JOURNAL OF DIGITAL SOCIAL RESEARCH

JDSR 6(4) (2024) 77-91 10.33621/jdsr.v6i440468

AI policymaking as drama

Stages, roles, and ghosts in AI governance in the United Kingdom and Canada

Alison Powell[1] and Fenwick McKelvey[2]

[1] London School of Economics, United Kingdom
[2] Concordia University, Canada

✉ fenwick.mckelvey@concordia.ca

Abstract

As two researchers faced with the prospect of still more knowledge mobilisation, and still more consultation, our manuscript critically reflects on strategies for engaging with consultations as critical questions in critical AI studies. Our intervention reflects on the often-ambivalent roles of researchers and 'experts' in the production, contestation, and transformation of consultations and the publicities therein concerning AI. Although 'AI' is increasingly becoming a marketing term, there are still substantive strategic efforts toward developing AI industries. These policy consultations do open opportunities for experts like the authors to contribute to public discourse and policy practice on AI. Regardless, in the process of negotiating and developing around these initiatives, a range of dominant publicities emerge, including inevitability and hype. We draw on our experiences contributing to AI policy-making processes in two Global North countries. Resurfacing long-standing critical questions about participation in policymaking, our manuscript reflects on the possibilities of critical scholarship faced with the uncertainty in the rhetoric of democracy and public engagement.

Keywords: policy; drama; AI governance; Canada, United Kingdom; critical policy studies; hauntology

1. Introduction

The Honorable Minister of Innovation, Science and Industry François-Philippe Champagne (2023) rose to introduce Canada's new AI regulation, the Artificial Intelligence and Data Act, before a Canadian parliamentary committee. Champagne extolled its virtues and the process: "I think the bill shows that we have listened to Canadians and parliamentarians, as we have done for other legislation. I gather there were more than 300 meetings or consultations with individuals who provided input into the process." Many questioned when and where those meetings happened. Critical technologist Andrew Clement finally received the list – only to discover that the consultations mostly happened with industry (and after the Bill's release) (Castaldo, 2023). One of the authors appeared on the list of consultees – but for discussions that took place after the Act had launched. So much for being an influential expert. Instead, what had been perceived as a friendly meeting became part of a spectacular number, part of the opening act in the Act's dramatic passage. This scene is one of many that have caused to your authors to wonder:

How can two policy scholars make sense of AI governance, of technology policy, if the consultations seem so staged?

What began in our self-reflection has become a deeper attempt to understand the particular complexity of AI as technology governance. Your authors started to reflect on the roles afforded to academics and critical technology policy scholars asked to participate in the emerging field of AI governance, when our expertise in policymaking also rests on our capacity to be continuously up-to-date with emerging technologies that are often being promoted in ways that mystify, rather than clarify their social impact (Powell, forthcoming). Our collaboration – emerging from shared debriefs about being played-out or outplayed in what-can-seem-like consultation theatre – turned toward a broader reflection on what AI governance is as a certain mode of publicity, of a way of public-sense making about a new technology. Rather than focusing on matters of framing or problematization (Bacchi, 2009; van Hulst & Yanow, 2016), we have instead considered the dramatic function of policy debates to specifically question the roles and opportunities provided to scholars participating in these debates.

Our contribution to the special issue is a reframing of the terms of the AI policy debate from two technology policy scholars, on two sides of the Atlantic contending with wave after wave of AI controversy and consultation. Our tempest is the drama that has become AI governance. Our reflections take place amidst growing demand for consultations to legitimate artificial intelligence and specifically generative AI by governments and industry (cf. Palmås & Surber, 2022). Our reflection concerns AI's policy publicities as well as our sense of a growing ambivalence in the face of an expectation to play one's part in the legitimation of technoscientific capitalism (Birch, 2017; Nagy & Neff, 2024). Our article unfolds as follows: we first theorize our position on policy making as drama, then discuss how we can use this to understand current AI policy research agendas. Next, using examples from AI policy-making dramas in both the UK and Canada, we discuss how scenes are set through reference to different issues, giving benefit to some players over others. We then outline the roles that academics are invited to play as consultants, 'good academics,' maybe even as policy operators. This offers a new way to understand AI governance by attending to the drama as process, the staging of scenes that coordinate participation, the roles afforded to academics, and, oh yes, the ghosts. The stages where policy dramas play out can include ghosts - unacknowledged presences of past and active exclusions. We conclude by reflecting on how our actions as policy actors can be haunted by our past efforts, and what we might gain (or lose) from ghosting the process.

2. Understanding AI policy as drama

We theorize (along with many others) that policy making, like so many human endeavours, is dramatic. There are policy actors, there is a stage, and there are different scenes that unfold as the narrative progresses. We find ourselves on these stages from time to time. Different scenarios form set piece scenes in the narrative: public consultations where the chorus is invited to recite lines drawn from the rough-and-tumble world of experience, or closed, Chatham-House-Rules meetings where side conversations in the corridors reveal or obscure the potential for agreements, misalignments, or capitulations. Sometimes, metaphorically speaking, someone is concealed behind an arras. If occasionally the drama bursts on to screens, more often the stages seem mundane.

Drama is an antagonist for conventional policy scholars. Harold Lasswell co-constructed the policy sciences as expert-led, scientific project for governments and academics to govern societies (Turnbull, 2008). Merely suggesting that policy making might involve power plays breaks character with this part of the field's foundations. Certainly, we risk tarnishing our reputation as competent policy scholars by emphasizing the drama, but you can only watch so many AI executives act out their version of Oppenheimer-like regret before questioning the policy sciences. Instead, we observe how Schön and Rein (1994) argue that making better policy does not require always making 'better' rational decisions, but instead involves the policy participant having a more refined awareness of their own position in the

https://doi.org/10.33621/jdsr.v6i440468

drama. Stone (1997) argues that policy making always involves conflicts over ideas, which are enacted through staged arguments. These orientations towards the dramatic, along with the acceptance of the diverse and often competing roles that play out across multiple scenes, help to clarify what policy making - and policy conflicts - are doing.

Our approach then differs from much of the other policy literature on AI in considering AI governances as a dramatic process. Much of the existing literature focuses on the consequences of policies about AI (Acemoglu, 2021; Angwin et al., 2016; Buolamwini & Gebru, 2018). On one hand, research has focused on AI governance as addressing existential challenges (not risk) where research aims to define key problems (e.g., explainability, traceability, or bias) that necessitate governance reforms (Büthe et al., 2022; Gasser & Almeida, 2017; Knight et al., 2024). At the most abstract level, AI has prompted a fruitful debate over the nature of contemporary governance and governmentality (Amoore, 2022; Fourcade & Gordon, 2020; Yeung, 2018). On the other, there is a growing turn toward understanding the discourses of national AI strategies as evidence of empirical policy trends. Studies range from comparative evaluation to more critical work describing the underpinning imaginaries that legitimate these polices (Bareis & Katzenbach, 2021; Hansen, 2021; Vicente & Dias-Trindade, 2021). We find inspiration in both joining to a class of spectators wondering how long AI government can maintain the status quo while claiming AI is revolutionary while also attending to the everyday efforts of AI governance. Amidst these two trends, our approach joins a growing interdisciplinary interest in the practical roles of AI governance (Sætra et al., 2022; Widder, 2024).

In this paper, we position our lived experiences and academic engagement in relation to what might be called a performative turn in critical policy studies (Clarke et al., 2015; Lea, 2020; Lipsky, 2010). John Clarke, Dave Bainton, Noémi Lendvai, and Paul Stubbs inspired us in their writing on policy as: "a way of imagining the world as an object of intervention; as a way of enrolling subjects into a process of acting; and as a practice that seeks to produce effects, including the act of 'taking the politics out of things'" (2015, p. 34). This allows policy to be understood as something that itself moves or contains movement and dynamism, "unlocking the narrow confines within which 'policy' is conventionally enclosed - the rational, technocratic world" (2015, p. 33). Through this we can explore how contradictory entities tussle between and within themselves to create ideas that position and discipline different groups of people in different ways. We aim to provide a language to understand AI governances as lived drama, beginning with, an assessment of how policy is staged.

2.1 Breaking the fourth estate / wall: Policy, stages and scenes

"Mr. ChatGPT goes to Washington" read the headline of the CNN title about Sam Altman's testimony before Congress (Fung, 2023a). Mr. Altman showed up before the television cameras to play the role of a good CEO. Why not compare the testimony to a movie? Altman had to wear a suit – a costume his now-jailed contemporary Sam Bankman-Fried only started to put on when he, too, went on trial. Altman exceeded expectations in his role. One glowing review of Altman's performance reads "OpenAI CEO Sam Altman seems to have achieved in a matter of hours what other tech execs have been struggling to do for years: He charmed the socks off Congress" (Fung, 2023b). What a way to win over the audience of regulators. Altman's success in staging gives us a good reason to understand AI policy as drama. The suit helped put him the role as an ethical technoscientific capitalist. Regardless of what might have transpired inside OpenAI, Altman was obliged to enact a responsible proxy for the firm. Altman's successful performance – one that a critic might argue displaced attention from deep concerns about the direction of the industry – is presently what has us wondering about the extent to which policy success is in the staging, and how this serves to conform to expectations of the performance – and thereby to influence the outcome.

If critical policy studies have a public role in responding to AI (Paul, 2022), then we encourage a reflection on how policy consultations are staged – often as media events – and to what extent this staging

https://doi.org/10.33621/jdsr.v6i440468

displaces attention that might otherwise focus on the "epistemic forgeries" (Katz, 2020), at the heart of AI. These forgeries include assumptions that AI innovation is inevitable – and that it can be made unproblematic or even 'safe' if only policy processes can be properly focused. The staging of policy dramas reflects this effort to contain and constrain policy making. Consider some of the settings such as the computer science conferences run by the ACM and the closed meetings hosted by technology companies, venerable Institutions, or government agencies where the drama unfolds away from the public eye. Of course, some actors have the capacity to move across various settings, shaping how the dramas are staged: OpenAI, Meta, Alphabet, Microsoft. If you watch long enough, you might see actors change into corporate costumes after playing other roles - perhaps in research, or within regulatory bodies – revolving doors operate on these stages too (Bates, 2013).

The AI policy making stage has also been growing, and the players are striding out into a runway cutting all the way through the theatre. The fourth wall is breaking. Members of the audience are leaping up and declaring interests in being part of this drama, too, proclaiming an interest in the proceedings. But to be sure, this is not Brechtian theatre where the audience can play along. A friction endures between the construction or instantiation of a public and policy formation (Marres, 2005). There is growing attention to the "consultation theatre" around AI, a charade where the consultation serves more as a chance for governments to perform their role as stewards of public opinion and showcase their ability to respond, but not actually to affect to shape policy (M. Jones & McKelvey, 2024). Not being able to look away might be the best defense of democracy these days when the public's only motivation is to watch the drama as a spectator.

Policy drama might also be an effect of how the mass media plays expectations on the policy process, what Maarten Hajer (2009) identifies as a 'mediatized' policy making. From this perspective, the cinematic aspect of Altman's US Congress appearance is not accidental. Steven Coleman (2013) describes how politics construct drama, and how cultural performance applies not only to the micro-interactions of interpersonal relationships but also to macro-political organizations like television. People justify their interactions through drama, or what Isabela and Norman Fairclough (2013) call the 'practical reasoning' behind these processes. A reasoning that, we find elsewhere, journalism has largely failed to provide, getting publics to worry about existential risks and taking the promise of large-language models as the next step to general artificial intelligence at face value (McKelvey & Roberge, 2023). But we are foreshadowing too much; let us show how AI policy scenes are set, and how the change of these scenes influences which voices are heard on stage.

2.2 From ethics to existential risk to safety - the UK policy scenes 2016-2024

One of your authors has watched many of these scene changes in the UK over the past decade. Concern about ethics and AI started as a problem for the 'good society' (Cath, 2018; Floridi, 2014). The policy scene expected players to describe AI's ethics and values (Jobin et al., 2019; Mittelstadt et al., 2016). Action in the mid-2010s focused on ethical principle statements for AI (European Union, 2019; UK Government, 2019), and subsequent discussions of how to put them into practice (Stark & Hoffmann, 2019). For a brief period, corporate actors played along, issuing statements and constituting company ethics boards (Taylor & Dencik, 2020). By 2020, however, AI harms were accruing regardless of these ethics statements (Morley et al., 2023).

The academic aspects of the policy scene started to consolidate into discussions drawing on disciplinary expertise, for example, involving technical concerns about model design or auditing capacity (Sloane, Moss, Awomolo, et al., 2022; Sloane, Moss, & Chowdhury, 2022) – changing the language and behavior of the scene in ways that sometimes left players like your authors in the wings (Gansky & McDonald, 2022).

Then the scene changed. Different stories tried to nuance the discussion of ethics, expanding it to cover various areas of practice, and then to attempt to connect these suggested areas of practice to regulatory

https://doi.org/10.33621/jdsr.v6i440468

capacities (Sætra et al., 2022). In the UK, these efforts drew from traditions of technology policy-making focused on responsibility (Owen et al., 2013; Stilgoe, 2018, 2019) and engineered-in trustworthiness (Naiseh et al., 2022; Schneiders et al., 2023).

These scenes structured what was focused on, by whom, and at what time. A mismatch with the unfolding scene might create confusion or result in being moved off stage as we have experienced. One author was part of the JUST AI project, a pilot hosted by then-recently-founded Ada Lovelace Institute. JUST AI employed feminist lab methods to convene conversations and hold space for dissent and foregrounded the inclusion of arts and artistic methods in reframing how the conversations about AI ethics were occurring – away from the statements of principle and corporate boards and in a collaborative, collegial, and slow-moving manner. JUST AI's networked approach was a backdrop against which to choreograph the entries of new actors whose script and position could generate research, conversation, and action on racial justice, repairability, and environmental sustainability as aspects of AI ethics work. The scenes in play in the early 2020s UK ethics drama complicated this. A JUST AI Fellowship focused on racial justice in AI supported artists and creative researchers interested in the use of AI in border technologies, the policy gaps relating to differential impacts of AI on racialized populations, and how the white supremacist origins of AI might be re-imagined (Chander & Nation, 2021), which challenged some of the more reductive discussions about bias mitigation (Buolamwini & Gebru, 2018; Gong et al., 2020). The unfolding conversations on these topics joined other working group discussions of the environmental ethics of AI (including its repairability), and the a disability-led discussion of the potential for both rights of access to AI and rights to its refusal (Hickman & Serlin, 2018; Newman-Griffis et al., 2023). These ethical conversations were bracing, contentious, and sometimes difficult to place within the policy scene. JUST AI events convened wide public audiences and influential academic participants (JUST-AI, 2023), but the project's focus on systemic injustice and large-scale environmental impact sometimes felt out of synch with other work at Ada Lovelace that engaged directly in the policy scenes attracting attention at the time.

The scene changed again; the revolving door whirled, and by 2023 the focus was on risk and safety, rather than ethics. This was solidified in the European Union's AI Act, which rested on an assumption about managing and regulating riskier applications of AI. But this script was influenced by actors with particular interests (Hendrycks & Mazeika, 2022; Turchin & Denkenberger, 2020) and increased attention to the capacities of Large Language Models, especially ChatGPT. A new set of actors introduced the idea of existential risks to human society that could result from significant technical developments in AI. Led by Nick Bostrom in Oxford and backed by significant funding from the technology industry, the 'long-termist' approach focused on the existential risks to the entire human race of unimpeded, speculative AI development (Gebru & Torres, 2024). By late-2023 the language had shifted from existential risk to safety – with the initial discussion still focused on large-scale threats and 'doomsday' scenarios now also accompanied by concerns about the development of multi-purpose or 'foundation' models in response to the commercialization of ChatGPT. In the dramatic and media-friendly AI Safety Summit convened by the UK Prime Minister in November 2023 highly visible actors called for the testing of AI 'foundation' models and the management of AI "for good" in a 'pro-innovation' framework (Hawes & Hall, 2023). The structure of the AI Safety Summit as an invitation-only meeting of many industry representatives, a handful of high-level politicians and selected academics (not your authors), foregrounded a cheery, media-friendly narrative of 'safety'. Very few of the UK's diverse range of civil society stakeholders were invited. A parallel set of events branded as the AI Fringe (2023) was organised by a research consultancy and bundled together sector-specific discussions of including AI and work, access rights, and public accountability.

The resulting "Bletchley Declaration" focused on the safety of 'foundation models' and called for the establishment of AI safety institutions – but might have had negligible impact on the UK's current AI regulatory frame, which focuses on principles of safety, transparency and explainability, fairness and accountability, and contestability and redress (UK Government, 2024). Results from the AI Fringe have

https://doi.org/10.33621/jdsr.v6i440468

included the TUC's AI Regulation and Employment Right Bill (Allen & Masters, 2024). Globally, more attention has been paid to the environmental impact of the use of AI systems, sometimes as an aspect of safety (Falk & van Wynsberghe, 2023) – demonstrating how a shift in the scene can bring an idea in from the wings.

These shifting scenes and focal points have required participants in policy dramas to develop a kind of improvisational agility, moving between scenes along with actors whose positions might also be changing. In the case of JUST AI, for example, the project's funding and institutional support rested on a partnership between the Ada Lovelace Institute, its background funder the Nuffield Foundation, and the UK's Arts and Humanities Research Council. The think-tank, the funders and the project all maneuvered across a rapidly changing policy scene, where reflections about ethics and the creation of new ways of networking across disciplines ceded space to direct interventions in government policy and technically-driven discussions about audit or features design (see Groves et al., 2024). JUST AI's contributions, including its Fellowship and commissioning of creative work, advanced ideas such as environmental sustainability that were not, at the time of the project, direct objects of policy development, although they have subsequently emerged as such. However, efficacy in policy making is often judged against experiences in the present rather than potential in the future, especially in a reactive context where researcher and other experts are called upon to continue to define or defend the 'public interest'. JUST AI's focus on environmental sustainability and repairability took some time to become salient within UK policy drama. To date, its work on rights, access, and refusal of AI systems from the perspective of critical disability studies has not yet become so. This is testament to what we've explored through this section – the ways that policy ideas, practices and potentials are captured within shifting scenes. Of course, within each of these scenes, scholars like us are called on to perform specific roles.

3. Some tentative roles for aspiring AI policy scholars

Your authors have been called to stage to perform as expert witnesses in Canada and the United Kingdom on a range of technology-influenced social issues, over the past decade and more. Now though, the intensity of AI consultations – this moment when governments and corporations are trying to advance AI's acceptability – have us questioning what it means to play a role to participate in AI policymaking.

Roles are both individual and collective. The role "refers to a cluster of behaviors and attitudes that are thought to belong together, so that an individual is viewed as acting consistently when performing the various components of a single role and inconsistently when failing to do so." (R. H. Turner, 2001, p. 233). Within these roles, Peter Burke claims that actors (or agents), exercise their agency in ways that are proper to the scenes they are participating in. "It is a principle of drama that the nature of acts and agents should be consistent with the nature of the scene" (Burke, 1945, p. 3). As your authors don and doff various costumes and move between various roles, we keep in mind these insights. Language matters in these settings, as Coleman identifies, but performance does too. Performing is "a culturally pragmatic strategy for the production of shared meaning" (Coleman, 2013, p. 330). Policy making, especially technology policy making, contends with both increasingly mediatized aspects of decision making (where strategic media visibility can mean leveraging policy positions) and with the influence of entrepreneurial innovation contexts (see Irani, 2015), which have also impacted the roles that professors and university researchers occupy. This generates conflicts between roles that different actors can play, and the extent and quality of their agency.

Based on our shared experience, we offer a few possible roles a policy scholar might play. In the next section, we draw on our experiences and observations as experts in the spaces to identify where academics, researchers, and intellectuals – domain experts, in other words – are invited and expected to act in AI policy-making dramas. We also explore the space of ambivalence or haunting that can accompany this performance of expertise and the strange experiences that can result when our actions

https://doi.org/10.33621/jdsr.v6i440468

seem out of line with roles, are haunted by what came before, or suggest potential futures that jar with the scenes at play.

3.1 The Consultant / Hired Gun / The Ronan

The first role is that of an intellectual worker as a consultant-for-hire. Getting paid is not the issue (participation should always be fairly compensated); rather, the concern relates to the point at which expertise becomes sponsored research (see Goldenfein & Mann, 2023). The role of the consultant acknowledges that much of academia maintains relationships with corporate partners that might aid or support broader lobbying agendas. Academics have been accused of being too close to the AI industry, or what Meredith Whittaker describes as "the steep cost of capture" where "the tech industry's dominance in AI research and knowledge production puts critical researchers and advocates within, and beyond, academia in a treacherous position" (2021, pp. 51–52). The issue has become particularly pronounced in regards to sponsorship from artificial intelligence firms, where the Washington Post finds that 6 out of 10 AI researchers with disclosures of funding come from AI (Menn & Nix, 2023). Dependence does not, necessarily, determine participation in policy, but more structurally, these findings establish a synergy between corporate and academic institutions that may translate into tacit support. Luke Stark at Western University made an important statement in 2021 when he turned down $60,000 of Google funding after Google fired many of its AI ethics teams (Metz, 2021). Stark's decision illustrates the rarity and potential of refusal, highlighting the fact that more commonly decisions are made that involve accepting funding and then later navigating fraught relationships.

3.2 The Good Academic

A proper academic would never have written this section. They would review submissions to policy consultations to identify emergent frames, ideally using double-blind coding, and then present a weighted summary of positions and key clusters. We did not do that. Instead, we quickly popped backstage and glanced over the script to see what scene we were playing now based on a story in the media. We saw a ghost in the corner, a memory of how this played out last time. We know how to play the "good academic" we are just not doing it right now.

One role for "the good academic" is to act as an independent voice, playing a reliable role as a good policy actor. This role is part of the foundation of policy studies, which Laswell argues are "concerned with knowledge of and in the decision processes of the public and civic order" (1971, p. 1) . Lasswell called for these positions to be defended, to be independent, and capable of being objective. This means that the 'good academic' is meant to be doing the work of doing research, in line with the expectations and norms of their field. The 'good academic' role differs from an on-demand consultant because their obligations come from the field, empiricism, and objectivity, rather than the contract.

A final consideration and one that is a looming challenge for objective AI research in general, is how "good academics" perform objectivity amid debates over methodology and critiques of their approaches. Presenting technical expertise in policy-making settings can put great strain on the role. An example of such a strain came from a performance one of the authors made during policy debates about Net Neutrality – a precursor to AI debates in requiring proof of the presence of bias in an infrastructural technology (McKelvey, 2018). Fenwick has an embarrassing tale of working with computer scientist Dave Choffnes to prove that one of Canada's mobile service providers had been modifying some of its telecommunication equipment to prioritize certain applications over others (Li et al., 2016). Done at the last minute, this transcript is the only record remaining of the brief exchange:

> 103 McKelvey: I understand some of the – my argument here will introduce new evidence and I'm wondering if I have permission to introduce this evidence in support of our submissions earlier?
>
> 104 THE CHAIRPERSON: Why were you unable to provide this evidence earlier?

https://doi.org/10.33621/jdsr.v6i440468

105 McKelvey: It's a great question. Working with a – by any serendipity honestly, I'm working in a new project working in measurement in Canada trying to understand and apply new internet measurement techniques to understand the operations of systems such as differentiate service provisions.

106 And it was only in early September that I knew of a methodology that is going to be presented at the internet measurement conference this November 2016 that I was – found out about it and it was basically over the past last month that I was able to actually deploy and use this measurement and try to support arguments we've made previously as part of our submission.

107 THE CHAIRPERSON: Okay. Well, we'll take your arguments under advisement but go ahead and ---

108 McKelvey: Okay.

109 THE CHAIRPERSON: But that doesn't mean ultimately we will necessarily admit it on the record.

With the evidence presented, questions turned to questions of telecommunication policy, and the infrastructure operator rejected Fenwick's assertion. Another opportunity for leveraging evidence in policymaking passed. The stage did not want to discuss academic methods, now was the time for deliberation of the law.

The example provides a warning to policy scholars of the limits of the "good academic" role here is not likely to be consensus on what counts as safe or risky AI, and no specific piece of evidence is going to be accepted as objective. Indeed, one strategy in climate science communication is for lobbyists to introduce facts as controversial, or debatable - just to raise enough doubt to thwart the efforts of the policy sciences (Hess, 2014). Fenwick's story should act as a corrective for any "good academic" hoping that some data and a novel method will expose bias and convince regulators. Instead, the method was not legible, indeed the whole project of offering "objective" Internet Measurement has become something of a hobby project since its only large Internet Service Providers willing to invest in the methods – maybe a warning to who is going to fund and legitimate the next AI audit?

3.3 The Policy Operator

The strains of the 'university in ruins' (Readings, 1997) drive the emergence of the policy operator, our third role. In this mode "the concept of scholarship as cultural production . . . is articulated through ideas about economically impactful deliverables, including a narrowly defined set of public engagement and outreach activities" (Luka et al., 2015, p. 177) . This role can involve working for hire as contractor, or alternatively an embrace of these norms to become something of a policy entrepreneur, speculating on the potential to develop new spaces of intervention. This can happen when "visibility and self-promotion naturally accompany a corporatized culture, requiring deliverables and impact to stand in for a broader conception of scholarship as cultural production" (Luka et al., 2015, p. 191). What could be more measurable than policy impact? The policy operator combines the pressures of the neoliberal university, the need for self-branding and greater visibility with the demand for good-faith participants in government policy. The mix, when successful, can be remarkable, producing a scholar able to manufacture a brand of authority, steer the policy agenda, and gain prominence through the new metrics of knowledge mobilisation. The policy operator then becomes a name who must be consulted. Your authors recognize our own ambivalence towards the policy operator that matches a sense of the position's power. We also recognize a concern over the demands to keep up, to be in the game, and to support or even benefit from a concentration of power in academia (noting its attendant inequalities, especially for those of us also tasked with managing students, programs, and departments alongside our performances in policymaking).

Canada's AI policy has been driven by policy operators, whose positions can grow to shape domestic and international policy scenes. Recently we have witnessed a strange performance by two of its three "godfathers of AI". These figures had occupied an oversized role in Canada's media as evangelists (Dandurand et al., 2023). Geoffrey Hinton and Yoshua Benigo then broke with the third 'godfather' Yann LeCun to publicly worry over AI's existential risks. Their rise to celebrity can matched 2023's summer blockbuster, Oppenheimer – or what was called the "Oppenheimer moment" (Tharoor, 2023). Hinton and

https://doi.org/10.33621/jdsr.v6i440468

Benigo filled the role of today's Oppenheimers for AI – a role that had them make frequent testimony about the risks of AI.

The Oppenheimer moment helped legitimate Canada's AI policy agenda. Canada tabled its first AI legislation, the Artificial Intelligence and Data Act (AIDA), in Spring 2022. AIDA attracted little public attention until Benigo signed the Institute for the Future's Open Letter to Pause AI Experiments a year later, Spring 2023. As the letter attracted international attention, Bengio joined most of Canada's AI establishment to sign a second letter calling for the swift passage of AIDA as the kind of urgently needed regulation called for in the first Open Letter. The combination of letters became a mediated spectacle. The once-AI-evangelists offered a pious second act of regret and reflection, all to compel the bill's swift passage – a mix of personal drama and public pressure that seems a hallmark of the policy operator. Champagne (2023) in the speech mentioned above expressly mentions the first letter, claiming "You don't need to take it from me, but I would advise you to read the letter that was signed by Yoshua Bengio, and hundreds of people from around the world, warning us that we need to take action." The policy operator is one who knows how to command the drama to advance their policy agenda. (We say this as an observation, not commentary. Off-stage, the two letters reminded one author of what it means to not be a policy operator after signing a civil society letter calling for reforms to AIDA not mentioned in the Act's introduction.)

Our first three roles: the consultant, the good academic and the policy operator, create mechanisms to carry micropolitical sensibilities into the macropolitical, providing relatively well-specified opportunities for academics to leverage their sense of public accountability and influence. However, we notice that there are other, more ambivalent positions that involve divergent interpretations of the public interest. These other positions might engage what we refer to, following Tess Lea, as policy hauntology. This can occur through echoes of past policy paths not taken or through the framing of potential future policy objects. The unstable temporality of interventions can make these efforts seem haunted by ghosts of past or future.

4. A policy hauntology: Ghosts, ghosting and ambivalence

Critical policy theorist Tess Lea understands policy hauntology, as "the ways policies past and present are physically incorporated, having (insidiously or noisily), seeped into lives, affecting probable destinies and shaping overall circumstances, if not immediately then certainly as a powerful stimulus" (2020, p. 117). Jacques Derrida introduced hauntology at one of Marx's many wakes. Hauntology, to Derrida, meant, something else, "this element itself is neither living nor dead, present nor absent: it spectralizes" (1994, p. 63). Derrida uses the figure of the ghost as an object of hauntology "because it points toward a thinking of the event that necessarily exceeds a binary or dialectical logic, the logic that distinguishes or opposes effectivity or actuality" (1994, p. 78). The purpose here is to surface our shared collective frustration in having to perform AI as a new problem, as requiring "new" innovations, while haunted by past failed technology policies. Sometimes the whole drama seems just like another mediocre Hollywood update in the tired franchise of public interest technology regulation.

Technology policy is haunted by past technology. In Canada, this was illustrated by an attempt to AI and smartphones to automate COVID-19 contact tracing by applying AI models to tracking smartphone proximity. To work, the system needed a certain threshold of public adoption – a novel idea, but one haunted by Canada's ongoing challenges to extend smartphone and internet connectivity to all, and the mobile services regulatory frameworks that have resulted in unaffordable mobile services. The problem was not AI, but what happened before AI. Similarly, the difficulty in advancing issues of the environmental cost of AI in the UK was related to a policy context wherein the government at the time focused on extending oil and gas exploration rather than discussing climate commitments. Our sense of haunting has a lot to do with critical scholar Eve Tuck and C. Rae's interpretation of the function of horror in understanding cultures of justice. They write,

https://doi.org/10.33621/jdsr.v6i440468

> The difference between notions of justice popularized in US horror films and notions of justice in these examples of horror films from Japan is that in the former, the hauntings are positioned as undeserved, and the innocent hero must destroy the monster to put the world in balance again (though predictably, several of the hero's companions who are women or people of color will likely be sacrificed along the way). In the latter, because the depth of injustice that begat the monster or ghost is acknowledged, the hero does not think herself to be innocent, or try to achieve reconciliation or healing, only mercy, often in the form of passing on the debt. (Tuck & Ree, 2013, p. 642)

The haunting is a force and a demand - to acknowledge past failures, or to begin at a place of humility. At the very least it is a rejection of the hype. So, we welcome the horror of haunting - that danger that threatens to bring the curtains down - because the other option is to have the same old tropes performed once again.

Hauntology helps us understand what the silence threatens. Critic Mark Fisher (2012) discussed hauntology to understand the ways that innovation cannot escape prior social forms. The relationship between technology governance and industrial strategy has always been awkward. Technology in a Western context is often tied to economic growth and innovation economies. This alignment obfuscates human rights and social impacts or (perhaps at best) offers a calculus of trade-offs of how to manage innovation responsibly. These challenges predate artificial intelligence but resonate as soon as claims of a 'fourth industrial revolution' or a transformation in employment enter the discussion. That histories, those pasts, become what Tess Lea sees as a policy hauntology – "the deeply saturated effects of past policies." (2020, p. 30).

Ghosts, however, are part of hauntology as observation, and hauntology as strategy. Writer and scholar Billy-Ray Belcourt wrote his Poltergeist Manifesto to acknowledge that "there are forms of life abandoned outside modernity's episteme whose expressivities surge with affects anomalous within the topography of settler colonialism" (2016, p. 24). The poltergeist is the "noisy ghost" that shows up when you build your house on the old graveyard. The noise speaks "speaking into existence an anti-subjectivity that emerges in the aftermath of death or murder" (2016, p. 26). A poltergeist is one ghost, a strategy for Belcourt, to find a queer indignity. Their ghost inspires us. The ghosts on stage are frankly unsettling to the whole performance. Climate is an excellent example. In a climate emergency, AI power demand is not only firing up old coal plants, it is diverting energy that might otherwise be used for housing (Simpson, 2024). The play now is to keep talking around the climate issue (Hogan, 2020) without disrupting anything fundamental. "Keep going, can't you, keep going!" Hamm yells in Beckett's play Endgame. Ghosts keep the play moving.

Maybe we can just up and ghost the process - not participate? Another kind of haunting is to be present in absence, or to settle into ambivalence. As Kearnes and Wynne argue, ambivalence can be interpreted as an "engaged – rather than passive – mode of relating to technological determinism" (2007, p. 131). Our efforts have generated such ambivalence. Perhaps we can see this kind of haunting in relation to how the JUST AI project engaged with AI policy issues, revealing temporal and conceptual ruptures in the policy through its attempt to reshape the models and practices through which AI policy was considered. JUST AI's lab models and slow discussions formed part of variously composed struggles against the automatic embedding of technologically-determinist understandings of innovation in technology policy, and the related expectation that social scientist policy scholars should be consistently up to date on specific technical risks. Working to frame issues before time and to resist constant extraction played at the ambivalent edge of participating and exiting.

To haunt could be to stay in the room but break the part by rejecting the frame, or to stay in the room and to operate rigorously in the interests of the most marginalised. Tess Lea foregrounds this approach when she describes the role of the 'institutional killjoy' who points out the future hauntings that the policy might produce, while also insisting on the integrity of the policy process. As Lea writes, "institutional kill-joys are those who, working hard to pull benefits from policy opportunities, insist that contracts be honoured, promises be kept, accounts be held, corruptions be resisted, projects have fidelity, and people

https://doi.org/10.33621/jdsr.v6i440468

be respectfully supported" (2020, p. 14). The institutional killjoy may not be a good academic; but they are a good advocate and are often found acting stringently in the public interest.

A question raised by the killjoy's role concerns how and in which ways these goals contradict the pace and scope of the policy drama as it unfolds. As such, this fourth role opens a question that may also haunt the policy actors as they enter the stage: what happens when policy processes become obscure, improvisational, or difficult to square with the previously considered notions of the 'public interest'? What happens when an actor's micro-sensibilities meet the messy meso-level of policy drama?

5. Ghosting the future: Ambivalent and generative outcomes

In determining whether our policy performances have been effective, we often examine the scene itself. Was there a Colombo moment, where the regulator was so swayed by evidence that we could yell, 'aha!'? Unlikely. Did my performance sound confident? Did I read the room? What is more common is that the ghosts of the past haunt the present and prefigure the future. For drama to make sense we need to understand both the stage and what haunts it.

In this article we have consciously employed the terms of drama - staging, scenes, roles, and theatrical ghosts – to illuminate some of the complexities of academic participation in technology policy making, especially AI policy making. We have outlined how the staging of policy making can favour performances from already-powerful (often corporate) actors, and how shifting policy scenes and themes challenges the performance of expertise. We have briefly explored some of the roles that policy actors play in sustaining AI policy's publicities, and the way that these roles are haunted by our past experiences and our anticipations of how developing technology in the public interest might, should, or could unfold. In our experience, we have both observed and occasionally played several of these roles, although we also observe how the intensification of AI policy-making's publicities at present involves processes that produce erasures of policy topics and policy perspectives. The implications of these erasures include the potential undermining of epistemic justice and the exclusion of marginalised voices, or the thinning out of democratic oversight processes in favour of policy capture by louder (often commercial or corporate) voices. We contend that participation in the AI policy drama and its attendant publicities produces ambivalent positions, as represented by our ghosts on the stage of the drama. These ghosts remind us of what could have been as well as what still might become. Accepting this ambivalence, including its dynamism, improvisational and oppositional nature may be important to feed, nourish and develop the practice of the public interest, which lives in the messy meso-level between our own sense of justice and the macro structures of the world.

Our aims feel unfulfilled, and ambivalences remain. Yet scenes do change – perhaps not in the ways we expect. In Canada, the current AI bill seems dead in the water, likely to die with the end of Parliament all the while attention has shifted to AI safety (probably as part of Canada's own performance on a global stage). For JUST AI, considerations of racial justice and ecological sustainability of AI systems initially slid out of view. This created a policy blind spot, since the results of the project were not easy to place within the shifting policy drama unfolding at the time. While the project's efforts to bring strong standpoint epistemology and epistemic justice perspectives (see Fricker, 2007) into AI policy making may not have been remarked on the main stage, they may remain as ghosts to haunt future projects, or even to step into future policy scenes (Powell, forthcoming). One year after the project's close, environmental sustainability has subsequently become a policy frame, while racial justice is addressed functionally in discussions of bias rather than through systemic efforts to consider the connections between innovation, exploitation, and systemic injustice. Disability-led design and the necessity for justified refusal of AI systems remain marginal policy topics that killjoys struggle to place in full view (Barron, 2024). Meanwhile, although large-scale strategic funding reframes the role of arts and humanities research as in service to, rather than critical of innovation (a recent strategic project "is dedicated to integrating Arts, Humanities and Social Science research more fully into the Responsible AI

https://doi.org/10.33621/jdsr.v6i440468

ecosystem" (BRAID, 2023)), other funding calls focus on 'a new model of research leadership and teamworking" and invites pilot experiments in organizing research more collaboratively (UKRI, 2014). Prototyping policy options may feel ambivalent, but the spectres it raises glimmer in the corners of future policy stages.

References

Acemoglu, D. (2021). *Harms of AI* (Working Paper No. 29247). National Bureau of Economic Research. https://doi.org/10.3386/w29247

AI Fringe. (2023). *AI Fringe*. https://aifringe.org/

Allen, R., & Masters, D. (2024, April 18). *Artificial Intelligence (Regulation and Employment Rights) Bill*. https://www.tuc.org.uk/research-analysis/reports/artificial-intelligence-regulation-and-employment-rights-bill

Amoore, L. (2022). Machine learning political orders. *Review of International Studies*, 1–17. https://doi.org/10.1017/S0260210522000031

Angwin, J., Larson, J., Kirchner, L., & Mattu, S. (2016, May 23). *Machine Bias*. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Bacchi, C. L. (2009). *Analysing Policy: What's the Problem Represented to Be?* Pearson.

Bareis, J., & Katzenbach, C. (2021). Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics. *Science, Technology, & Human Values*, 01622439211030007. https://doi.org/10.1177/01622439211030007

Barron, D. (2024, June 3). *A radically inclusive approach to digital society*. Promising Trouble. https://www.promisingtrouble.net/blog/a-radically-inclusive-approach-to-digital-society

Bates, J. (2013). The Domestication of Open Government Data Advocacy in the United Kingdom: A Neo-Gramscian Analysis. *Policy & Internet*, *5*(1), 118–137. https://doi.org/10.1002/poi3.25

Belcourt, B.-R. (2016). A poltergeist manifesto. *Feral Feminisms*, *6*, 22–32.

Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts.: Hearing before the Standing Committee on Industry and Technology, House of Commons of Canada 1 (2023). https://www.ourcommons.ca/DocumentViewer/en/44-1/INDU/meeting-86/evidence

Birch, K. (2017). Techno-economic Assumptions. *Science as Culture*, *26*(4), 433–444. https://doi.org/10.1080/09505431.2017.1377389

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

Burke, K. (1945). *Grammar of motives*. Prentice-Hall. http://archive.org/details/grammarofmotives1945burk

Büthe, T., Djeffal, C., Lütge, C., Maasen, S., & Ingersleben-Seip, N. von. (2022). Governing AI – attempting to herd cats? Introduction to the special issue on the Governance of Artificial Intelligence. *Journal of European Public Policy*, *29*(11), 1721–1752. https://doi.org/10.1080/13501763.2022.2126515

Castaldo, J. (2023, December 10). *Federal consultations on AI regulations heavily skewed toward businesses, industry groups, say critics*. The Globe and Mail. https://www.theglobeandmail.com/business/article-canada-ai-law/

Cath, C. (2018). Governing artificial intelligence: Ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2133), 20180080. https://doi.org/10.1098/rsta.2018.0080

Chander, S. D., & Nation, S. (2021, June 24). *Who gets to write the future?* Ada Lovelace Institute. https://www.adalovelaceinstitute.org/blog/who-write-future/

Clarke, J., Bainton, D., Lendvai, N., & Stubbs, P. (2015). *Making policy move: Towards a politics of translation and assemblage* (1st ed.). Bristol University Press; JSTOR. https://doi.org/10.2307/j.ctt1t89fpb

Coleman, S. (2013). How to Make a Drama Out of a Crisis. *Political Studies Review*, *11*(3), 328–335. https://doi.org/10.1111/1478-9302.12024

Dandurand, G., McKelvey, F., & Roberge, J. (2023). Freezing out: Legacy media's shaping of AI as a cold controversy. *Big Data & Society*, *10*(2), 20539517231219242. https://doi.org/10.1177/20539517231219242

Derrida, J. (1994). *Specters of Marx: The State of the Debt, the Work of Mourning, and the New International*. Psychology Press.

European Union. (2019, April 8). *Ethics guidelines for trustworthy AI | Shaping Europe's digital future*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Fairclough, I., & Fairclough, N. (2013). *Political Discourse Analysis: A Method for Advanced Students*. Routledge.

Falk, S., & van Wynsberghe, A. (2023). Challenging AI for Sustainability: What ought it mean? *AI and Ethics*. https://doi.org/10.1007/s43681-023-00323-3

Fisher, M. (2012). What Is Hauntology? *Film Quarterly*, *66*(1), 16–24. https://doi.org/10.1525/fq.2012.66.1.16
Floridi, L. (2014). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press.
Fourcade, M., & Gordon, J. (2020). Learning Like a State: Statecraft in the Digital Age. *Journal of Law and Political Economy*, *1*(1). https://doi.org/10.5070/LP61150258
Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001
Fung, B. (2023a, May 16). *Mr. ChatGPT goes to Washington: OpenAI CEO Sam Altman testifies before Congress on AI risks | CNN Business*. CNN. https://www.cnn.com/2023/05/16/tech/sam-altman-openai-congress/index.html
Fung, B. (2023b, May 17). *How the CEO behind ChatGPT won over Congress | CNN Business*. CNN. https://www.cnn.com/2023/05/17/tech/sam-altman-congress/index.html
Gansky, B., & McDonald, S. (2022). CounterFAccTual: How FAccT Undermines Its Organizing Principles. *2022 ACM Conference on Fairness, Accountability, and Transparency*, 1982–1992. https://doi.org/10.1145/3531146.3533241
Gasser, U., & Almeida, V. A. F. (2017). A Layered Model for AI Governance. *IEEE Internet Computing*, *21*(6), 58–62. https://doi.org/10.1109/MIC.2017.4180835
Gebru, T., & Torres, É. P. (2024). The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence. *First Monday*. https://doi.org/10.5210/fm.v29i4.13636
Goldenfein, J., & Mann, M. (2023). Tech money in civil society: Whose interests do digital rights organisations represent? *Cultural Studies*, *37*(1), 88–122. https://doi.org/10.1080/09502386.2022.2042582
Gong, S., Liu, X., & Jain, A. K. (2020). *Mitigating Face Recognition Bias via Group Adaptive Classifier* (No. arXiv:2006.07576). arXiv. https://doi.org/10.48550/arXiv.2006.07576
Groves, L., Metcalf, J., Kennedy, A., Vecchione, B., & Strait, A. (2024). *Auditing Work: Exploring the New York City algorithmic bias audit regime* (No. arXiv:2402.08101). arXiv. https://doi.org/10.48550/arXiv.2402.08101
Hajer, M. A. (2009). *Authoritative Governance: Policy Making in the Age of Mediatization*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199281671.001.0001
Hansen, S. S. (2021). Public AI imaginaries: How the debate on artificial intelligence was covered in Danish newspapers and magazines 1956–2021. *Nordicom Review*, *43*(1), 56–78. https://doi.org/10.2478/nor-2022-0004
Hawes, B., & Hall, D. W. (2023, December). *After the Summit: Progress in public policy on AI* [Monograph]. University of Southampton. https://doi.org/10.5258/SOTON/WSI-WP007
Hendrycks, D., & Mazeika, M. (2022). *X-Risk Analysis for AI Research* (No. arXiv:2206.05862). arXiv. https://doi.org/10.48550/arXiv.2206.05862
Hess, D. J. (2014). When Green Became Blue: Epistemic Rift and the Corralling of Climate Science. In *Fields of Knowledge: Science, Politics and Publics in the Neoliberal Age* (Vol. 27, pp. 123–153). Emerald Group Publishing Limited. https://doi.org/10.1108/S0198-871920140000027012
Hickman, L., & Serlin, D. (2018). Towards a crip methodology for critical disability studies. In *Interdisciplinary Approaches to Disability*. Routledge.
Hogan, M. (2020). The Data Center Industrial Complex (2021). *Saturation: An Elemental Politics (Duke University Press)*. https://www.academia.edu/39043972/The_Data_Center_Industrial_Complex_2021_
Irani, L. (2015). The cultural work of microwork. *New Media & Society*, *17*(5), 720–739. https://doi.org/10.1177/1461444813511926
Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), Article 9. https://doi.org/10.1038/s42256-019-0088-2
Jones, M., & McKelvey, F. (2024). Deconstructing public participation in the governance of facial recognition technologies in Canada. *AI & SOCIETY*. https://doi.org/10.1007/s00146-024-01952-w
JUST-AI. (2023). *Prototyping Ethical Futures Series: 2021 Convening on Futures*. JUST AI. https://www.just-ai.net/copy-of-2021-convening-futures
Katz, Y. (2020). Artificial Whiteness: Politics and Ideology in Artificial Intelligence. In *Artificial Whiteness*. Columbia University Press. https://doi.org/10.7312/katz19490
Kearnes, M., & Wynne, B. (2007). On Nanotechnology and Ambivalence: The Politics of Enthusiasm. *NanoEthics*, *1*(2), 131–142. https://doi.org/10.1007/s11569-007-0014-7
Knight, S., Shibani, A., & Vincent, N. (2024). Ethical AI governance: Mapping a research ecosystem. *AI and Ethics*. https://doi.org/10.1007/s43681-023-00416-z
Lasswell, H. D. (1971). *A pre-view of policy sciences*. American Elsevier Publishing Company. http://archive.org/details/previewofpolicys0000lass
Lea, T. (2020). *Wild policy: Indigeneity and the unruly logics of intervention*. Stanford University Press.
Li, F., Kakhki, A. M., Choffnes, D., Gill, P., & Mislove, A. (2016, 16). Classifiers Unclassified: An Efficient Approach to Revealing IP Traffic Classification Rules. *Proceedings of the 16th ACM Internet Measurement Conference*. Internet Measurement Conference, Santa Monica, CA. http://david.choffnes.com/pubs/ClassifiersUnclassified-IMC16.pdf
Lipsky, M. (2010). *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service, 30th Anniversary Expanded Edition* (30th Anniversary, Expanded ed. edition). Russell Sage Foundation.

Luka, M. E., Harvey, A., Hogan, M., Shepherd, T., & Zeffiro, A. (2015). Scholarship as Cultural Production in the Neoliberal University: Working Within and Against 'Deliverables.' *Studies in Social Justice*, *9*(2), Article 2. https://doi.org/10.26522/ssj.v9i2.1138

Marres, N. (2005). *Issues Spark a Public into Being: A Key But Often Forgotten Point of the Lippmann-Dewey Debate* (B. Latour & P. Weibel, Eds.; pp. 208–217). MIT Press.

McKelvey, F. (2018). *Internet daemons: Digital communications possessed*. University of Minnesota Press.

McKelvey, F., & Roberge, J. (2023). Recursive Power. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 21–32). Edward Elgar Publishing.

Menn, J., & Nix, N. (2023, December 7). Big Tech funds the very people who are supposed to hold it accountable. *Washington Post*. https://www.washingtonpost.com/technology/2023/12/06/academic-research-meta-google-university-influence/

Metz, R. (2021, March 24). *Google offered a professor $60,000, but he turned it down. Here's why*. CNN. https://www.cnn.com/2021/03/24/tech/google-ai-ethics-reputation/index.html

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., & Floridi, L. (2023). Operationalising AI ethics: Barriers, enablers and next steps. *AI & SOCIETY*, *38*(1), 411–423. https://doi.org/10.1007/s00146-021-01308-8

Nagy, P., & Neff, G. (2024). Conjuring algorithms: Understanding the tech industry as stage magicians. *New Media & Society*, *26*(9), 4938–4954. https://doi.org/10.1177/14614448241251789

Naiseh, M., Bentley, C., Ramchurn, S., Williams, E., Awad, E., & Alix, C. (2022). Methods, Tools and Techniques for Trustworthy Autonomous Systems (TAS) Design and Development. *Companion of the 2022 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, 66–69. https://doi.org/10.1145/3531706.3536459

Newman-Griffis, D., Rauchberg, J. S., Alharbi, R., Hickman, L., & Hochheiser, H. (2023). Definition drives design: Disability models and mechanisms of bias in AI technologies. *First Monday*. https://doi.org/10.5210/fm.v28i1.12903

Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., & Guston, D. (2013). A Framework for Responsible Innovation. In *Responsible Innovation* (pp. 27–50). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch2

Palmås, K., & Surber, N. (2022). Legitimation crisis in contemporary technoscientific capitalism. *Journal of Cultural Economy*, *15*(3), 373–379. https://doi.org/10.1080/17530350.2022.2065331

Paul, R. (2022). Can critical policy studies outsmart AI? Research agenda on artificial intelligence technologies and public policy. *Critical Policy Studies*, *16*(4), 497–509. https://doi.org/10.1080/19460171.2022.2123018

Readings, B. (1997). *The University in Ruins*. Harvard University Press.

Sætra, H. S., Coeckelbergh, M., & Danaher, J. (2022). The AI ethicist's dilemma: Fighting Big Tech by supporting Big Tech. *AI and Ethics*, *2*(1), 15–27. https://doi.org/10.1007/s43681-021-00123-7

Schneiders, E., Chamberlain, A., Fischer, J. E., Benford, S., Castle-Green, S., Ngo, V., Kucukyilmaz, A., Barnard, P., Row Farr, J., Adams, M., Tandavanitj, N., Devlin, K., Mancini, C., & Mills, D. (2023). TAS for Cats: An Artist-led Exploration of Trustworthy Autonomous Systems for Companion Animals. *Proceedings of the First International Symposium on Trustworthy Autonomous Systems*, 1–5. https://doi.org/10.1145/3597512.3597517

Schön, D. A., & Rein, M. (1994). *Frame reflection: Toward the resolution of intractable policy controversies*. New York : BasicBooks. http://archive.org/details/framereflectiont00dona

Simpson, J. (2024, May 4). Capacity crunch: Why the UK doesn't have the power to solve the housing crisis. *The Observer*. https://www.theguardian.com/business/article/2024/may/04/capacity-crunch-why-the-uk-doesnt-have-the-power-to-solve-housing-crisis

Sloane, M., Moss, E., Awomolo, O., & Forlano, L. (2022). Participation Is not a Design Fix for Machine Learning. *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–6. https://doi.org/10.1145/3551624.3555285

Sloane, M., Moss, E., & Chowdhury, R. (2022). A Silicon Valley love triangle: Hiring algorithms, pseudo-science, and the quest for auditability. *Patterns*, *3*(2), 100425. https://doi.org/10.1016/j.patter.2021.100425

Stark, L., & Hoffmann, A. L. (2019). Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture. *Journal of Cultural Analytics*. https://doi.org/10.22148/16.036

Stilgoe, J. (2018). Machine learning, social learning and the governance of self-driving cars. *Social Studies of Science*, *48*(1), 25–56. https://doi.org/10.1177/0306312717741687

Stilgoe, J. (2019). *Who's Driving Innovation?: New Technologies and the Collaborative State*. Springer International Publishing.

Stone, D. (1997). *Policy paradox: The art of political decision making* (Revised Edition). New York : W.W. Norton & Co. http://archive.org/details/policyparadoxart0000ston

Taylor, L., & Dencik, L. (2020). Constructing Commercial Data Ethics. *Technology and Regulation*, *2020*, 1–10. https://doi.org/10.26116/techreg.2020.001

Tharoor, I. (2023, July 26). Analysis | The world reckons with a new 'Oppenheimer moment.' *Washington Post*. https://www.washingtonpost.com/world/2023/07/26/oppenheimer-nolan-ai-artificial-intelligence/

Tuck, E., & Ree, C. (2013). A Glossary of Haunting. In S. H. Jones, T. E. Adams, & C. Ellis (Eds.), *Handbook of autoethnography* (pp. 639–657). Left Coast Press.

Turchin, A., & Denkenberger, D. (2020). Classification of global catastrophic risks connected with artificial intelligence. *AI & SOCIETY*, *35*(1), 147–163. https://doi.org/10.1007/s00146-018-0845-5

Turnbull, N. (2008). Harold lasswell's "problem orientation" for the policy sciences. *Critical Policy Studies*, *2*(1), 72–91. https://doi.org/10.1080/19460171.2008.9518532

Turner, R. H. (2001). Role Theory. In J. H. Turner (Ed.), *Handbook of Sociological Theory* (pp. 233–254). Springer US. https://doi.org/10.1007/0-387-36274-6_12

UK Government. (2019, June 10). *A guide to using artificial intelligence in the public sector*. GOV.UK. https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector

UK Government. (2024, February 6). *A pro-innovation approach to AI regulation: Government response*. GOV.UK. https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response

van Hulst, M., & Yanow, D. (2016). From Policy "Frames" to "Framing": Theorizing a More Dynamic, Political Approach. *The American Review of Public Administration*, *46*(1), 92–112. https://doi.org/10.1177/0275074014533142

Vicente, P. N., & Dias-Trindade, S. (2021). Reframing sociotechnical imaginaries: The case of the Fourth Industrial Revolution. *Public Understanding of Science (Bristol, England)*, *30*(6), 708–723. https://doi.org/10.1177/09636625211013513

Whittaker, M. (2021). The steep cost of capture. *Interactions*, *28*(6), 50–55. https://doi.org/10.1145/3488666

Widder, D. G. (2024). Epistemic Power in AI Ethics Labor: Legitimizing Located Complaints. *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, 1295–1304. https://doi.org/10.1145/3630106.3658973

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, *12*(4), 505–523. https://doi.org/10.1111/rego.12158

