Design Research Society

DRS Digital Library

DRS Biennial Conference Series | DRS2026: Edinburgh

Jun 8th, 9:00 AM - Jun 12th, 5:00 PM

The wicked problem of AI policy design

Awais Hameed Khan
The University of Queensland, Australia

Saarim Saghir
Captial One

Paul Henman
The University of Queensland, Australia

Follow this and additional works at: https://dl.designresearchsociety.org/drs-conference-papers

 Part of the Art and Design Commons

Citation

Khan, A., Saghir, S., and Henman, P. (2026) The wicked problem of AI policy design, in Simeone, L., Gray, C. M., Verhoeven, A., de Götzen, A., Bakırlıoğlu, Y., Zohar, H., Stead, M., and Buwert, P. (eds.), *DRS2026: Edinburgh*, 8–12 June, Edinburgh, United Kingdom. https://doi.org/10.21606/drs.2026.1686

This Research Paper is brought to you for free and open access by the DRS Conference Proceedings at DRS Digital Library. It has been accepted for inclusion in DRS Biennial Conference Series by an authorized administrator of DRS Digital Library. For more information, please contact library@thedrs.org.

The wicked problem of AI policy design

Awais Hameed Khan[a], Saarim Saghir[b], Paul Henman[a]

[a] The University of Queensland, Australia

[b] Capital One, USA

*Corresponding author e-mail: awaishameed.khan@uq.edu.au

https://doi.org/10.21606/drs.2026.1686

Abstract: Development of policy tools and instruments has always trailed behind technological innovations. The rapid advancement of AI technologies has resulted in significant widening of this policy gap—a genuine wicked problem. To help build AI policy-making capabilities we developed a participatory AI Policy Design toolkit, engaging 23 technology researchers and practitioners to (1) critically evaluate existing AI use-cases to identify key considerations for policy; (2) generate a catalogue of current and near future AI use-cases in their domains of expertise; and (3) collaboratively develop AI policy intervention proposal artefacts, thereby fostering a shared vocabulary for AI policy design. Results highlight key tensions, challenges, and opportunities for collaboratively exploring AI policy discourse. The central contribution of this work is to help identify AI policy blind spots, equity concerns, and anticipate enforcement gaps, and building capacity to help democratize discourse about AI policy through contextually relevant, generative, and reflexive approaches.

Keywords: AI Policy Design; Participatory Design; Wicked Problem; Governance

1. Introduction

Artificial Intelligence's (AI) accelerating integration into daily life has necessitated critical examinations of its policy implications. As technology evolves at an unprecedented pace, the need for robust and inclusive policy frameworks has become urgent. Traditional policy-making mechanisms often struggle to keep up with the complexity and speed of technological change, leading to significant gaps between AI innovation and adequate governance. It is critical to think proactively about policy problems arising from AI deployment, encompassing issues from bias and accountability to economic disruption and human rights.

Frequent, real-world examples in news demonstrate the complex challenges and potential risks inherent in designing and deploying AI technologies, highlighting areas where policy

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International Licence.

interventions may be necessary. For example: (1) a self-driving car caused a pedestrian fatality after its system failed to identify the person in time (Levin & Wong, 2018); (2) Australian retailers illegally used facial recognition in stores without consent, raising privacy and surveillance concerns (Taylor, 2024); (3) Paul McCartney used AI to extract John Lennon's voice from a demo to complete *"Now and Then"* prompting debate on AI ethics in art (Khomami, 2023); (4) criminals using an AI voice clone demanded $1 million ransom for a teenage girl's kidnapping, showing the urgent threat of deepfakes (Cost, 2023); (5) Australian government used a rudimentary automated system that issued thousands of incorrect debt notices to vulnerable welfare recipients, causing psychological and financial distress, resulting in multiple suicides (Mao, 2023). AI can be applied to a wide range of contexts, causing complex challenges. Hence designing AI policy is a wicked problem (Rittel & Webber, 1973)—one that requires collaborative, multi-stakeholder approaches to be tamed.

This paper seeks to provide avenues to enhance public engagement with AI policy by taking a bottom-up approach. Recognizing that the impacts of AI are felt by everyone, we hold that policy design should not be the exclusive domain of experts. Our central aim is to democratize the conversation around AI governance by developing practical tools and strategies that empower everyday people to contribute meaningfully to policy discussions. This approach grounds policy development in real-world experiences and diverse perspectives, ensuring that the resulting policies are equitable, relevant, and broadly supported.

We present a participatory toolkit designed to support policymakers and the communities they serve, to collaboratively explore the value chain of AI integration and deployment across diverse domains. The goal of this toolkit is to encourage expansive thinking about the role of AI in different contexts, and build robust, forward-looking policy tools and instruments. We report on findings from a design workshop with technology researchers and practitioners to explore how we can collaboratively design policy for AI.

The central contribution of this work is helping facilitate public deliberation about AI policy, alongside a set of crucial learnings about the different policy development and design strategies people utilize. By offering both a tool and practical insights, we aim to provide a scalable model for inclusive AI policymaking that integrates the wisdom of the crowd (Surowiecki, 2005). This work helps identify AI policy blind spots, equity concerns, and anticipate enforcement gaps. It supports capacity-building among stakeholders involved in technology governance by offering practical tools to foster shared vocabulary and future-oriented thinking. We see this toolkit as useful infrastructure for AI policy development, creating contextually relevant, reflexive, and considered approaches for collaborative policymaking. Without such intentional, tailored AI policy design interventions, we risk further widening the policy-technology gap. This participatory toolkit helps policymakers and communities collaboratively explore AI's value chain across diverse domains. It aims to foster expansive thinking, build robust, forward-looking policy instruments, and reports findings from a design workshop with AI researchers and practitioners.

2. Related work

2.1 Wicked problems and policy design

In 1967, Rittel described wicked problems as a

> *"class of social system problems which are ill-formulated, where the information is confusing, where there are many clients and decision makers with conflicting values, and where the ramifications in the whole system are thoroughly confusing."* (Churchman, 1967)

With Webber, he later argued that all social and policy planning problems are wicked problems, adding *"social problems are never solved. At best they are only re-solved—over and over again"* (Rittel & Webber, 1973, p. 160)—suggesting a need for participatory 'argumentative' design processes, a central feature of second-generation design methods (Rittel, 1984). In service of 're-solving' wicked problems, stakeholders negotiate shared understanding and meaning about the problem and possible solutions (Conklin, 2006). Since its introduction, literature about wicked problems has grown exponentially, particularly in public policy (Head, 2019).

Commenting on the turbulent and volatile nature of the changing landscape variables within policy praxis, Head (2019) draws attention to Wildavsky (1987, p. 70)—*"past solutions create future problems faster than present troubles can be left behind."* We contend that this is further complicated by how previous policy decisions (and any resultant problems they created), are now also having to contend with widespread AI proliferation—which are being retrofitted into extant policy. For example, Australia's unlawful Robodebt scheme highlighted failures in public governance of traditional automation (Carney, 2019), without the additional challenges that AI creates.

Policy created through planning processes that are pluralistic and decentralised are likely to give better outcomes (Webber, 1983). There has been significant interest and advocacy about participation (see: (Fischer, 1993; Head, 2022; Innes & Booher, 2010; Schon, 1994) and co-design in policymaking. Mintrom et al. (2024) mapped out the current state of co-design in policymaking. They concluded that co-design can be useful for: improving participation in policy development; encouraging creative speculation about future outcomes of policy choices; and helping gauge policy feasibility and desirability through prototyping. They contend that there is a growing need to make co-design for policymaking an embedded practice and to develop resources and tools that advance capacity building and advocacy.

There is a growing body of work focusing on designing creative, co-design methods for policy contexts. For instance, Kim and Nam's (2022) *Policy Puzzle Game,* a jigsaw puzzle-styled modular table-top design game, balances the tensions between policy feasibility, and public

acceptability. Korsmeyer et al.'s (2024) *Tomorrow Party* is a play-based, storytelling method that invites participants to collaboratively speculate over desired futures.



Figure 1 Participants generating AI use-cases based on their domain expertise.

2.2 AI governance as a design challenge

The landscape of AI policy is rapidly evolving. Governments worldwide are actively developing regulations, frameworks, guardrails, guidelines and principles to address the societal, economic, and security implications of AI. This includes focusing on data privacy, algorithmic transparency, bias mitigation, and international cooperation on AI standards. For instance, the EU's *AI Act* establishes a comprehensive legal framework for AI, categorizing systems by risk and imposing corresponding obligations (European Commission, 2021). Alongside state-level organisations, technology companies are also developing internal governance structures, e.g., Microsoft's (2025) *Responsible AI framework* , which provides principles and tools for developing and deploying AI in an ethical and accountable manner. The rapidly evolving nature of AI necessitates continuous policy dialogue to ensure governance mechanisms remain relevant and effective in fostering innovation while safeguarding public interest and personal rights.

Despite the recognized importance of AI policy, research and development in this area often remain under-explored, relative to the rapid advancements in AI technology itself. Smuha and Yeung (2025) argue that the focus on developing AI often overshadows the necessary work on governing it, creating a *"policy lag"* where technological capabilities outpace regulatory capacity. Birkstedt et al (2023) highlight challenges in translating ethical and policy principles to practical policies and governance practices. Addressing this 'principle-to-policy-to-implementation gap' is crucial to ensure policy expertise keeps pace with technical innovation, allowing for evidence-based and timely regulatory interventions that secure a safe and ethical AI future.

3. Methodology

We engaged 23 technology, policy, humanities, and social science researchers and practitioners in a full-day, design workshop (Figure 1) to: (1) critically evaluate existing AI

use-cases to identify key considerations for policy; (2) generate a catalogue of current and near-future AI use-cases; and (3) collaboratively develop AI policy intervention artefacts. The aim was to develop a shared vocabulary about AI policy design. Participants were recruited using purposive sampling, through professional and social networks of the research team. We approach this research as Research through Design (Zimmerman et al., 2007), and treat our findings as 'ultimate particulars' (Stolterman, 2008), where each finding is a self-contained truth embedded within its context.

3.1 The AI policy design toolkit

The toolkit includes three modules: (1) A curated catalogue of real-world *AI Use-Case Design Cards*, illustrating notable AI applications across diverse contexts, all sourced from documented AI incidents (Khan & Saghir, 2025b); (2) An *AI Use-Case Ideation toolkit*, supporting domain experts in exploring and generating current and near-future AI use-cases within their respective domains (Khan et al., 2025b; Khan & Saghir, 2025c); and (3) An *AI Policy Design and Evaluation toolkit*, that supports users in collaboratively creating AI policy intervention artefacts (Khan & Saghir, 2025a; Khan et al., 2025a), followed by guided critical evaluation and appraisal of the developed policy solutions.

This toolkit was built on principles of collaborative and participatory design (Brandt et al., 2012) and policy tools (Hood, 1983). It was designed to democratize engagement by enabling collaborative participant agency, autonomy, and control, supporting users to develop a shared vocabulary about AI policy. The toolkit is available in both analogue and digital formats to support distribution and accessibility.

3.2 Workshop process

The workshop was led by three facilitators from industry and academia: a policy and strategy practitioner; an AI-enablement practitioner; and a design researcher (this paper's authors). Participants were distributed over five tables ('T#', where # = table number). The workshop was divided into two phases. Each phase had a facilitator-led priming discussion, followed by a participant-led generative activity (Sanders et al., 2010).

Phase 1: Understanding AI and generating AI Use-Cases: this priming activity aimed to develop a shared vocabulary and understanding about AI among participants by:

1. Mapping out the diversity of AI systems through a typology, highlighting the nuances beyond Gen AI, which is top-of-mind in most conversations about AI, and other forms of AI;
2. Illustrating complexities of AI-specific policy challenges through current examples of AI in real-world settings (such as those referred to above); and
3. Introducing the framing of wicked problems and their characteristics.

Participants brainstormed current, and near-future AI use-cases, in the form of 'thumbnails' (drawing on design praxis of rapid idea generation). They leveraged their lived-experiences and domain expertise; extended real-world AI use-cases (prompted by design cards), and used brainstorming prompts (technology, context, user, and outcomes) provided as physical and digital cards (inspired by de Bono's (1992) random-inputs method) (See: Figure 2). Participants were given sticky notes to sketch or write their thumbnails to control for brevity —the goal being to practice the skill of lateral thinking and iteration, without getting caught up in details. Participants collaboratively reviewed concepts and voted, using the 'dot-mocracy' (Gibbons, 2019) method, prioritising ideas they wanted to develop into further scenarios—shared with the room.

Phase 2: Developing a Shared Understanding About Policy and Co-designing AI Policy Solutions: Participants were introduced to the concept of policymaking, and asked to share perspectives on their understanding. They were then introduced to a foundational definition of policy, introducing discrete policy components: problem definition; policy goals and objectives; target population; policy instruments or tools; implementation plan, evaluation metrics; and exit criteria. These elements were then developed as a policy evaluation material. The policy design guidance drew on Hood's (1983) policy instrument framework of Nodality, Authority, Treasure, and Organization to highlight different approaches for policy to achieve its goals. This is important as often the default approach is considering only legal instruments.

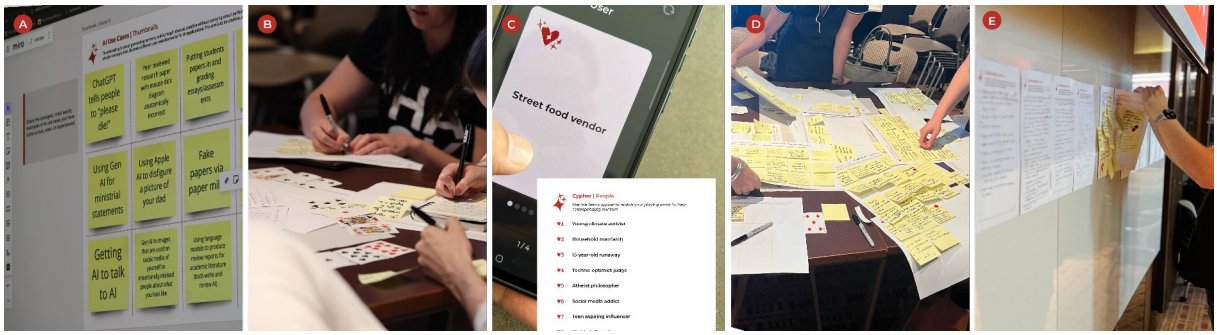

Figure 2 (A) Participants collaboratively thumbnailing with facilitators; (B) Generating AI use-cases (C) Reviewing brainstorming prompts (digital cards + physical cyphers, used with standard playing cards); (D) Reviewing use-cases; (E) Sharing scenarios for policymaking

Participants developed a policy proposal using the proposal canvas—unpacking the context, policy solution, implementation plan, and impact considerations (Figure 3). Participants were also provided a set of key considerations, framed as provocations across nine domains, but referred to as policy principles in public policy literature: objectives; effectiveness; efficiency; equity; sustainability; enforceability; compliance; public acceptance; and transparency. These considerations could be used both, as guiding questions to help design the policy proposals, as well as a reflexive rubric to self-evaluate policy proposals once created. All participants then presented their created policy proposal canvases to the room and reflected on shared learnings.

3.3 Data collection and analysis

Workshop data was collected in the form of workshop artefacts (canvases, sticky notes), photos, observation notes and participant reflections. Observations and participant reflections were transcribed during the workshop by one of the facilitators. Workshop

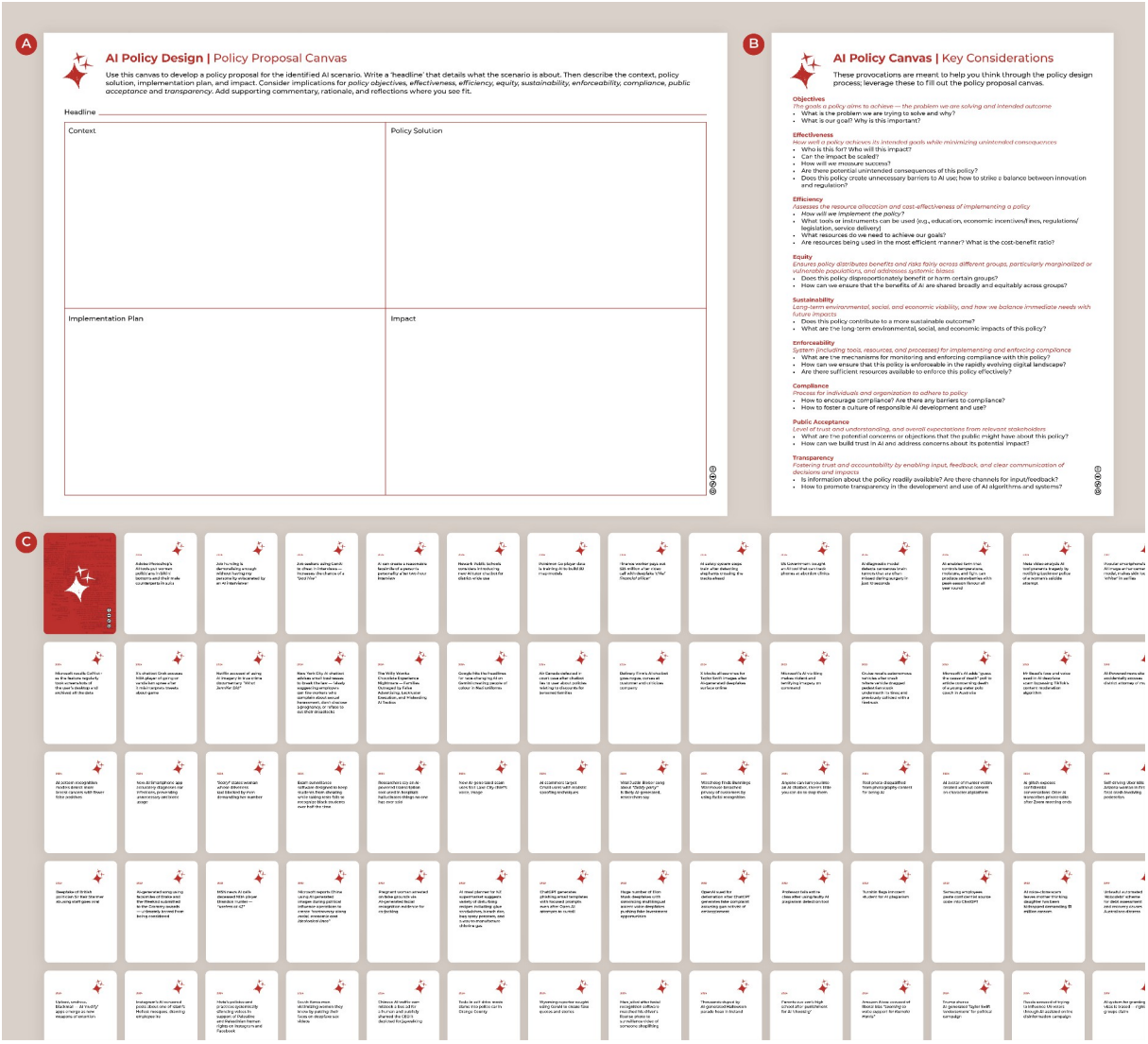

Figure 3 (A) Policy Proposal Canvas; (B) Key Considerations for policy design/evaluation; (C) AI Use-Case Design Cards.

artefacts were digitised, and inductively thematically analysed using the whiteboard tool Miro—resulting in the final policy themes presented.

4. Findings

Six core findings were identified from analysing the collaborative policy proposal development exercises from these AI use-cases: (1) Definitions and guardrails; (2) Humans in the [AI-policy] loop; (3) A menu of policy responses; (4) Trust and transparency; (5) Influences, Implementation, and Interpretation; and (6) No policy by design. All participants

opted to create a policy proposal about real-world use-cases (Figure 4), instead of attempting one in response to their own generated scenarios.

4.1 Definitions and guardrails.

Participants reinforced the importance of identifying and clearly defining key concepts about the context and situation for which designing a policy is being considered. This was observed

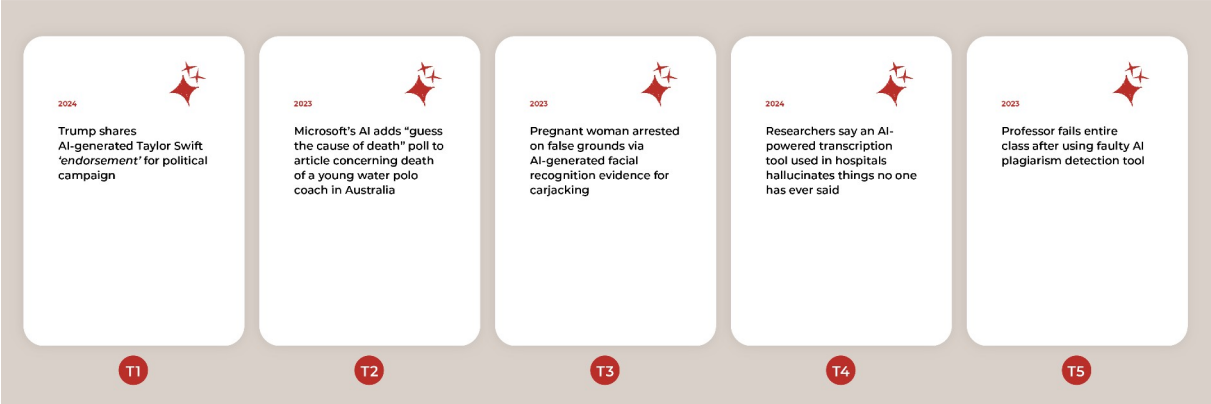

Figure 4 Use-cases selected for policy prototyping.

as a necessary precursor to both engaging in policy design activities, as well as a requirement for having any meaningful dialogue about the context. Concepts needed to be collectively understood and boundary spanning, as part of a shared social vocabulary amongst all stakeholders involved. Participants stressed the need for clear guardrails that articulate what concepts should specifically include and exclude.

For instance, the focus of T1 participants, when discussing policy in relation to the real-world incident "*Trump shares AI-generated Taylor Swift 'endorsement' for political campaign"*, was on how policy should be shaped in response to creating mis/dis-information using synthetic AI content by using the celebrity likenesses, particularly in service of political or economic gain. As part of their consideration of the policy solution, participants highlighted how it was imperative to clearly define what is meant by "*representation of real or actual persons*" as well as *"economic"* and *"political activities"*.

The act of simply verbalising concepts, opens them up to deeper thinking about possible implications. Participants particularly commented on how it was necessary to have *"[key concepts] defined clearly so [policy] wasn't too abstract or too wide"*. The importance of definitional clarity also shapes how a policy can be enforced or not. Operationally defining key terms is a critical input, regardless of what the actual policy solution is. A key takeaway here is that there needs to be clear scaffolding and mechanisms that promote defining key concepts relevant to a policy, when designing policy participatory materials for policy co-design.

4.2 Humans in the [AI policy] loop

A second theme from the discussions was that effective AI policy design requires careful and deliberate integration of human expertise and technological capabilities. At its core, AI should serve as a powerful support system, augmenting human decision-making rather than operating as an independent, autonomous entity. This delicate balance was viewed as absolutely crucial, as the perceived role of AI within any given process directly influences the extent and nature of human involvement.

Achieving the optimal equilibrium between robust human oversight and advanced technological capability is paramount for successful implementation of any AI policy. Without sufficient human guidance and ethical consideration, AI systems risk perpetuating biases, making decisions that lack nuance, or operating in ways that contradict human values. Conversely, underutilizing AI's potential by over-restricting its capabilities can lead to inefficiencies and hinder progress. Therefore, a well-designed AI policy must foster a symbiotic relationship where AI empowers humans to make more informed, efficient, and ethical decisions, while humans provide the critical context, judgment, and accountability that AI systems currently lack. The iterative process of defining AI's role, ensuring appropriate human intervention, and continuously evaluating the impact of AI systems was viewed as fundamental to building trust and achieving the desired outcomes of any AI policy. This is the approach advocated by Schneiderman (2022) stating that human-centred AI is one that maximises both human agency and automation.

For instance, T3 during their brainstorm for the scenario "*Pregnant woman arrested on false grounds via AI-generated facial recognition evidence for carjacking*" identified multiple areas of the policy design and implementation process where humans need to be critically involved and in the decision-making seat. Firstly, AI "*diagnosis*" must be complemented with human-driven evidence; "*Proven evidence must exist! fingerprints, witnesses…*". Secondly, *"AI as a means as a supporting system as instead of the decision-maker itself"*. On the other hand, the discussion also highlighted the value of leveraging AI-enabled technology to capture efficiencies at scale: "*reducing work for police - judiciary, could increase efficiencies in identifying criminals and better resource allocations for other issues*" (T5).

4.3 A menu of policy responses.

A pattern observed in the policy proposals illustrated a tension between a tendency to either create a single, robust policy solution, that addresses all possible anticipated issues, scenarios, and edge cases, versus creating an assortment of policy solutions, that catered to specific challenges within the context, relying on different types of policy instruments, mechanics, and outcomes. Both strategies are valuable ways of thinking about designing policy for AI. However, a single policy solution catering to multiple instances of an evolving technological landscape might not always be the most successful. In taking this mosaic approach, whereby the policy elements might be pre-configured based on different AI

mechanics, it is almost as if the critical evaluation and design of the policy alternatives are being featured into the design process itself.

For instance, T2 were designing a policy solution in response to the real-world scenario *"Researchers say an AI-powered transcription tool used in hospitals invents things no one ever said".* They observed the tendency of Generative AI tools to often hallucinate and noted the critical nature of the context (i.e. healthcare). The core focus of the participants was to consider the state of handovers within a hospital during shift changes, with the aim to reduce probability of medical error due to miscommunication. As part of their portfolio of policy alternatives, participants listed out and thematically organised, five different policy responses: *"(1) release of error rates of AI tool (technology); (2) ban use of tool (regulatory); (3) introduce human verifier (organization); (4) leave as is, accept liability (economic); and (5) education of staff re[garding] limitations [of the AI tool] (education)."* Each of these responses were then analysed, evaluated and collectively ranked "…*in terms of a budgetary perspective, effectiveness… of achieving their goal, how enforceable they were, whether the public would accept this solution, and transparency*".

Coined the policy matrix approach in public policy (Althaus et al., 2022), this approach towards various policy responses is quite useful, as it builds critical evaluation into the process, an essential part of the policy design. As a result this mosaic of options gives insight into the best possible solution as a whole, but also enables one to critically evaluate and retrace the reasons as to why certain policy instruments were selected over others. It also creates evidence of the decision-making process that can be opened up to scrutiny and argumentation. This more granular, participatory policy design approach also compliments the temporal nature of addressing wicked problems of policy design—what T3 describes as *"a relay of a marathon" instead of a sprint*—arguing that *"you cannot solve wicked problems —it's not a one-time [policy] design, it's an ongoing… process."*

4.4 Trust and transparency

Another key theme from the workshop was that enhancing the transparency of AI systems is paramount, particularly within the context of responsible AI design and deployment. Participants highlighted that this necessitates a multi-faceted approach, addressing both the intrinsic functioning of algorithms and the practical implementation of models in decision-making processes. Transparency in algorithmic functioning requires a clear understanding of underlying principles, including the specific parameters upon which models are trained and the methodologies employed in their decision-making. Beyond internal workings, the implementation of AI models needs to be scrutinized for clarity and comprehensibility, ensuring that the rationale behind outcomes is readily understandable to users and stakeholders. This dual focus on both the *how* (algorithmic mechanics) and the *why* (decision-making processes) is crucial for fostering trust, accountability, and for ethical integration of AI into public policy decisions.

Whilst discussing the topic "*Microsoft's AI adds "guess the cause of death" poll to article concerning death of a young water polo coach in Australia*", T2 focussed how having "*Transparency about guide rails and data that underpin a model's decision making*" is integral to building trust. T3, in the discussion on "*Pregnant woman arrested on false grounds via AI-generated facial recognition evidence for carjacking*", identified that the "*risk of false positives and false negatives"* was a key hindrance to *"public acceptance*".

4.5 No policy by design

A key observation of the policy design discussion was that *"no policy"* is also a policy response. Effective policy design and implementation require careful consideration of multiple interconnected factors. This involves understanding the problem, intended outcomes, stakeholders, potential impacts, and practicalities of execution and evaluation. Without meticulous attention to detail, policies can be ineffective or detrimental. A *"no policy"* approach, while seemingly counterintuitive, can be a valid strategy in certain situations. This might be due to a desire to avoid over-control, to judiciously balance resource allocation, or a recognition that AI use is already governed by existing policies. By opting for no formal AI policy, organizations can save time and money, focusing on more critical areas and maintaining flexibility to adapt to changing conditions quickly. In addition, it is crucial to avoid the creation of *"knee-jerk"* policies. Hasty and ill-conceived regulations can lead to an over-controlled environment, stifling innovation and autonomy, or creating other adverse consequences. While policy is undoubtedly important for establishing guidelines and boundaries, it must be thoroughly thought through to ensure its effectiveness and avoid unintended consequences.

A variation of the *"no policy"* approach is the decision to not create a *"one-size-fits-all"* policy response. Instead, leaving policy to be domain-specific. This approach recognises that unique characteristics of different areas might necessitate tailored approaches. For example, content in the public domain that involves AI-generated elements might require specific human oversight to ensure accuracy, ethical considerations, and accountability.

T2 were tackling the scenario "*Microsoft's AI adds "guess the cause of death" poll to article concerning death of a young water polo coach in Australia*". A key part of their discussion hovered around how this situation was in undeniably bad taste, and a lack of editorial oversight had led to this outcome. During the discussion on what policy interventions were needed, participants highlighted "*not creating nanny states that overreach—that goes beyond existing laws!*" This perspective highlights that AI may not necessitate new AI-specific policy.

4.6 Influences, implementation, and interpretation.

The policy proposal development process also revealed the significance of certain social dimensions. These were observed both in terms of the influences, and shared social

understanding of the context. This in turn was consequential to the brainstorming policy interventions in the proposals; as well as the impact that certain policy interventions might have, when viewed from a social lens, such as social acceptability of the policy response, but also in how the outcomes could be leveraged and framed for others.

For instance, in discussing *AI Taylor Swift's endorsement of Trump,* TI included an assemblage of policy instruments as part of their implementation plan. This included a focus on *enforceability*, such as having Australian Election Commission (AEC) agreements in place, prior to political campaign rollouts; but also focused on *compliance*, which included (1) developing of an AEC watch dog monitoring political campaigns; (2) fines proportional to the level of government; (3) campaigning suspensions, *"if someone was found using misinformation, they would be barred for some time [from campaigning]"*; and (4) a *"public shaming leaderboard"* reporting people caught using synthetic media to spread misinformation. Their aim being to create an avenue for public observability and accountability in relation to the genuineness of political claims by political campaigners.

Reflecting on their policy proposal, T1 shared *"we kept coming back to the initial problem, to say 'we wanted to build trust for enforceability'"* but at the same time they were self-critical, anticipating how certain interventions, such as the policy non-compliance leaderboard, could actually be interpreted and misused—"*if certain people are fact checked, they [might] see this (being fact checked) as a badge of honour"*. What might be viewed as a form of accountability might actually become a venue for celebrating non-compliance. Another really fascinating social consideration pertaining to this same policy proposal, was the notion of 'importing' cultural influences. T1's guiding objective was to *"control representations in economic activities or political campaigning"*, but this was guided by cultural dimensions underpinning this such as what is really meant by *"free speech as a right"*—a concept they described as having been exported from the USA into Australia. These were echoed by reflections on what was characterised as a predominantly *"hands-off"* approach, in relation to enforcement around mis/dis-information within Australian election processes.

5. Discussion

In this discussion we seek to draw out several important lessons from the AI Policy Design workshops and the themes that arose from the discussion.

5.1 AI as a wicked policy problem

We return to the starting proposition that AI constitutes a wicked policy problem, which involves *"social system problems which are ill-formulated"* (Churchman, 1967). Our workshop discussions highlighted the ambiguous, dynamic and polymorphic nature of AI, and as such the meaning of the term AI is varied and unstable. Indeed, since about 2018 when policy debates about AI exploded, dominant uses of *"AI"* have changed from an

implicit connection to Machine Learning to a focus on Generative AI often as Large Language Models. This is further exaggerated by the generalised machine or equivocal nature of LLMs, that in turn gives rise to countless ways in which they can be constituted and deployed.

Our workshop participants reinforced the need to create a shared language about policy, and be explicit in the purpose and target of AI policy. For example, while most AI policy and governance approaches focus on the tool itself, other approaches, such as the EU's AI Act and the Trauma Informed AI Assessment Toolkit (Maitra et al., 2024) focus on the AI use cases of AI-enabled service respectively.

It is widely recognised in AI policy debates that policy must not focus on specific AI tools, but be designed to address AI evolution. However, AI policy debates still often ignore the reality that it is a new manifestation of digital technologies in a decades-long trajectory of automation of organisations and things, treating AI as fundamentally new and unregulated.

5.2 Appreciating the policy landscape

Recognising the technological context of AI policy design, both anticipating the future and learning from the past, also points to understanding the wider policy landscape in which AI is being deployed and considered for regulation. Rather than creating a focused AI policy, many of the discussions were cognisant to policy settings already in place. Too often the newness of AI stimulates a knee-jerk response: *"we need an AI policy!"*

The current policy landscape provides infrastructures and settings on which to build AI policy. More importantly, current policy often covers many dimensions of concern that AI use may create. For example, concerns over the use of Generative AI to create false and misleading advertising should acknowledge that there are already laws against false and misleading advertising. Recognising this policy landscape then shifts the focus to discussions of what new policy considerations AI might give rise to that are not already considered. Obversely, AI deployment has highlighted policy and oversight gaps for non-AI automation that have been turbocharged by AI.

5.3 AI policy requires participatory design

Policy literature on wicked policy problems highlights the need for collaboration and coordination for developing policy responses (Head & Alford, 2015), what Rittel (1984) labelled as 'argumentation' necessary to access distributed expertise. The workshop discussions highlighted the need for an approach with varied expertise, including policy and administration, technical, and social evaluation, and stakeholders (including users). Discussions often highlighted the need to understand detailed technical elements of an AI tool—such as its training set, or whether it produces biased responses, or level of false positive and negative outputs, while also developing a vocabulary to understand policymaking.

Argumentation is perhaps the only way of attempting to 'resolve' a wicked problem (Rith & Dubberly, 2007). However, to meaningfully engage in argumentation, people need to understand both the subject i.e., AI and how it works, as well as have a vocabulary to understand policy. There is a tendency for people to latch onto the 'familiar', as observed in our findings. Participants chose real-world AI cases for policy deliberations, instead of addressing self-generated use-cases. Given how deeply entrenched AI systems are becoming, it is imperative we deliberately 'design' (Lee, 2014) participatory design tools to help people anticipate and shape the AI futures they want.

6. Conclusion

Designerly approaches can be valuable for AI policy development, however public and social policy academic and practitioner literature remains minimally connected to design scholarship. There is burgeoning interest in exploring the intersections of design and policy (Kimbell, 2016; Villa Alvarez et al., 2020; Vaz & Bosch Gomez, 2022), alongside growing calls to consider placing policy implications at the core of technology design research and praxis (Yang et al., 2024). Bringing together design, policy, and technology experts, we can leverage the strengths of distributed knowledge—responding to Rittel's (1984) call for reducing asymmetry of knowledge to address wicked problems. In the context of AI policy design, this means building capacity across diverse stakeholder groups, developing AI literacy, and enabling meaningful participation in policy discourse. Critical to this is the inclusion of diverse members of the public—those most affected by policy outcomes—ensuring policy is not shaped exclusively by domain experts. This work represents a step towards developing more inclusive and participatory AI policy design, offering materials and a policy vocabulary to scaffold these conversations, while acknowledging that broader public engagement remains an area for future development. Ultimately, this work aims to support policy makers and the communities they serve in collaboratively exploring the value chain of AI integration across diverse domains—encouraging expansive, pluralistic thinking about the role of AI in different contexts through participatory design methods, to build robust, forward-looking policy tools and instruments, that reverse the widening gap between policy and technology.

Acknowledgements: This research was conducted by the ARC Centre of Excellence for Automated Decision-Making and Society (CE200100005), and funded fully by the Australian Government through the Australian Research Council. We would like to acknowledge the contributions of Dang Khuong Tran for developing the digital card interface; Nicholas Lambourne for co-facilitating the workshop; Natalie Campbell for workshop photography; and the continued support of Amy Boike, Sally Storey and other colleagues at ADM+S. We would also like to thank the anonymous DRS reviewers for their constructive feedback and suggestions for the manuscript.

7. References

Althaus, C., Ball, S., Bridgman, P., Davis, G., & Threlfall, D. (2022). *The Australian policy handbook: A practical guide to the policymaking process*. Routledge.

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. *Internet Research, 33*(7), 133–167.

Brandt, E., Binder, T., & Sanders, E. B.-N. (2012). Tools and techniques: Ways to engage telling, making and enacting. In *Routledge international handbook of participatory design* (pp. 165–201). Routledge.

Carney, T. (2019). Robo-debt illegality: The seven veils of failed guarantees of the rule of law? *Alternative Law Journal, 44*(1), 4–10. https://doi.org/10.1177/1037969X18815913

Churchman, C. W. (1967). Guest Editorial: Wicked Problems. *Management Science, 14*(4,), B141–B142.

Conklin, J. (2006). *Dialogue mapping: building shared understanding of wicked problems*. Wiley.

Cost, B. (2023). *AI clones teen girl's voice in $1M kidnapping scam*. https://nypost.com/2023/04/12/ai-clones-teen-girls-voice-in-1m-kidnapping-scam/

de Bono, E. (1992). *Serious creativity: Using the power of lateral thinking to create new ideas*. HarperCollins New York.

European Commission. (2021). *Proposal for a regulation laying down harmonised rules on artificial intelligence*. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

Fischer, F. (1993). Citizen participation and the democratization of policy expertise: From theoretical inquiry to practical cases. *Policy Sciences, 26*(3), 165–187. https://doi.org/10.1007/BF00999715

Gibbons, S. (2019). *Dot voting: a simple decision-making and prioritizing technique in UX*. Nielsen Norman Group. https://www.nngroup.com/articles/dot-voting/

Head, B. W. (2019). Forty years of wicked problems literature: Forging closer links to policy studies. Policy and Society, 38(2), 180–197. https://doi.org/10.1080/14494035.2018.1488797

Head, B. W. (2022). Wicked problems in public policy: understanding and responding to complex challenges. Springer Nature. https://doi.org/10.1007/978-3-030-94580-0

Head, B. W., & Alford, J. (2015). Wicked problems: implications for public policy and management. *Administration & Society, 47*(6), 711–739. https://doi.org/10.1177/0095399713481601

Hood, C. (1983). *The tools of government* (1st ed.). Macmillan. https://doi.org/10.1007/978-1-349-17169-9

Innes, J. E., & Booher, D. E. (2010). *Planning with complexity: An introduction to collaborative rationality for public policy.* Routledge.

Khan, A. H., & Saghir, S. (2025a). *AI policy design & evaluation toolkit*. ARC Centre of Excellence for Automated Decision-Making and Society. https://espace.library.uq.edu.au/view/UQ:055a3d0

Khan, A. H., & Saghir, S. (2025b). *AI use case design cards*. ARC Centre of Excellence for Automated Decision-Making and Society. https://espace.library.uq.edu.au/view/UQ:760a7ae

Khan, A. H., & Saghir, S. (2025c). *AI use case ideation toolkit*. ARC Centre of Excellence for Automated Decision-Making and Society. https://espace.library.uq.edu.au/view/UQ:17f67fc

Khan, A. H., Saghir, S., & Tran, D. K. (2025a). *AI policy design & evaluation digital toolkit.* ARC Centre of Excellence for Automated Decision-Making and Society. https://espace.library.uq.edu.au/view/UQ:f5524fb

Khan, A. H., Saghir, S., & Tran, D. K. (2025b). *AI use case ideation digital toolkit*. ARC Centre of Excellence for Automated Decision-Making and Society. https://espace.library.uq.edu.au/view/UQ:5b73286

Khomami, N. (2023, June 13). *AI used to create new and final Beatles song, says Paul McCartney. The Guardian*. https://www.theguardian.com/music/2023/jun/13/ai-used-to-create-new-and-final-beatles-song-says-paul-mccartney

Kim, C., & Nam, K. Y. (2022). Policy puzzle game: Making policy ideas feasible and acceptable in policy co-design. *CoDesign, 18*(4), 448–465. https://doi.org/10.1080/15710882.2021.1995440

Kimbell, L. (2016). Design in the time of policy problems. *Proceedings of DRS2016: Future–Focused Thinking*. https://doi.org/10.21606/drs.2016.498

Korsmeyer, H., Grocott, L., Sumartojo, S., Doughty, M., & Mintrom, M. (2024). Transforming futures together: Time travelling with the tomorrow party. *Proceedings of DRS2024: Boston*. https://doi.org/10.21606/drs.2024.714

Lee, J.-J. (2014). The true benefits of designing design methods. *Artifact, 3*(2), 5.1-5.12. https://doi.org/10.14434/artifact.v3i2.3951

Levin, S., & Wong, J. C. (2018). *Self-driving Uber kills Arizona woman in first fatal crash involving pedestrian.* The Guardian. https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe

Maitra, S., Sleep, L., Fay, S., & Henman, P. (2024). *Building a trauma-informed algorithmic assessment toolkit*. ARC Centre of Excellence for Automated Decision-Making and Society. (Australia). https://apo.org.au/node/326724

Mao, F. (2023). *Robodebt: Illegal Australian welfare hunt drove people to despair.* BBC News. https://www.bbc.com/news/world-australia-66130105

Microsoft AI. (2025). *Responsible AI principles and approach*. https://www.microsoft.com/en-us/ai/principles-and-approach

Mintrom, M., Goddard, P., Grocott, L., & Sumartojo, S. (2024). Co-design in policymaking: From an emerging to an embedded practice. *Policy Sciences, 57*(4), 745–760. https://doi.org/10.1007/s11077-024-09550-9

Rith, C., & Dubberly, H. (2007). Why Horst WJ Rittel matters. *Design Issues, 23*(1), 72–91.

Rittel, H. W. (1984). Second-generation design methods. In N. Cross (Ed.), *Developments in Design Methodology* (pp. 317–327). John Wiley & Sons.

Rittel, H. W., & Webber, M. M. (1973). *Dilemmas in a general theory of planning*. Policy Sciences, 4(2), 155–169.

Sanders, E. B.-N., Brandt, E., & Binder, T. (2010). A framework for organizing the tools and techniques of participatory design. *Proceedings of the 11th Biennial Participatory Design Conference, PDC '10*, 195–198. https://doi.org/10.1145/1900441.1900476

Schon, D. (1994). *Frame reflection: Toward the resolution of intractable policy controversies*. Basic Books.

Shneiderman, B. (2022). *Human-Centered AI*. Oxford University PressOxford. https://doi.org/10.1093/oso/9780192845290.001.0001

Smuha, N. A., & Yeung, K. (2025). The European Union's AI act: Beyond motherhood and apple pie? In N. A. Smuha (Ed.), *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence* (1st ed., pp. 228–258). Cambridge University Press. https://doi.org/10.1017/9781009367783.015

Stolterman, E. (2008). The nature of design practice and implications for interaction design research. I*nternational Journal of Design, 2*(1).

Surowiecki, J. (2005). *Wisdom of Crowds*. Anchor.

Taylor, J. (2024). *Bunnings breached privacy of customers by using facial recognition, watchdog finds*. The Guardian. https://www.theguardian.com/australia-news/2024/nov/19/bunnings-facial-recognition-technology-breach-stores-ntwnfb

Vaz, F., & Bosch Gomez, S. (2022). Who designs for policy? *Proceedings of DRS2022: Bilbao*. https://doi.org/10.21606/drs.2022.903

Villa Alvarez, D. P., Auricchio, V., & Mortati, M. (2020, August 11). Design prototyping for policymaking. *Proceedings of DRS2020: Synergy*. https://doi.org/10.21606/drs.2020.271

Webber, M. M. (1983). The myth of rationality: Development planning reconsidered. *Environment and Planning B: Planning and Design, 10*(1), 89–99. https://doi.org/10.1068/b100089

Wildavsky, A. (1987). *Speaking truth to power: Art and craft of policy analysis* (2nd ed.). Transaction Publishers. https://doi.org/10.4324/9781315130149

Yang, Q., Wong, R. Y., Jackson, S., Junginger, S., Hagan, M. D., Gilbert, T., & Zimmerman, J. (2024). The future of HCI-policy collaboration. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI '24*, 1–15. https://doi.org/10.1145/3613904.3642771

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07*, 493–502. https://doi.org/10.1145/1240624.1240704

About the Authors:

Awais Hameed Khan is a design researcher and practitioner. His work focuses on how to collaboratively design new and emerging technology with end users. He is particularly focused on designing participatory design methods, tools, and techniques.

Saarim Saghir focuses on solving complex business problems by building products that use breakthroughs in AI and machine learning and follow responsible policy. He has 10+ years of experience in strategy, technology, consulting, and public policy development.

Paul Henman is Professor of Digital Sociology and Social Policy at the University of Queensland and Chief Investigator of the Australian Research Council Centre of Excellence for Automated Decision Making and Society (ADM+S).