees mentioned only one cluster. Melissa and Cynthia's interviews did not receive any cluster label, since they did not describe any policy efforts in sufficient detail. Cluster 5: *Multinational and Global Efforts* was mentioned the most (7 of 15 respondents), and clusters 1, 3, and 7 were each mentioned by 6 interviewees. Cluster 2: *US Congressional Efforts* was discussed only by James.

Table 1 Thematic cluster labeling by interviewee

Interviewee	1. US Policies	2. US Congressional Efforts	3. Agency Working Groups	4. Non-US National Efforts	5. Multinational and Global Efforts	6. Private Sector Efforts	7. Nonprofit and Academic Efforts
Julia	x						
Benjamin	x		x		x		x
Emma				x	x		
David			x		x		
James	x	x					x
Gary					x		
Vera			x				x
Daley			x				
Melissa							
Elise	x		x		x		x
Stephen	x			x	x		
Jeff			x				x
Otis						x	
Cynthia							
Irene	x				x	x	x
Total	6	1	6	2	7	2	6

4.3 Specific Policy Documents

Table 2 lists all policy documents or specific policy efforts referenced by interviewees, grouped by thematic cluster. In total, 21 documents and efforts were identified by 9 of 15 interviewees. 5 of those 9 respondents mentioned two or more documents. One interviewee alone mentioned 10 documents. The US *Blueprint for an AI Bill of Rights*, the NIST *AI Risk Management Framework*, and the EU *AI Act* were mentioned the most often, while 14 documents were only mentioned by a single respondent.

4.4 Common Perceptions

Below, we summarize overlapping perceptions among interviewees across four major topic areas: general US regulation, the US Congress, non-US regulation (including multinational efforts), and the private and nonprofit sectors.

1. ***General Regulation in the United States*** — There was general agreement among many interviewees who spoke about AI regulation in the US that regulatory work is *currently* underway. The interviewees underscored the enormity of this task by saying, for example, "I believe people are working furiously on [AI policy]" (Emma) and "all that process, that policymaking process, is happening right now" (Benjamin). While there was strong agreement among interviewees about the timeliness of AI policy development in the US, there was less clarity on *who* (which governmental actors) are driving this development. Interview responses on this subject included a large list of different executive branch offices and agencies (e.g. "you have everything from DOD [Department of Defense], … DOE

Table 2 Specific policy documents or efforts referenced by interviewees, sorted by thematic cluster and number of references

Thematic Cluster	Document or Effort	Number of References
1. US Policies	Blueprint for an AI Bill of Rights	5
	NIST AI Risk Management Framework	4
	USAID AI Action Plan	1
	Department of Defense Responsible AI Strategy	1
2. US Congressional Efforts	S. 1394: Block Nuclear Launch by Autonomous Artificial Intelligence Act of 2023	1
4. Non-US National Efforts	Canada Artificial Intelligence and Data Act	1
	China Provisions on Deep Synthesis Technology	1
	China Provisions on Management of Generative AI Services	1
	Italy Ban on ChatGPT	1
	UK AI Strategy	1
5. Multinational or Global Efforts	EU AI Act	4
	Council of Europe Draft Framework Convention on AI	2
	OECD AI Principles	2
	US-EU Joint Statement of the Trade and Technology Council	2
	African Union Draft Strategy for AI	1
	ASEAN Guide on AI Governance and Ethics	1
	G7 Statement on Hiroshima AI Process	1
	Montreal Declaration	1
	UN Global Digital Compact	1
6. Private Sector Efforts	IBM's Principles for Trust and Transparency[1]	1
	Microsoft Responsible AI Standard	1

[1]See clarification on the identification process for this document in Sect. 3.3

[Department of Energy], FTC [Federal Trade Commission] … OSTP [Office of Science and Technology Policy]…") as well as ambiguous references to general "conversations" or working groups.

The policy experts also consistently noted the United States' delayed action on AI oversight compared to other countries and offered two different but not necessarily contradictory explanations. The first explanation is that of a purposeful 'wait-and-see' stance. As Emma formulates it, "In America … we'll … let the rest of the world play around with stuff and we'll see how it goes, and then America will then formulate its policy." Other experts argue that the delay in policymaking is due to being awestruck by the capabilities of emerging AI technology. Julia describes, "AI is in this kind of pathway, where it's cool, it's shiny, it's pretty, it's new, we're not being honest enough about what its complications are and the consequences…" However, both Julia and James indicated that the AI governance landscape is not even across the federal government. Julia noted that

there is always a "big divide" between "discretionary spending agencies" and defense-related agencies. James' interview responses highlighting the AI regulatory work within the Department of Defense seem to substantiate this divide. James also referenced specific actions to *implement* Department policy in existing systems (e.g. regulating access to sensitive data, system validation, etc.).

2. ***US Congress*** — The US Congress was not mentioned as often as the executive branch, and references to Congress generally lacked specificity (e.g. "there are lots of initiatives in Congress right now"). As noted above, there was only a single concrete Congressional policy effort identified through the interviews (S. 1394). The lack of discussion of the legislative branch highlights both the nascency of AI policy discussions in the US and that the executive branch is taking a much more active role in defining US government perspectives on AI governance. Julia's responses on Congressional activity on AI are the most informative in this sense. She notes that Congress "can't even agree that people should have a right to privacy. So we're not anywhere near having rules and regulations around AI." She goes on to say that, "in fact, the field does not want Congress to be the one making large, sweeping things. And they better hope they don't get to the point where Congress feels like it has to because things will be very wrong and will go very draconian." In other words, "when Washington doesn't fully understand something, their answer is to ban it, right?"
3. ***Non-US Regulation*** — Discussion of AI policy efforts outside of the United States generally took the shape of the interviewee listing countries they knew had done *something* concerning AI. "I think [in] France, [and Canada], there are policies that are being developed in terms of what are the ethical uses for AI," (David), and "I think also South Korea, [and] Japan—there are a lot of places in the world that seem to be a little bit further ahead on these [AI] policy issues," (Emma) are typical responses in this regard. Other interviewees were able to name or describe specific policy efforts from other countries, and even more mentioned multinational or global policy efforts such as the OECD *AI Principles* or Council of Europe *Draft Framework Convention on AI*. Many experts framed international policy developments as a step ahead of those in the US, often listing specific foreign policies before US policies (if at all). But by far the most common discussion point when talking about non-US policy efforts was the EU AI Act, which was referred to by Irene as the "elephant in the room," because it is "trying to legislate, in an a priori way, risk mitigation measures for [a] huge variety of … systems." Interestingly, Benjamin emphasized the influence of the EU AI Act on *US* policy by noting that, "There are also lots of conversations right now happening within the US government about the EU AI Act and what should be the US government's response to [the Act]."
4. ***Private and Nonprofit Sectors*** — Notwithstanding the two private sector policy documents mentioned by Irene, the discussion of the private and nonprofit or academic sectors was the least specific. Like with US regulatory efforts, there was some consensus that companies have "internal policies about the use of their technologies and development" concerning AI (Elise), but very little mention of what those policies are. Beyond Irene's mention of Microsoft and IBM, only Otis added Apple and Tesla as additional private sector actors engaging in the AI ethics policy space. Daley underscored the reason why these companies are engaged in AI ethics conversations, noting that companies are "trying to make policies to ensure that users trust their products to avoid regulatory scrutiny." This proactive

private sector approach is the flipside of the 'wait-and-see' regulatory approach from the US government, as companies who rely on AI technology try to avoid provoking regulatory intervention from the federal government (especially Congress). There was similarly little discussion on the role of the nonprofit and academic sectors in AI policy development. Cynthia, who works at a research institute, noted that the institute's current "AI projects are … providing the documentations [*sic*] for making good decisions about how AI should be utilized by various public institutions, industry institutions, or academic institutions. We're contributing to the wealth of knowledge to help support good decision-making."

5 Discussion

While 13 of the 15 we experts interviewed were generally familiar with various public and private sector AI ethics policy efforts, only 9 experts mentioned specific policy documents or initiatives, either in the US or abroad. Among these, the *Framework for an AI Bill of Rights*, the OECD *AI Framework*, and the EU *AI Act* were mentioned the most.

5.1 US Policy

For the most part, each expert was knowledgeable of policy developments within their government agency or subdomain. There was broad agreement that regulators were actively working on policy frameworks to govern the use of AI in the government and the private sector, and more of these efforts have come to light in recent years with the publication of policies (e.g. NIST *Risk Management Framework*), guidance (e.g. *Framework for an AI Bill of Rights*), and legislation (e.g. EU *AI Act*). Regarding legislative efforts in the US, there was a sense among most interviewees that Congress is working on passing laws to deal with AI. However, discussions of Congressional oversight efforts lacked the detail that the experts could easily provide when talking about executive (administrative or regulatory) policies. Only one interviewee named a specific piece of Congressional legislation, namely a bill to prohibit the use of AI to launch nuclear weapons without human authorization (S. 1394).

This finding underscores the extent to which legislative efforts on AI governance in the US currently lag behind those abroad and even the policy initiatives in the US executive branch. We posit that this delay is furthered, if not caused by, still disparate ideological views among members of Congress on central principles that would form the foundation of potential AI legislation, mirroring the disagreements found in statehouse discussions about AI regulation (Parinandi et al., 2024). One expert noted that disagreements within Congress over privacy rights and laws are preventing Congress from confronting more complicated issues over AI regulation. The expert went on to say that "the field [of AI developers and regulators] does not want Congress to be the one making large, sweeping [legislation]" because there is a sense that when Congress does step in it does so in a "draconian" way. This attitude is perhaps a consequence of the American regulatory system fostered by the separation of powers; Congress prefers to exercise hands-off 'fire-alarm oversight', only intervening when egregious mismanagement is brought to its attention (McCubbins & Schwartz, 1984). Thus, minimal (and often nonbinding) executive policy

is seen a sufficient during this nascent stage of AI innovation. Congressional legislative action is seen as premature and likely to be overly restrictive.

That the future of US AI regulation is a topic of current discussion in the halls of the US Capitol is evidenced by the over 500 bills from the 118th Congress (2023–2024) alone that are returned by the Congress.gov legislative search engine when searching for "artificial intelligence," including over three dozen with "artificial intelligence" in the title. Yet no single effort has consolidated policymakers' attention. And as the 118th Congress draws to a close, only three of the over 500 bills have become law, including the omnibus National Defense Authorization Act and FAA Reauthorization Act. In the absence of a standard-bearer, federal legislation is subject to disproportionate influence by a variety of political interests seeking to shape what aspects of AI are regulated and how. The exact dynamics of this AI lobbying should be subject to further research exploration. However, experimental evidence from state legislatures in a study by Schiff and Schiff (2023) shows that both expert opinion and narratives ("stories involving characters, contexts, plots, and morals") are highly effective at engaging legislators on AI policy. Even in a highly technical policy domain, the authors find that "'passion' can be just as important as 'reason' in policy influence efforts," (D.S. Schiff & Schiff, 2023). This suggests that, in contrast to the expert-driven approach of the European Commission in developing the EU AI Act (Justo-Hanani, 2022), US AI legislation may be driven as much by the societal 'hype' created by new AI technologies. Yet whether this hype causes Congress to continue its 'hands-off' regulatory approach, or to intervene on a much larger scale remains to be seen.

5.2 Overlapping US Efforts

The variety of policy efforts listed by the experts shows that there is perhaps no single catch-all US policy on AI that is broadly identifiable across government and industry. The *Framework for an AI Bill of Rights* could be considered the most widely known effort, and we might expect its name recognition to grow in the coming months, especially in light of President Biden's October 2023 executive order (which had not been published at the time we conducted our interviews). Abroad, the EU AI Act performs the function of standard-bearer for all-encompassing AI regulation, and its influence is evident in our interviewees' responses. Its name recognition, we hypothesize, can be attributed to its wide remit to deal with multiple challenges of ethical AI use (foundation models, biometric data, etc.) and its large impact as a law that will impact over 700 million people across 27 EU nations.

US legislative and regulatory efforts are perhaps not as advanced as those abroad (particularly in the EU and China), as noted by multiple experts and confirmed by literature on this subject. In the meantime, while viable Congressional legislation is still lacking, the fabric of disparate and overlapping agency policies, recommendations, and opinions paints a cloudy picture of how AI will be regulated in the future. This lack of unity and clarity was epitomized by the response from an expert from the National Institute for Science and Technology (NIST), who told us of his desire for "NIST to be able to think about a supercluster that would then be able to set up the framework by which we would be thinking about AI". However, exactly such a document (the *AI Risk Management Framework*) already existed at the time of the interview and was published by NIST itself in 2022. That an employee in the originating authority can be oblivious to a major guiding document underscores the need to unify disparate federal efforts to regulate AI.

5.3 Non-US and Multinational Efforts

Even though all experts we interviewed worked for the US government or within US companies or organizations, many nevertheless had knowledge of non-US policies governing ethical AI deployment, both at the level of national legislation or regulation and multinational policies or frameworks. The EU AI Act was mentioned broadly, but interviewees also named additional national and multinational policies that establish guidance or implement regulation for AI. Here we draw connections between the knowledge of these experts beyond the US context and the "cleaving" power to decouple "law and territoriality" described by Floridi (2017). Not only are the EU's AI Act and General Data Protection Regulation (GDPR) examples of cases where technology's power to transcend national borders necessitated a *regional* regulatory approach, but in the case of the GDPR, the difficulty (or cost) of tailoring digital services to different countries means that some GDPR provisions apply *de facto* to non-EU territories, including the US. Increasingly, websites of EU and non-EU companies require users to select cookie settings upon loading their website, even if the user is located outside of the European Union. In this way, knowledge of consequential non-US technology policy efforts is important for policymakers not only because of the potential *influence* of these policy antecedents on future policy, but also because of the current *impact* of foreign regulation on US user and company behavior. While we do not yet know whether and how the EU AI Act may influence US AI legislation, we should still expect to soon see the impact of the Act on US technology companies also operating in the EU.

Looking beyond major multinational policy developments, knowledge of non-US national-level policy efforts was less widely distributed. Of the five non-US national efforts mentioned, four were mentioned by the same interviewee. Nevertheless, it is promising to hear mentions of AI strategies from the African Union and ASEAN given the current dominance of Western and Global North countries in discussions around ethical AI implementation.

5.4 Engaging the AI Policy Literature

On the one hand, the responses of our participants confirm many of the findings from existing AI policy literature. Among these are the predominance of Northern and Western countries in AI governance discussions; the importance of the US, EU, and China as leaders in the AI policy space; and the underdevelopment of US AI policy compared to the EU and China. The latter point in particular was underscored by the experts both explicitly and implicitly. Moreover, the EU AI Act and US Blueprint for an AI Bill of Rights were both referenced the most by experts, who all highlighted these documents as seminal for their respective jurisdiction's approach to AI governance.

At the same time, however, the responses of our participants differ from the policy documents scholars highlighted as important to the development of AI regulation. For example, the two Chinese AI policy documents highlighted in the literature—the AIDP and PIPL—were not mentioned by the participant who was knowledgeable of two more recent legislative efforts in China surrounding deepfakes and generative AI. This difference in the mentioned documents is, in one sense, understandable given that the experts we interviewed were from the US and not China. Yet this also indicates that the

experts may be more focused on *outcomes* (through regulation and legislation) than on *strategy*, especially beyond the US. For the policy experts currently grappling with how to develop AI regulation in the US, what matters is not how other comparable countries arrived at their outcomes, but simply what those outcomes are. Relating to China in particular, the relative obscurity of the AIDP may also underscore the realities of "fragmented authoritarianism" (Roberts et al., 2022); because it is unclear what national CCP goals will cut through competing priorities and be implemented by local authorities, what matters to foreign observers is what regulation becomes implemented.

Regarding the processes for AI policy development, the insights from the experts support the trends we highlight in Sect. 2.1, especially with respect to the US. Since none of the experts we interviewed worked outside of the US, discussion of non-US policy processes was necessarily limited. However, the predominance of the EU AI Act in discussions on AI governance within the European Union does emphasize that the EU itself is, for the most part, taking the lead on AI regulation. Thus, European AI policy can be characterized, more or less, as a single policy, not 27 different policies.

With respect to the US, the responses from the experts support the conclusions that (1) the US has been slower to develop and execute policies on AI regulation, and (2) the executive branch, not Congress, is currently leading these regulatory efforts. However, the variety of responses on executive branch AI policy efforts indicates that US AI 'policy' is more diverse and disjointed than a review of US AI policy literature would suggest. We offer two reasons for this. First, the reality of policy work for these experts is often siloed, with policymakers primarily concerned with how AI is regulated within their agency's remit, and not necessarily how a particular department's AI policy fits with the overall AI strategy of the administration. In other words, we hear from these experts a different perspective on AI policy than the top-down view presented in the literature. Second, US efforts to regulate AI have advanced significantly since conducting our interviews. This is evident not only in the signing of Executive Order 14110 in October 2023, but also in President Biden's nationally-televised called to "Ban A.I. voice impersonation and more" during the 2024 State of the Union address (Biden, 2024).

6 Conclusion

In this paper, we present a qualitative analysis of the knowledge and perceptions of fifteen AI experts in policy-related roles regarding AI ethics policy and governance. We find that most experts are knowledgeable about policy efforts that relate to their agency or field of work. For example, an expert at the US State Department whose work focuses on trade policy would be familiar with statements from the US-EU Trade and Technology Council regarding AI ethics, while an expert at the Department of Defense could be expected to have knowledge of efforts in Congress to regulate AI in defense settings. However, we find that expert knowledge does not uniformly extend far beyond the walls of the expert's agency. While we documented 21 specific policy efforts referenced in our interviews, 12 of those efforts came from only four experts. Knowledge of non-US AI policy efforts is particularly concentrated and tends to focus on Northern and Western countries, with the US, European Union, and China consistently recognized as the primary actors in the AI policy space. The minimal mention of countries in the Global South emphasizes the need

for increased focus on how these states approach AI governance, and to what extent they are influenced in this endeavor by countries like the US or China.

Overall, the responses from the expert interviews corroborate the importance of policy documents like the US *Blueprint for an AI Bill of Rights* and EU *AI Act*, which are widely discussed in existing literature. At the same time, however, the discussion on Chinese AI regulation differed from the literature, with the expert who mentioned Chinese policy efforts citing two different regulatory documents than those included in our literature review. Likewise, the discussion of US policy efforts was more diffuse than might be expected. This was epitomized in the response of the expert from NIST who was unaware of the already-published NIST *AI Risk Management Framework*. Crucially, many of the experts who mentioned the US *Blueprint for an AI Bill of Rights* highlighted the document's limitations and argued that a new policy document is needed to formalize and operationalize US policy on AI governance. We posit that President Biden's October 2023 executive order (E.O. 14110) will take on this role, at least in the near future.

6.1 Limitations

We acknowledge that, in addition to only representing the US perspective on AI ethics policy, our experts also form a relatively small sample. For this reason, we placed emphasis on the qualitative evaluation of expert interview responses, and used our thematic cluster coding to emphasize similarities in the areas of AI ethics policy intervention mentioned across multiple experts. Although a larger sample of policy experts is likely to cover a wider number of policy efforts and enable robuster conclusions, we also note that policy work is not a simple 'numbers game'. A closer analysis of the position, scope, and influence of a policy expert is needed to assess the impact that their perspective may have on policy development. In this sense, we provide the summaries and analysis of our experts' perspectives to elucidate the wide range of policy efforts connected (sometimes only tangentially) to AI regulation in the US and around the world. The key takeaway here is that 'AI ethics policy,' especially in the US context, is not as simple as pointing to a single law or executive order 'on AI'. Indeed, the breadth of policy efforts underscores the difficulties government agencies and private industry face in trying to comply with the intent, if not the letter, of still-developing policies.

6.2 Future Work

Our findings also indicate promising directions for further research on AI ethics policy development. In particular, a similar study to this one should be conducted over a longer time period to investigate the growth in awareness and influence of newer policy efforts such as the EU AI Act (which was passed by the European Parliament after we finished our interviews) and Executive Order 14110 (which was released after our interviews). Findings from such a study could be compared to this paper to better understand the development over time of *how* policy experts think about AI regulation. Moreover, a quantitative analysis of Congressional attitudes towards AI regulation building on the work of Parinandi et al. (2024) would be timely given increasing Congressional activity on AI, and would further understanding of the *politics* of AI regulation on a federal level.

Appendix A: Referenced Policy Documents

Tables A1 and A2 contain all AI policy documents referenced in the literature review and by the policy experts we interviewed.

Table A1 List of referenced policy documents

Document Title	Country or Organization	Originating Authority	Abbreviation
Executive Order 13859: "Maintaining American Leadership in AI"	United States	White House	E.O. 13859
Blueprint for an AI Bill of Rights	United States	Office of Science and Technology Policy	"Blueprint"
Artificial Intelligence Risk Management Framework	United States	National Institute of Standards and Technology	RMF
Executive Order 14110: "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence"	United States	White House	E.O. 14110
Artificial Intelligence Action Plan	United States	US Agency for International Development	
Responsible Artificial Intelligence Strategy and Implementation Pathway	United States	Department of Defense	
S. 1394: Block Nuclear Launch by Autonomous Artificial Intelligence Act of 2023	United States	Congress	S. 1394
General Data Protection Regulation	European Union	European Commission	GDPR
Artificial Intelligence Act	European Union	European Commission	"EU AI Act"
A New Generation Artificial Intelligence Development Plan	China	State Council	AIDP
Personal Information Protection Law	China	National People's Congress Standing Committee	PIPL
Administrative Provisions on Deep Synthesis in Internet-Based Information Services	China	Cyberspace Administration of China, Ministry of Industry and Information Technology, Ministry of Public Security	
Interim Measures for the Management of Generative Artificial Intelligence Services	China	Cyberspace Administration of China[2]	
Artificial Intelligence and Data Act[3]	Canada	Canadian Parliament	
National Artificial Intelligence Strategy	United Kingdom	Department for Digital, Culture, Media & Sport	
Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law	Council of Europe	Committee on Artificial Intelligence	
Recommendation of the Council on Artificial Intelligence	OECD	Committee on Digital Economy Policy	"OECD AI Principles"
Joint Statement of the Trade and Technology Council (May 2023)	US/EU	Trade and Technology Council	

[2]Jointly released along with the National Development and Reform Commission, Ministry of Education, Ministry of Science and Technology, Ministry of Industry and Information Technology, Ministry of Public Security, and National Radio and Television Administration

[3]Part 3 of the Digital Charter Implementation Act of 2022

Table A2 List of referenced policy documents (continued)

Document Title	Country or Organization	Originating Authority	Abbreviation
Draft Conceptual Framework of the Continental Strategy on Artificial Intelligence[4]	African Union	Specialised Technical Committee on Communication and Information Communications Technology	
Guide on AI Governance and Ethics	Association of Southeast Asian Nations	ASEAN Secretariat	
G7 Leaders' Statement on the Hiroshima AI Process	G7	G7 Leaders	
Montreal Declaration for a Responsible Development of Artificial Intelligence	Voluntary Declaration	University of Montreal	"Montreal Declaration"
Global Digital Compact	United Nations	Office of the Secretary-General's Envoy on Technology	"UN Global Digital Compact"
Principles for Trust and Transparency	IBM	IBM	
Responsible AI Standard	Microsoft	Microsoft	

[4]Text of the framework could not be found

Acknowledgments This work has been supported by the + Policy Research Fellowship of the Policy Destination Area (PDA) of the Institute for Society, Culture, and Environment (ISCE), the Diversity Inclusion Seed Investment Grant funded by the Institute for Critical Technology and Applied Science (ICAT) at Virginia Tech, and the National Science Foundation Grant No. 2412398. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding programs.

Data Availability Full pseudonymized interview transcripts are available upon reasonable request made to the corresponding author.

Declarations

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Biden, J. (2024, March). *Remarks of President Joe Biden — State of the Union Address As Prepared for Delivery*. The Capitol. Retrieved from https://www.whitehouse.gov/briefing-room/speeches-remarks/2024/03/07/remarks-of-president-joe-biden-state-of-the-union-address-as-prepared-for-delivery-2/

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications.
Coeckelbergh, M. (2020). *AI Ethics*. The MIT Press.
Corrêa, N. K., Galvão, C., Santos, J. W., Del Pinto, C., Pinto, E. P., Barbosa, C., … de Oliveira, N. (2023). Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance. *Patterns*, *4* (10). https://doi.org/10.1016/j.patter.2023.100857
Council of the European Union (2022, December). *Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights* (Press release No. 1008/22). Council of the European Union. Retrieved from https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/pdf/
Ding, J. (2018, March). *Deciphering China's AI Dream: The Context, Components, Capabilities, and Consequences of China's Strategy to Lead the World in AI* (Tech. Rep.). Future of Humanity Institute. Retrieved from https://www.fhi.ox.ac.uk/wp-content/uploads/DecipheringChinasAI-Dream.pdf
Exec. Order No. 14110 (2023). *Safe, secure, and trustworthy development and use of artificial intelligence*. Retrieved from https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence
Floridi, L. (2017, May). Digital's cleaving power and its consequences. *Philosophy and Technology*, *30*, 123–129. https://doi.org/10.1007/s13347-017-0259-1
Floridi, L. (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
Gianni, R., Lehtinen, S., & Nieminen, M. (2022). Governance of responsible AI: From ethical guidelines to cooperative policies. *Frontiers in Computer Science*, *4*. https://doi.org/10.3389/fcomp.2022.873437
Harris, L., & Jaikaran, C. (2023, November). *Highlights of the 2023 executive order on artificial intelligence for Congress* (Tech. Rep. No. CRS Report No. R47843). Congressional Research Service. Retrieved from https://crsreports.congress.gov/product/pdf/R/R47843
Hine, E., & Floridi, L. (2022, June). Artificial intelligence with American values and Chinese characteristics: A comparative analysis of American and Chinese governmental AI policies. *AI and Society*, *39*, 257–278. https://doi.org/10.1007/s00146-022-01499-8
Hine, E., & Floridi, L. (2023, January). The blueprint for an AI bill of rights: In search of enaction, at risk of inaction. *Minds and Machines*, *33*, 285–292. https://doi.org/10.1007/s11023-023-09625-1
Hoofnagle, C. J., von der Sloot, B., & Zuiderveen Borgesius, F. (2019, February). The European Union general data protection regulation: What it is and what it means. *Information & Communications Technology Law*, *28* (1), 65–98. https://doi.org/10.1080/13600834.2019.1573501
Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*, 389–399. https://doi.org/10.1038/s42256-019-0088-2
Justo-Hanani, R. (2022, March). The politics of Artificial Intelligence regulation and governance reform in the European Union. *Policy Sciences*, *55*, 137–159. https://doi.org/10.1007/s11077-022-09452-8
Krane, D. (1993, June). American federalism, state governments, and public policy: Weaving together loose theoretical threads. *PS: Political Science and Politics*, *26* (2), 186–190. https://doi.org/10.2307/419826
Madiega, T. (2024, March). *Artificial Intelligence Act* (Tech. Rep. No. PE 698.792). Members' Research Service. Retrieved from https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792EN.pdf
McCubbins, M. D., & Schwartz, T. (1984, February). Congressional oversight over-looked: Police Patrols versus Fire Alarms. *American Journal of Political Science*, *28* (1), 165–179. https://doi.org/10.2307/2110792
National Institute of Standards and Technology (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. U.S. Department of Commerce.
Office of Science and Technology Policy (2022). *Blueprint for an AI Bill of Rights*. The White House. Retrieved from https://www.whitehouse.gov/ostp/ai-bill-of-rights/
Parinandi, S., Crosson, J., Peterson, K., & Nadarevic, S. (2024, March). Investigating the politics and content of US State artificial intelligence legislation. *Business and Politics*, 1–23. https://doi.org/10.1017/bap.2023.40
Ponzano, P., Hermanin, C., & Corona, D. (2012, February). *The power of initiative of the European Commission: a progressive erosion?* (Tech. Rep. No. 2012/89). Notre Europe. Retrieved from https://hdl.handle.net/1814/69959
Roberts, H. (2022, February). Informational privacy with Chinese characteristics. *Digital Ethics Lab Yearbook 2021*, Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstractid=3988745

Roberts, H., Cowls, J., Hine, E., Mazzi, F., Tsamados, A., Taddeo, M., & Floridi, L. (2021). Achieving a 'Good AI Society': Comparing the aims and progress of the EU and the US. *Science and Engineering Ethics*, *27*. https://doi.org/10.1007/s11948-021-00340-7

Roberts, H., Cowls, J., Hine, E., Morley, J., Wang, V., Taddeo, M., & Floridi, L. (2022, September). Governing artificial intelligence in China and the European Union: Comparing aims and promoting ethical outcomes. *The Information Society*, *39* (2), 79–97. https://doi.org/10.1080/01972243.2022.2124565

Rose-Ackerman, S. (2022). Democracy and executive power: Administrative Policymaking in comparative perspective. *Revista de Derecho Público: Teoría Y Método*, *6*, 155–182. https://doi.org/10.37417/RPD/vol_6_2022_978

Schiff, D., Biddle, J., Borenstein, J., & Laas, K. (2020, February). What's Next for AI Ethics, policy, and governance? A global overview. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 153–158). New York: Association for Computing Machinery.

Schiff, D. S., & Schiff, K. J. (2023, July). Narratives and expert information in agenda-setting: Experimental evidence on state legislator engagement with artificial intelligence policy. *Policy Studies Journal*, *51* (4), 817–842. https://doi.org/10.1111/psj.12511

Tabassi, E., Carnahan, L., Hogan, M., & Heyman, M. (2019, August). *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools* (Tech. Rep.). National Institute of Standards and Technology. Retrieved from https://www.nist.gov/system/files/documents/2019/08/10/aistandardsfedengagementplan

Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, *22* (4), 97–112. https://doi.org/10.9785/cri-2021-220402

Zhu, Q., Kim, D., Eldardiry, H., & Ausman, M. C. (2023). A preliminary investigation of the ethics policy concerns of artificial intelligence: Insights from AI professionals working in policy-related roles. *Proceedings of the 2023 IEEE Frontiers in Education Conference*, College Station, TX. https://doi.org/10.1109/FIE58773.2023.10343253