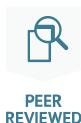




RESEARCH
ARTICLE



OPEN
ACCESS



PEER
REVIEWED

Troubling translation: Sociotechnical research in AI policy and governance

Serena Oduro *Data & Society Research Institute*

Alice E. Marwick *Data & Society Research Institute* alice@datasociety.net

Charley Johnson *Data & Society Research Institute*

Erie Meyer *Georgetown Institute for Technology & Law*

DOI: <https://doi.org/10.14763/2025.4.2043>

Published: 18 November 2025

Received: 7 March 2025 **Accepted:** 25 July 2025

Funding: This work was made possible, in part, with support from the Annie E. Casey Foundation and the David and Lucile Packard Foundation.

Competing Interests: The author has declared that no competing interests exist that have influenced the text.

Licence: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. <https://creativecommons.org/licenses/by/3.0/de/deed.en>
Copyright remains with the author(s).

Citation: Oduro, S., Marwick, A.E., Johnson, C., & Meyer, E. (2025). Troubling translation: Sociotechnical research in AI policy and governance. *Internet Policy Review*, 14(4). <https://doi.org/10.14763/2025.4.2043>

Keywords: Policy translation, Artificial intelligence, Sociotechnical research, Knowledge transfer, Evidence-based policy

Abstract: As technology companies develop and incorporate artificial intelligence (AI) systems across society, calls for a sociotechnical approach to AI policy and governance have intensified. Sociotechnical research emphasises that understanding the efficacy, harms, and risks of AI requires attention to the cultural, social, and economic conditions that shape these systems. Yet development and regulation often remain split between technical research and sociotechnical work rooted in the humanities and social sciences. Bridging not only disciplinary divides but the gap between research and AI regulation, policy, and governance is crucial to building an AI ecosystem that centres human risk. Researchers are frequently urged to “translate” findings for policymakers, with translation framed as a pathway to interdisciplinary collaboration and evidence-based governance. This paper troubles this notion of translation by examining two case studies: the National Institute of Standards and Technology’s US AI Safety Institute Consortium and the Public Technology Leadership Collaborative. We show that spanning research and policy requires more than simplified communication, but depends on building relationships and navigating the uneven terrain where academics, policymakers, and practitioners meet. We conclude with recommendations for existing governmental mechanisms for incorporating researchers into policymaking, which may not be widely known to academics.

This paper is part of **The craft of interdisciplinary research and methods in public interest cybersecurity, privacy, and digital rights governance**, a special issue of *Internet Policy Review*, guest-edited by Adam Molnar, Diarmaid Harkin, and Urs Hengartner.

Introduction

As technology companies develop and incorporate artificial intelligence (AI) systems across society, calls for a sociotechnical approach to AI policy and governance have intensified (Dean et al., 2021; Sartori & Theodorou, 2022). Sociotechnical research emphasises that understanding the efficacy, harms, and risks of AI systems, and technology in general, requires assessing the human factors that shape them, including cultural, social, and economic influences (Oduro & Kneese, 2024). AI development and regulation is currently bifurcated between *technical* and *sociotechnical* research, the latter requiring expertise in the humanities and social sciences. Bridging the gap not just between these epistemically and disciplinarily distinct paradigms but to the realities of AI regulation, policy, and governance is crucial to building an AI ecosystem that centres grounded human risk.

Most commonly, researchers are urged to “translate” their research findings to be usable to policymakers (California Policy Lab, 2023; Rushmer et al., 2019). This translation process supposedly serves as a tool to unite interdisciplinary groups of researchers and policy experts, fostering opportunities for evidence-based policymaking. This paper explores two case studies, the National Institute of Standards and Technology’s US AI Safety Institute Consortium and the Public Technology Leadership Collaborative (PTLC) to trouble the concept of translation (Grimshaw et al., 2012a; Tseng, 2012). We draw on our professional experiences as researchers and civil servants deeply engaged with policymaking to argue that spanning research and policy requires building meaningful relationships. Simply advocating for translation ignores the much rockier terrain academics and policy experts must cross to successfully incorporate research into policy. In order to meaningfully influence policy, researchers must work directly with government agencies, navigating bureaucracy and potential political hostility. By analysing the PTLC and the Consortium, we explore the challenges of bridging empirical research and policy, and outline exactly what is needed to do this well, based on our deep immersion in both communities. We conclude with recommendations for governmental mechanisms for incorporating researchers into policymaking, which may not be widely known to academics.

Literature review

Evidence-based policy and translation

Scientific advances in the late 19th and early 20th centuries – from refrigeration to insulin – often had such immediate and obvious impacts that rigorous assessment of public implementation seemed unnecessary (Baron, 2018). However, as scientific progress became more incremental, policymakers began seeking systematic ways to evaluate the effectiveness of interventions. In the 1960s and 1970s, social scientists across the US, Canada, Europe, and Australia were deeply involved in the design and implementation of large-scale public social service programmes, such as education, prisons, and social welfare (Head, 2010). Subsequent implementation and evaluation studies show that these initiatives produced mixed and often modest impacts, with recurrent obstacles in delivery and unanticipated effects, even though many programmes endured and became foundational to contemporary governance (Pressman & Wildavsky 1973; Rossi 1987).

In response, the 1980s saw the emergence of the “evidence-based policy” movement, which originated in the United Kingdom within the health sciences and gradually expanded to political science and other social sciences (Bowers & Testa, 2019; Nutley & Webb, 2000; Saltelli & Giampietro, 2017). This movement sought to incorporate rigorous empirical research into policymaking, by ensuring that government decisions were informed by reliable data rather than tradition, ideology, or anecdotal evidence. In the United States, this trend gained institutional support with the establishment of the Commission on Evidence-Based Policymaking in 2016. This body was tasked with encouraging the use of “rigorous evidence” to inform public policy and assess government operations (Bowers & Testa, 2019).

Despite its promise, evidence-based policymaking faces significant challenges. One of the primary debates revolves around what qualifies as “rigorous” evidence. Randomised control trials (RCTs) are widely considered the gold standard in health policy and have increasingly been used in social science research to evaluate programmes such as welfare-to-work initiatives (Baron, 2018). However, RCTs are not always appropriate for many social science disciplines, where ethical constraints, logistical difficulties, and the complexity of social phenomena make experimental methods impractical.¹ In other cases, there may not *be* a clear evidence base, or

1. While the “replication crisis” has undermined confidence in some strands of social and behavioural science, particularly small-sample or novel experimental studies, most evidence used in policymaking comes from larger-scale evaluations, systematic reviews, and administrative data analyses. Its primary impact on evidence-based policy has been to raise the standards for evidence used to influence decisions, encouraging replication and transparency, and tempering the rapid scaling of in-

scientists may not be in consensus (Rushmer et al., 2019, p. 130). As a result, some scientists have moved towards “evidence-informed” policy, which recognises that knowledge is not only the result of RCTs, but may include “professional knowledge and expertise, tacit knowledge, situated knowledge and collective and organisational memory” (Boaz et al., 2019; Rushmer et al., 2019, p. 131).

Another challenge is the difference in timeframes between research and policymaking (Ellen et al., 2014; Haynes et al., 2018; Natow, 2022). Policymaking often requires quick responses, while scholarly research is inherently slow, involving lengthy processes of study design, data collection, peer review, and publication. This misalignment frequently results in policymakers relying on outdated or incomplete research. Moreover, there is a persistent communication gap between researchers and policymakers due to differences in professional cultures, technical language, and publication practices (Ellen et al., 2014; Farley-Ripple et al., 2018; Haynes et al., 2018; Natow, 2022). Many policymakers lack training in interpreting academic research, and researchers often struggle to translate their findings into practical policy recommendations. Additionally, paywalls and the technical nature of academic publications further restrict policymakers’ access to relevant studies (Haynes et al., 2018).

Concerns about the misuse of research also present challenges. Researchers may fear that their findings will be misused or politicised by policymakers, especially when policymakers lack training in handling empirical research (Bowers & Testa, 2019; Ellen et al., 2014). On the policymaker side, a “crisis of generalisability,” the concern that studies conducted in specific contexts may not be applicable elsewhere, can discourage the adoption of research-based recommendations (Bowers & Testa, 2019). This concern is often tied to policymakers’ preference for quantitative, positivist research over qualitative studies, as they perceive the latter to be anecdotal and lacking in rigor (Natow, 2022). In other cases, policymakers embrace splashy studies funded by industry or pre-prints published before peer-review, rather than more sober independent research, especially during politically heightened times (Every-Palmer & Howick, 2014; Flanagan, Fontanarosa & Bauchner, 2020). Scholars have proposed various knowledge-translation activities to bridge the gap between research generation and its incorporation into practice and policy (Rushmer et al., 2019). One approach is to leverage storytelling in qualitative research to make findings more compelling and accessible to policymakers, who often rely on narratives to support their arguments (Crow & Jones, 2018; Natow, 2022). Another recommendation is for researchers to refine their communication

terventions supported by single, unreplicated studies.

techniques by diversifying research outputs, publicising findings through multiple channels, and creating clear, accessible summaries outside of academic journals, such as policy briefs or evidence briefings (Haynes et al., 2018; Oliver et al., 2014). The technique of “evidence synthesis” maintains that rather than attempting to translate individual studies, scholars should work from systematic reviews or other syntheses, so that policymakers are exposed to a larger evidence base (Grimshaw et al., 2012b; Rushmer et al., 2019).

Collaboration between researchers and policymakers is also crucial. Scholars suggest that researchers should engage directly with policymakers to understand legislative and regulatory processes that foster mutual understanding (Oliver & Cairney, 2019). Additionally, co-developing research goals and expectations with policymakers can lead to more practical and impactful policy applications, though this requires flexibility and compromise on both sides (Bowers & Testa, 2019). In sum, while evidence-based policymaking and knowledge translation have the potential to improve public policy, they face significant structural, cultural, and methodological challenges. Addressing these issues will require ongoing efforts to enhance communication, accessibility, and collaboration between the research and policy communities.

Evidence in AI policy

Many have also called for an evidence-based approach to artificial intelligence regulation (Bommasani et al., 2024). While this may seem obvious, it must be put in the proper context alongside the dominance of “existential risk,” or “x-risk,” in discussions of AI safety in the early 2020s (Ahmed et al., 2024; Beard & Torres, 2024; Jecker et al., 2024). X-risk is a calamitous narrative in which artificial intelligence achieves “superintelligence,” causing the end of human life or civilisation. There are three dominant narratives in which fictional superintelligent AI poses an existential risk: AI pursues its own goals which conflict with human life; it seizes control of key elements of society such as capital, infrastructure, or energy; or it is used to supercharge biological or nuclear attacks (Webb & Schönberger, 2025). However fanciful, x-risk has taken hold among major tech leaders, research centres, and online forums, and is heavily funded, with considerable public influence (Ahmed et al., 2024). Ahmed et al. argue that the field of “AI Safety” constitutes an “epistemic community” which has “translated [its] shared moral and normative claims into technical solutions and recommendations for AI policy that may have lasting, global implications” (Ahmed et al., 2024). This agenda can be summed up in the 22-word “Statement on AI Safety” released by the San Francisco-based Center for AI Safety, signed by the CEOs of Google DeepMind, Anthropic, and OpenAI

and highly-regarded computer science professors: “Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war” (Center for AI Safety, 2023).

Importantly, there is no evidence for x-risk, scientific or otherwise, as it relies on speculation and futurism (Westerstrand et al., 2024). As anthropologist Jenna Burrell points out, “this future focus distracts us from the present while also absolving technologists of responsibility for the here and now” (Burrell, 2023). The *present* risks of AI are innumerable, including deskilling and algorithmic control of workers (Nguyen & Mateescu, 2024), racial and gender bias (Ferrara, 2025; Shah & Sureja, 2025), environmental harms (Kneese, 2024), diminishing privacy and increasing surveillance (Lee et al., 2024; Saheb, 2023), and increasing disinformation and anti-democratic threats (Jungherr, 2023; Kreps & Kriner, 2023). However, many of these phenomena, like implicit bias and unprecedented impacts, are very hard to study, especially using quantitative methods.

Current AI research is dominated by wealthy, industry-connected scholars (Ahmed et al., 2023; Septiandri et al., 2023), and the social impacts of global technologies can be hard to understand in full from Western countries (Casper et al., 2025). This is why *sociotechnical* research is needed to inform AI policy (Bogen & Winecoff, 2024; Chen & Metcalf, 2024; Dobbe & Wolters, 2024). As Brian Chen and Jake Metcalf write, “Treating technology and society as one coherent unit means that developers, deployers, and regulators are accountable not just for the mechanical workings of technical machines, but for how those machines integrate with, reshape, and sometimes harm social systems” (2024, p. 6). However, sociotechnical research is primarily qualitative, and sometimes even humanistic, making it even more difficult to translate into policy regimes where quantification is prioritised.

The case studies in this paper are first-person reflections on our involvement in two attempts to incorporate sociotechnical perspectives into AI governance, with varying degrees of success; they reflect our own experiences and points of view. First, we discuss our involvement in the US AI Safety Consortium, which demonstrates some of the pitfalls involved when researchers engage in policymaking (we note that no confidentiality agreements, NDAs, or similar restrictions apply to the material presented). Then, with the PTLC, we present a more successful case study. We discuss what we learned from our involvement in both institutions, and close with concrete recommendations for better ways for policymakers and researchers to work together.

Case study No. 1: US AI Safety Consortium

Established in November 2023 by the Biden-Harris administration through the Department of Commerce, the Artificial Intelligence Safety Institute Consortium (AISIC) is a part of the National Institute of Standards and Technology's (NIST) United States AI Safety Institute (USAISI). AISIC tasks approximately 280 members of industry, academia, and civil society organisations to advance the science of AI safety. In its first year, the Consortium focused on fulfilling the responsibilities outlined in President Biden's Executive Order on AI, which included developing guidelines for managing the risks associated with generative AI and dual-use foundation models (Wörsdörfer, 2024).

Although AISIC's language suggested a straightforward collaboration among industry, academia, and civil society to advance AI safety, successfully engaging with it was far more complex due to confusion about its structure and power dynamics. AISIC's focus, scope of projects, and working group goals often excluded or failed to address civil society concerns. As an AISIC member, it was not enough to simply bring research to AISIC: one needed to know how to navigate the turbulent bureaucratic shifts and processes. Members would wait for months for work to begin, then face a swiftly-approaching deadline to present a research tool that they had never heard of before. Ultimately, being a researcher or advocate in an interdisciplinary and cross-sectoral group such as the AISIC required relationships, resources, and a keen understanding of the government organisation in which one was operating. In the end, how and whether research was integrated or informed the AISICs work was nearly impossible to predict. It is especially difficult to know the internal and external politics that could be influencing an agency or government's willingness to adopt or integrate research.

Soon after the inception of the USAISI, NIST held a public workshop on "Collaboration to Enable Safe and Trustworthy AI". After hearing remarks from NIST leadership and leaders in safe and trustworthy AI, the NIST team led a discussion on the Consortium's working groups and invited feedback from participants. The working groups proposed at the meeting are shown in Table 1.

TABLE 1: AISIC working groups, proposed November 2023

WORKING GROUP	FOCUS
Generative AI	• Risk and capabilities
Synthetic Content	• Authentication, detection, labelling • Deepfakes
Evaluating AI Capabilities	• AI red-teaming and other testing methodologies
	• Pre-deployment testing and post-deployment monitoring
Society and Technology	• Standards • Operationalise the Artificial Intelligence Risk Management Framework

During that meeting, the first author asked publicly how AISIC would ensure that civil society concerns would be addressed, such as AI's impact on marginalised communities. These concerns were not, and, as the first author noted, should not be limited to the remit of a single working group. Instead, they spanned many areas, whether establishing commitments to investigate how bias and discrimination arise, or which AI safety tools and methods can be used to address them. Addressing bias and discrimination is often considered to be an issue best solved by a subset of people, rather than fundamental conditions that impact how we should holistically consider AI solutions (Hanna et al., 2020). This point was ultimately not integrated, and the NIST team noted that the Society and Technology group could address these issues. However, when the final working groups of the AISIC were released, the Society and Technology working group had been removed (Table 2).²

2. It was not communicated to the first author why the working groups were changed.

TABLE 2: Final AISIC working groups, February 2024

Working Group 1	Risk Management for Generative AI
Working Group 2	Synthetic Content
Working Group 3	Capability Evaluations
Working Group 4	Red-Teaming
Working Group 5	Safety & Security

Shifting WG 5 from “Society and Technology” to “Safety & Security” is emblematic of the shift in AI from “responsible AI” to “AI safety,” which has a markedly more technical, national security bent. Over the next year, member organisations and the first author would try to work within the structure of the working groups to advance their mutual aims of ensuring that AI safety efforts incorporated the methods, tools, and approaches needed to ensure that AISIC recommendations were in the public interest. However, this was difficult when the working group that NIST had indicated was most appropriate for discussing issues of bias and discrimination had been redesignated.

During the first six months of 2024, while AISIC was adding members to its working groups and building out their draft plans, it became clear that we would have to influence AISIC’s efforts through an inside-outside strategy. In other words, we would use publications and communications to create external resources and pressure on AISIC. Data & Society, the Center for Democracy and Technology, and many other organisations addressed a letter to then Department of Commerce Secretary Gina Raimondo. We wanted to ensure that AI safety efforts were rooted in addressing immediate algorithmic harms, rather than the x-risk touted by Silicon Valley insiders (Data & Society et al., 2024). While the rise of generative AI and fear of x-risk drove the creation of new AI safety practices, we believed they had to be rooted in existing real-life harms, not science fiction fantasies. Moreover, over many years, technologists had developed practices that could be applied to generative AI, such as “predeployment testing, explainability, assessment of impact, ongoing measurement, and harm remediation” (Data & Society et al., 2024, p. 3).

Along with the joint letter submitted to Secretary Raimondo, we spent much of 2024 working to ensure that the vast body of research on AI’s public harms and how to remediate them through a *sociotechnical approach* were not ignored in favor of the more technical, national security focus on AI safety that was taking over policy conversations. We published two policy briefs: *AI Governance Needs So-*

ciotechnical Expertise and *A Sociotechnical Approach to AI Policy*, and a reading list, “Why AI Safety Requires a Sociotechnical Approach: Our Top Ten Reads,” to show the range of policy and academic literature that demonstrates the need for a sociotechnical approach, or the usefulness of sociotechnical approaches in established areas of science, such as safety engineering. The letter, policy briefs, and reading lists represented a larger fight over political memory and narrative, one that often shifted and morphed depending on who had capital and political power and where those powers were located.

Because NIST was housed in the Department of Commerce and focused on measurement science, a sociotechnical academic lens, and focus on AI’s impacts on historically marginalised communities, often felt at odds with the foci of the AISIC and NIST at large. This is not because they seemed disinterested, but because a sociotechnical approach radically shifts the process of evaluating and creating AI safety institutions, which is very difficult to rectify *in practice*. During the “AI Ethics” era, many researchers sought to address racism in algorithmic systems through simple solutions like removing racial data from datasets, which, of course, did not work (Gilis & Spiess, 2019; Horsfall et al., 2025). Yet, as we see it, the field of AI safety and AI governance more broadly still views racial discrimination and other systemic social issues, and the solutions needed to address them, as singular fixes or practices that can simply be implemented in one portion of AI development, evaluation, or oversight. Similarly, from our perspectives, many policymakers hold the false assumption that to achieve equity or fairness, civil society organisations, academics, and the public can just be added to government programmes and practices without questioning whether these processes and problem scopes even allow for differing voices to be heard or adequately integrated. These limitations stifle the ability for coordinating bodies that are supposed to be interdisciplinary, like AISIC, to truly build robust solutions that incorporate multi-disciplinary expertise to address AI safety issues that threaten the public.

We attempted to inform NIST’s AI Risk Management Framework for Generative AI Profile, mandated by President Biden’s Executive Order on AI, by providing detailed line edits to the 70+ page document to ensure that sociotechnical methods, participatory methods, and bias mitigation tactics were integrated throughout the document. The first author of this paper (Oduro) decided to provide detailed line edits instead of a more general comment calling for the integration of sociotechnical methods throughout the framework. As a policy expert aiming to use research to inform policy, the first author understands the factors that can prevent research from influencing policy. These include recommendations that are too broad or dif-

difficult to translate into policy documents, as well as delays in adapting research within the tight timelines required for rulemaking and publication. The request for comments on the draft profile was released on 29 April 2024 and open for comment until 2 June 2024. Since the Executive Order on AI required NIST to finalise its work within 270 days of its publication on 30 October 2023, we anticipated that the final draft would be released in the summer. Given this timeline, we expected that direct line edits incorporating sociotechnical methods and bias mitigation techniques would be easier for government personnel to integrate. In addition to submitting a public comment, AISIC members were also expected to influence the document. To elevate these recommendations, the first author emailed their comments to NIST AISIC leadership. Despite these efforts, most of our recommendations were not incorporated into the final document. The first author expressed our disappointment in an op-ed, noting the lack of stronger actions to address harmful bias in the final version (Oduro, 2024).

Confusing government processes, limited influence over AISIC's focus, and significant disciplinary differences in interpreting sociotechnical approaches and cross-sector collaboration made it easier for us to shape the broader political conversation on AI safety in 2024 than to influence AISIC's internal work. While consortiums and other government-led bodies should provide opportunities for researchers, industry experts, and advocates to collaboratively advance AI science or other objectives, the reality of influencing these spaces is often far more limited. Amid these challenges, the first author and Borhane Bili-Hammelin created a space for researchers and civil society experts, many of whom were also AISIC members, to strategise how best to engage with AISIC. This valuable initiative met once or twice a month, providing a sense of stability in an often-unpredictable environment. For researchers aiming to translate their work into policy within government spaces, AISIC serves as a case study in the bureaucratic, political, and disciplinary barriers that must be navigated to have even a chance of influencing policy outcomes.

Case Study No. 2: Public Technology Leadership Collaborative (PTLC)

The PTLC is a peer learning collective anchored by 10 research institutions and led by the Data & Society Research Institute, an independent nonprofit. It includes more than 500 scholars, practitioners, and government leaders who cultivate trust-based relationships to ensure that data and technology serve the public interest. Before the PTLC was created, its director interviewed 60 participants, half from

government, and half from academia, to ensure that it met the needs of both groups. Government leaders spoke about the desire for closed spaces where they could take off their “role identity hat” and not feel pressure to represent the equities of their institution. Our interviews showed that government leaders are trained and incentivised to speak from pre-determined talking points, and tend to share less. Speaking on behalf of a government agency carries significant weight and risk, so communication is over-architected and planned, leaving little room for the dynamic back-and-forth required for learning. Academics spoke of the flip side of this equation. They saw the US government as a black box, and found it challenging to know who to talk to about their research. Even if they were able to identify the right person, researchers didn’t feel that government leaders shared enough context to help inform how they might position or communicate their research. There wasn’t enough context sharing, mutual understanding, or trust to facilitate meaningful translation.

This is the challenge the PTLC sought to solve. It attempts to provide a space where government leaders, practitioners, and scholars can convene to share pressing concerns and internal constraints and challenges, free of the limitations inherent when representing their institution publicly. For example, a policymaker can admit they don’t understand a particular AI-related concept, and a top researcher can walk them through it, while also learning through the conversation how to better explain their own research for a policymaker audience. The PTLC was founded based on the belief that building these relationships is essential for moving toward collective sensemaking and collective action. The question “how to translate across difference” is at the centre of the PTLC’s work. Typical notions of “research translation” presume that researchers are the experts and that they can share key information and knowledge that can then be applied by policymakers in their context, but the PTLC’s work troubles this notion in showing that the typical notion of research translation rarely plays out in practice.

First, effective translation does not rest on an expert conveying information in a particular fashion. Instead, *it is an ongoing, dynamic conversation amongst peers that takes time and iteration. It is a looping, mutually shaping process, not a linear one.* Importantly, effective translation does not start when research is complete and final. The PTLC invites researchers to share drafts of initial research findings with their government counterparts to imagine and co-create potential policy recommendations. This not only encourages academic researchers to consider the practical applications of their research, but also helps them develop recommendations that are more likely to be implemented. This goes both ways. For example, in the

spring of 2024, the PTLC hosted an event with the FTC's Office of Technology on the kinds of research questions they should prioritise in the context of AI concentration and exclusivity, AI-enabled frauds and scams, commercial surveillance, sensitive data tracking, and discrimination in algorithmic decision making. In response, scholars Alice Marwick and Lana Swartz initiated research into AI-enabled frauds and scams, and shared the initial findings with the PTLC community to co-create next steps and recommendations. In 2023, the PTLC hosted academic Tamera K. Nopper to present her in-progress paper, *Medicalizing Inequity: The Risks of Financial Wellness for Workers* (Nopper, 2024). Staffers at the Consumer Financial Protection Bureau (CFPB) attended the session and a working relationship grew from there. Dr. Nopper later briefed the entire CFPB financial education team on her pre-publication work, which facilitated a dialogue that allowed both the policy-makers and the researchers to refine forthcoming projects.

Second, how different groups see a problem, and in turn, what they view as potential solutions, is informed by different expertise, epistemologies, frames, and language. It is not as simple as establishing what is "true" to develop a shared objective reality, and then acting accordingly. In contrast to linear positivist approaches to communication, systems work assumes many things may be true at once for different people within different parts of the system. This often looks like letting go of being correct while leaning into finding shared stakes and priorities, toward collaborating on a shared vision of the future. It is not just that stakeholders from different vantage points might see a given challenge differently, but that through a sociotechnical lens, we do not "see" problems, but *make them*. Whether we individualise a problem or emphasise its structural elements; whether we consider it a technological problem or a social one; what frames we use and whose experience is legitimised, all contribute to the "making" of a given problem. This requires a collective sensemaking process. In this way, we can "get on the same page" using slightly different maps, while using each others' perspectives to discover additional entry points for action and intervention within the system. In the PTLC, this effort starts with meeting individually with every participant in a planned salon. In this interaction, the facilitator is trying to culturally onboard the participant into the spirit and norms of the PTLC, but also understand their perspective on the problem of the salon series in order to help connect dots that are not being connected or give extra context to someone's intervention. The PTLC uses structured questions to reveal the assumptions and prior knowledge(s) that contour a participant's conception of the problem (e.g. How might your role and background contribute to how you understand the problem? How might your organisation's structure and/or incentives shape how you understand the problem? How might your

societal position shape what you see and don't?)

In a PTLC working session about cybersecurity, policy staffers from the government asked security researchers why the recommendations section of a recent paper had not included using phishing-resistant multi-factor authentication, updating software to ensure it contains the latest patches, and engaging in active credential and password management to prevent the re-use of compromised credentials. The security researchers were surprised to hear that the government would be interested in such well-known and basic protections. The policymakers candidly shared that in fact, none of the government workers on the call that day were allowed to use these basic protections, including frequently updated software, let alone modern cloud-based password management tools. The government staffers reported regularly being sent shared spreadsheets titled “passwords,” because that was the current password management process. They were amidst internal battles over expanding the use of phishing-resistant multi-factor authentication. The government workers were not only trying to convince their own IT staff to expand these protections, but to also include these things as a matter of policy for regulated entities. Recent, clear recommendations as part of top security research would be immensely helpful in those debates. The security researchers were shocked, and happy to help.

Third, effective translation requires a deep understanding of the organisational context in which research will be understood and used, which is shaped by its attendant norms and incentives. The PTLC uses community guidelines to encourage stakeholders to take off their role-identity hats and show up as people first, as peer-learners and co-problem solvers. And it dedicates significant time to participating stakeholders (government leaders, academics, and practitioners) getting to know one another as people, and sharing one another's context so that participants understand each others' constraints, the incentives that guide their actions, and their broader context. Policymakers are not paid for their participation, but, in our experience, commit time because they believe that trust-based relationships are key to effecting meaningful change. The policymakers who choose to participate in the PTLC seem to understand how challenging it is to build the relationships required for translating across differences and creating effective policies and programs; thus, they recognise the value of a third space that de-centres role identities and centres relationships.

The PTLC also prioritises events that seek to cultivate a shared understanding of one's context. For example, the PTLC hosts informal events where government stakeholders share the organisational dynamics that shape their ability to design

effective programmes. These deep dives into hiring, procurement, and incentive structures, in turn, help academics better understand how to frame and tailor their research translation attempts. It also hosts “hidden curriculum” workshops with academics, helping to demystify how the government works, and how to think about framing their work to better align with contextual nuances. Finally, we hold events with various enforcement agencies (DoJ, CFPB, FTC, etc.) in order to more directly explain how academic research can impact enforcement actions.

It is common practice for advocacy groups to submit “complaints” to regulatory enforcement bodies that oversee industry, essentially outlining potential lawsuits by presenting allegations of illegal activity and detailing the conduct at issue. These complaints are often directed to the head of the enforcement agency or the executive secretary. However, agencies tasked with regulating large tech companies frequently face resource constraints, with a stark imbalance in the legal resources available to each party. While tech companies may have virtually unlimited budgets for legal defence, government agencies often have a single person handling complex investigations. In this context, public regulatory enforcement and private enforcement of civil wrongs would greatly benefit from academic researchers submitting well-drafted complaints or briefing documents that detail technical concerns related to complex technical questions, such as AI.

These submissions can aid in uncovering issues that might otherwise be overlooked, leveraging the researchers’ technical expertise to identify hard-to-detect problems and potentially prevent significant harm. However, the downside for researchers is that regulatory agencies are typically unable to confirm or deny the existence of an investigation until it becomes public through a company’s Security and Exchange Commission filings or other legal actions. As a result, while a researcher’s submission could lead to relatively swift enforcement and litigation, they may remain unaware of the impact of their work for years. This underscores the importance of reciprocal working relationships between researchers and policymakers, facilitating collaborative efforts to address complex issues while managing the challenges of transparency and timing.

Recommendations

We recommend three approaches to encourage collaboration, knowledge exchange, and relationship building between researchers and policymakers. These approaches leverage existing mechanisms in law and policy to bridge the gap between AI research and policy.

Temporary placement of experts in government roles

Many governments maintain mechanisms for secondments, fellowships, or short-term placements that allow academic and technical experts to serve within public agencies. In the United States, for example, stakeholders can use the Intergovernmental Personnel Act (IPA) Mobility Program to place academic experts into government roles on short-term assignments, commonly known as secondments, detail assignments, or rotations (IPA Mobility Program, 2025). The IPA allows for temporary assignments of experts to and from a variety of non-federal organisations, including universities, state and local governments, and tribal entities, without loss of employee rights and benefits. Comparable programmes exist elsewhere, including civil service exchange schemes, research council fellowships, and sector-specific rotations. These arrangements allow agencies to draw on specialised expertise while fostering enduring professional relationships. Positions may be funded by universities, research institutes, or the agencies themselves, and are often advertised through official job boards or professional networks.

This mechanism can enable valuable knowledge transfer and relationship-building. For example, academic researchers with expertise in AI training practices can be integrated into a government agency to inform policy decisions on issues like abusive data practices; in 2021, Lina Khan hired four academic experts on technology policy to help advise AI strategy (Office of Public Affairs, 2021).

Using compelled research authority for access to corporate data

We recommend that researchers and policymakers work together using existing compelled research authority. In the United States, different divisions of the federal government have compelled research authority, which allows them to order companies to grant access to corporate data that is typically unavailable to external researchers. For example, the Federal Trade Commission (FTC) uses its 6b authority, and the Consumer Financial Protection Bureau (CFPB) uses 1022 authority, to compel companies to provide detailed data that can then be used internally to inform public policy and be published with research in the aggregate.³ In 2025, for example, the FTC published a report investigating three AI partnerships between major cloud service providers and genAI companies, such as Microsoft-OpenAI, Amazon-Anthropic, and Google-Anthropic. They used their 6b authority to request

3. The FTC's "6(b) authority," derived from Section 6(b) of the FTC Act, permits the agency to compel companies to provide information for market studies outside of law-enforcement actions. The CFPB has a parallel power under Section 1022(c)(4) of the Dodd–Frank Act, sometimes called its "1022 authority," which allows it to issue similar compulsory orders to support supervision, rulemaking, and market monitoring.

information, including non-public information, from five respondents, which they used to analyse these partnerships (Federal Trade Commission, Office of Technology Staff, 2025). Elsewhere, competition regulators, consumer protection bodies, and sectoral agencies may hold similar authorities.

Researchers can approach policymakers to suggest meaningful questions to be posed to relevant companies. For instance, a researcher might ask, “For companies A, B, and C, how many individuals are using their retail AI tools? What is the monthly profit generated from these tools? How is the training data accessed?” This approach could develop a collaborative relationship where the researcher is able to use their insight and expertise to suggest work that could only be done with access to data that is otherwise unavailable to independent researchers, and the government agency could use their authority to obtain and analyse the data. Alternatively, as with the IPA model above, the researcher could engage in a term-limited assignment with the government agency to work directly on the data and publish findings in collaboration with the agency.

Joining or building collaborative spaces for researcher–government engagement

In addition to formalised mechanisms above, we recommend that stakeholders participate in or create collaborative spaces designed to facilitate sustained interaction and knowledge exchange. Examples of such spaces in the United States include initiatives like the aforementioned Public Technology Leadership Collaborative (PTLC) and the Knight-Georgetown Institute (KGI), both of which aim to bridge the gap between independent research and technology policy. Internationally, similar functions are performed by multi-stakeholder forums, research–policy networks, and communities of practice. These environments encourage dialogue, mutual learning, and the building of professional networks that extend beyond formal projects.

In the same vein, technical and policy conferences can be valuable spaces for open exchanges and relationship-building. DefCon, once known for its tongue-in-cheek “spot the fed” game in which attendees tried to identify undercover government participants, now hosts formal policymaker roundtables and collaborative “villages” that encourage open exchange. These spaces cultivate understanding and deepen relationships between the two communities. By joining or creating these types of collaborative spaces, both researchers and government employees can break down barriers, better understand one another’s perspectives, and create lasting, productive relationships that enhance policy development, technology gover-

nance, and public good.

Conclusion

The task of bridging the gap between AI research and policy requires much more than “evidence-based” policy or a simple translation of findings from academia to the policy arena. As illustrated through case studies of the US AI Safety Institute Consortium and the Public Technology Leadership Collaborative (PTLC), effective translation requires rethinking the traditional, top-down approach in which researchers are seen as experts imparting knowledge to policymakers. Instead, the process must be viewed as a collaborative, iterative exchange that fosters mutual understanding and respect between interdisciplinary actors.

At the same time, we recognise that sociotechnical perspectives are often marginalised for explicitly political reasons, including shifting priorities or competing agendas within government. Mitigating these dynamics requires institutional safeguards, such as embedding sociotechnical expertise into advisory boards and formal review processes, and cultivating broad coalitions of academic, civil society, and practitioner voices that can sustain pressure and visibility even when political winds change.

The PTLC model highlights the value of building trust-based relationships and cultivating a shared context to navigate the complex sociotechnical issues inherent in AI governance. By promoting open dialogue among scholars, practitioners, and government leaders, PTLC suggests that successful translation involves co-creating knowledge and engaging in continuous sensemaking rather than simply delivering findings. Meanwhile, the AISIC’s work underlines the limitations of multi-stakeholder efforts when they fail to successfully engage with disciplinary and political differences. This critique underscores the importance of designing AI policy not only for technical safety, but also for social accountability and inclusivity.

Finally, for academics who want to become more involved in policymaking, we recommend three key approaches: taking advantage of programmes that place academic experts in government roles, using compelled research authority to gain access to corporate data for policy-relevant research, and fostering engagement through collaborative spaces like research institutes and conferences. These mechanisms facilitate knowledge exchange, build long-term relationships, and enhance the integration of AI research into policy development.

Ultimately, effective knowledge translation requires policymakers and researchers

to recognise each other's contexts, constraints, and conceptual frameworks. For translation efforts to succeed, they must consider the relational and organisational factors that shape how knowledge is interpreted and applied. By reframing translation as an inclusive, collaborative process grounded in empathy and shared goals, we can better align AI policy and governance with the human realities and risks that these technologies entail. This shift is essential for building an AI ecosystem that not only mitigates risks but advances the public interest in meaningful and equitable ways.

References

- Ahmed, N., Wahed, M., & Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, 379(6635), 884–886. <https://doi.org/10.1126/science.ade2420>
- Ahmed, S., Jaźwińska, K., Ahlawat, A., Winecoff, A., & Wang, M. (2024). Field-building and the epistemic culture of AI safety. *First Monday*. <https://doi.org/10.5210/fm.v29i4.13626>
- Baron, J. (2018). A brief history of evidence-based policy. *The ANNALS of the American Academy of Political and Social Science*, 678(1), 40–50. <https://doi.org/10.1177/0002716218763128>
- Beard, S. J., & Torres, É. (2024). Ripples on the great sea of life: A brief history of existential risk studies. In S. J. Beard & T. Hobson (Eds), *An Anthology of Global Risk* (pp. 23–74). Open Book Publishers. <https://doi.org/10.11647/OBP.0360>
- Boaz, A., Huw, D., Fraser, A., & Nutley, S. (2019). *What works now?: Evidence-informed policy and practice*. Policy Press.
- Bogen, M., & Winecoff, A. (2024). *Applying sociotechnical approaches to AI governance in practice*. Center for Democracy and Technology. <https://cdt.org/wp-content/uploads/2024/05/2024-05-23-AI-Gov-Lab-applying-sociotechnical-approaches-to-ai-governance-in-practice-final.pdf>
- Bommasani, R., Arora, S., Choi, Y., Fei-Fei, L., Ho, D., Jurafsky, D., Koyejo, S., Lakkaraju, H., Narayanan, A., Nelson, A., Pierson, E., Pineau, J., Varoquaux, G., Venkatasubramanian, S., Stoica, I., Liang, P., & Song, D. (2024). *A path for science- and evidence-based ai policy*. <https://understanding-ai-safety.org/>
- Bowers, J., & Testa, P. F. (2019). Better government, better science: The promise of and challenges facing the evidence-informed policy movement. *Annual Review of Political Science*, 22(1), 521–542. <https://doi.org/10.1146/annurev-polisci-050517-124041>
- Burrell, J. (2023, June 6). *Artificial intelligence and the ever-receding horizon of the future*. Tech Policy Press. <https://techpolicy.press/artificial-intelligence-and-the-ever-receding-horizon-of-the-future>
- California Policy Lab. (2023). *Policy briefs 101: A guide for translating research findings into policy impact*. University of California, Los Angeles. <https://www.capolicylab.org/wp-content/uploads/2023/05/Policy-Brief-Guide.pdf>
- Casper, S., Krueger, D., & Hadfield-Menell, D. (2025). *Pitfalls of evidence-based AI policy* (No. arXiv:2502.09618). arXiv. <https://doi.org/10.48550/arXiv.2502.09618>
- Center for AI Safety. (2023, May 30). *Statement on AI risk*. <https://www.safe.ai/work/statement-on-ai->

risk

Chen, B. J., & Metcalf, J. (2024). Explainer: A sociotechnical approach to AI policy. *Data & Society*. <https://datasociety.net/library/a-sociotechnical-approach-to-ai-policy/>

Crow, D., & Jones, M. (2018). Narratives as tools for influencing policy change. *Policy & Politics*, 46(2), 217–234. <https://doi.org/10.1332/030557318X15230061022899>

Data & Society, Center for Democracy and Technology, Accountable Tech, AFL-CIO, Technology Institute, AI Now Institute, AI Risk and Vulnerability Alliance, American Civil Liberties Union, Center for AI and Digital Policy, Center on Race and Digital Justice, Communications Workers of America, Consumer Reports, Electronic Privacy Information Center, Fight for the Future, Government Information Watch, GovTrack.us, Kapor Center, Mozilla, National Employment Law Project, & New America's Open Technology Institute. (2024, March 18). *Ensuring "AI safety" into the future begins with how we address algorithmic harms now*. Data & Society. <https://datasociety.net/wp-content/uploads/2024/03/Civil-Society-Letter-to-Secretary-Rai-mondo-2024.pdf>

Dean, S., Gilbert, T. K., Lambert, N., & Zick, T. (2021). Axes for sociotechnical inquiry in AI research. *IEEE Transactions on Technology and Society*, 2(2), 62–70. <https://doi.org/10.1109/TTS.2021.3074097>

Dobbe, R., & Wolters, A. (2024). Toward sociotechnical AI: Mapping vulnerabilities for machine learning in context. *Minds and Machines*, 34(2), 12. <https://doi.org/10.1007/s11023-024-09668-y>

Ellen, M. E., Lavis, J. N., Sharon, A., & Shemer, J. (2014). Health systems and policy research evidence in health policy making in Israel: What are researchers' practices in transferring knowledge to policy makers? *Health Research Policy and Systems*, 12(1), 67. <https://doi.org/10.1186/1478-4505-12-67>

Every-Palmer, S., & Howick, J. (2014). How evidence-based medicine is failing due to biased trials and selective publication. *Journal of Evaluation in Clinical Practice*, 20(6), 908–914. <https://doi.org/10.1111/jep.12147>

Farley-Ripple, E., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018). Rethinking connections between research and practice in education: A conceptual framework. *Educational Researcher*, 47(4), 235–245. <https://doi.org/10.3102/0013189X18761042>

Federal Trade Commission & Office of Technology Staff. (2025). *Partnerships between cloud service providers and AI developers: FTC staff report on AI partnerships & investments 6(b) study*. Federal Trade Commission.

Ferrara, E. (2025). Addressing racial bias in AI: Towards a more equitable future. In I. Czarnowski, R. J. Howlett, & L. C. Jain (Eds), *Intelligent Decision Technologies* (Vol. 411, pp. 227–235). Springer Nature Singapore. https://doi.org/10.1007/978-981-97-7419-7_20

Flanagin, A., Fontanarosa, P. B., & Bauchner, H. (2020). Preprints involving medical research—Do the benefits outweigh the challenges? *JAMA*, 324(18), 1840. <https://doi.org/10.1001/jama.2020.20674>

Gilis, T. B., & Spiess, J. L. (2019). Big data and discrimination. *The University of Chicago Law Review*, 86(2), 459–487.

Grimshaw, J. M., Eccles, M. P., Lavis, J. N., Hill, S. J., & Squires, J. E. (2012). Knowledge translation of research findings. *Implementation Science*, 7(1), 50. <https://doi.org/10.1186/1748-5908-7-50>

Hanna, A., Denton, R., Smart, A., & Smith-Loud, J. (2020). Towards a critical race methodology in algorithmic fairness. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*,

501–512. <https://doi.org/10.1145/3351095.3372826>

Haynes, A., Rowbotham, S. J., Redman, S., Brennan, S., Williamson, A., & Moore, G. (2018). What can we learn from interventions that aim to increase policy-makers' capacity to use research? A realist scoping review. *Health Research Policy and Systems*, 16(1), 31. <https://doi.org/10.1186/s12961-018-0277-1>

Head, B. W. (2010). Reconsidering evidence-based policy: Key issues and challenges. *Policy and Society*, 29(2), 77–94. <https://doi.org/10.1016/j.polsoc.2010.03.001>

Horsfall, L., Bondaronek, P., Ive, J., & Poduval, S. (2025). Clinical algorithms and the legacy of race-based correction: Historical errors, contemporary revisions and equity-oriented methodologies for epidemiologists. *Clinical Epidemiology*, Volume 17, 647–662. <https://doi.org/10.2147/CLEP.S527000>

IPA Mobility Program. (2025). *Intergovernment Personnel Act*. US Office of Personell Management. <https://www.opm.gov/policy-data-oversight/hiring-information/intergovernment-personnel-act/>

Jecker, N. S., Atuire, C. A., Bélisle-Pipon, J.-C., Ravitsky, V., & Ho, A. (2024). AI and the falling sky: Interrogating x-risk. *Journal of Medical Ethics*, 50(12), 811–817. <https://doi.org/10.1136/jme-2023-109702>

Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social Media + Society*, 9(3), 20563051231186353. <https://doi.org/10.1177/20563051231186353>

Kneese, T. (2024, February 12). *Measuring AI's environmental impacts requires empirical research and standards*. Tech Policy Press. <https://techpolicy.press/measuring-ais-environmental-impacts-require-s-empirical-research-and-standards>

Kreps, S., & Kriner, D. (2023). How AI threatens democracy. *Journal of Democracy*, 34(4), 122–131. <https://doi.org/10.1353/jod.2023.a907693>

Lee, H.-P. (Hank), Yang, Y.-J., Von Davier, T. S., Forlizzi, J., & Das, S. (2024). Deepfakes, phrenology, surveillance, and more! A taxonomy of AI privacy risks. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3613904.3642116>

Natow, R. S. (2022). Policy actors' perceptions of qualitative research in policymaking: The case of higher education rulemaking in the United States. *Evidence & Policy*, 18(1), 109–126. <https://doi.org/10.1332/174426420X16047229138089>

Nguyen, A., & Mateescu, A. (2024). Generative AI and labor: Power, hype, and value at work. *Data & Society Research Institute*. <https://datasociety.net/library/generative-ai-and-labor/>

Nopper, T. K. (2024). Medicalizing inequity: The risks of financial wellness for workers. *Data & Society Research Institute*. <https://datasociety.net/library/medicalizing-inequity/>

Nutley, S., & Webb, J. (2000). Evidence and the policy process. In *What works? Evidence based policy and practice in public services* (pp. 13–41). <https://bristoluniversitypressdigital.com/display/book/9781847425119/ch002.xml>

Oduro, S. (2024). Without accounting for bias and discrimination, AI safety standards will fall short. *Tech Policy Press*, September 16. <https://www.techpolicy.press/without-accounting-for-bias-and-discrimination-ai-safety-st>

Oduro, S., & Kneese, T. (2024). *AI governance needs sociotechnical expertise*. Data & Society. <https://datasociety.net/library/ai-governance-needs-sociotechnical-expertise/>

Office Public Affairs. (2021). *FTC chair Lina M. Khan announces new appointments in agency leadership positions*. Federal Trade Commission. <https://www.ftc.gov/news-events/news/press-releases/2021/11/ftc-chair-lina-m-khan-announces-new-appointments-agency-leadership-positions>

Oliver, K., & Cairney, P. (2019). The dos and don'ts of influencing policy: A systematic review of advice to academics. *Palgrave Communications*, 5(1), 21. <https://doi.org/10.1057/s41599-019-0232-y>

Oliver, K., Innvar, S., Lorenc, T., Woodman, J., & Thomas, J. (2014). A systematic review of barriers to and facilitators of the use of evidence by policymakers. *BMC Health Services Research*, 14(1), 2. <http://doi.org/10.1186/1472-6963-14-2>

Pressman, J. L., & Wildavsky, A. B. (1973). *Implementation: How great expectations in Washington are dashed in Oakland; or, why it's amazing that federal programs work at all*. University of California Press.

Rossi, P. H. (1987). The Iron Law of Evaluation and other metallic rules. *Research in Social Problems and Public Policy*, 4(1), 3–20.

Rushmer, R., Ward, V., Nguyen, T., & Kuchenmüller, T. (2019). Knowledge translation: Key concepts, terms and activities. In M. Verschuuren & H. Van Oers (Eds), *Population Health Monitoring* (pp. 127–150). Springer International Publishing. https://doi.org/10.1007/978-3-319-76562-4_7

Saheb, T. (2023). Ethically contentious aspects of artificial intelligence surveillance: A social science perspective. *AI and Ethics*, 3(2), 369–379.

Saltelli, A., & Giampietro, M. (2017). What is wrong with evidence based policy, and how can it be improved? *Futures*, 91, 62–71. <https://doi.org/10.1016/j.futures.2016.11.012>

Sartori, L., & Theodorou, A. (2022). A sociotechnical perspective for the future of AI: Narratives, inequalities, and human control. *Ethics and Information Technology*, 24(1), 4. <https://doi.org/10.1007/s10676-022-09624-3>

Septiandri, A. A., Constantinides, M., Tahaei, M., & Quercia, D. (2023). WEIRD FAccTs: How Western, Educated, Industrialized, Rich, and Democratic is FAccT? *2023 ACM Conference on Fairness, Accountability and Transparency*, 160–171. <https://doi.org/10.1145/3593013.3593985>

Shah, M., & Sureja, N. (2025). A comprehensive review of bias in deep learning models: Methods, impacts, and future directions. *Archives of Computational Methods in Engineering*, 32(1), 255–267. <https://doi.org/10.1007/s11831-024-10134-2>

Tseng, V. (2012). The uses of research in policy and practice. *Sharing Child and Youth Development Knowledge*, 26(2), 1–24.

Webb, L., & Schönberger, D. (2025). Generative AI and the problem of existential risk. In P. Hacker, A. Engel, S. Hammer, & B. Mittelstadt (Eds), *Oxford Handbook on the foundations and regulation of generative AI*.

Westerstrand, S., Westerstrand, R., & Koskinen, J. (2024). Talking existential risk into being: A Habermasian critical discourse perspective to AI hype. *AI and Ethics*, 4(3), 713–726. <https://doi.org/10.1007/s43681-024-00464-z>

Wörsdörfer, M. (2025). Biden's executive order on AI: Strengths, weaknesses, and possible reform steps. *AI and Ethics*, 5(2), 1669–1683. <https://doi.org/10.1007/s43681-024-00510-w>

Published by



ALEXANDER VON HUMBOLDT
INSTITUTE FOR INTERNET
AND SOCIETY



RESEARCH
FOR THE
DIGITAL AGE

in cooperation with



CREATE



centre
— internet
et societe



R&I

IN3

Internet
interdisciplinary
Institute

Universitat Oberta de Catalunya



UNIVERSITY OF TARTU
Johan Skytte Institute of
Political Studies