

A Replica for our Democracies? On Using Digital Twins to Enhance Deliberative Democracy

Claudio Novelli¹, Javier Argota Sánchez-Vaquerizo², Dirk Helbing^{2,3}, Antonino Rotolo⁴, Luciano Floridi^{1,4}

¹ Digital Ethics Center, Yale University, 85 Trumbull Street, New Haven, United States

² ETH Zürich, Computational Social Science, Stampfenbachstrasse 48, 8092 Zürich, Switzerland

³ Complexity Science Hub Vienna, Metternichgasse 8, 1030 Vienna, Austria

⁴ University of Bologna, Alma AI and Department of Legal Studies, Via Zamboni 22, Bologna, Italy

Abstract. Deliberative democracy depends on carefully designed institutional frameworks — such as participant selection, facilitation methods, and decision-making mechanisms — that shape how deliberation occurs. However, determining which institutional design best suits a given context often proves difficult when relying solely on real-world observations or laboratory experiments, which can be resource-intensive and hard to replicate. To address these challenges, this paper explores Digital Twin (DT) technology as a regulatory sandbox for deliberative democracy. DTs enable researchers and policymakers to run “what-if” scenarios on varied deliberative designs in a controlled virtual environment by creating dynamic, computer-based models that mirror real or synthetic data. This makes systematic analysis of the institutional design possible without the practical constraints of real-world or lab-based settings. The paper also discusses the limitations of this approach and outlines key considerations for future research.

1. Introduction

Pluralist societies naturally experience significant disagreement. Disagreements can be amplified in virtual environments, where individuals often encounter beliefs reinforcing their preexisting views. This polarization distorts perceptions of division, weakens civic dialogue, and reduces political engagement. Perspectives that could once have been seen as reasonably compatible are now frequently regarded as mutually exclusive. Democracies, at present, struggle to address this challenge effectively.

Deliberative democracy has long been proposed as a solution, emphasizing communication and reasoned debate to improve civic engagement. Unlike aggregative democracy, which focuses on vote-counting and advancing interests, deliberative democracy prioritizes public reasoning and collective decision-making (Andre Bächtiger et al. 2018).¹

Deliberative democracy is built upon reasoned debate and mutual justification, highlighting the importance of argumentation and dialogue in public decision-making. While there is broad consensus on the necessity of justifying laws and fostering respectful civic discourse, the effectiveness of deliberative democracy as a viable alternative to traditional aggregative approaches largely

¹ In practice, this conceptual opposition does not manifest in such a sharp way, having liberal modern liberal democratic a mix of processes of deliberations and aggregations

depends on careful institutional design. This design includes structural and procedural elements such as participant recruitment (who should take part), training and information provided to participants, debate structure, timing (frequency of deliberative rounds), decision-making, and monitoring (Fung 2007).

However, selecting the most suitable institutional design to fulfill specific deliberative goals is challenging. Testing and refining such institutional designs—and assessing their impact on the quality of deliberation—poses significant challenges in both real-world settings and lab experiments, mainly due to the generalizability and impartiality of the results.

To address these limitations, this paper explores how Digital Twin (DT) technology, a dynamic computational modeling framework, can serve as a regulatory sandbox for deliberative democracy. DTs mirror deliberative communities, whether real or synthetic, by integrating real-world data (e.g., demographic, behavioral, or social data) and advanced computational techniques (e.g., agent-based modeling, machine learning, or network analysis). Institutional designers, including scientists, policymakers, and public or private advocates, can use this experimental tool to simulate deliberative processes, test different design options, generate predictive insights, and refine deliberation procedures in controlled virtual environments.

The paper is structured as follows. Section 2 outlines the foundational principles of deliberative democracy. Section 3 categorizes the procedural rules of deliberation and explores challenges in modeling and testing these rules using real-world and lab-controlled methods (Section 3.1). Section 4 introduces Digital Twin (DT) technology's core features and potential for replicating complex social systems. Section 5 examines its application to deliberative democracy as a sandbox for testing procedural rules. Section 5.1 outlines the development of a deliberative community DT, using mini-publics as a case study. Section 5.2 demonstrates how the DT can be used to test and refine the institutional design. Section 6 concludes the article by discussing the limitations of using DTs in this area and highlights key considerations for future research.

2. Deliberative democracy: the cogency of the best argument

Traditionally, the shared core of what is called deliberative democracy is the belief that the legitimate source of law derives not merely from a straightforward aggregation of people's preferences in the face of conflicting interests but rather from the exercise of reason and mutual respect directed toward the common good (Mansbridge 2007; Gutmann and Thompson 2009). Consequently, a key aspect of deliberative democracy involves viewing civic engagement as a communal commitment to mutual justification and openness to persuasion. Open discussion

is frequently recognized as one of the most effective ways to challenge our (instinctive) presuppositions and assumptions about the world (Haidt 2012).

Deliberative democracy differs from traditional aggregative democracy in various ways. From a foundational perspective, it does not rely on any single, fixed principle to determine whether a procedure or law is justified; rather, its principles are inherently dynamic and subject to ongoing revision in light of new moral and political arguments (Gutmann and Thompson 2000).

From a procedural perspective, deliberation represents a distinct decision-making paradigm. Unlike aggregative democracy, which reduces decision-making to the mere mechanistic expression of preferences through voting, deliberative democracy emphasizes argumentative complexity and aims to achieve higher standards of rationality—both substantively and procedurally (Dryzek 2002).² Substantively, it seeks outcomes reflecting the common good, informed by individuals advancing their interests and collective ends. Procedurally, it prioritizes inclusive and open discussion, brainstorming, information-pooling, and collaborative problem-solving (Fung 2007).

Scholars debate deliberation's ideal outcomes: Habermas seeks rational consensus (Habermas 1998), Elster emphasizes preference transformation through reasoned argument (Elster 1998), and Cohen supports rationally motivated consensus but acknowledges the need for voting when consensus is unattainable (Bächtiger et al. 2018, 7). Participatory Budgeting illustrates this by enabling citizens to collectively deliberate on budget allocations while relying on aggregative voting mechanisms at key stages (preference aggregation). Deliberation thus serves multiple goals, including epistemic improvements in collective knowledge (Estlund 1993; Nino 1996; Marti and Besson 2006) — all of which share the idea that outcomes should be the most rationally justified under conditions of disagreement.

Where is deliberative democracy practiced? Formal state institutions—parliaments, executives, and courts—contain deliberative elements in the form of debate and discussion. However, they are not the purest exemplars of deliberative democracy, being more marked by strategic communication and the pursuit of individual or group interests than by genuine reason-based mediation. In such settings, the best-positioned argument still prevails over the most rationally robust one (Cohen 2021).

More explicitly inspired by deliberative standards are non-partisan forums—often called mini-publics when they involve relatively few participants, and citizens' assemblies or deliberative polls when larger numbers are involved. In both formats, participants meet face-to-face and typically develop a

² Real-world democratic practices often blend these two models rather than adhering strictly to one. A notable example is Participatory Budgeting, which enables citizens to collectively deliberate on budget allocations while relying on aggregative voting mechanisms at key stages (preference aggregation).

recommendation or report on a policy issue (André Bächtiger et al. 2018; Fournier 2011). Topics these bodies address include the risks and promises of new technologies, climate change, electoral system reform (e.g., the British Columbia Citizens' Assembly), urban development, and fiscal policies (Fishkin 2013; Landemore 2021; Willis, Curato, and Smith 2022).

2.1. Deliberation in complex societies

Modern, complex societies operate under the shared principle—across democracy, capitalism, and science—that “the best idea(s)” should prevail. Achieving this, however, requires establishing suitable settings, ensuring good opportunities for everyone to (co-)create ideas, fair competition among ideas that will not be biased by interests, and suitable procedures to identify and select the best ideas.

However, no single idea can fully address the needs and interests of all members in a pluralistic society. This is because modern societies thrive based on a division of labor and roles mutually complementing each other. Consequently, people have different lives as well as diverse interests and needs. Such societies must balance multiple objectives simultaneously rather than pursue a single overriding goal. Importantly, societies need many kinds of minorities to thrive, such as innovators. “The best ideas” will, therefore, have to consider and address a suitable scope of diversity to allow all members of society to thrive and contribute to its functioning with their various talents, skills, knowledge, and resources. Implementing “the best ideas” should act as a catalyst that brings the potential of diverse actors and groups of society to fruition.

Deliberative democracies aim to create a framework that supports all of this. Hereby, deliberation identifies, considers, supports, and unleashes the required diversity and its potential by promoting constructive dialogue that brings different perspectives and ideas together. Furthermore, it aims to develop the integrated ideas resulting from this process. Because of human and social complexity, but also to allow for flexibility, implemented ideas should not seek to micro-manage people or societies but to assist their self-organization while considering societal and democratic values.

Against this background, deliberative democracy still offers several advantages over aggregative democracy, particularly in focusing on the public power exercised through reasoning under idealized conditions of information and equality (Cohen 2007). By fostering deliberative practices, individuals are encouraged to present reason-based arguments that can be justified on shared grounds. Exposure to such arguments may shift participants' preferences—potentially reducing polarization (Dryzek et al. 2019)—and ultimately strengthen the legitimacy of laws as they emerge from inclusive and rational public debate rather than power struggles.

At the same, empirical research indicates that while deliberative democracy has the potential to produce positive outcomes, its success is not guaranteed (Cohen 2007, 230 ff.). Effective implementation depends on a well-designed institutional framework that promotes diverse participation and aligns with the broader socio-political context. This will be the focus of the next section.

3. The Institutional Design of Deliberative Democracy: A Three-Rule Framework

Deliberation differs from free discussion, as its success heavily depends on its institutional design (Grönlund and Herne 2022, 170). Institutional design is shaped by a series of rules that govern different stages of the procedure: before, during, and after the deliberation. Procedural rules determine critical aspects, such as who participates, how the process unfolds, and what is done with the outcomes (Parkinson and Mansbridge 2012; Fishkin 2009; Gutmann and Thompson 2000). Together, they create the framework within which deliberative democracy operates:

- a) *Pre-deliberation rules* set up the deliberation process. They define the composition of participants —ranging from random selection (Gastil and Richards 2013) or targeted stakeholder inclusion to purely voluntary participation – and address agenda setting, which can be predetermined, defined by facilitators, or collaboratively group-defined. Equally important is the informational preparation of participants, who may be equipped with expert presentations, background reading materials on the policy problem (Weeks 2000), or access to open data to ensure they are well informed about the topic under discussion (Ruijter et al. 2024). This preparatory phase lays the foundation for meaningful deliberation.
- b) *Discussion rules* govern how the process is conducted and how participants engage with one another. For example, the debate format may consist of a single session, multiple rounds, or iterative discussions that build on previous exchanges. Speaking turns can be carefully structured—whether controlled by a facilitator, constrained by time limits in open-floor discussions, or managed through queued requests. Similarly, decision-making approaches vary, ranging from consensus-building and majority voting (Cohen 2021) to methods like deliberative polling, which emphasize informed opinion (Fishkin and Luskin 2005). The role of facilitators — who guide the deliberation—can differ depending on the chosen approach, whether they aim for neutral mediation, prioritize specific agenda items, or rotate facilitation duties among participants (Escobar 2019; Moore 2012). Other aspects, such as how arguments and rebuttals are handled –

e.g., open or paired rebuttals – mechanisms for resolving conflicts – e.g., structured dialogue or majority rule in stalemates – and criteria for determining “winning” arguments – e.g., iterative refinement or argument scoring/ranking – are all carefully designed to ensure a productive process (Bächtiger and Parkinson 2019; Moshman 2020).

- c) *Post-deliberation rules* address what happens after the deliberation, focusing on outcomes and their follow-up (Gutmann and Thompson 2000). These include documenting results, such as summary reports or actionable recommendations, and gathering participant feedback through surveys or anonymous evaluations (Hartz-Karp 2005). Depending on the objectives and privacy considerations, the resulting reports may be made public or kept confidential. Monitoring the implementation of decisions is also crucial and can consist of public updates, audits, or similar accountability measures. These steps ensure that deliberation has a tangible impact on policy or practice (Stark, Thompson, and Marston 2021).

The content of these rules and the choice between different options are heavily influenced by (i) the model of deliberative democracy being considered and (ii) the quality assessment metric of deliberation. For instance, a more communitarian model might prioritize consensus-building and mutual understanding, whereas a liberal model might emphasize fair representation and structured debate (Forst 2001).

Regarding the quality of deliberation, various metrics have been proposed, all broadly inspired by core values historically associated with deliberative democracy, such as inclusiveness, respect, equality, and reason-giving. However, each metric emphasizes these values differently. Some of the most widely recognized metrics include discourse quality indices (Steenbergen et al. 2003; Bächtiger, Gerber, and Fournier-Tombs 2022), the group deliberative reason index (Niemeyer and Veri 2022), levels of justification (Bächtiger et al. 2010), the listening quality index (Scudder 2022), and the extent of respectful engagement (Mansbridge et al. 2012).

A full exploration of these evaluation metrics and their ideal applications lies beyond the scope of this paper. Instead, we address a distinct, cross-cutting issue: testing the feasibility of different deliberative institutional designs. The outcomes of such feasibility tests should inform which procedural rules are most suitable among the available options. For instance, we might ask whether a particular setting calls more for a single debate session or multiple rounds. This inquiry applies to various expected deliberation goals and the various metrics used to evaluate them.

3.1. The challenges of modeling and testing deliberative democracy

Because deliberative democracy is primarily a normative theory and lacks a single universal model for deliberations, testing its pre-deliberation, discussion, and post-deliberation rules is challenging. Nevertheless, empirical approaches remain relevant (Grönlund and Herne 2022, 166).

One way to study deliberation is to analyze real-world settings where an actual community deliberates on a concrete issue. Alternatively, researchers can employ controlled experimental designs—such as lab-in-the-field experiments, in which a sample of the population interacts (online or in presence) in a staged deliberative format, or scenario experiments (Werner and Muradova 2022). A third alternative would be to use traditional simulation techniques that we shall not discuss here, as they typically rely more heavily on theoretical assumptions than empirical data.

These methods enable researchers to study causal mechanisms — such as how participation in a mini-public influences individual opinions (Setälä, Grönlund, and Herne 2010) or how different voting methods affect the perceived legitimacy of collective decision-making (Hausladen et al. 2024) — by isolating and testing specific variables, such as discussion format and participant composition, while holding other factors constant. This approach helps determine the causal impact of deliberation (Grönlund and Herne 2022; Kingzette and Neblo 2022).³ Thus, the key advantage of these experimental methods is that they allow researchers to focus on individual components of the deliberations—and their governing rules—to measure their effects more precisely. This, in turn, enhances the replicability of research across different contexts, including various countries. Nevertheless, both real-world and lab-in-field experiments have their limitations:

- 1) *Replicability*. Real-world deliberative settings often involve unique socio-political, cultural, and institutional features that make them difficult to reproduce and impede the testing of causal mechanisms across multiple contexts. In addition, real-world deliberations depend heavily on political will, funding, or community enthusiasm, which may not be consistently available for subsequent iterations. In contrast, lab-in-the-field experiments offer greater replicability by allowing researchers to follow standardized procedures across different samples (Grönlund and Herne 2022). However, this advantage comes with a trade-off: controlled experiments often simplify or strip away the complexity of real-world contexts to ensure replicability. While this enables researchers to isolate and test specific variables, it raises questions about the extent to which

³ A 2012 Finnish study on enclave deliberation and group polarization tested how like-minded versus mixed-opinion discussion groups affected opinion shifts on immigration policy (Grönlund, Herne, and Setälä 2015).

the findings are reliable enough to be applied to real-world settings. Deliberation within complex socio-technical systems involves numerous intricate causal relationships among participants, institutions, and broader societal influences, as well as emergent behaviors. These complexities inherently make full replication or control challenging in both real-world and experimental settings.

2) *Generalizability*. Both real-world deliberations and lab experiments encounter significant challenges in generalizing their results—that is, in applying insights from specific cases or studies to broader or more diverse contexts. Real-world deliberations are typically unique, context-specific events, making it challenging to transfer lessons across different policy domains (Levine 2005; Parkinson 2006)—for instance, an environmental policy issue in one region may not translate well to a healthcare policy in another. Likewise, experimental designs often rely on small, carefully selected samples that, while representative in some respects, may not capture the full diversity of perspectives, interests, and backgrounds found in larger populations. This limitation underscores scalability challenges, as expanding the number of participants or iterations is often impractical within these methods (Friedman 2006). Finally, experimental studies tend to focus on short-term outcomes – such as immediate opinion changes – due to the constrained timeframes inherent in experimental designs, whereas measuring deliberative democracy requires observing long-term dynamics.

3) *Flexibility to iterations and follow-ups*. Identifying the "best fit" for a deliberation process often involves multiple testing cycles, learning from feedback, and refining the approach iteratively. A key challenge in real-world testing is that outcomes are only observable after the process is complete, leaving no opportunity to refine the procedure and assess how those refinements might improve the deliberation. Once a real-world deliberation begins, pausing, modifying, and restarting the process is difficult.

In theory, this issue may be less pronounced in laboratory experiments. However, researchers often maintain fixed procedures to ensure the reliability and consistency of their results. This rigidity can limit opportunities for iterative improvements or real-time adaptations (Lee et al. 2022). Furthermore, if participants become aware of changes or adaptations during testing, their behavior may be influenced—a concern that leads directly to the next problem.

4) *Observer Effect*. Compared to real-world scenarios, a specific challenge of lab-in-the-field experiments is the so-called observer effect (or

Hawthorne effect) (Oswald, Sherratt, and Smith 2014). In lab or online experiments, participants are typically more aware that they are part of a study, which can compromise the authenticity of the deliberation process. This heightened awareness may lead them to behave differently than in real-world political or civic contexts, where the pressures and dynamics of observation are often less explicit. As a result, observation may also change the outcomes of deliberation, e.g., inducing conservative decision-making. Accordingly, participants in a deliberative democracy experiment may alter their behavior when they know peers, media, or researchers are observing them (Gastil 2000; Carpini, Cook, and Jacobs 2004). For example, individuals may feel less inclined to express dissent or voice unpopular opinions and more inclined to agree with dominant opinions or authority.

- 5) *Participant compliance, retention, and engagement.* Ensuring that participants (or facilitators) adhere to agreed-upon protocols and remain engaged throughout the process is a significant challenge in real-world deliberations and experimental settings. Logistical disruptions or a lack of perceived benefits can lead to protocol deviations or high dropout rates, jeopardizing data integrity and representativeness.

Time constraints often exacerbate this challenge. While it is generally easier to enforce protocols (e.g., turn-taking rules, discussion time limits) and maintain engagement in lab or online settings, not all participants may fully comply, particularly if the task is lengthy or perceived as low-stakes. Since participants often receive only a one-time incentive, dropout rates can be high after initial recruitment, thereby endangering sample sizes and the representativeness of the study.

- 6) *Resource and time constraints.* Conducting repeated or large-scale experiments to refine procedural rules is resource-intensive, requiring substantial financial and human investment (Iyengar et al. 2003). Effective testing involves organizing deliberative events, recruiting and compensating participants (if applicable), training facilitators, securing venues, providing expert materials, collecting and analyzing data, and managing logistics. Deliberative processes often span multiple sessions and require follow-ups, which can be challenging given the limited availability of participants and organizers. Lab experiments demand funding for recruitment, platform development, and participant incentives, particularly when conducted across diverse contexts.

The following section examines how digital twin technology can address these limitations and help model and test deliberative democracy practices.

4. The Digital Twin Technology

Multiple definitions of Digital Twins (DTs) exist in the literature. However, they generally align on a core idea: a DT is a dynamic, computer-based model that replicates a physical entity—such as an object, process, person, or human interactions—using real-time data to mirror its behavior, performance, and evolution (Barricelli, Casiraghi, and Fogli 2019; L. Zhang, Zhou, and Horn 2021). Echoing early visions of digital worlds (Gelernter 1991), the concept was coined initially for life cycle management in manufacturing and aerospace (Grieves 2015). Over time, DTs have gained interest in more complex settings that are not as easily predictable: e.g., manufacturing, healthcare, and smart cities.

DTs stand apart from static models by continually integrating data from their physical counterparts and surrounding environments through IoT, sensors, AI, and predictive analytics (Fuller et al. 2020). Constantly synchronized with its physical twin through bidirectional data flows and feedback loops, a DT monitors ongoing processes and anticipates future trends, including potential damages and failures. The continual update cycle—often referred to as the ‘twinning rate’—involves measuring the real-time state of the physical entity and replicating those parameters in the virtual environment, and vice versa, enabling the virtual environment to inform and change the physical environment so that both states remain as close to ‘equal’ as possible (Jones et al. 2020, 42–43). This bidirectional coupling differentiates DTs from preexisting digital models (Thelen et al. 2022; Argota Sánchez-Vaquerizo 2024). Ultimately, through these capabilities, DTs enable scenario testing, inform decision-making, and support proactive interventions to enhance the real-world system they represent (Barricelli, Casiraghi, and Fogli 2019, 167656).

The success of DTs relies on serving a purpose, being trustworthy, and functioning effectively. Therefore, the first step is to define the purpose of the DT, which can range from real-time monitoring and predictive maintenance to more exploratory what-if analyses. Next, robust data infrastructure planning and collection are essential for maintaining reliable, real-time information flow. In this phase, practitioners identify and gather data sources—such as historical data, real-time sensor readings, or external datasets—depending on the physical or social entity (Jones et al. 2020).

Social Digital Twins require altogether the simulation of the states of the individuals, objects, and systems considered, and also the activities, interactions, and mechanisms that shape the environment.⁴ Therefore, modeling a complex social system—such as a deliberative democracy—often requires data drawn from social media, online forums, administrative records (e.g., demographic data), and qualitative surveys or interviews (Franco-Guillén, Laile, and Parkinson 2022).

⁴ Fujitsu has developed a Policy Twin, a new digital twin technology to maximize effectiveness of local government policies for solving societal issues. <https://www.fujitsu.com/global/about/resources/news/press-releases/2024/1126-01.html>

Because social data is typically heterogeneous and may include text, images, and geospatial information, it often requires preprocessing with techniques like natural language processing (for text) or image recognition (for visuals). All this data can be used to profile behaviorally and psychologically simulated agents. This complexity reflects humans' inherently multi-layered nature and interactions, creating a multi-level complex system (Helbing and Argota Sánchez-Vaquerizo 2023, 83). After cleansing and normalizing these varied inputs, the data undergoes Extract, Load, Transform (ELT) processes to prepare it for storage and analysis. This final step ensures that the DT's software can interpret and use the data effectively, continuously refining the model and delivering actionable insights into the real-world system it mirrors through interface, human-mediated action, or even actuators (Jones et al. 2020).

Once the primary objective is defined and the relevant data is collected, the next step is to develop the core model. The specific configuration of this model depends on the nature of the replicated entity, including its components, interactions, and constraints. This explains the wide range of modeling techniques enabling DTs. Geometric modeling (e.g., CAD, point clouds) is key for representation and spatially supporting interaction layers. Behavior and dynamics can be modeled in a continuum of approaches from purely physics-based to purely data-driven. Physics-based modeling derives from first principles and classical mechanics, typically governed by equations and relying on high-fidelity methods (e.g., Finite Element Modeling, FEM, Computational Flow Dynamics, CFM) (Hinchy et al. 2020; Thelen et al. 2022). Equally important are the social and behavioral sciences, as human interactions and decision-making processes can critically shape system outcomes. Computational social science combines these different modeling approaches, bridging social processes with physics-based and data-driven frameworks. Complexity science provides theoretical tools for capturing emergent phenomena, feedback loops, and multi-scale interactions.

Data-driven approaches use Machine Learning (ML), Deep Learning (DL), and statistical methods to map unknown relations between inputs and outputs from historical data—without encoding explicitly physics-based equations—. Physics and data-driven modeling merge in hybrid approaches that combine the advantages of both (e.g., Physics-informed ML), reduce the complexity of physics-based modeling (e.g., Reduced-Order Modeling and Surrogate Models), enhance real-time operations, and continuous adaptive modeling or deal with incomplete, noisy and changing data through probabilistic and uncertainty modeling (Thelen et al. 2022).

If we aim to twin multi-instance settings (e.g., agents, states, interactions, levels, scales), the modeling approach might focus on relations and interactions such as Agent-Based Models (ABMs), Graph-based Models, Discrete Event Simulations (DES), System Dynamics (SD), or their combination (Orozco-Romero, Arias-Portela, and Saucedo 2020; Qiu et al. 2023). In an ABM, autonomous

“agents” are defined with particular behaviors, rules, or preferences, and their micro-level interactions can lead to emergent, bottom-up, system-wide phenomena—such as the formation of coalitions (Bonabeau 2002). Graph-based modeling, by contrast, centers on the ties among instances, making it possible to encode domain-specific knowledge suitable for hybrid modeling techniques (knowledge graphs or Bayesian networks) and making semantic representations possible (Zheng, Lu, and Kiritsis 2022; Listl et al. 2024). Also, graph representation makes it possible to identify key influencers, detect subgroups, or reveal communication patterns in social structures through Social Network Analysis (SNA) (Borgatti et al. 2009) or social influence models for opinion formation (Chacoma and Zanette 2015). DES simulates discrete events over time—often under assumptions like rational choice or bounded rationality—to examine how incremental changes (e.g., a new message or vote) can influence the system’s overall trajectory (Goldsmann and Goldsmann 2015). Finally, SD can handle stocks, flows, time delays, and feedback loops at an aggregate level to analyze how collective behaviors or broader interventions, such as policy changes, may affect the entire system.

At a higher level, from a systems modeling perspective, developers can adopt a multi-model architecture or “co-simulation” framework to integrate multiple modeling approaches (Borshchev and Filippov 2004) within a single DT environment (Wang et al. 2023) for different instances. In this setup, each instance (e.g., element, level, interactions) of the virtual environment can be modeled differently and may run independently while exchanging relevant state updates, thus offering a comprehensive, unified simulation simultaneously orchestrated by an integration layer.

Once the core model is built, its performance is typically refined through an iterative calibration loop in its life cycle: outputs are compared to real-world results, and parameters are adjusted based on any discrepancies. In many cases, developers will add predictive modeling or simulation social tools (e.g., DES) to forecast future states, test various “what-if” scenarios, anticipate outcomes, and inform decision-making (Agalinos et al. 2020). These models incorporate real-life data—such as traffic flow—to calibrate and adjust simulated behaviors in real-time, as part of the twinning process, where sensor fusion, adaptive, probabilistic, and uncertainty modeling play a crucial role (Thelen et al. 2022). The overarching aim is to transform incoming data (e.g., from sensors, social media, and databases) into actionable outputs like predictions, insights, or recommended actions for decision-making or to (semi-)automate changes in the mirrored physical environment. The final step involves completing the user interface—enabling users to issue commands or make adjustments in the digital environment—including mixed-reality settings (Argota Sánchez-Vaquerizo 2024).

Throughout the DT development process, AI can play numerous roles, from data processing and model-building – e.g., using Machine Learning to map unknown relations between inputs and outputs from historical data – to

predictive tasks – e.g., time-series analysis – and scenario testing or optimization – e.g., sensitivity analysis and reinforcement learning. AI also supports continuous monitoring of system states, further enhancing the adaptability and utility of the digital twin (Rathore et al. 2021). More recently, new foundation models trained using Deep Learning (DL) in vast amounts of data open new possibilities not only for analysis and modeling. They outperform in generative tasks that can ease the retrieval of data, their interpretation, modeling, and interfacing with digital twins. However, due to their generic training, these models must be adjusted (e.g., fine-tuning, prompt-engineering, Retrieval Augmented Generation—RAG) for specific applications (Ali, Arcaini, and Arrieta 2025). Despite their limitations, they promise to leverage challenges to model and simulate complex systems, particularly involving human cognition through agentic AI.

5. Testing Deliberative Democracy Through Digital Twins

Having briefly introduced DT technology, we now examine its potential application to deliberative democracy. Specifically, we aim to determine whether DTs can more effectively test the procedural rules introduced in Section 3 than real-life or lab-based experiments. We propose to use a DT of a deliberative community as a regulatory sandbox to test procedural rules and identify those best suited to achieving the desired outcomes.

Digital twins are virtual replicas that model actual processes, allowing researchers to recreate complex environments and interactions at scale. By mirroring authentic conditions and participant behavior, these simulations enable systematic testing of procedural rules without the costs and risks inherent in real-life experiments. Given the challenges of replicating, iterating, and scaling deliberative experiments in physical or lab contexts, DT technology presents a significant new opportunity for empirical inquiry into procedural rules. In what follows, we first outline how one might develop a DT of a deliberative community (Section 5.1), then discuss how it could—and, in our view, should—be employed (Section 5.2).

5.1. How to develop a DT of a deliberative community: the case of mini-publics

A DT of a deliberative community is a computational model designed to replicate the structure, internal dynamics, and behaviors of a deliberative community, whether hypothetical or actual. Unlike a basic simulation model, a DT is a dynamic, "living" virtual counterpart continuously updated with real-world data. This real-time linkage to operational data enables the DT to test scenarios, predict

outcomes, and support decision-making processes (Boschert and Rosen 2016, 61).

Once a specific conception of deliberative democracy and associated quality metrics have been defined and the DT's objectives are set, the next step is data collection. This would require socio-demographic information (such as age, education, ethnicity, and other socio-economic indicators) and critical behavioral and interaction data. Sources range from mainstream social media platforms (e.g., Facebook Groups, LinkedIn, X, Threads, Reddit, or even Wikipedia for tracking community interactions) to specialized civic engagement tools like Decidim, Citizen Lab, Polis, or Ethelo (Shin et al. 2024). While the former often provides more data, the latter yields more structured and purpose-focused interactions. However, specialized platforms—like social media—can introduce biases by attracting users who are already civically engaged.

Additional data can be gathered from official government channels—such as parliamentary or congressional debate transcripts, voting records, press releases, and party communications. For an overview of how big data analysis might be applied to deliberative democracy experiments, see (Franco-Guillén, Laile, and Parkinson 2022).

Data for a deliberative community DT can be collected in several ways. Many civic engagement platforms provide APIs or export features for user interactions, discussion posts, and voting records. Where no official API or export option exists—typically forums and social media—web crawlers or scrapers may be used to collect and periodically update posts, comments, and participation metrics (Franco-Guillén, Laile, and Parkinson 2022, 236). This approach may raise legal concerns, necessitating compliance with relevant regulations and the terms of service of the targeted websites. Any non-digital data must be gathered manually.

Once collected, these data streams feed into the DT to keep it current. The frequency of updates should be balanced against computational costs, and calibration and validation—particularly challenging for social systems—must be done carefully.⁵ In some instances, frequent updates may be unnecessary if the DT's primary goal is to test procedural rules under controlled conditions.

A major concern in computer-assisted text analysis for the social and political sciences is ensuring the data are as clean and bias-free as possible—or at least understanding how missing or misreported data could affect the DT's

⁵ Calibration and validation are distinct processes and should not be confused. Calibration involves adjusting a model's parameters to align its outputs with observed data from the physical system in its current state. This process fine-tunes the model to represent better the system being studied. Conversely, validation assesses the model's overall accuracy and reliability by comparing its outputs to independent datasets not used during calibration. This step evaluates whether the model can predict the system's behaviour under various conditions. For example, validation might involve testing the model's performance against new operational data to ensure it accurately forecasts outcomes in different scenarios, such as changes in the deliberative community composition or variations in a voting system.

predictions. Equally important is considering the context and interpretation of the data. Mitigation strategies should be documented and implemented to address these issues (Lucas et al. 2015).

In this context, Structured Topic Modeling (STM) offers a robust method for analyzing large-scale textual data. By incorporating relevant metadata—such as age, education, platform type, or time—alongside textual content, STM allows researchers to systematically identify and track discussion topics across different subgroups or contexts (Franco-Guillén, Laile, and Parkinson 2022, 238). Similarly, argument mining techniques like Argument Structure Analysis have been developed to extract debate structures in online deliberation. These techniques use natural language processing (NLP) to analyze how discussions evolve, identify controversial issues, and pinpoint conflicting viewpoints within deliberative exchanges (Lawrence et al. 2017).

These methods offer insights into which topics resonate with specific demographic segments or user communities and how discourse shifts over time. Developers can refine calibration and validation processes by integrating these findings into the DT. For instance, if real-world data show that specific topics appear more frequently than the DT predicts, the parameters can be calibrated to align more closely with STM results. The ensuing iterative feedback loop helps reduce bias, address blind spots, and ultimately produce a more accurate representation of collective deliberation.

The modeling component of a DT is among the most complex stages. A DT system must integrate multiple modeling paradigms – as emphasized by the pluralistic modeling approach (Helbing 2010) – to effectively simulate deliberative democracy scenarios. Such plurality also expands to develop multi-instance architectures where different modules or components are modeled and coordinated through different simulation paradigms. These paradigms include simulations based on existing (historical) data and predictive models that assess how changes in variables or rules are expected to impact the outcomes of deliberation.

Agent-Based Modeling (ABM) may be beneficial. It has been studied in the broader democracy and decision-making literature (Qiu and Phang 2020) and has received some attention in deliberative democracy (Lustick and Miodownik 2000; Lee et al. 2022; Butler, Pigozzi, and Rouchier 2019). In ABM, each community member or facilitator is represented as an “agent” characterized by key sociodemographic attributes such as age, income, and education; these agents “behave” according to decision rules, preferences, psychological traits, and engagement levels (including the propensity to vote). ABM further models interactions among agents, capturing processes like opinion diffusion, coalition formation, and resource allocation. Its ability to represent heterogeneity among agents and simulate emergent phenomena makes ABM a powerful tool for studying the dynamics of deliberation.

ABM capabilities are mainly boosted by embedding complex behavior into its agents to perceive, process, and react more effectively and realistically in the environment. Recently, the development of Large Language Models (LLMs) opened new possibilities for improving knowledge retrieval, adaptability, and reasoning of agentic AI, approximating behavior expected in humans (Park et al. 2023), particularly when combined with more reliable information access (e.g., RAG) (Zhang et al. 2024). The ABM is populated with autonomous agents that encapsulate multiple modules for mirroring different aspects of human cognition (e.g., short and long-term memory, social and action awareness, talking, etc.) and can exhibit plausible human behavior. The resulting virtual social system shows emergent features and evolution that can be related to human societies (Altera et al. 2024), including task-solving through collaboration (Qian et al. 2025).

A different approach to modeling deliberative processes focuses on aggregate relationships and dynamic interactions. System Dynamics (SD) is well-suited for exploring macro-level trends, as it models complex systems through feedback loops and time-dependent behaviors (Bala, Arshad, and Noh 2017).⁶ At the same macro-level, Social Network Analysis (SNA) maps social relationships and communication flows, making it particularly useful in deliberative democracy contexts for identifying influential community members or subgroups that shape debates (da Silva, Ribeiro, and Higgins 2022).

Complementarily to traditional social dynamics studying opinion formation, recent developments with LLMs have also been used to study the emergent properties of social networks regarding information propagation, opinion dynamics, and echo chambers (Zheng and Tang 2025). LLMs have facilitated exploring how processes, such as voting rules, can be improved for participatory budgeting. Some of these approaches allow for multiple, complementary solutions rather than a one-size-fits-all solution, which can increase satisfaction among participants. It may also be beneficial to replace classical deliberation formats with new forms of co-creation, such as “re-mixing” (Carpentras, Hänggli Fricker, and Helbing 2024). LLMs have also been proposed to simulate the decision-making and interactions of diverse personas (Yang et al. 2024). Furthermore, LLMs can be used as moderators, support the co-creation of integrated ideas, or illustrate the consequences of certain decisions.

Another technique, Discrete-Event Simulation (DES), simulates how specific events within a deliberation process—such as the emergence of consensus or voting phases—impact on community outcomes (Charalabidis 2011). While DES is unconventional in this field (typically used in logistics and service operations), its application to deliberative processes is theoretically feasible and could provide novel insights.

⁶ SD typically abstracts away micro-level details—individual agents, local network structures, etc.—so when applying it to deliberation or opinion dynamics, modelers must be comfortable with the loss of fine-grained resolution.

Regardless of the chosen modeling approach—or a combination of techniques—the model must be dynamic, updating in response to agent interactions and external influences (e.g., using RAG). For example, it should predict how often and to what degree agents change their opinions based on peer interactions, exposure to news, or official statements. Machine Learning (ML) or other AI techniques can be integrated to forecast sentiment or voter turnout changes based on historical patterns.⁷ These predictions can then be fed back into the model (e.g., an ABM model) to refine its accuracy and alignment with human behavior and psychology.

Once developed, domain experts, community leaders, or academic researchers must calibrate and validate the model. This step is critical for establishing the model's credibility. Iterative refinement may be necessary to incorporate new data, address feedback loops, and improve the model's precision over time. The “black box” character of state-of-the-art AI based on DL challenges efforts for explainability. However, combining these models with cognitive architectures designed to align with human cognitive processes can help to mitigate these issues (Bickley and Torgler 2023).

5.1.1. Twinning mini-publics

To illustrate how this approach could work in practice, consider developing a digital twin (DT) of a mini-public—a small group of city residents convened to deliberate on a policy issue to produce recommendations, such as how to allocate funding for environmental initiatives. Real mini-publics of this nature exist in practice; for example, the UK Climate Assembly in 2020 (Willis, Curato, and Smith 2022) is a well-known case. These deliberative forums are often chosen for experimental methodologies (Grönlund and Herne 2022) due to their manageable size and the potential for structured evaluation.

Ensuring the demographic representativeness of the mini-public is complex (Germann 2025). DT designers can gather institutional data (for population attributes) and engagement platform data (for behaviors and interactions) to perform what-if testing of different compositions to see if they yield outcomes like those you would expect from the broader population. When building an ABM or other simulation techniques, each agent is randomly assigned demographic attributes that align with official city-wide distributions (if direct participant-level data are unavailable). Engagement-platform metrics (e.g., number of posts, comments, votes, timestamps, and discussion content) help distinguish between (the spectrum of) more active and passive agents. Sentiment analysis or topic modeling can reveal whether participants favor or oppose particular policies.

⁷ It is not strictly necessary that “all” what-if simulations must use ML to predict changes—one could also use rule-based approaches or simpler parameter sweeps.

Based on these data, agent profiles can be constructed to represent participants' behaviors and general orientation. Crucially, the simulation does not need to replicate a specific community exactly, as the goal is not twinning or replacing human constituents; it can be a hypothetical or randomized (i.e., synthetic) group whose characteristics are derived from real-world distributions. Also, this facilitates sensitivity studies about potential risks and biases caused by the selection process of participants in mini-publics in real life. Finally, following data protection regulations and ethical standards (Bäumer et al. 2024), personal data must be anonymized or pseudonymized so individuals cannot be re-identified, while retaining the demographic and behavioral variations necessary for an accurate simulation.

Once the data have been cleaned, they are integrated into the mini-public's DT. The core model may use ABM, in which each agent (representing a participant) has specific attributes such as age, income, and education. Decision rules that govern agent behavior must be defined to mirror real-world participant dynamics, such as the probability of posting versus lurking or the likelihood of shifting positions when presented with well-argued evidence.

The ABM also requires the design of interaction rules based on observed behaviors from digital platforms, such as posting and commenting patterns. The model can simulate how agents communicate in discussion threads, form temporary coalitions, or change their opinions. For instance, if an agent with moderate environmental concern encounters a well-supported argument from a highly respected participant or facilitator, their level of concern may increase slightly. Similarly, if real-world data show that participants from certain demographic groups tend to cluster or share similar views, the ABM can replicate this behavior by increasing the likelihood of aligning with similar agents.

The same cleaned data—particularly those about participation trends—can also inform a System Dynamics (SD) model. This is not trivial, as in SD, individual-level data become stocks and flows that capture system-wide dynamics. This approach is helpful in modeling higher-level feedback loops, such as how increased dissatisfaction can reduce future participation or how extended debate might simultaneously promote consensus and lead to participant fatigue.⁸

Real-world data on (social) interactions—such as mentions, replies, or co-annotations—can be used to construct adjacency matrices for Social Network Analysis (SNA). These matrices represent the structure of a network, where nodes represent individuals and edges represent interactions. Once the network structure is established, it can be analyzed to measure centrality (identifying key influencers), detect subgroups (e.g., communities or clusters), and track idea propagation (e.g., how information spreads). SNA can help identify potential

⁸ In practice, going from raw, cleaned data to a fully calibrated ABM or SD model often requires more intermediate steps: e.g., sensitivity analysis to test how changes in parameters affect outcomes (Ligmann-Zielinska et al. 2014).

influencers within a group and reveal how ideas propagate between subgroups, highlighting phenomena like bridging connections (where subgroups interact) or echo chambers (where subgroups become isolated). Also, from an aggregated point of view, SNA can be used in combination with other NLP techniques for content, emotion, and sentiment analysis, both at the level of specific agents or the general environment, for which specific, independent LLMs could be used (Krugmann and Hartmann 2024).

DES can be employed to model how specific events alter the state of the mini-public. For example, the timing of real-world actions—such as when participants post, vote, or change their stances—can be used to define events and event triggers in the simulation. Observed frequencies from actual data (e.g., “a major influencer emerges in 1 out of 3 debates”) can be translated into probabilities within the DES. This allows the simulation to replicate realistic patterns of behavior and interaction. Predictions about future community states—such as who is likely to drop out, who will submit or rebut arguments, or how likely a proposal will succeed—often require Machine Learning, Natural Language Processing, and sentiment analysis. Sentiment analysis, in particular, can quantify changes in support, neutrality, or opposition (Liu 2022).⁹

If the dialogue in the mini-public has been constructive can be measured by the depth of argumentation (including evidence or references) and balance in speaking time (Steenbergen et al. 2003; Bächtiger, Gerber, and Fournier-Tombs 2022; Scudder 2022). Accordingly, agent behavior in the DT simulation might reflect the likelihood of participants bringing new evidence, responding to opposing views, or remaining silent. Meanwhile, SNA indicators can gauge how ideas spread and identify dominant groups or influencers, assessing whether this dominance affects argumentation depth or pluralism (Musso and Helbing 2024) positively or negatively. The constructiveness of a debate may be assessed through sentiment analysis.

Finally, the DT must be calibrated to ensure the agents’ collective behavior matches real-world patterns. One way to calibrate it is to simulate a known past event (e.g., last year’s participatory budgeting process) and compare the predicted outcomes (such as final proposal support rates) to the actual historical results. This calibration step helps validate the model and improve its predictive accuracy.

In short, following state-of-the-art societal and behavioral simulation, one of the most granular possibilities for generating a DT for a mini-public can result in a multi-stage pipeline integrating structured socio-demographic data, behavioral modeling, and LLM-based personas for the AI agents. These autonomous agents

⁹ Because this is still a simulation (even in “digital twin” form), it must be calibrated to ensure the agents’ collective behavior closely matches real-world patterns. One way to achieve this is to simulate a known past event (e.g., last year’s participatory budgeting process) and compare the predicted outcomes (such as final proposal support rates) to the actual historical results. This calibration step helps validate the model and improve its predictive accuracy.

can populate a dynamic environment in which they will interact and exchange information with other agents and with a changing context (e.g., policies, protocols, news), as shown in Figure 1. These changing conditions will also inform their personas by updating their memory and opinion based on actions and answers. Aggregated metrics in the emerging network can be extracted and used to calibrate and evaluate the performance of the DT historical data. Nonetheless, in some settings or for testing, some hypotheses may not be necessary to deploy a fully multi-agent and multi-model DT, and a different modeling approach may be sufficient for some scenarios.

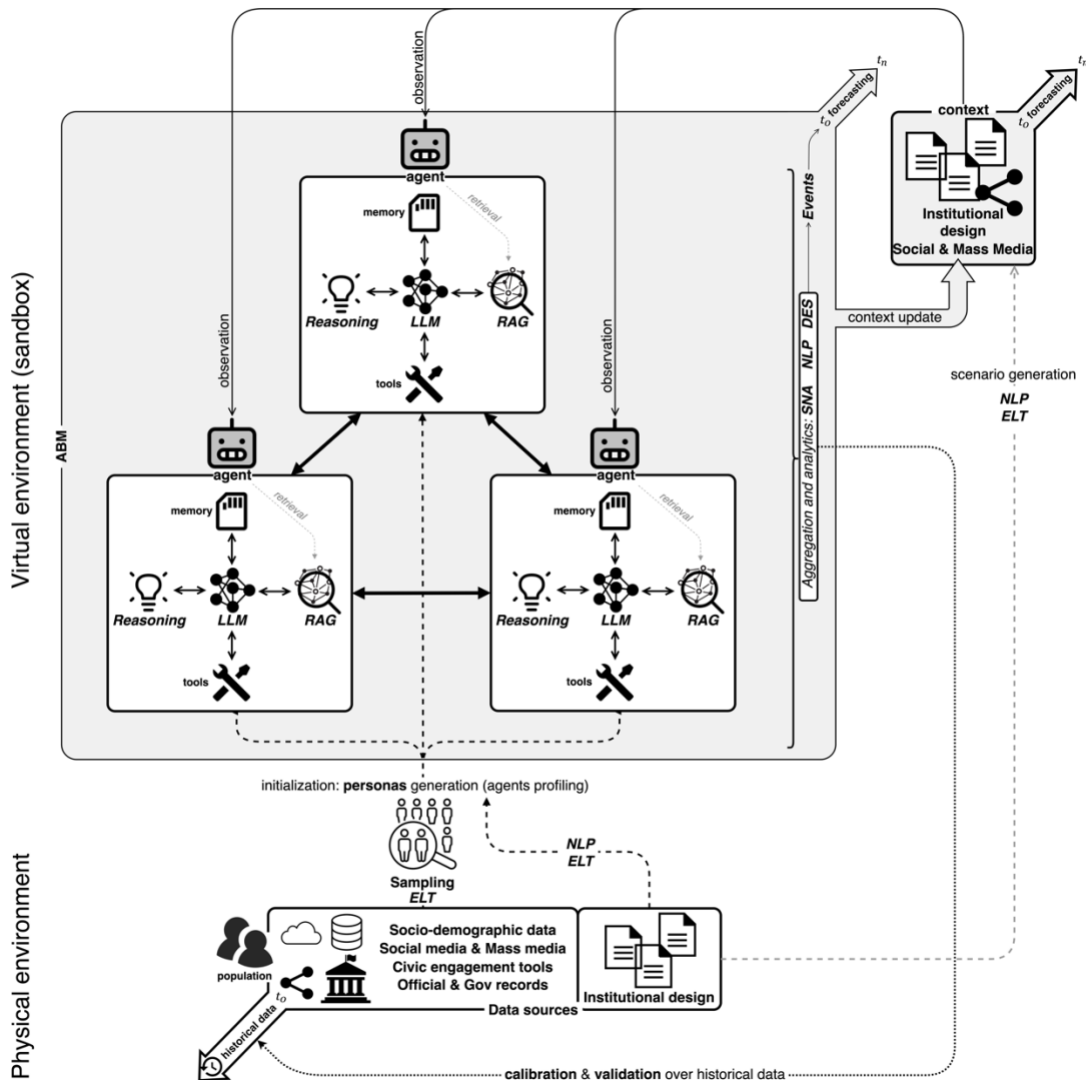


Figure 1. Proposed Multi-agent-based model (ABM) architecture for a democracy deliberation DT, based on AI-powered agents profiled as different personas from actual socio-demographic, behavioral, and public records data, which is pre-processed through extract, load, and transform (ELT) processes and using Natural Language

Processing (NLP) techniques. Each agent is equipped with multiple modules able to perform different tasks (reasoning, information retrieval, memorizing, action planning) that allow them to communicate and interact with other agents and the context (e.g., institutional design and news), and therefore, obtaining feedback to adjust to the changing environment. The virtual environment exhibits emerging social network features that can then be related to social network analytics (SNA) and analyzed using NLP. The aggregation of these social dynamics can be used for higher-scale modeling, such as Discrete-Event Simulation (DES), to calibrate and evaluate the performance of the DT based on actual historical data (as layers of data sources) and to update the context over time, forecasting new scenarios.

5.2. Using the DT to test (and refine) deliberative democracy rules

Our objective is not to replace real deliberative communities with their digital replicas. Substituting human agents with virtual approximations would lead to a computationally deterministic scenario fraught with oversimplifications. The goal is instead to use the DT as a sandbox for experimenting with procedural rules. We can examine how institutional designs unfold and assess whether they achieve their stated outcomes.

DT may help deliberative democracy designers to conduct experiments that would be impractical or unethical to carry out in the real world (Helbing and Sánchez-Vaquerizo 2023, 65), especially given the constraints discussed in Section 3.1. To illustrate this, we turn to the three main sets of procedural rules outlined in Section 3.

- a) *Pre-deliberation rules.* A foundational aspect of pre-deliberation rules is determining who participates. Here, a DT can be employed to evaluate the impact of different recruitment strategies on inclusivity, diversity, and the overall quality of deliberation. Through ABM—simulating virtual participants with varied demographics, interests, and needs—designers can explore how different recruitment strategies (e.g., random vs. weighted sampling or voluntary participation) influence group composition and discussion dynamics. Simulations may reveal whether purely random selection yields a diverse group or whether specific demographics are likelier to drop out due to low interest or external constraints, thereby reducing actual diversity in participation. If the goal is to maximize diversity, results might indicate that simple random selection is inadequate because of high opt-out rates or systematic underrepresentation of minorities, thus favoring weighted or stratified sampling (Rowe and Frewer 2000; Smith 2009). Combining representative sampling with minority-focused panels can address challenges related to small sample sizes. Conversely, if the goal is to increase public buy-in, simulations can be used to evaluate voluntary participation. They might show that volunteers are more motivated to become informed and engage in reasoned discussion; however, they could also reveal that this approach

disproportionately attracts individuals with strong pre-existing opinions, indicating a need for outreach campaigns to favor neutrality and listening.

Through the same simulation approaches, we can compare stakeholder-targeted recruitment (i.e., inviting specific groups such as activists or business leaders) with more open or random selection to see whether it promotes balanced participation or leads to dominance of specific interests or increased polarization. DES can further reveal how financial incentives or mandatory participation policies influence retention rates.

Interestingly, the quality of deliberation can serve as a proxy for evaluating how well different groups are represented. For instance, measuring “argument diversity” — using metrics such as entropy indices or the number of distinct argument types (Palau and Moens 2009; Grimmer and Stewart 2013) — can help determine whether including minority voices meaningfully shapes discussion content or if they remain overshadowed in practice.

Another key set of pre-deliberation rules involves informational preparation on the topics to be discussed. ABM and SNA can be used to understand how participants absorb information and how it spreads within subgroups. Specifically, SNA could construct interaction graphs to track information flow and identify influential participants. The objective should not be to suppress influential voices — like in totalitarian regimes or non-pluralistic platforms — but rather to reduce influence concentration. Promoting a more balanced influence distribution can help prevent uncritical epistemic deference and encourage more independent, reasoned participation.

Likewise, DES can help determine how long the preparatory phase should be: e.g., simulations might show that expert presentations increase overall knowledge but risk creating excessive deference to authority (although deference may be tricky to quantify in the simulation), whereas encouraging participants to do background readings may foster deeper, self-driven learning. The optimal approach will depend on the primary goal—whether it is factual accuracy, broad participation, or some balance of the two—and might include evidence-oriented measures to mitigate the risk of undue deference.

Representativeness is often an ideal rather than a fully attainable standard. Designers must balance the desire for diverse participation with real-world constraints like budget, recruitment time, and participant availability. Here, a DT can simulate how these constraints shape different selection rules. DES or system dynamics can project how much recruitment time or funding each rule requires, helping designers refine their choices to maximize representativeness without exceeding practical limits.

Finally, agendas could also be mapped using computational methods. Sentiment analysis, for example, can gauge how different agendas might influence participant engagement or emotional reactions by analyzing data from simulated interactions, while ABM can reveal whether a collaboratively developed agenda generates greater ownership despite taking longer to finalize.

- b) *Discussion rules.* The design of discussion formats can be enhanced with DT simulations to test various session structures. For example, simulations can help determine whether deeper argumentation is best achieved in a single session or multiple rounds (e.g., iterating homogeneous and heterogeneous panels/sessions), examining how argument complexity evolves. Argument complexity can be measured using metrics such as layers of reasoning (e.g., the number of premises and rebuttals in an argument) or the degree of justification (e.g., the extent to which arguments are backed by evidence and acknowledged counterpoints) (Walton, Reed, and Macagno 2008; Prakken and Vreeswijk 2002; Lippi and Torroni 2016).

Simulations might reveal that single-session formats work better if the goal is efficient deliberation. However, if the goal is to reduce polarization or bias, simulations may suggest that iterative deliberation—with repeated exposure to opposing views—is more effective. That said, recent findings with ABM for deliberative democracy (Lee et al. 2022) show that while multiple rounds of discussion are more impactful than simply providing information, their incremental utility decreases over time, with limited effects on reducing polarization.

Simulations can also test how participants present opinions and rebut arguments and how these modalities affect deliberation outcomes such as fairness or engagement. For example, DES can model how different time constraints impact argument complexity or completeness. Additionally, ABM combined with NLP can analyze whether rebuttals deepen discussions or lead to repetition. SD can further simulate how rebuttal styles (e.g., iterative step-by-step refinements vs. one-time paired rebuttals) influence opinion shifts and consensus-building.

Moreover, DT simulations can evaluate which facilitation/moderation styles best promote argument diversity and fairness and whether frequent moderator interventions enhance or hinder the natural flow of discourse. SNA can help detect power imbalances, such as a facilitator who repeatedly calls on the same participants.

A critical aspect of discussion rules concerns how final decisions are reached. Decision-making strategies can be tested via ABM, possibly supported by game-theoretic models that treat participants as rational or boundedly rational agents employing strategic behavior rather than simply

sharing sincere opinions. These models can show how power imbalances and knowledge asymmetries shape coalition success (Parsons, Gymtrasiewicz, and Wooldridge 2012). More specifically, we may simulate whether participants adopt cooperative or aggressive strategies when debating contentious issues (the so-called hawk-dove game) and evaluate whether conflict resolution rules (structured mediation, facilitation styles) encourage cooperation (Amadae and Watts 2023).

Through these simulation techniques, designers can draw insights about the trade-offs of different decision-making rules. So, DT simulations may reveal that consensus-building, while fostering broader participation and depth of dialogue, can also slow down decision-making. Conversely, majority voting may speed up processes but risks systematically disadvantaging certain groups—perhaps due to unequal cooperation or strategic abilities—resulting in disengagement and reduced overall effectiveness of deliberation.

- c) *Post-deliberation rules.* The deliberative process also includes rules for monitoring and implementing decisions after deliberation, though these are often less emphasized than pre-deliberation or in-process rules. Unlike the other two sets of procedural rules, which can be tested using simulated data from online interactions, post-deliberation rules mostly require real-world data sources. This focus on real-world data is necessary because post-deliberation interactions are highly context-specific – lacking reliable proxies in other datasets – and challenging to replicate in simulations. Accordingly, the emphasis shifts to data analytics more than simulation. These data include surveys, voting records, and historical records of participatory policymaking. Such data help researchers understand who engages in deliberation and how participants react to post-deliberation processes.

In this context, data analytics tools and predictive modeling techniques can help determine the most feasible documentation and reporting mechanisms. For example, these tools can help evaluate whether open-access reports increase public trust or lead to selective misinterpretation compared to restricted reports. Additionally, they can analyze historical feedback data to assess how feedback collection methods—such as anonymous vs. named feedback, immediate vs. delayed feedback, or facilitated group reflections—impact participant engagement and depth of reflection.

Sentiment Analysis and Natural Language Processing can provide insights into participant engagement, though qualitative analysis should supplement their interpretations. These tools can reveal how participants feel about the deliberation process and what they focus on when giving feedback or discussing outcomes. However, qualitative analysis should

supplement their interpretations. So, if people who saw the open-access reports consistently express more positive sentiment (e.g., relief, trust, willingness to engage) than those who saw the restricted reports, it suggests open access may foster greater public confidence. Conversely, if negative sentiments (e.g., anger, skepticism) dominate when reports are fully open, it might indicate issues like data overload or misunderstandings.

This section shows how a DT can help institutional designers test and compare different procedural rules in deliberative democracy, ultimately guiding them toward rules that better support their desired outcomes. Before we move on to the potential advantages and limitations of this “sandboxing” strategy, it is crucial to note that although we used an illustrative example of a comprehensive DT replicating an entire deliberative community and its institutional design, in practice, this may not be the most feasible approach. A DT may be used to test only specific aspects that designers wish to prioritize. For instance, if the recruitment strategy is already well-defined, designers might use the DT solely to test how final decisions should be made rather than modeling every phase of the process.

6. Limitations and future research

Digital twins offer a promising avenue for empirical research into procedural rules, particularly given the challenges of replicating, iterating, and scaling deliberative experiments in physical or lab settings. However, their effectiveness depends on the accuracy of behavioral assumptions, the quality of input data, and their capacity to generalize to real-world democratic practice. For instance, oversimplified assumptions about specific agent behaviors and decision-making rules (e.g., how and when participants disengage) can introduce biases and lead to misleading conclusions. Moreover, simulating future outcomes may be challenging because of the inherent complexity involved in human-based and social systems. For example, the complex dynamics observed in networked systems are always challenging to model, fit, and validate (Caldarelli et al. 2023).

These limitations highlight trade-offs and challenges of creating a digital twin of a city or society, including its people, to explore and assess different democratic deliberation and voting formats. A complete digital replica of a society may be infeasible, calling to scale down and develop issue-specific digital twins. Also, AI-powered agents aiming at rendering people’s behavior may exhibit human and non-human psychological biases (Rossetti et al. 2024). Despite the potentially limited alignment with real human psychology and behavior, simulation experiments can offer controlled experimentation and reproducibility of social interactions, and they can foster the use of adaptive experimental design to deal with vast spaces of possible variables and interventions (Offer-Westort,

Coppock, and Green 2021). By exploring the differences between agentic AI and humans engaging in the same issues and challenges, it would be possible to assess and improve the ecological validity of such digital twins.

Digital twins have been in the making for some time. In principle, these can support various data-driven, AI-managed, cybernetic forms of society, ranging from digital versions of feudalism, communism, and fascism to digital democracies and other digital societies.

Similarly, as already observed in DTs for cities (Batty 2021), they exist on a range between strict mirroring of the physical environment, which limits speculative exploration, and decoupled simulation, which enables testing of alternative scenarios. This means that for democratic deliberation, DTs can act as predictive models running before any phase in the deliberation or real-time interactive systems that assist and evolve alongside the physical, human-driven processes. This also enables such frameworks to support different types of agencies, processes, and involvement levels from humans, including hybrid, collective, and symbiotic intelligence of humans and AI (e.g., Human-in-the-Loop).

Key steps for future research to ensure that the use of digital twins aligns with the values of (liberal and deliberative) democracies are the following:

- *Avoid surveillance.* Effective measures should be taken to prevent personal surveillance, scoring, and targeting. This means that one should not strive to create and use as-identical-as-possible digital twins of people but “noisy” digital twins, i.e., to work with samples of hypothetical personas having statistically representative or theoretically assumed characteristics. This should also serve the goal of privacy protection.
- *Ensure transparency and accountability.* Ensuring transparency in the data, software, and procedures is crucial to getting people's trust, public support, and encouraging participation. Otherwise, there could be hidden manipulation. Transparency is also essential for ensuring accountability in all aspects of digital deliberation—whether in the design and implementation of deliberative practices or in addressing concerns and making necessary corrections.
- *Open, plural, and fair discussion.* The deliberation platform must offer a level playing field such that the competition of ideas and the consideration of justified individual interests is fair. It should promote the mutual understanding of different perspectives. In doing so, it should support a respectful exchange of ideas and a constructive dialogue, which allows people to voice diverse opinions, interests, and needs without fear.
- *Avoid mis- and disinformation.* It should promote accurate, verifiable information rather than mis- or disinformation. It should support an

evidence-based approach emphasizing critical evaluation, hypothesis verification, and falsification rather than relying solely on raw data, which can be misinterpreted or manipulated without proper contextual analysis.

- *Human-centered approach.* It should have a human-centered approach that serves the interests and needs of people (rather than merely optimizing efficiency, automation, or institutional control). Given that AI often functions as an opaque “black box” and social systems are inherently complex, a human-centered approach must integrate cognitive architectures that mirror human thought processes (instead of relying solely on data-driven methods).
- *Respect multi-dimensional values.* It should consider legal, ethical, societal, and cultural values and qualities people care about, such as human dignity, friendship, love, trust, creativity, beauty, etc.
- *Enhance human agency.* It should support human agency and offer participatory opportunities that allow people to co-create solution ideas.

In short, digital twin technology should serve as an open and inclusive platform rather than for “control room” approaches that aim at socially engineering peoples’ settings, environments, and behaviors. This way, multiple stakeholders can openly exchange perspectives, collaboratively address conflicts, and simulate solutions in a virtual space before implementing concrete actions in the physical world—aligning with the ‘peace room’ approach (Helbing and Seele 2017).

DT technology can ultimately help find alternatives to the traditional models of deliberative democracy that we have discussed in this paper, i.e., based on structured debates and formal decision-making processes. These conventional methods often follow linear, rule-based, and argumentative structures. In contrast, DTs and AI may enable more dynamic, flexible approaches, such as re-mixing. In fact, instead of locking participants into strict plans and fixed positions, re-mixing is an iterative decision-making process that allows people to submit modular elements of their views and combine and modify them also based on real-time feedback. In short, ideas evolve based on feedback and experimentation, avoiding premature commitment to a single plan. DTs offer an environment where re-mixing is simulated, adjusting the decision collaboratively before real-world implementation.

Bibliography

Agalianos, K., S. T. Ponis, E. Aretoulaki, G. Plakas, and O. Efthymiou. 2020. ‘Discrete

Event Simulation and Digital Twins: Review and Challenges for Logistics'. *Procedia Manufacturing*, 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021), 51 (January):1636–41. <https://doi.org/10.1016/j.promfg.2020.10.228>.

Ali, Shaukat, Paolo Arcaini, and Aitor Arrieta. 2025. 'Foundation Models for the Digital Twins Creation of Cyber-Physical Systems'. In *Leveraging Applications of Formal Methods, Verification and Validation. Application Areas*, edited by Tiziana Margaria and Bernhard Steffen, 9–26. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-75390-9_2.

Altera, AL, Andrew Ahn, Nic Becker, Stephanie Carroll, Nico Christie, Manuel Cortes, Arda Demirci, et al. 2024. 'Project Sid: Many-Agent Simulations toward AI Civilization'. arXiv. <https://doi.org/10.48550/arXiv.2411.00114>.

Amadae, S. M., and Christopher J. Watts. 2023. 'Red Queen and Red King Effects in Cultural Agent-Based Modeling: Hawk Dove Binary and Systemic Discrimination'. *The Journal of Mathematical Sociology* 47 (4): 283–310. <https://doi.org/10.1080/0022250X.2021.2012668>.

Argota Sánchez-Vaquerizo, Javier. 2024. 'Urban Digital Twins and Metaverses towards City Multiplicities: Uniting or Dividing Urban Experiences?' *Ethics and Information Technology* 27 (1): 4. <https://doi.org/10.1007/s10676-024-09812-3>.

Bächtiger, Andre, John S. Dryzek, Jane Mansbridge, and Mark Warren. 2018. 'Deliberative Democracy: An Introduction'. In *The Oxford Handbook of Deliberative Democracy*, edited by Andre Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark Warren, 0. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198747369.013.50>.

Bächtiger, André, John S. Dryzek, Jane Mansbridge, and Mark E. Warren. 2018. *The Oxford Handbook of Deliberative Democracy*. Oxford University Press.

Bächtiger, André, Marlène Gerber, and Eléonore Fournier-Tombs. 2022. 'Discourse Quality Index'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press. <https://doi.org/10.1093/oso/9780192848925.003.0006>.

Bächtiger, André, Simon Niemeyer, Michael Neblo, Marco R. Steenbergen, and Jürg Steiner. 2010. 'Disentangling Diversity in Deliberative Democracy: Competing Theories, Their Blind Spots and Complementarities'. *Journal of Political Philosophy* 18 (1): 32–63. <https://doi.org/10.1111/j.1467-9760.2009.00342.x>.

Bächtiger, André, and John Parkinson. 2019. *Mapping and Measuring Deliberation: Towards a New Deliberative Quality*. Oxford University Press.

Bala, Bilash Kanti, Fatimah Mohamed Arshad, and Kusairi Mohd Noh. 2017. *System Dynamics*. Springer Texts in Business and Economics. Singapore: Springer. <https://doi.org/10.1007/978-981-10-2045-2>.

Barricelli, Barbara Rita, Elena Casiraghi, and Daniela Fogli. 2019. 'A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications'.

IEEE Access 7:167653–71. <https://doi.org/10.1109/ACCESS.2019.2953499>.

Batty, Michael. 2021. 'Multiple Models'. *Environment and Planning B: Urban Analytics and City Science* 48 (8): 2129–32. <https://doi.org/10.1177/23998083211051139>.

Bäumer, Frederik Simon, Sergej Schultenkämper, Michaela Geierhos, and Yeong Su Lee. 2024. 'Mirroring Privacy Risks with Digital Twins: When Pieces of Personal Data Suddenly Fit Together'. *SN Computer Science* 5 (8): 1109. <https://doi.org/10.1007/s42979-024-03413-z>.

Bex, Floris, Henry Prakken, Chris Reed, and Douglas Walton. 2003. 'Towards a Formal Account of Reasoning about Evidence: Argumentation Schemes and Generalisations'. *Artificial Intelligence and Law* 11 (2): 125–65. <https://doi.org/10.1023/B:ARTI.0000046007.11806.9a>.

Bickley, Steve J., and Benno Torgler. 2023. 'Cognitive Architectures for Artificial Intelligence Ethics'. *AI & SOCIETY* 38 (2): 501–19. <https://doi.org/10.1007/s00146-022-01452-9>.

Bonabeau, Eric. 2002. 'Agent-Based Modeling: Methods and Techniques for Simulating Human Systems'. *Proceedings of the National Academy of Sciences* 99 (suppl_3): 7280–87. <https://doi.org/10.1073/pnas.082080899>.

Borgatti, Stephen P., Ajay Mehra, Daniel J. Brass, and Giuseppe Labianca. 2009. 'Network Analysis in the Social Sciences'. *Science* 323 (5916): 892–95. <https://doi.org/10.1126/science.1165821>.

Borshchev, Andrei, and A Filippov. 2004. *From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools*.

Boschert, Stefan, and Roland Rosen. 2016. 'Digital Twin—The Simulation Aspect'. In *Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and Their Designers*, edited by Peter Hehenberger and David Bradley, 59–74. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-32156-1_5.

Butler, George, Gabriella Pigozzi, and Juliette Rouchier. 2019. 'Mixing Dyadic and Deliberative Opinion Dynamics in an Agent-Based Model of Group Decision-Making'. *Complexity* 2019 (1): 3758159. <https://doi.org/10.1155/2019/3758159>.

Caldarelli, G., E. Arcaute, M. Barthelemy, M. Batty, C. Gershenson, D. Helbing, S. Mancuso, et al. 2023. 'The Role of Complexity for Digital Twins of Cities'. *Nature Computational Science* 3 (5): 374–81. <https://doi.org/10.1038/s43588-023-00431-4>.

Carpentras, Dino, Regula Hänggli Fricker, and Dirk Helbing. 2024. 'Empowering Minorities and Everyone in Participatory Budgeting: An Agent-Based Modelling Perspective'. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 382 (2285): 20240090. <https://doi.org/10.1098/rsta.2024.0090>.

Carpini, Michael X. Delli, Fay Lomax Cook, and Lawrence R. Jacobs. 2004. 'PUBLIC DELIBERATION, DISCURSIVE PARTICIPATION, AND CITIZEN ENGAGEMENT: A

Review of the Empirical Literature'. *Annual Review of Political Science* 7 (Volume 7, 2004): 315–44. <https://doi.org/10.1146/annurev.polisci.7.121003.091630>.

Chacoma, Andrés, and Damián H. Zanette. 2015. 'Opinion Formation by Social Influence: From Experiments to Modeling'. *PLOS ONE* 10 (10): e0140406. <https://doi.org/10.1371/journal.pone.0140406>.

Chambers, Simone. 2018. 'The Philosophic Origins of Deliberative Ideals'. Edited by Andre Bächtiger, John S. Dryzek, Jane Mansbridge, and Mark Warren, September, 54–69. <https://doi.org/10.1093/oxfordhb/9780198747369.013.1>.

Charalabidis, Yannis. 2011. 'Enhancing Participative Policy Making through Modeling and Simulation: A State of the Art Review'. *European, Mediterranean & Middle Eastern Conference on Information Systems, Athens, Greece*, January. https://www.academia.edu/2852608/Enhancing_participative_policy_making_through_modeling_and_simulation_a_state_of_the_art_review.

Cohen, Joshua. 2007. 'Deliberative Democracy'. In *Deliberation, Participation and Democracy*, edited by Shawn W. Rosenberg, 219–36. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230591080_10.

———. 2021. 'Procedure and Substance in Deliberative Democracy'. In *Democracy and Difference: Contesting the Boundaries of the Political*, edited by Seyla Benhabib, 95–119. Princeton University Press. <https://doi.org/10.1515/9780691234168-006>.

Curato, Nicole, and Marit Böker. 2016. 'Linking Mini-Publics to the Deliberative System: A Research Agenda'. *Policy Sciences* 49 (2): 173–90. <https://doi.org/10.1007/s11077-015-9238-5>.

Dryzek, John S. 2002. *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*. Oxford University Press.

Dryzek, John S., André Bächtiger, Simone Chambers, Joshua Cohen, James N. Druckman, Andrea Felicetti, James S. Fishkin, et al. 2019. 'The Crisis of Democracy and the Science of Deliberation'. *Science*, March. <https://doi.org/10.1126/science.aaw2694>.

Escobar, Oliver. 2019. 'Facilitators: The Micropolitics of Public Participation and Deliberation'. In *Handbook of Democratic Innovation and Governance*, 178–95. Edward Elgar Publishing. <https://www.elgaronline.com/edcollchap/edcoll/9781786433855/9781786433855.00022.xml>.

Estlund, David. 1993. 'Making Truth Safe for Democracy'. In *The Idea of Democracy*, edited by David Copp, Jean Hampton, and John E. Roemer, 71–100. Cambridge University Press. <https://philarchive.org/rec/ESTMTS>.

Fishkin, James. 2009. *When the People Speak: Deliberative Democracy and Public Consultation*. OUP Oxford.

Fishkin, James S. 2013. 'When Citizens Decide: Lessons from Citizens' Assemblies on Electoral Reform; Designing Deliberative Democracy: The British Columbia Citizens Assembly'. *Perspectives on Politics* 11 (2): 670–72. <https://doi.org/10.1017/S1537592713000674>.

Fishkin, James S, and Robert C Luskin. 2005. 'Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion'. *Acta Politica* 40 (3): 284–98. <https://doi.org/10.1057/palgrave.ap.5500121>.

Fishkin, Professor James S. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. First Edition. New Haven: Yale University Press.

Forst, Rainer. 2001. 'The Rule of Reasons. Three Models of Deliberative Democracy'. *Ratio Juris* 14 (4): 345–78. <https://doi.org/10.1111/1467-9337.00186>.

Fournier, Patrick. 2011. *When Citizens Decide: Lessons from Citizen Assemblies on Electoral Reform*. Oxford University Press.

Franco-Guillén, Núria, Sebastian De Laile, and John Parkinson. 2022. 'Big Data Analysis'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press. <https://doi.org/10.1093/oso/9780192848925.003.0016>.

Friedman, Will. 2006. 'Deliberative Democracy and the Problem of Scope'. *Journal of Deliberative Democracy* 2 (1). <https://doi.org/10.16997/jdd.42>.

Fuller, Aidan, Zhong Fan, Charles Day, and Chris Barlow. 2020. 'Digital Twin: Enabling Technologies, Challenges and Open Research'. *IEEE Access* 8:108952–71. <https://doi.org/10.1109/ACCESS.2020.2998358>.

Fung, Archon. 2007. 'Minipublics: Deliberative Designs and Their Consequences'. In *Deliberation, Participation and Democracy: Can the People Govern?*, edited by Shawn W. Rosenberg, 159–83. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230591080_8.

Ganuza, Ernesto, and Francisco Francés. 2012. 'The Deliberative Turn in Participation: The Problem of Inclusion and Deliberative Opportunities in Participatory Budgeting'. *European Political Science Review* 4 (2): 283–302. <https://doi.org/10.1017/S1755773911000270>.

Gastil, John. 2000. *By Popular Demand: Revitalizing Representative Democracy Through Deliberative Elections*. First Edition. Berkeley, CA: University of California Press.

Gastil, John, and Robert Richards. 2013. 'Making Direct Democracy Deliberative through Random Assemblies'. *Politics & Society* 41 (2): 253–81. <https://doi.org/10.1177/0032329213483109>.

Gelernter, David. 1991. *Mirror Worlds or the Day Software Puts the Universe in a Shoebox: How Will It Happen and What It Will Mean*. USA: Oxford University Press, Inc.

Germann, Micha. 2025. 'Mini-Publics, (Lack of) Representativeness, and Legitimacy Beliefs'. *British Journal of Political Science* 55 (January):e11. <https://doi.org/10.1017/S0007123424000322>.

Goldsman, David, and Paul Goldsman. 2015. 'Discrete-Event Simulation'. In *Modeling and Simulation in the Systems Engineering Life Cycle: Core Concepts and Accompanying Lectures*, edited by Margaret L. Loper, 103–9. London: Springer. https://doi.org/10.1007/978-1-4471-5634-5_10.

Grieves, Michael. 2015. 'Digital Twin: Manufacturing Excellence through Virtual Factory Replication', March.

Grimmer, Justin, and Brandon M. Stewart. 2013. 'Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts'. *Political Analysis* 21 (3): 267–97. <https://doi.org/10.1093/pan/mps028>.

Grönlund, Kimmo, and Kaisa Herne. 2022. 'Experimental Methods'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press. <https://doi.org/10.1093/oso/9780192848925.003.0011>.

Grönlund, Kimmo, Kaisa Herne, and Maija Setälä. 2015. 'Does Enclave Deliberation Polarize Opinions?' *Political Behavior* 37 (4): 995–1020. <https://doi.org/10.1007/s11109-015-9304-x>.

Gutmann, Amy, and Dennis Thompson. 2000. 'Why Deliberative Democracy Is Different'. *Social Philosophy and Policy* 17 (1): 161–80. <https://doi.org/10.1017/S0265052500002570>.

Gutmann, Amy, and Dennis F. Thompson. 2009. *Democracy and Disagreement*. Harvard University Press.

Habermas, Jurgen. 1998. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Reprint edition. Cambridge, Mass.: The MIT Press.

Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. The Righteous Mind: Why Good People Are Divided by Politics and Religion. New York, NY, US: Pantheon/Random House.

Hartz-Karp, Janette. 2005. 'A Case Study in Deliberative Democracy: Dialogue with the City'. *Journal of Deliberative Democracy* 1 (1). <https://doi.org/10.16997/jdd.27>.

Hausladen, Carina I., Regula Hänggli Fricker, Dirk Helbing, Renato Kunz, Junling Wang, and Evangelos Pournaras. 2024. 'How Voting Rules Impact Legitimacy'. *Humanities and Social Sciences Communications* 11 (1): 1–10. <https://doi.org/10.1057/s41599-024-03056-8>.

Helbing, Dirk. 2010. 'Pluralistic Modeling of Complex Systems'. arXiv. <https://doi.org/10.48550/arXiv.1007.2818>.

Helbing, Dirk, and Javier Argota Sánchez-Vaquerizo. 2023. 'Digital Twins: Potentials, Ethical Issues and Limitations'. In *Handbook on the Politics and Governance of Big Data and Artificial Intelligence*, 64–104. Edward Elgar Publishing. <https://www.elgaronline.com/edcollchap-0a/book/9781800887374/book-part-9781800887374-13.xml>.

Helbing, Dirk, and Peter Seele. 2017. 'Turn War Rooms into Peace Rooms'. *Nature* 549 (7673): 458–458. <https://doi.org/10.1038/549458c>.

Hinchy, E. P., C. Carcagno, N. P. O'Dowd, and C. T. McCarthy. 2020. 'Using Finite Element Analysis to Develop a Digital Twin of a Manufacturing Bending Operation'. *Procedia CIRP*, 53rd CIRP Conference on Manufacturing Systems 2020, 93 (January):568–74. <https://doi.org/10.1016/j.procir.2020.03.031>.

Holman, Bennett, William J. Berger, Daniel J. Singer, Patrick Grim, and Aaron

- Bramson. 2018. 'Diversity and Democracy: Agent-Based Modeling in Political Philosophy'. *Historical Social Research / Historische Sozialforschung* 43 (1 (163)): 259–84.
- Iyengar, Shanto, Robert Luskin, Austin James, and S. Fishkin. 2003. 'Facilitating Informed Public Opinion: Evidence from Face-to-Face and Online Deliberative Polls', January.
- Izquierdo, Luis R., Segismundo S. Izquierdo, and William H. Sandholm. 2019. 'An Introduction to *ABED*: Agent-Based Simulation of Evolutionary Game Dynamics'. *Games and Economic Behavior* 118 (November):434–62. <https://doi.org/10.1016/j.geb.2019.09.014>.
- Jones, David, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. 2020. 'Characterising the Digital Twin: A Systematic Literature Review'. *CIRP Journal of Manufacturing Science and Technology* 29 (May):36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>.
- Kingzette, Jon, and Michael Neblo. 2022. 'Deliberative Field Experiments'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press. <https://doi.org/10.1093/oso/9780192848925.003.0012>.
- Landemore, Hélène. 2021. '2. Open Democracy and Digital Technologies'. In *Digital Technology and Democratic Theory*, edited by Lucy Bernholz, Hélène Landemore, and Rob Reich, 62–89. University of Chicago Press. <https://doi.org/10.7208/9780226748603-003>.
- Lawrence, John, Joonsuk Park, Katarzyna Budzynska, Claire Cardie, Barbara Konat, and Chris Reed. 2017. 'Using Argumentative Structure to Interpret Debates in Online Deliberative Democracy and eRulemaking'. *ACM Trans. Internet Technol.* 17 (3): 25:1-25:22. <https://doi.org/10.1145/3032989>.
- Lee, Bokwon, Yohan Kim, K.M Lee, and Jae-Suk Yang. 2022. 'An Agent-Based Model of Deliberative Democracy and Polarization'. *The Journal of Mathematical Sociology* 46 (2): 173–94. <https://doi.org/10.1080/0022250X.2020.1836625>.
- Levine, John Gastil and Peter, ed. 2005. *The Deliberative Democracy Handbook: Strategies for Effective Civic Engagement in the Twenty-First Century*. 1st edition. San Francisco, Calif: Jossey-Bass.
- Ligmann-Zielinska, Arika, Daniel B. Kramer, Kendra Spence Cheruvelil, and Patricia A. Soranno. 2014. 'Using Uncertainty and Sensitivity Analyses in Socioecological Agent-Based Models to Improve Their Analytical Performance and Policy Relevance'. *PLOS ONE* 9 (10): e109779. <https://doi.org/10.1371/journal.pone.0109779>.
- Lippi, Marco, and Paolo Torroni. 2016. 'Argumentation Mining: State of the Art and Emerging Trends'. *ACM Trans. Internet Technol.* 16 (2): 10:1-10:25. <https://doi.org/10.1145/2850417>.
- Listl, Franz Georg, Daniel Dittler, Gary Hildebrandt, Valentin Stegmaier, Nasser Jazdi, and Michael Weyrich. 2024. 'Knowledge Graphs in the Digital Twin: A Systematic Literature Review About the Combination of Semantic Technologies

and Simulation in Industrial Automation'. *IEEE Access* 12:187828–43. <https://doi.org/10.1109/ACCESS.2024.3514923>.

Liu, Bing. 2022. *Sentiment Analysis and Opinion Mining*. Springer Nature.

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. 'Computer-Assisted Text Analysis for Comparative Politics'. *Political Analysis* 23 (2): 254–77. <https://doi.org/10.1093/pan/mpu019>.

Lustick, Ian S., and Dan Miodownik. 2000. 'Deliberative Democracy and Public Discourse: The Agent-Based Argument Repertoire Model'. *Complexity* 5 (4): 13–30. [https://doi.org/10.1002/1099-0526\(200003/04\)5:4<13::AID-CPLX3>3.0.CO;2-G](https://doi.org/10.1002/1099-0526(200003/04)5:4<13::AID-CPLX3>3.0.CO;2-G).

Mansbridge, Jane. 2007. "“Deliberative Democracy” or “Democratic Deliberation”?" In *Deliberation, Participation and Democracy*, edited by Shawn W. Rosenberg, 251–71. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230591080_12.

Mansbridge, Jane, James Bohman, Simone Chambers, Thomas Christiano, Archon Fung, John Parkinson, Dennis F. Thompson, and Mark E. Warren. 2012. 'A Systemic Approach to Deliberative Democracy'. In *Deliberative Systems: Deliberative Democracy at the Large Scale*, edited by Jane Mansbridge and John Parkinson, 1–26. Theories of Institutional Design. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139178914.002>.

Marti, Jose Luis, and Samantha Besson, eds. 2006. 'The Epistemic Conception of Deliberative Democracy Defended Reasons, Rightness and Equal Political Autonomy'. In *Deliberative Democracy and Its Discontents*. Routledge.

Moore, Alfred. 2012. 'Following from the Front: Theorizing Deliberative Facilitation'. *Critical Policy Studies* 6 (2): 146–62. <https://doi.org/10.1080/19460171.2012.689735>.

Moshman, David. 2020. *Reasoning, Argumentation, and Deliberative Democracy*. New York: Routledge. <https://doi.org/10.4324/9780429316029>.

Musso, Andrea, and Dirk Helbing. 2024. 'How Networks Shape Diversity for Better or Worse'. *Royal Society Open Science* 11 (5): 230505. <https://doi.org/10.1098/rsos.230505>.

Niemeyer, Simon, and Francesco Veri. 2022. 'Deliberative Reason Index'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press. <https://doi.org/10.1093/oso/9780192848925.003.0007>.

Nino, Carlos Santiago. 1996. *The Constitution of Deliberative Democracy* /. Yale University Press,.

Orozco-Romero, Areli, Claudia Yohana Arias-Portela, and Jose Antonio Marmolejo- Saucedo. 2020. 'The Use of Agent-Based Models Boosted by Digital Twins in the Supply Chain: A Literature Review'. In *Intelligent Computing and Optimization*, edited by Pandian Vasant, Ivan Zelinka, and Gerhard-Wilhelm Weber, 642–52. Cham: Springer International Publishing.

https://doi.org/10.1007/978-3-030-33585-4_62.

Oswald, David, Fred Sherratt, and Simon Smith. 2014. 'Handling the Hawthorne Effect: The Challenges Surrounding a Participant Observer'. *Review of Social Studies* 1 (1): 53–73.

Palau, Raquel Mochales, and Marie-Francine Moens. 2009. 'Argumentation Mining: The Detection, Classification and Structure of Arguments in Text'. In *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, 98–107. ICAIL '09. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/1568234.1568246>.

Park, Joon Sung, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. 'Generative Agents: Interactive Simulacra of Human Behavior'. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, 1–22. UIST '23. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3586183.3606763>.

Parkinson, John, ed. 2006. *Deliberating in the Real World: Problems of Legitimacy in Deliberative Democracy*. Oxford University Press.

Parkinson, John, and Jane Mansbridge. 2012. *Deliberative Systems: Deliberative Democracy at the Large Scale*. Cambridge University Press.

Parsons, Simon D., Piotr Gymtrasiewicz, and Michael Wooldridge. 2012. *Game Theory and Decision Theory in Agent-Based Systems*. Springer Science & Business Media.

Prakken, Henry, and Gerard Vreeswijk. 2002. 'Logics for Defeasible Argumentation'. In *Handbook of Philosophical Logic*, edited by D. M. Gabbay and F. Guenther, 219–318. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-0456-4_3.

Qian, Chen, Zihao Xie, YiFei Wang, Wei Liu, Kunlun Zhu, Hanchen Xia, Yufan Dang, et al. 2025. 'Scaling Large-Language-Model-Based Multi-Agent Collaboration'. arXiv. <https://doi.org/10.48550/arXiv.2406.07155>.

Qiu, Hongbin, Yong Chen, Huaxiang Zhang, Wenchao Yi, and Yingde Li. 2023. 'Evolutionary Digital Twin Model with an Agent-Based Discrete-Event Simulation Method'. *Applied Intelligence* 53 (6): 6178–94. <https://doi.org/10.1007/s10489-022-03507-2>.

Qiu, Lin, and Riyang Phang. 2020. 'Agent-Based Modeling in Political Decision Making'. In *Oxford Research Encyclopedia of Politics*. <https://doi.org/10.1093/acrefore/9780190228637.013.913>.

Rathore, M. Mazhar, Syed Attique Shah, Dharendra Shukla, Elmahdi Bentafat, and Spiridon Bakiras. 2021. 'The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities'. *IEEE Access* 9:32030–52. <https://doi.org/10.1109/ACCESS.2021.3060863>.

Rowe, Gene, and Lynn J. Frewer. 2000. 'Public Participation Methods: A Framework for Evaluation'. *Science, Technology, & Human Values* 25 (1): 3–29. <https://doi.org/10.1177/016224390002500101>.

Ruijter, Erna, Carmen Dymanus, Erik-Jan van Kesteren, Laura Boeschoten, and Albert Meijer. 2024. 'Open Data Work for Empowered Deliberative Democracy: Findings from a Living Lab Study'. *Government Information Quarterly* 41 (1): 101902. <https://doi.org/10.1016/j.giq.2023.101902>.

Scudder, Mary F. 2022. 'Listening Quality Index'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press. <https://doi.org/10.1093/oso/9780192848925.003.0008>.

Setälä, Maija, Kimmo Grönlund, and Kaisa Herne. 2010. 'Citizen Deliberation on Nuclear Power: A Comparison of Two Decision-Making Methods'. *Political Studies* 58 (4): 688–714. <https://doi.org/10.1111/j.1467-9248.2010.00822.x>.

Shin, Bokyong, Jacqueline Floch, Mikko Rask, Peter Bæck, Christopher Edgar, Aleksandra Berditchevskaia, Pierre Mesure, and Matthieu Branlat. 2024. 'A Systematic Analysis of Digital Tools for Citizen Participation'. *Government Information Quarterly* 41 (3): 101954. <https://doi.org/10.1016/j.giq.2024.101954>.

Silva, Eduardo M. da, Antônio Carlos A. Ribeiro, and Silvio S. Higgins. 2022. 'Social Network Analysis'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press. <https://doi.org/10.1093/oso/9780192848925.003.0015>.

Smith, G. 2009. *Democratic Innovations: Designing Institutions for Citizen Participation*. Cambridge: Cambridge University Press. <https://westminsterresearch.westminster.ac.uk/item/90zq3/democratic-innovations-designing-institutions-for-citizen-participation>.

Stark, Alastair, N. K. Thompson, and Greg Marston. 2021. 'Public Deliberation and Policy Design'. *Policy Design and Practice* 4 (4): 452–64. <https://doi.org/10.1080/25741292.2021.1912906>.

Steel, Daniel, Naseeb Bolduc, Kristina Jenei, and Michael Burgess. 2020. 'Rethinking Representation and Diversity in Deliberative Minipublics'. *Journal of Deliberative Democracy* 16 (1). <https://doi.org/10.16997/jdd.398>.

Steenbergen, Marco R, André Bächtiger, Markus Spörndli, and Jürg Steiner. 2003. 'Measuring Political Deliberation: A Discourse Quality Index'. *Comparative European Politics* 1 (1): 21–48. <https://doi.org/10.1057/palgrave.cep.6110002>.

Thelen, Adam, Xiaoge Zhang, Olga Fink, Yan Lu, Sayan Ghosh, Byeng D. Youn, Michael D. Todd, Sankaran Mahadevan, Chao Hu, and Zhen Hu. 2022. 'A Comprehensive Review of Digital Twin — Part 1: Modeling and Twinning Enabling Technologies'. *Structural and Multidisciplinary Optimization* 65 (12): 354. <https://doi.org/10.1007/s00158-022-03425-4>.

Walton, Douglas, Christopher Reed, and Fabrizio Macagno. 2008. *Argumentation Schemes*. Cambridge University Press.

Wang, Zijin, Ou Zheng, Liangding Li, Mohamed Abdel-Aty, Carolina Cruz-Neira, and Zubayer Islam. 2023. 'Towards Next Generation of Pedestrian and Connected Vehicle In-the-Loop Research: A Digital Twin Co-Simulation Framework'. *IEEE*

- Transactions on Intelligent Vehicles* 8 (4): 2674–83.
<https://doi.org/10.1109/TIV.2023.3250353>.
- Weeks, Edward C. 2000. 'The Practice of Deliberative Democracy: Results from Four Large-Scale Trials'. *Public Administration Review* 60 (4): 360–72.
<https://doi.org/10.1111/0033-3352.00098>.
- Werner, Hannah, and Lala Muradova. 2022. 'Scenario Experiments'. In *Research Methods in Deliberative Democracy*, edited by Selen A. Ercan, Hans Asenbaum, Nicole Curato, and Ricardo F. Mendonça, 0. Oxford University Press.
<https://doi.org/10.1093/oso/9780192848925.003.0013>.
- Willis, Rebecca, Nicole Curato, and Graham Smith. 2022. 'Deliberative Democracy and the Climate Crisis'. *WIREs Climate Change* 13 (2): e759.
<https://doi.org/10.1002/wcc.759>.
- Yang, Joshua C., Damian Dailisan, Marcin Korecki, Carina I. Hausladen, and Dirk Helbing. 2024. 'LLM Voting: Human Choices and AI Collective Decision-Making'. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 7 (1): 1696–1708. <https://doi.org/10.1609/aies.v7i1.31758>.
- Yang, Joshua C., Carina I. Hausladen, Dominik Peters, Evangelos Pournaras, Regula Hnggli Fricker, and Dirk Helbing. 2024. 'Designing Digital Voting Systems for Citizens: Achieving Fairness and Legitimacy in Participatory Budgeting'. *Digit. Gov.: Res. Pract.* 5 (3): 26:1-26:30. <https://doi.org/10.1145/3665332>.
- Zhang, Lin, Longfei Zhou, and Berthold K. P. Horn. 2021. 'Building a Right Digital Twin with Model Engineering'. *Journal of Manufacturing Systems* 59 (April):151–64. <https://doi.org/10.1016/j.jmsy.2021.02.009>.
- Zhang, Nan, Christian Vergara-Marcillo, Georgios Diamantopoulos, Jingran Shen, Nikos Tziritas, Rami Bahsoon, and Georgios Theodoropoulos. 2024. 'Large Language Models for Explainable Decisions in Dynamic Digital Twins'. arXiv. <https://doi.org/10.48550/arXiv.2405.14411>.
- Zheng, Wenzhen, and Xijin Tang. 2025. 'Simulating Social Network with LLM Agents: An Analysis of Information Propagation and Echo Chambers'. In *Knowledge and Systems Sciences*, edited by Xijin Tang, Van Nam Huynh, Haoxiang Xia, and Quan Bai, 63–77. Singapore: Springer Nature.
https://doi.org/10.1007/978-981-96-0178-3_5.
- Zheng, Xiaochen, Jinzhi Lu, and Dimitris Kiritsis. 2022. 'The Emergence of Cognitive Digital Twin: Vision, Challenges and Opportunities'. *International Journal of Production Research* 60 (24): 7610–32.
<https://doi.org/10.1080/00207543.2021.2014591>.