Politics and Governance
2026 • Volume 14 • Article 10632
https://doi.org/10.17645/pag.10632

ARTICLE

Open Access Journal

From Efficiency to Deliberation: Rethinking AI's Role in Institutionalizing Democratic Innovations

Arild Ohren [1,2], Elizabeth Calderón Lüning [2], Čedomir Markov [3], and Max Stearns [4]

[1] Norwegian University of Science and Technology, Norway
[2] The Democratic Society, Belgium
[3] Institute for Philosophy and Social Theory, University of Belgrade, Serbia
[4] Independent Researcher, France

Correspondence: Elizabeth Calderón Lüning (elizabeth.cl@demsoc.eu)

Submitted: 30 April 2025 **Accepted:** 19 August 2025 **Published:** 22 January 2026

Issue: This article is part of the issue "Towards an Innovative Democracy: Institutionalizing Participation in Challenging Times" edited by Irena Fiket (Institute for Philosophy and Social Theory – University of Belgrade), Giovanni Allegretti (Centre for Social Studies – University of Coimbra), and Gazela Pudar Drasko (Institute for Philosophy and Social Theory – University of Belgrade), fully open access at https://doi.org/10.17645/pag.i479

Abstract

As AI becomes increasingly embedded in democratic innovation (DI), critical questions arise about how these technologies shape deliberative quality, civic agency, and institutional design. While AI promises to expand and scale deliberative mini-publics, it also risks undermining the democratic goods that make such processes meaningful, particularly inclusiveness, popular control, considered judgment, and transparency. This article introduces the democracy-in-the-loop (DITL) framework as both a normative and practical approach to integrating AI into democratic settings. Building upon and expanding models like human-in-the-loop and society-in-the-loop, DITL embeds contestation, reflexivity, and participant agency into the operation and governance of AI systems. A key feature of the DITL approach is the intentional use of "meaningful frictions" (disruptions designed to slow down interaction, surface assumptions, and invite critical engagement). We explore the DITL model through the Digital Democracy Lab, a series of four experimental workshops held in 2024 in Brussels, Madrid, Kraków, and Dublin as part of the EU-funded Knowledge Technologies for Democracy project. Each workshop combined a purpose-built AI Demonstrator platform with facilitated deliberation to explore how AI can support, rather than supplant, democratic reasoning. Findings suggest that AI-enabled DIs should focus on flexibility, contestability, and democratic oversight, not merely technical efficiency. Institutionalizing DIs in the age of AI requires more than simply scaling tools; it calls for embedding democratic values into the design, deployment, and evolution of socio-technical systems.

Keywords
algorithmic accountability; AI; deliberative democracy; deliberative mini-publics; democracy-in-the-loop; democratic innovations; digital deliberation; human–AI interaction

1. Introduction

Democratic innovation (DI) has emerged as a response to the growing disconnect between citizens and institutions of representative democracy (Courant, 2022; Newton, 2012; Pogrebinschi, 2018). Drawing from traditions of participatory, deliberative, and direct democracy, DIs present new processes and institutions aimed at extending citizen involvement beyond electoral cycles into various stages of political decision-making (Pogrebinschi, 2018; Theuwis et al., 2025). Their novelty is strictly contextual: A process may be considered innovative in one policy or governance setting even if it has long existed elsewhere. What sets DIs apart from more conventional forms of political participation is that they are purposefully designed to deepen and reimagine the role of citizens, seeing them not as passive consumers, but as co-creators of public policy (Elstub & Escobar, 2019; Smith, 2009). Because of this focus, DIs are often associated with effects that enhance democratic legitimacy, including increased political knowledge, reasoning skills, internal efficacy, and institutional trust (e.g., Lacelle-Webster & Warren, 2021; Theuwis et al., 2025).

While DIs have shown promise in enhancing democratic legitimacy and citizen engagement, their institutional impact remains limited. Many still exist as one-off experiments, constrained by high costs, unclear mandates, and limited scalability (Bussu et al., 2022). In response, scholars have explored strategies for institutionalization that maintain the experimental spirit of DIs while improving stability, reach, and long-term impact (e.g., Courant, 2022; Escobar, 2022; Pogrebinschi, 2018; Randma-Liiv, 2023). Information and communication technologies, and more recently, AI systems, are proposed as part of this solution, displaying potential to reduce organizational burdens, support large-scale participation, and synthesize citizens' inputs (McKinney, 2024; Randma-Liiv, 2023; Revel & Penigaud, 2025; Smith, 2009). The main challenge, however, is to ensure that the efficiency gains promoted by AI do not come at the expense of democratic quality, particularly regarding inclusiveness, deliberative depth, and popular control.

Against this backdrop, we introduce the democracy-in-the-loop (DITL) approach as a response to these tensions. Building on existing models such as human-in-the-loop (HITL) and society-in-the-loop (SITL; Rahwan, 2018), DITL calls for real-time, iterative, and participatory oversight of AI systems, emphasizing reflexivity, contestation, and civic agency. This is primarily achieved by introducing meaningful frictions (intentional design elements that slow down interaction, encourage reflection, and break the seamlessness typically associated with AI systems).

We examined the possibilities and limitations of this approach through four experimental workshops held in Brussels, Madrid, Kraków, and Dublin in 2024. These workshops formed the Digital Democracy Lab (DDL), part of the Knowledge Technologies for Democracy (KT4D) project. With the DDL, the DITL approach was operationalized by embedding a custom-built AI Demonstrator platform into facilitated deliberative processes. Rather than optimizing for speed or consensus, the workshops deliberately introduced disruptions and transparency mechanisms to determine whether AI could support deeper deliberation, strengthen democratic literacy, and increase civic agency.

We argue that institutionalizing DIs in the age of AI requires more than just technical tools; it calls for a shift in how we design and govern participatory systems. The DITL framework offers a way to rethink institutionalization not merely as standardizing procedures, but as the dynamic embedding of democratic reflexivity, contestation, and transparency within socio-technical systems.

While the DI ecosystem varies widely in participatory logics, actor constellations, institutionalization levels, policy-cycle stages, and democratic aims (Elstub & Escobar, 2019), our focus is on deliberative mini-publics (DMPs), or "carefully designed forums where a representative subset of the wider population come[s] together to engage in open, inclusive, informed, and consequential discussions" (Curato et al., 2021, p. 3). The principles we propose, however, extend to the broader DI landscape.

The article is structured as follows. First, we reflect on the current debate surrounding DI institutionalization. Drawing on Smith's (2009) framework for assessing DIs, we then discuss how AI can contribute to or impede both institutional and democratic goods, with a special focus on DMPs. In Section 4, we introduce the DITL model and distinguish it from other similar approaches, such as HITL and SITL. Section 5 then presents the DDL and its four workshops conducted in 2024 as examples of the DITL framework and how meaningful frictions are applied within an AI system. In the concluding section, we highlight the implications of our findings for institutional and democratic goods, emphasizing the importance of incorporating reflexivity, friction, and responsiveness into future institutionalization efforts.

2. AI and the Institutionalization of DI

For decades, DIs like participatory budgeting and DMPs have been used worldwide to address democratic deficits and complex social challenges. This "experimentation phase" involved trial and adaptation by governments, civil society, and academia within varied institutional and cultural contexts, contributing to the global "deliberative wave" marked by the rise of DMPs (Bussu et al., 2022; OECD, 2020).

As these mechanisms spread, scholars increasingly ask how to move beyond ad hoc experiments toward institutionalization (Bussu et al., 2022; Courant, 2022; Escobar, 2022; Pogrebinschi, 2018; Randma-Liiv, 2023). The experimentation phase is typically characterized by one-off initiatives with varied designs and uncertain mandates. Institutionalization seeks to reduce arbitrariness in initiation and design, and so potentially enhance legitimacy, predictability, and long-term impact (Pogrebinschi, 2018).

Pogrebinschi (2018) identifies five dimensions shaping institutionalization: formalization, representativeness, scope, scale, and decisiveness. She argues that only institutionalized DIs have the potential to reform governance systems meaningfully and contribute to democratic quality, provided they meet these criteria.

Some scholars distinguish between institutionalization and embeddedness. Bussu et al. (2022) suggest that embeddedness involves a deeper normative and practical integration of participatory innovations into governance systems. While institutionalization may simply refer to formalizing a mechanism, embeddedness reflects how deeply these innovations are rooted in institutional routines, supported by political will, and aligned with democratic values. Institutionalization alone does not guarantee meaningful participation. Courant (2022) contrasts tamed consultation, where DMPs provide symbolic legitimacy for predetermined decisions, with radical democracy or representative klerocracy models that grant citizens genuine decision-making power through sortition.

Despite their promise, DIs often remain isolated, experimental, and episodic, with limited impact on mainstream policy (Escobar, 2022; Pogrebinschi, 2018; Smith, 2009). However, efforts to institutionalize DIs are often hampered by constraints that are structural (e.g., centralized decision-making, bureaucratic resistance, etc.), normative (e.g., fear of hindering experimentation), and operational (e.g., cost and scalability; Bussu et al., 2022; Courant, 2022; Randma-Liiv, 2023). Scaling deliberative practices to broader publics remains particularly challenging.

As Goodin (2000, p. 82) famously put it: "The challenge facing deliberative democrats is thus to find some way of adapting their deliberative ideals to any remotely large-scale society, where it is simply infeasible to arrange face-to-face discussions across the entire community." As deliberative processes scale, they become more difficult to maintain through conventional, labor-intensive methods, due to the demands of facilitation and synthesizing complex, long-form participant input and outputs. Information and communication technologies have long been seen as potential remedies to some of these constraints (e.g., Smith, 2009), and recently, scholars have explored AI's potential to address practical and institutional scaling challenges, especially for DMPs (Landemore, 2024). This article primarily examines the institutionalization push associated with government-led DMPs, while recognizing that many of the AI-related arguments are equally relevant to bottom-up, citizen-led, or movement-led DMPs.

3. Evaluating AI's Impact on Democratic Processes

The introduction of AI into DMPs has elicited both enthusiasm and caution. Proponents highlight AI's potential to scale participation, reduce costs, and facilitate institutionalization. Skeptics, however, raise concerns that efficiency gains may come at the expense of deliberative depth and democratic trust. In light of these tensions, evaluating AI's role requires more than just assessing technical feasibility or efficiency; it must account for AI's impact on core democratic values. To guide this evaluation, we draw on Smith's (2009) framework of democratic and institutional goods.

The democratic goods include inclusiveness, popular control, considered judgment, and transparency; the institutional goods are efficiency and transferability.

Regarding democratic goods, inclusiveness involves both presence and voice, whether democratic processes include participants from diverse backgrounds and whether all participants have meaningful and equal opportunities to speak, be heard, and influence outcomes. Popular control refers to the genuine influence citizens can exert over decision-making, from shaping agendas to impacting final outcomes. Considered judgment emphasizes the need for thoughtful, informed, and reflective deliberation, where decisions come from reasoned exchange and mutual learning rather than raw preferences. Transparency involves clarity about how decisions are made and communicated. This applies both to participants, who must understand the purpose, structure, and consequences of their involvement, and to the broader public, which needs accessible information about how the process works and why it matters.

Turning to institutional goods, efficiency assesses the demands a process places on citizens and organizers alike. It asks whether a process can be run effectively without overburdening those involved. Transferability, finally, refers to how easily a model can be adapted across different political, cultural, or institutional contexts—how well it travels beyond its place of origin.

It cannot be expected that the introduction of AI into a deliberative process will enhance all democratic goods equally. DIs vary in focus and design, often strengthening some goods while leaving others unchanged (Elstub & Escobar, 2019; Smith, 2009). The key issue, however, is whether AI risks weakening any core democratic goods, as this could undermine the legitimacy and effectiveness of DI.

There is growing interest in how AI can address challenges related to the scale and efficiency in DMPs. This section explores how current and proposed AI deployments affect democratic goods (building on Smith's, 2009, framework). It examines efficiency gains first, then considers contributions to other democratic goods, and finally highlights potential risks and trade-offs.

3.1. AI and Institutional Goods

Emerging research highlights an increasing variety of AI applications in DMPs (Flanigan et al., 2021; Landemore, 2024; McKinney, 2024). Many aim to address the resource-intensive nature of deliberation and promote participation at scale. Consequently, most AI implementations emphasize Smith's (2009) "efficiency good," aiming to scale deliberative processes significantly.

Large language models (LLMs) assist with translation, summarizing materials, clustering perspectives, and generating questions, counterarguments, and consensus statements. In multi-lingual settings, AI-powered translation reduces translation costs, especially in transnational DMPs.

Platforms like Pol.is and Remesh, while not using AI, demonstrate how algorithmic filtering and clustering synthesize large volumes of citizen input into coherent summaries (Revel & Penigaud, 2025; see also Bussu et al., 2022). AI-facilitated synthesis can similarly produce collective narratives that promote reflexivity and acknowledge plural viewpoints (Revel & Penigaud, 2025).

While AI cannot achieve the ideal of universal deliberation, it can expand engagement. Landemore (2024) identifies two scalable models: mass online deliberation (gathering large virtual publics) and multiple rotating mini-publics (engaging many citizens' assemblies simultaneously and with rotating participants). In both models, AI aids scalability through facilitation, translation, fact-checking, clustering, and synthesis. Emerging systems like the Habermas Machine (Tessler et al., 2024, as cited in Revel & Penigaud, 2025, p. 12) or the Stanford Online Deliberation Platform (Gelauff et al., 2023) illustrate how AI tools can lower participation thresholds, enhance perceived fairness, and work with existing representative structures.

3.2. Potential Democratic Benefits

While the focus has been on efficiency, AI implementation can also promote democratic goods. For example, algorithmic tools have been used for participant selection in DMPs, not only to increase efficiency but also to enhance fairness and representation. This directly relates to the aspect of inclusiveness, which concerns who gets a seat at the table. Flanigan et al. (2021) developed a selection algorithm that balances demographic quotas with the principle that everyone should have an equal chance of being chosen. Aiming for "maximal fairness," they sought to make selection probabilities as equal as possible while meeting representativeness goals (Flanigan et al., 2021, p. 549).

A primary area where AI has been introduced is facilitation, with the aim of reducing costs while also yielding democratic benefits. Landemore (2024) argues that AI could serve as an impartial facilitator, mitigating biases often associated with human moderators. Moreover, AI-facilitated discussions may augment the voice dimension of inclusion, improving gender equity in discussions compared to traditional, in-person formats (Gelauff et al., 2023). Taken together, these developments could help make DMPs more inclusive.

AI is also being explored as a tool for increased popular control. LLMs can process and cluster large amounts of citizen input, synthesizing views and identifying shared concerns. This can help ensure that diverse voices are reflected in agenda-setting and that influence is spread more evenly across the population, including the so-called "maxi-public" (McKinney, 2024).

AI may also contribute to considered judgment. By acting as accessible sources of contextual information or real-time Q&A systems, tools like LLMs can support participants in navigating complex topics, especially when expert input is limited or unavailable (Landemore, 2024; McKinney, 2024). Beyond potentially lowering costs, this could help reduce informational inequalities among participants and strengthen the epistemic foundations of deliberation.

3.3. Democratic Trade-Offs and Risks

While AI may enhance inclusion, popular control, and considered judgment in DMPs, its integration also introduces significant democratic risks. As Landemore (2024) and McKinney (2024) warn, AI systems can reproduce input biases and fabricate misleading content, while overreliance on them could diminish popular control and human agency. The opacity of algorithmic reasoning (the "black box" problem) makes it difficult to trace or contest how decisions are made, which are qualities essential to both institutional and democratic legitimacy (McKinney, 2024; Smith, 2009).

One specific concern with AI-driven facilitation is its lack of human sensitivity. As McKinney (2024) notes, AI cannot easily navigate nuanced communication, build trust, or address group inequalities—essential skills that, if absent, risk undermining both inclusion and the quality of deliberation. Similarly, Alnemr (2020) warns that AI systems, by enforcing language norms like grammar correction, exclude certain forms of expression, such as storytelling or non-standard language. LLMs also inherit bias from training data, often overrepresenting dominant (e.g., English-speaking, Western, and male) perspectives (Binns, 2018), which may silence or distort minoritized voices and fail to reflect local contexts. Moreover, technical and cultural barriers, such as digital literacy, cognitive accessibility, and dominant assumptions about user experience, may further exclude participants (Costanza-Chock, 2020; Eubanks, 2018). Together, these issues limit inclusiveness and diversity, while AI's difficulty in handling real-world complexity may also limit quality deliberation by reducing the depth of deliberation.

Similarly, the use of AI in the selection process, despite claims of “maximal fairness,” can obscure political decisions. As Alnemr (2021) argues, fairness is very contextual. In the case of the Global Assembly, an algorithm decided on the exclusion of participants from Pacific Island states—amongst the most affected by climate change—on the grounds of fairness. This illustrates how algorithmic selection can sideline vulnerable groups. As Ohren (2024) notes, selection is often framed as a neutral, technical step but is in fact shaped by implicit political decisions—about which categories matter, which criteria are prioritized, and how representation is defined.

Most fundamentally, the growing reliance on AI in deliberative processes may risk undermining popular control. With AI curating information, citizens may lose the ability to scrutinize or challenge the content provided. Alnemr (2020) notes that, without human facilitators, citizens’ agency is reduced, undermining their emancipation and the freedom to critically engage with the process. This directly diminishes popular control, and the deliberative process becomes increasingly more mediated by technology rather than by its participants. Without participants’ power to determine what should be discussed, the system tends to favor certain framings or knowledge systems over others. This shift creates the illusion of influence (Sloane et al., 2022): although participants interact with AI tools, the underlying decision-making remains opaque and technocratic, often leaving citizens unable to see how their inputs affect outcomes.

These challenges are further compounded by the lack of public oversight in how AI tools are developed, updated, and deployed (McKinney, 2024). Even if AI is meant only to be a supportive tool, institutional reliance on it without democratic safeguards risks eroding accountability. At the same time, AI models often oversimplify complex political questions, reducing rich and contested deliberations into neat outputs that fail to capture the nuance, ambiguity, and dissent essential to democratic reasoning (DiSalvo, 2022).

These risks coalesce into a broader structural concern: that the very design of AI-mediated deliberation may exclude citizens from shaping their own democratic environment. As Alnemr (2021, p. 71) writes: “expert-imposed design can undermine mini-publics’ democratic and emancipatory potential since citizens are excluded from creating the conditions for deliberation.”

3.4. The Need for Human Oversight in AI-Mediated Democracy

In summary, introducing AI into deliberative processes not only presents technical challenges but also democratic ones—such as who determines the terms of public reasoning, whose voices are heard, and how legitimate authority is built. Concerns about representational bias, value alignment, and the potential for technocratic capture are particularly salient when AI tools are embedded into formal governance systems without safeguards for transparency, accountability, and public oversight (Randma-Liiv, 2023). As McKinney (2024) emphasizes, the goal is not to automate decision-making, but to use AI as a compass for collective self-understanding (supporting institutional design while maintaining democratic ownership of the process). Across the literature, there is a growing consensus that human oversight is essential to preserving democratic values in AI-supported deliberation (e.g., McKinney, 2024; Randma-Liiv, 2023; Revel & Penigaud, 2025).

4. Beyond Current AI Governance Models

In evaluating how AI can or should be introduced as a potential scaling mechanism for DIs, it is crucial to consider the different models by which AI technologies are designed and governed. Consequently, this section offers a consideration of two existing AI governance models, which lack the dynamic and iterative capacity to ensure AI usage aligns with effective democratic function (HITL and SITL), and proposes a third model that can achieve this: DITL.

4.1. HITL: Individual Oversight and Its Limitations

In an HITL system, a human operator plays a vital role within an automated control process. The human handle specific tasks such as supervision, exception control, optimization, and maintenance (Rahwan, 2018). This paradigm emerged from the field of supervisory control, where human operators intermittently program and continuously receive information from autonomous systems (Rahwan, 2018). However, while HITL offers necessary safeguards, it has notable limitations in democratic contexts. First, it places a single human as the sole representative of broader societal interests. Second, the human operator often lacks the structural power to challenge the system's fundamental assumptions or values. Third, HITL systems typically constrain the human's role to ensuring efficiency or preventing catastrophic errors, rather than promoting democratic deliberation.

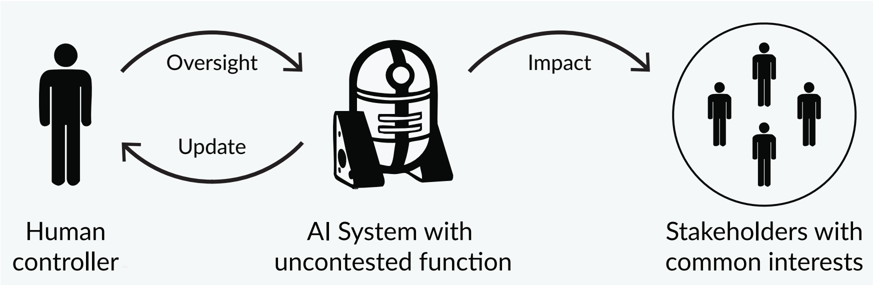
4.2. SITL: Collective Values and Their Limitations

SITL emerged as an extension of HITL, recognizing that AI systems with broad societal impact require oversight beyond individual operators to include society as a whole (Rahwan, 2018). While HITL focuses on embedding individual human judgment in optimization of narrow-impact AI systems, SITL attempts to embed collective societal values in algorithmic governance systems with broad implications (see Figure 1). However, despite its advancements over HITL, SITL still faces notable limitations. It often assumes a single, coherent "society" with shared values, overlooking deep social divisions. It may simply reflect existing power structures rather than fostering meaningful debate. Additionally, standard SITL implementations tend to emphasize gathering input rather than enabling ongoing deliberation.

4.3. DITL: A New Framework for Democratic AI

DITL addresses the limitations of both HITL and SITL by embedding democratic feedback loops—such as contestation, deliberation, and participatory decision-making—directly into the technology and functional use of AI systems. While HITL emphasizes static human oversight and SITL expands oversight to include societal values, neither fully captures the dynamic and iterative nature of democratic deliberation. DITL creates spaces for expressing dissent, makes visible the values embedded in systems, and empowers participants to continuously calibrate the technology, accommodating a plurality of viewpoints rather than assuming consensus (see Figure 2). Moreover, DITL goes beyond simply collecting societal input; it creates mechanisms within AI use for democratic exchanges that support ongoing contestation, deliberation, and reconfiguration of such AI systems.

HITL



SITL

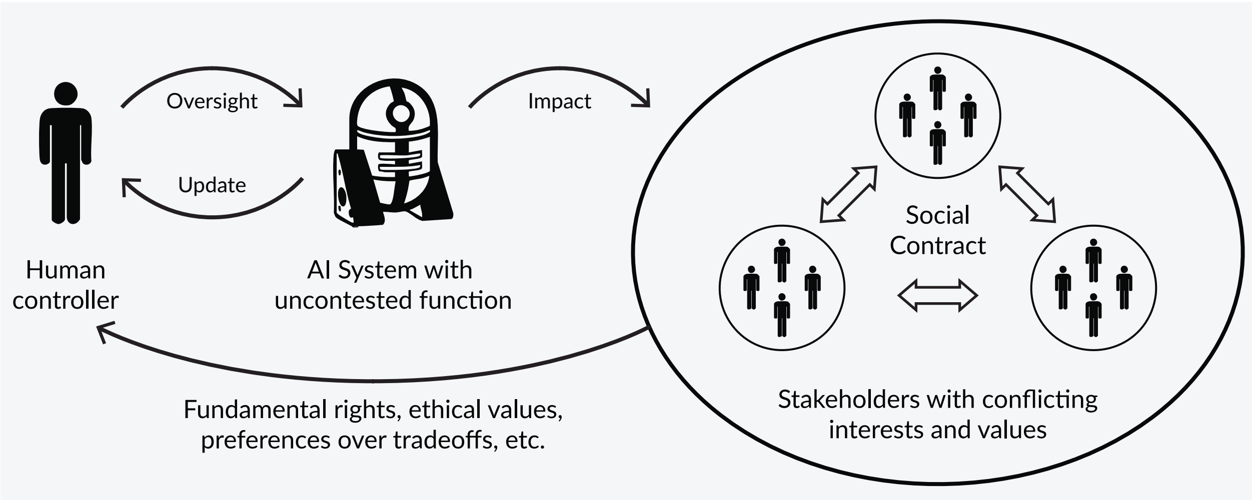


Figure 1. Diagram of the SITL model and illustration of the differentiation between the HITL and SITL models from Rahwan's (2018) introduction.

DITL

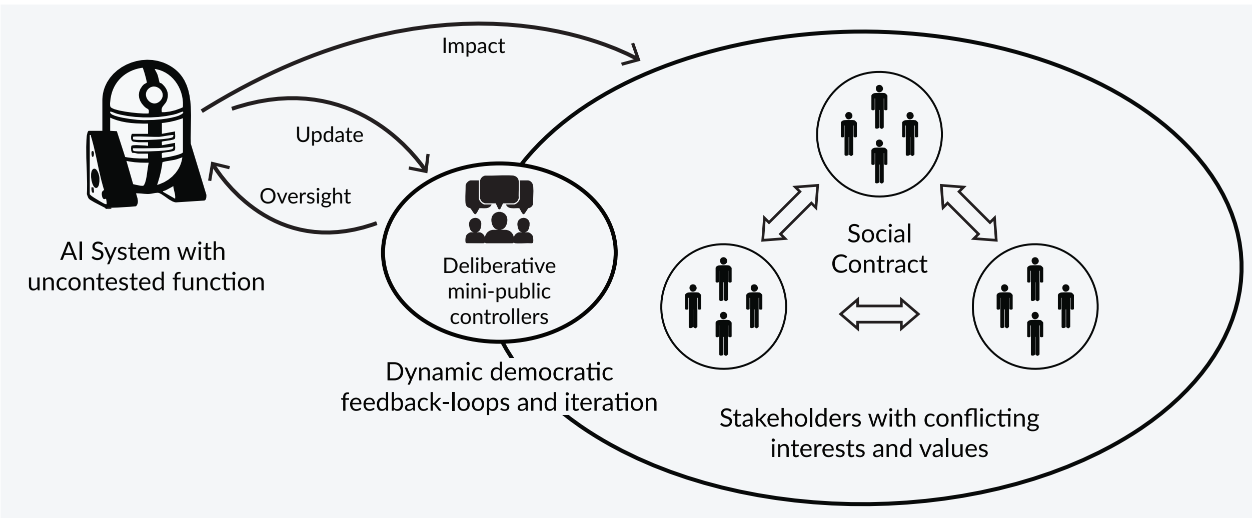


Figure 2. Diagrammatic representation of the DITL model.

4.3.1. Meaningful Frictions: Designing for Deliberation and Contestation

Among the mentioned mechanisms is the intentional design of what Gordon and Walter (2019) call "meaningful inefficiencies," also termed mindful or meaningful frictions (Cox et al., 2016; Pierri et al., in press; Ratto, 2007). Meaningful frictions are intentionally designed points of resistance that disrupt automatic interactions, prompting more deliberate engagement. They create "microboundaries" (Cox et al., 2016) that shift people from automatic to reflective thinking, enabling citizens to consider their actions. They foster collective deliberation, reveal system values, enable contestation of decisions, and help adapt systems to community needs (Elkin-Koren & Perel, 2023).

By creating space for deliberation and contestation, the meaningful frictions within the DITL model help ensure that AI systems promote rather than undermine democratic values. As automated decision-making becomes increasingly relevant to institutionalizing DIs, designing AI to support and not supplant democratic processes is essential. The aim is not to make democratic systems more efficient, but more democratic.

A central part of the DITL approach is revealing the underlying power structures within technology. This involves not only potential biases in datasets or corporate intentions behind the technology, but also transparency about decision-making in its design—clarifying when humans made key choices and when and how machines operate autonomously. While the overall design of the technology, including the placement of friction points, cannot be fully open for participant modification, opening the black box begins with making these design choices visible and understandable.

5. The DDL: Testing DITL in Practice

The DDL, developed under the EU-funded KT4D project, offers an empirically grounded inquiry into how AI can be integrated into DIs. As AI increasingly shapes how knowledge is produced and used, it is vital to assess whether these technologies support rather than compromise participatory governance.

The DDL was a one-day deliberative workshop that explored how AI systems can be integrated into DIs like DMPs. While it was not tested within a DMP as defined in the literature, it was implemented in a deliberative setting. It was facilitated by KT4D partners and involved three different participant groups (see Section 5.1, for details), offering valuable insights into how AI might function in such contexts. Specifically, it aimed to test whether AI could be used to support deliberation, increase democratic and digital literacy, and generate deeper participant engagement, without sacrificing transparency, inclusion, or public trust. Rather than an external decision tool, AI was embedded in the deliberative process to observe its impact on collective reasoning.

Crucially, this experiment marked the first operational iteration of the DITL framework. As described in Section 4, DITL extends previous models such as HITL and SITL by embedding AI systems into live democratic settings in ways that are iteratively contestable, responsive to deliberative input, and sensitive to democratic values. The goal was not just to evaluate AI's functionality but to explore whether and how AI systems can be shaped by—and made accountable to—the democratic practices in which they are embedded.

Over four workshops, the DDL combined a purpose-built digital Demonstrator platform (employing GPT-4, a LLM by OpenAI that generates human-like text, and Retrieval-Augmented Generation that improves accuracy by combining model responses with relevant information retrieved from external databases or documents) with structured in-person facilitation. This hybrid format allowed participants to engage with the “metabolism” of the AI system—its profiling logic, curated data sources, and content generation processes—while reflecting on how it shaped deliberative exchange.

The DDL was not meant to replicate a traditional DMP but served as a research and design tool to assess whether deliberative processes can accommodate AI to strengthen democratic goods like transparency, popular control, and considered judgment. It provides a helpful starting point for theorizing DI amid digital transformation.

5.1. Methodological Approach and Research Setting

The DDL workshops were developed collaboratively within the KT4D project by a democratic design team (The Democratic Society) and a technical development team (Hybridcore). Through an iterative co-design process, the teams created the Demonstrator platform—a transparent, participatory, and deliberately frictional AI tool intended to support, not substitute, democratic reasoning. Its design was guided by three principles: (a) adding value to democratic exchange, (b) overcoming opacity through inspectable and augmentable system design, and (c) operationalizing DITL through embedded frictions and intervention points. All workshops focused on the same topic: how the EU might support a just labor market transition in the approaching AI era.

In each setting, the Demonstrator included a purpose-built chatbot acting as an “expert,” drawing on curated policy, media, and academic sources via Retrieval-Augmented Generation. AI was integrated into the deliberative process as both an information source and a participant in joint reasoning. Deliberation was face-to-face in all four workshops. Participants interacted with the Demonstrator both individually (via computers, tablets, or personal devices) and through a single shared interface projected for the group. Individual tools, such as the profiling quiz and “metabolism” exploration, were used alone; AI prompts, however, were always collectively agreed upon. A facilitator entered the agreed-upon prompt into the system, outputs were displayed to everyone, and the group discussed, critiqued, and refined them as needed. This design avoided parallel private chats and kept the AI embedded in shared discussions.

The workshops took place in four socio-political contexts:

- Brussels (15 October 2024, 10:00–12:30): ~30 participants, mainly from policymaking bodies (50%), civil society organizations (40%), and other sectors (10%).
- Madrid (29 October 2024, 10:00–15:00): 12 participants (ages 18–70), recruited via a partnering NGO, with backgrounds in community education and lifelong learning.
- Kraków (25 October 2024, 10:00–15:00): 16 economics students from Kraków University of Economics.
- Dublin (13 November 2024, 18:00–20:30): 12 participants from the technology and design sectors as part of a digital innovation festival.

Local partners recruited participants: policymakers via organizational invitations (Brussels), citizens via partner networks (Madrid), students via university channels (Kraków), and industry actors via an open festival call (Dublin). The participants knew the general topic beforehand, but the deliberative activities and AI interactions occurred in real time. All sessions followed a shared structure and facilitation guide, with local adaptations for language and context. Data collection combined ethnographic field notes and 60-minute post-workshop facilitator debriefs (later transcribed), forming the empirical basis for our analysis of AI-enabled deliberative tools in practice.

5.2. The Demonstrator Platform and Lab: Hybridizing AI and Human Deliberation

The DDL recognized that AI often aims to optimize speed, scale, and cost, but such goals can undermine deliberative values like reflection, contestation, and inclusion. Instead of pursuing seamlessness, the DDL embedded intentional frictions—deliberate slowdowns and disruptions—to expose the limitations and politics of AI and prompt reflection. Drawing on critical design and science and technology studies traditions (Ananny, 2022; Ratto, 2007; Rettberg, 2022), these frictions operated on multiple levels:

- Technological friction: The Demonstrator used GPT-4 with Retrieval-Augmented Generation but constrained its behavior. Participants could only submit collective prompts, forcing negotiation and shared sense-making beforehand. The system's profiling mechanism was made visible and debatable, with participants invited to reflect on how their own digital personas were constructed. Data sources (10 academic papers, 10 policy documents, and 10 journalistic articles) were explicitly disclosed and selectable by participants, creating space for discussion on bias, credibility, and epistemic inclusion. Even the AI-generated explainer videos, designed to introduce the system's architecture, became a site of friction due to their flat, uncritical tone, prompting users to interrogate the nature of AI explanation itself.
- Facilitation friction: Facilitators guided participants through a structured deliberative arc, introducing a complex policy challenge (the EU's role in a just labor market transition), then slowing down to allow critique, adjustment, and iterative interaction with the Demonstrator. Instead of fixed roles or outcomes, facilitators fostered an open inquiry space where misunderstandings, disagreements, and discomfort could be productive.
- Socio-technical friction: The DDL positioned AI as a constructed, fallible actor, not an objective oracle. AI outputs served as provocations, sparking questions, revealing misalignments, and exposing limits. In this way, the DDL enacted the early principles of DITL—a framework for embedding real-time democratic oversight and reflection within AI-enabled participation.

Indeed, not all friction was planned. As a beta version—the first version of the Demonstrator to go live—it also experienced unexpected glitches and awkward interactions, which Ananny (2022) describes as "algorithmic errors." Some of these breakdowns, such as confusing user interfaces or limited responsiveness, were more frustrating than fruitful. But others generated "technical intuitions (that) function as an interface between technical and human cognizers" (Kronman, 2020, as cited in Rettberg, 2022, p. 4). These emergent flaws underscored a key insight: democratic deliberation benefits not from perfect systems, but from contestable ones.

5.3. Friction as Method: Embedding Contestability into the Digital Democracy Lab

The commitment to friction informed how the Demonstrator platform was developed and how participants engaged with it. By adding procedural hurdles and interpretive ambiguities, they were designed to challenge assumptions about what AI is and how it should operate in democratic settings.

5.3.1. Friction to Expose Assumptions: Profiling and Knowledge Bias

One key friction was the participant profiling mechanism. Early in the workshop, participants completed a 30-question quiz assessing their views on technology, society, and economics. This generated one of eight personas (market proponent, responsible innovator, technology skeptic, balanced regulator, technology enthusiast, social impact advocate, system critic, or worker advocate), which influenced how AI framed its responses.

In typical online systems, profiling is hidden. Here, it was made explicit to encourage reflection on how categorization affects digital interactions. Many participants criticized the fixed personas as polarizing and oversimplified, arguing that they reinforced divisions rather than building common ground. This design decision sparked debate about democratic exchange and the role of categorization in deliberation. Some expressed concerns that the personas could amplify echo chambers instead of promoting dialogue. Others suggested that future tools should focus on shared values rather than ideological divides.

Participants also examined the curated data sources (10 each from academic, policy, and media outlets on the topic of AI and labor markets). By limiting and disclosing these sources, the system made its boundaries visible. This raised questions about which knowledge is valid and how to ensure pluralism in AI training. It also unveiled the importance of local contexts and cultural specificities. While participants in Kraków were dismissive of integrating any data sources stemming from governmental agencies, in Madrid, they felt hesitant towards the integration of news media outlets. These reactions reinforced the importance of epistemic inclusivity and cultural context in AI-supported deliberation.

5.3.2. Friction to Slow Interaction: Collective Prompting

Another intentional friction was the input process. Instead of using individual prompts, participants had to agree on a single collective question. This constraint required negotiation and collaborative framing before using the AI.

During multiple sessions, the constraints to collectively negotiating the prompt for the Demonstrator's chatbot led participants to deliberate over their underlying questions on the topic, rather than solely concentrating on solutions or policy recommendations, which are often the focus in non-tech-augmented DMPs. This also prompted reflection on differences in phrasing, the assumptions behind preferred wordings, and the challenge of representing multiple perspectives. It slowed down the process intentionally, acting as a democratic bottleneck that emphasized deliberation over immediacy. The exercise underscored the value of disagreement and the challenge of reaching a consensus, reinforcing the idea that deliberation benefits from critical, collective engagement rather than seamless input.

5.3.3. Friction to Trigger Deliberation: System Outputs as Objects of Critique

The form of the AI-generated output was itself a deliberate design choice that aimed to create meaningful friction. The textual responses to the group's prompt, presented as pre-formatted policy one-pagers, were often flashpoints for critical engagement. The tone was intentionally flat, technocratic, and policy-like, mirroring the institutional voice often associated with expert systems. Rather than aiming for persuasive or engaging content, the system emphasized synthesis and neutrality. This sharply contrasted with the vivid, nuanced conversations taking place in the room. This apparent neutrality, however, quickly became a subject of critique. Participants questioned the omissions in the responses, debated the weight given to different source materials, and highlighted contradictions or oversimplifications. In doing so, they treated the outputs not as authoritative conclusions, but as starting points for further inquiry—precisely the kind of deliberative stance that DITL seeks to cultivate. Once again, local contexts became evident constraints for the Demonstrator system. In both Madrid and Kraków, participants sought answers to hyperlocal issues they observed in their cities, something the LLM in the Demonstrator could not provide. This was partly due to limited data sources, but LLMs often lack training on culturally and locally coded information, giving way to geo-biased outputs. This shortfall reflects a broader pattern in LLMs, which tend to overrepresent regions with high data availability and dominant cultural perspectives, while lacking adequate training on locally grounded, culturally specific knowledge (Blodgett et al., 2020). Even the AI-generated explainer videos—intended to clarify the system—were criticized for their simplistic delivery, sparking reflection on how AI “speaks” and how its design choices shape perception.

5.3.4. Concluding Insights: Cross-Case Patterns

Beyond the deliberately embedded frictions, the DDL also encountered unplanned disruptions, such as bugs, awkward transitions, and interface limitations. While some of these were simply frustrating, others functioned as what Rettberg (2022) calls productive failures or technical intuitions—moments when a system’s breakdown reveals its underlying logic.

Across the four DDL workshops, participants responded to frictions—both designed and emergent—not with frustration or disengagement, but with heightened curiosity, critique, and reflection. This critical stance should not be interpreted as general public skepticism over AI; research from the KT4D project indicated that citizens often exhibit excessive trust in AI outputs, especially when presented in authoritative “expert” language (Morisseau & Lima, 2024). The frictional design of the DDL—including its deliberately technocratic, policy-like outputs—was meant to counter this tendency, encouraging critical interrogation over passive acceptance. Despite differences in context and background, common themes emerged: a shared skepticism of AI-generated authority, a preference for transparent and inspectable systems, and a desire for democratic agency in shaping how digital tools are used. While the Demonstrator was not embedded in formal decision-making structures, participants approached it more as a civic learning tool: one that encouraged them to interrogate technology, deliberate collectively, and envision alternative uses of AI.

6. Reflections on Uses of AI and Implications for Institutionalization

In this section, we reflect on the implications of our experiment for AI use in DMPs, focusing on how the case relates to the democratic and institutional goods framework. We also offer broader reflections on AI’s promise and limitations in DMPs, particularly regarding the institutionalization debate.

6.1. Inclusiveness

Although we did not control participant selection in the DDL project, inclusiveness—especially epistemic inclusiveness—was central to both deliberation and technology design. This involved making sure that diverse knowledge forms, including marginalized local and experiential perspectives, were represented.

Participants discussed whose voices were excluded from the data sources and how biases might be embedded in LLMs. In Kraków, for example, they debated the need for new democratic institutions to guide training data decisions and amplify overlooked voices. This highlights a broader issue: inclusion is dependent not only on AI access but also on how AI is trained and designed.

To foster algorithmic accountability, we made design decisions, such as profiling quizzes and data selection, transparent through the demonstrator's "metabolism." Although full inclusiveness was not achieved, the DDL considered it a core concern. By revealing how algorithmic systems shape knowledge and participation, the lab encouraged imagining more inclusive data and design approaches.

6.2. Popular Control

A central aim of the DDL was to reconfigure the relationship between participants and technology by fostering popular control. Instead of letting technology steer deliberation, we sought to cultivate active agents who critically engaged with the tools provided. Drawing on Ratto's (2007) concept of "intra-action," we embedded frictions that encouraged participants to actively negotiate their relationship with the system, treating it as a contestable rather than seamless or neutral.

However, these frictions are never neutral. Who defines them—and why—holds power, and poorly designed frictions can become tools of manipulation. The DITL approach mitigates this by making design choices visible and, when possible, open to participant scrutiny and adjustment. This shifts control from the technology's hidden architecture to a shared space where participants can interrogate, influence, and reshape the interaction.

These frictions disrupted the default smooth, invisible tech integration, providing moments for participants to reassert control. This fostered agency within the participant–technology dynamic rather than imposing it externally.

6.3. Considered Judgement

Friction was central to encouraging thoughtful judgment. Instead of prioritizing speed or quick solutions, the DDL built in pauses and interruptions to prompt reflection, inquiry, and contestation. This aligns with Dewey's (2011) notion of civic efficiency, which values slow, collaborative, and messy collective action as the foundation for democratic learning.

Techniques like single-prompt exercises and iterative critiques of AI outputs shifted the focus from immediate solutions to deeper understanding. These moments allowed participants to question assumptions, deliberate across differences, and engage in the reflective judgment Dewey (2011) saw as vital to democratic life.

6.4. Transparency

In the DDL, transparency focused on enabling participants to understand how the technology was developed and used, not just on providing technical explanations. We made the Demonstrator's "metabolism" visible, showing decision points, roles, and how the system evolved. This allowed participants to see not only the

AI's outputs but also the human decisions behind them. The approach frames transparency in a democratic context, allowing participants to scrutinize and challenge the social and political aspects of how the technology was designed.

6.5. Efficiency

Efficiency was not the DDL's primary goal, but it remains relevant. While AI promises to streamline deliberation, traditional efficiency often conflicts with deliberative quality. By intentionally slowing down the process through friction, we prioritized effectiveness over speed, creating opportunities for more thoughtful democratic engagement.

The case also highlights the continued importance of human facilitators. Although automation is proposed for scaling DMPs cost-effectively, the DDL—especially the DITL model—shows that human facilitation remains essential. Guiding discussions, managing discomfort, and supporting iterative critique are tasks that AI cannot (yet) perform.

AI can provide information, suggest directions, or moderate basic exchanges, but it does not have the sensitivity to recognize when confusion is productive, disagreements need unpacking, or emotional and social reflection are necessary. These core democratic facilitation tasks require human sensitivity and improvisation, which AI cannot perform without risking a flattening of deliberative depth.

6.6. Transferability

When introducing AI into deliberative processes, a key question is how well it adapts to different contexts. In the DDL, the Demonstrator was translated into the three majority languages of participating countries. However, linguistic translation alone cannot guarantee transferability. As the case study shows, cultural and contextual differences shaped how participants engaged with the technology and data—participants in Kraków were wary of government-produced data, while some in Madrid were uncomfortable with specific news sources.

These differences underscore the risk of cultural blindness in technology development and deployment and the need for a strong cultural lens (Ahern et al., 2024) in both AI regulation and its integration into democratic processes across all geographic scales.

Although AI struggles with local contexts and languages, it does not limit its use to national-scale DMPs. Instead, it highlights the need for culturally sensitive adaptation to ensure relevance and acceptance at local, regional, and national levels.

6.7. Broader Reflections on Institutionalization

The introduction of AI into DMPs lies at the heart of the institutionalization debate. Scholars have long noted the tension between institutionalizing DIs and preserving flexibility and creativity. While institutionalization ensures sustainability and adoption by embedding processes within existing governance structures, it can also reduce experimentation and critical contestation, which are essential for healthy deliberation. AI heightens this tension, potentially turning participatory spaces into a technocratic mechanism that prioritizes efficiency over reflection and deliberation.

The DDL experiment illustrates this tension. It serves as a prototype for testing AI in deliberative processes while trying to preserve democratic goods. While AI promises efficiency and scalability, we argue that these benefits often diminish democratic engagement. In the DDL, AI was employed not for standardization or rapid decisions, but to support deliberation through deliberate friction—pauses, critiques, and iteration—ensuring space for critical engagement, reflection, and contestation. Accomplishing this, however, relied heavily on facilitation. The frictions embedded in the process occasionally made discussions slower or more complex, requiring facilitators to navigate participants through technical explanations, surface hidden assumptions, and maintain the productivity of the debates concerning the technology itself. This points to a broader design implication: a democratically effective application of AI in DMPs not only necessitates rethinking the deliberation format to integrate structured, critical reflection on the technology, but also requires facilitators with strong critical digital literacy. Such facilitators must be capable of questioning the technology as rigorously as the policy matter under discussion, ensuring that AI remains a tool for democratic reasoning rather than an invisible driver.

A central feature of the DDL was its flexible design, which adapted to cultural nuances and participants' specific interactions with the AI system. Rather than following a rigid script, the process evolved in a pre-considered direction, but dynamically through pauses, critiques, and iterative exercises. This flexibility afforded space for participant agency and intra-activity, fostering creativity and unanticipated outcomes essential to democratic goods in AI-supported deliberation.

Flexibility was protected by resisting early formalization. As DMPs scale, there is pressure to standardize them. The DDL pushed back against this, using AI to support rather than constrain deliberation. Continuous iteration kept the process adaptable, participatory, and resistant to being locked into a fixed model.

Finally, the DDL embedded mechanisms for reflection, contestation, and co-creation. Friction points enabled participants to assess AI outputs and challenge underlying assumptions, giving them agency to shape both the process and the technology. This ensured deliberation remained a participatory rather than technocratic process. Maintaining space for contestability kept the process responsive and creative, even with AI at the core.

6.8. Implications of the Digital Democracy Lab Case Study

Our case study suggests that while AI can help expand and institutionalize DMPs, it should be handled with caution. Implementing AI too quickly, without room for contestation, flexibility, and reflection, can harm deliberative quality. The DDL shows experimentation is essential to preserve democratic goods. Consequently, AI should not only speed up or optimize processes but also support deeper engagement that respects the complexity of deliberation and allows for ongoing adaptation. Balancing efficiency with democratic effectiveness is key to ensuring AI enhances rather than undermines fundamental democratic goods.

7. Conclusion

The introduction of AI into DIs presents both opportunities and challenges for institutionalization. Our DDL case shows AI can enhance the scale, efficiency, and accessibility of DMPs, but these benefits must be balanced against risks to the democratic goods that give such innovations their value. The DITL model—with its contestational design, meaningful frictions, and intra-activity—offers a promising way to advance technology without sacrificing democratic quality.

Our research indicates institutionalizing AI-embedded DIs requires more than just technical implementation; it demands thoughtful integration that maintains participant agency, critical reflection, contestation, and collective deliberation. Participants valued transparency and contestability over seamless efficiency and appreciated opportunities to scrutinize and shape AI systems. This highlights the need for designing democratic processes open to adaptation, critique, and ongoing negotiation of values.

Beyond DMPs, this research suggests a broader reconceptualization of institutionalization—not as standardization or formalization, but as embedding democratic reflexivity within socio-technical governance. This entails creating institutional arrangements that remain flexible, adaptable, and responsive to citizen input, not despite but because of their integration with AI. Meaningful frictions become essential features of effective, democratically legitimate institutions rather than obstacles to efficiency.

As DIs evolve alongside emerging technologies, further research should explore DITL principles across contexts and scales. Future work should examine how AI might be designed for more formal deliberative settings, how meaningful frictions can be made accessible to diverse participants, and how AI governance in democratic contexts can itself become more participatory and intra-active. These efforts are crucial for institutionalizing DIs that leverage technology while upholding democratic values.

Funding

This work was supported by the European Research Executive Agency under grant agreement No. 101094302 (KT4D), the EU's Horizon Europe research and innovation programme under grant agreement No. 101112280 (Nets4Dem), and the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia, in accordance with the Agreement on the Implementation and Financing of Scientific Research No. 451–03-136/2025–03/200025.

Conflict of Interests

The authors declare no conflict of interests.

References

Ahern, J., Edmond, J., Lima, E., Clarke, E., Osipov, A., Akar, F., & Savini, G. (2024). *Culture's role in navigating technological change: The KT4D perspective on recent developments in European AI policy*. Knowledge Technologies for Democracy. https://doi.org/10.5281/zenodo.14768234

Alnemr, N. (2020). Emancipation cannot be programmed: Blind spots of algorithmic facilitation in online deliberation. *Contemporary Politics*, *26*(5), 531–552. https://doi.org/10.1080/13569775.2020.1791306

Alnemr, N. (2021). *Citizens and intelligent machines: Algorithms in deliberative democracy* [Unpublished doctoral dissertation]. University of Canberra. https://doi.org/10.26191/e5jb-9746

Ananny, M. (2022). Seeing like an algorithmic error: What are algorithmic mistakes, why do they matter, how might they be public problems? *Yale Journal of Law & Technology*, *24*, 342–364.
Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of Machine Learning Research*, *81*. https://doi.org/10.48550/arXiv.1712.03586
Blodgett, S. L., Barocas, S., Daumé, H., III., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5454–5476). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.485
Bussu, S., Bua, A., Dean, R., & Smith, G. (2022). Introduction: Embedding participatory governance. *Critical Policy Studies*, *16*(2), 133–145. https://doi.org/10.1080/19460171.2022.2053179
Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. The MIT Press.
Courant, D. (2022). Institutionalizing deliberative mini-publics? Issues of legitimacy and power for randomly selected assemblies in political systems. *Critical Policy Studies*, *16*(2), 162–180. https://doi.org/10.1080/19460171.2021.2000453
Cox, A. L., Gould, S. J. J., Cecchinato, M. E., Iacovides, I., & Renfree, I. (2016). Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1389–1397). Association for Computing Machinery. https://doi.org/10.1145/2851581.2892410
Curato, N., Farrell, D. M., Geissel, B., Grönlund, K., Mockler, P., Pilet, J.-B., Renwick, A., Rose, J., Setälä, M., & Suiter, J. (2021). *Deliberative mini-publics: Core design features*. Bristol University Press.
Dewey, J. (2011). *Democracy and education*. Simon & Brown.
DiSalvo, C. (2022). *Design as democratic inquiry: Putting experimental civics into practice*. The MIT Press.
Elkin-Koren, N., & Perel, M. (2023). Democratic friction in speech governance by AI. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 643–655). Edward Elgar Publishing. https://doi.org/10.4337/9781803928562.00066
Elstub, S., & Escobar, O. (2019). Defining and typologising democratic innovations. In O. Escobar & S. Elstub (Eds.), *Handbook of democratic innovation and governance* (pp. 11–31). Edward Elgar Publishing.
Escobar, O. (2022). Between radical aspirations and pragmatic challenges: Institutionalizing participatory governance in Scotland. *Critical Policy Studies*, *16*(2), 146–161. https://doi.org/10.1080/19460171.2021.1993290
Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
Flanigan, B., Gölz, P., Gupta, A., Hennig, B., & Procaccia, A. D. (2021). Fair algorithms for selecting citizens' assemblies. *Nature*, *596*, 548–552. https://doi.org/10.1038/s41586-021-03788-6
Gelauff, L., Nikolenko, L., Sakshuwong, S., Fishkin, J., Goel, A., Munagala, K., & Siu, A. (2023). Achieving parity with human moderators: A self-moderating platform for online deliberation. In S. Boucher, C. A. Hallin, & L. Paulson (Eds.), *The Routledge handbook of collective intelligence for democracy and governance* (pp. 202–221). Routledge.
Goodin, R. E. (2000). Democratic deliberation within. *Philosophy and Public Affairs*, *29*(1), 81–108.
Gordon, E., & Walter, S. (2019). Meaningful inefficiencies: Resisting the logic of technological efficiency in the design of civic systems. In R. Glas, S. Lammes, M. de Lange, J. Raessens, & I. de Vries (Eds.), *The playful citizen: Civic engagement in a mediatized culture* (pp. 310–334). Amsterdam University Press.
Lacelle-Webster, A., & Warren, M. (2021). *Citizens' assemblies and democracy*. Politics. https://doi.org/10.1093/acrefore/9780190228637.013.1975

Landemore, H. (2024). Can artificial intelligence bring deliberation to the masses? In R. Chang & A. Srinivasan (Eds.), *Conversations in philosophy, law, and politics* (pp. 39–69). Oxford University Press. https://doi.org/10.1093/oso/9780198864523.003.0003

McKinney, S. (2024). Integrating artificial intelligence into citizens' assemblies: Benefits, concerns and future pathways. *Journal of Deliberative Democracy*, *20*(1). https://doi.org/10.16997/jdd.1556

Morisseau, T., & Lima, E. (2024). *KT4D social risk toolkit—Modules A, B, and C* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.11491487

Newton, K. (2012). Curing the democratic malaise with democratic innovations. In B. Geissel & K. Newton (Eds.), *Evaluating democratic innovations: Curing the democratic malaise?* (pp. 3–20). Routledge.

OECD. (2020). *Innovative citizen participation and new democratic institutions: Catching the deliberative wave*. https://doi.org/10.1787/339306da-en

Ohren, A. (2024). *The representative claims of deliberative mini-publics* [Unpublished doctoral dissertation]. Norwegian University of Science and Technology. https://hdl.handle.net/11250/3132395

Pierri, P., Calderón Lüning, E., & Stearns, M. (in press). The virtues of friction: Designing for democratic efficacy in an age of optimisation. In R. Cooper, S. Junginger & Y. Bakirlioglu (Eds.), *The handbook of design management*. Bloomsbury.

Pogrebinschi, T. (2018). Can participatory governance improve the quality of democracy? A response from Latin America. In H. Heinelt (Ed.), *Handbook on participatory governance* (pp. 94–116). Edward Elgar Publishing. https://doi.org/10.4337/9781785364358.00011

Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, *20*, 5–14. https://doi.org/10.1007/s10676-017-9430-8

Randma-Liiv, T. (2023). Adoption is not enough: Institutionalization of e-participation initiatives. *Public Policy and Administration*, *38*(3), 329–351. https://doi.org/10.1177/09520767211069199

Ratto, M. (2007). Ethics of seamless infrastructures: Resources and future directions. *International Review of Information Ethics*, *8*(8), 20–27.

Rettberg, J. W. (2022). Algorithmic failure as a humanities methodology: Machine learning's mispredictions identify rich cases for qualitative analysis. *Big Data & Society*, *9*(2). https://doi.org/10.1177/20539517221131290

Revel, M., & Penigaud, T. (2025). *AI-facilitated collective judgements*. SSRN. https://doi.org/10.2139/ssrn.5167340

Sloane, M., Moss, E., Awomolo, O., & Forlano, L. (2022). Participation is not a design fix for machine learning. In *EAAMO '22: Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (pp. 1–6). Association for Computing Machinery. https://doi.org/10.1145/3551624.3555285

Smith, G. (2009). *Democratic innovations: Designing institutions for citizen participation*. Cambridge University Press.

Theuwis, M., van Ham, C., & Jacobs, K. (2025). A meta-analysis of the effects of democratic innovations on participants' attitudes, behaviour and capabilities. *European Journal of Political Research*, *64*(2), 960–984. https://doi.org/10.1111/1475-6765.12722

About the Authors

Arild Ohren is a researcher at the Norwegian University of Science and Technology and a senior democratic expert at Democratic Society. His work focuses on democratic theory, democratic innovation, and citizen participation.

Elizabeth Calderón Lüning is mission lead at Democratic Society for Emerging Technologies and Democracy and associate researcher at the Weizenbaum Institute. Her work focuses on digital sovereignty, participatory democracy, and urban transformation, bridging civil society, politics, and academia through transdisciplinary research and teaching.

Čedomir Markov is a research fellow at the Institute for Philosophy and Social Theory, University of Belgrade. His research explores audience–media relations in hybrid regimes, focusing on trust, cynicism, and anti-press hostility. He examines the intersection of media and democratic innovation, particularly regarding the quality of deliberation and media accountability.

Max Stearns is co-founder of errantry media lab, where he works across disciplines to engage with problems that resist conventional framing. His work explores the intersection of civic innovation, democratic design, and emerging technologies, focusing on creating expansive rather than merely inclusive infrastructures that serve human agency and democratic values.