Testing Deliberative Democracy Through Digital Twins

Claudio Novelli[1], Javier Argota Sánchez-Vaquerizo[2], Dirk Helbing[2,3], Antonino Rotolo[3], Luciano Floridi[1,3]

[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 relies on well-designed institutional frameworks—like participant selection, facilitation, and decision-making. Yet identifying the best design for a given context is challenging, as real-world and lab-based studies are often costly, time-consuming, and difficult to replicate. This commentary proposes Digital Twin (DT) technology as a regulatory sandbox for deliberative democracy. By simulating dynamic, data-driven models of real or synthetic communities, DTs allow researchers and policymakers to test alternative designs through controlled "what-if" scenarios, free from real-world constraints. The commentary also examines limitations and key directions for future research.[i]

1. Introduction

Deliberative democracy emphasizes reasoned debate and mutual justification in public decision-making. While broadly valued for promoting just laws and civic discourse, its effectiveness over traditional aggregative methods depends on well-designed institutions. Key design elements include participant recruitment, training, debate structure, timing, decision-making processes, and monitoring (Fung 2007).

However, selecting and evaluating optimal designs for specific deliberative goals is challenging, as real-world and lab experiments often face issues of generalizability and impartiality.

This commentary proposes Digital Twin (DT) technology as a regulatory sandbox for deliberative democracy. DTs simulate real or synthetic deliberative communities by integrating real-world data (e.g., demographic, behavioral, social) with computational tools like agent-based modeling, machine learning, and network analysis. Policymakers, researchers, and advocates can use DTs to test institutional designs, generate predictive insights, and refine deliberative procedures in virtual environments.

2. The Institutional Design of Deliberative Democracy

Deliberative democracy differs from free discussion, as its success heavily depends on 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 (Parkinson and Mansbridge 2012; Fishkin 2009; Gutmann and Thompson 2000):

a) *Pre-deliberation rules* establish the framework for deliberation, determining participant selection—whether random, stakeholder-based, or voluntary—and agenda setting, which may be predefined, facilitator-led, or group-defined. They also provide participants with essential background information, expert input, or open data access, enabling them to make informed policy decisions (Ruijer et al. 2024).

b) *Discussion rules* govern the conduct of deliberation, including formats (single sessions, multi-round, or iterative) and speaking structures (facilitator-controlled, time-limited, or queued). Decision-making methods vary from consensus and majority voting (Cohen 2021) to deliberative polling focused on informed opinion (Fishkin and Luskin 2005). Facilitators may act neutrally, steer agendas, or rotate roles (Escobar 2019; Moore 2012). Additional rules address argument and rebuttal formats, conflict resolution mechanisms (e.g., structured dialogue or majority rule), and criteria for evaluating arguments, such as iterative refinement or scoring (Bächtiger and Parkinson 2019).

c) *Post-deliberation rules* focus on outcomes and accountability. These include documentation (e.g., summary reports, recommendations), participant feedback (e.g., surveys, evaluations) (Hartz-Karp 2005), publication choices, and follow-up measures like public updates or audits to track implementation.

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.

3. The Challenges of Modeling and Testing Deliberative Democracy

Deliberative democracy, as a normative theory without a universal model, makes testing its institutional design 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).

These methods allow for isolating variables—like discussion format or participant makeup—to examine causal effects, such as how voting methods shape legitimacy perceptions (Hausladen et al. 2024).[1] However, both real-world and experimental methods have limitations:

1) *Replicability* Real-world deliberations are hard to reproduce due to unique socio-political, cultural, and institutional contexts, as well as reliance on political will, funding, or community enthusiasm. Lab-in-the-field experiments improve replicability through standardized procedures (Grönlund and Herne 2022), but this comes at the cost of reducing real-world complexity, raising concerns about reliability.

2) *Generalizability*. Both real-world and lab-based deliberations struggle to generalize findings. Real-world events are context-specific and often don't translate across policy domains (Levine 2005; Parkinson 2006). Lab studies usually use small, selective samples that may not reflect broader populations, limiting scalability. Additionally, they tend to measure short-term effects, while deliberative democracy often depends on long-term observation (Friedman 2006).

3) *Flexibility to iterations and follow-ups*. Refining deliberative processes requires iterative testing, but real-world settings don't allow mid-course adjustments. Outcomes are only observable post-process, making revisions impractical. Lab experiments theoretically allow more flexibility yet often maintain fixed protocols for reliability, limiting adaptability. Moreover, visible changes during testing may influence participant behavior (Lee et al. 2022).

4) *Observer Effect*. In experimental settings, participants know they're being observed (the Hawthorne effect), which may distort behavior. They may conform to dominant views or avoid dissent to maintain social desirability (Oswald, Sherratt, and Smith 2014), reducing authenticity. Accordingly, participants in a deliberative democracy experiment may alter their behavior when they know peers, media, or researchers are observing them (Gastil 2000).

5) *Participant compliance, retention, and engagement*. Maintaining protocol adherence and engagement is challenging in both real-world and experimental settings. Time demands, low perceived value, and one-time

[1] 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).

incentives can lead to noncompliance or dropout, affecting data quality and representativeness. While labs can better enforce rules, they still face similar risks.

6) *Resource and time constraints*. Deliberative experiments are resource-intensive. They require funding, trained facilitators, participant recruitment, venues, expert materials, and data analysis (Iyengar et al. 2003). Multi-session formats and follow-ups add logistical burdens. Lab experiments also demand resources, especially for diverse or repeated trials.

4. The Digital Twin Technology

A DT is a dynamic, computer-based model replicating 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). 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). Ultimately, 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).

5. Testing Deliberative Democracy Through Digital Twins

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 conception of deliberative democracy and evaluation metrics is defined, data collection is next. 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 sources include official government data such as debate transcripts, voting records, and party communications. Data can be retrieved via APIs or exports; where unavailable, web scraping may be used to update interaction metrics (Franco-Guillén, Laile, and Parkinson 2022, 236).

A key concern is data quality and bias. Researchers must assess how incomplete or biased data may distort predictions and document mitigation strategies (Lucas et al. 2015). Interpretation and context also matter.

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, use NLP to map debate structures and identify controversial topics and conflicting views (Lawrence et al. 2017).

The modeling phase is incredibly complex. Following a pluralistic modeling approach (Helbing 2010), DTs often integrate multiple paradigms—using both historical simulations and predictive models—to simulate deliberative democracy and test how variable changes affect outcomes. Agent-Based Modeling (ABM) has proven helpful in studying democracy and decision-making (Qiu and Phang 2020) and has seen growing application in deliberative democracy (Lustick and Miodownik 2000; Lee et al. 2022). ABM represents each community member or facilitator as an "agent" defined by sociodemographic attributes (e.g., age, income, education) and guided by decision rules, preferences, psychological traits, and engagement levels (e.g., voting propensity). It simulates agent interactions to capture processes like opinion diffusion, coalition formation, and resource allocation.

A different but complementary approach to modeling 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). At the same macro-level, Social Network Analysis (SNA) maps social ties and communication flows, aiding in identifying influential actors in deliberation (da Silva, Ribeiro, and Higgins 2022).

Recent advances in Large Language Models (LLMs) have enhanced agentic AI by improving knowledge retrieval, adaptability, and reasoning, especially when paired with tools like Retrieval-Augmented Generation (RAG) (Zhang et al. 2024). ABMs can now include autonomous agents with cognitive modules (e.g., memory, social awareness, communication) capable of exhibiting human-like behavior. LLMs have also advanced the study of emergent properties in social networks—such as information spread, opinion dynamics, and echo chambers (Zheng and Tang 2025)—and have helped improve participatory budgeting by enabling multiple tailored solutions that boost participant satisfaction. These innovations support alternatives to traditional deliberation formats, such as co-creative methods like 're-mixing' (Carpentras, Hänggli Fricker, and Helbing 2024).

Lastly, though uncommon in this domain, Discrete-Event Simulation (DES) can model how events like consensus-building or voting phases influence outcomes in deliberative processes (Charalabidis 2011), offering a promising, underexplored tool.

Regardless of the modeling approach, the model should be able to predict how agents shift opinions due to peer influence, news, or official statements. AI techniques like Machine Learning can help forecast sentiment or voter turnout using historical data.

After development, domain experts, community leaders, or researchers must calibrate and validate the model. Ongoing refinement may be needed to integrate new data, account for feedback loops, and enhance accuracy over time.

5.1. Twinning Mini-Publics

To illustrate this approach, consider building a DT of a mini-public—a small, demographically representative group of city residents convened to deliberate on issues like environmental funding (Germann 2025). Ensuring representativeness is complex, but DT designers can use institutional data (e.g., demographics) and engagement platform data (e.g., behaviors, interactions) to test different group compositions and assess whether outcomes align with broader population expectations.

In an Agent-Based Model (ABM) or similar simulation, agents are assigned demographic attributes based on city-wide distributions if participant-level data are unavailable. Engagement metrics—such as posts, comments, votes, and timestamps—help distinguish active from passive participants. Sentiment analysis and topic modeling reveal policy preferences.

These inputs inform agent profiles that simulate behavior and orientation. The goal isn't to replicate a specific community but to model a synthetic group grounded in real-world data, enabling sensitivity analyses of participant selection biases in mini-publics. Finally, in line with data protection standards (Bäumer et al. 2024), all personal data must be anonymized or pseudonymized while preserving demographic and behavioral diversity for accurate modeling.

After data cleaning, information is integrated into the mini-public's DT. The core model may use ABM, where each agent represents a participant with age, income, and education attributes. Decision rules reflect real-world behaviors—e.g., likelihood to post vs. lurk or to shift views when presented with strong arguments. Interaction rules are based on digital platform behavior, simulating how agents communicate, form coalitions, or change opinions. Demographic clustering observed in actual data can be modeled by increasing the likelihood of alignment among similar agents or epistemic influence. For instance, a moderately concerned agent may become more environmentally engaged after interacting with a respected peer.

The same data, especially on participation trends, can support a System Dynamics (SD) model to capture macro-level feedback loops—e.g., how dissatisfaction reduces future participation or how prolonged debate fosters consensus and fatigue.

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 networks can be analyzed to identify influencers (via centrality), subgroups, and idea flow, revealing patterns like bridging connections (where subgroups interact) or echo chambers (where subgroups become isolated).

Discrete-Event Simulation (DES) can model how specific events impact the mini-public. Real-world actions—like posting, voting, or stance changes—define event triggers, with frequencies from observed data (e.g., "an influencer emerges in 1 of 3 debates") converted into probabilities. Machine Learning (ML), Natural Language Processing (NLP), and sentiment analysis help predict outcomes such as participant dropout, argument submission, or proposal success. Sentiment analysis can track shifts in support, neutrality, or opposition (Liu 2022).

Dialogue constructiveness can be measured by argument depth (e.g., evidence use) and balanced speaking time (Scudder 2022). DT agents can reflect behaviors like presenting evidence, engaging with opposing views, or staying silent. SNA indicators reveal idea flow and dominant groups, assessing how influence affects debate quality and pluralism (Musso and Helbing 2024).

Finally, the DT must be calibrated to align with real-world behavior. One method is simulating a known past event (e.g., a prior participatory budgeting process) and comparing predicted outcomes with actual results to validate and refine the model.

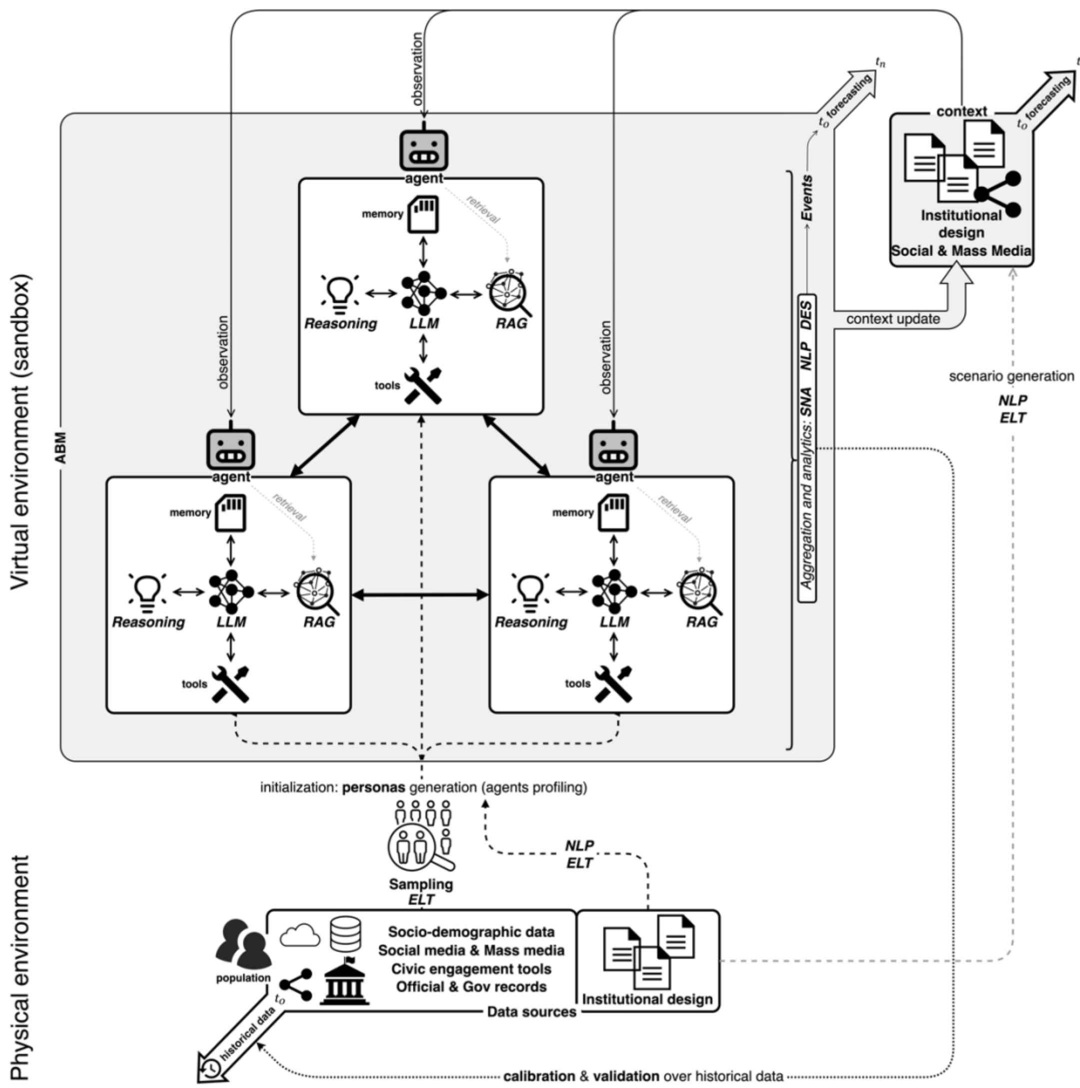

Figure 1. Proposed Multi-agent-based model (ABM) architecture for a democracy deliberation DT, based on AI-powered agents profiled as different personas from actual socio-demographic, behavioral, and public records data, which is pre-processed through extract, load, and transform (ELT) processes and using Natural Language Processing (NLP) techniques. Each agent includes modules for reasoning, information retrieval, memory, and action planning, enabling interaction with other agents and contextual elements (e.g., institutional design, news) and adapting through feedback. The virtual environment reflects emergent social network structures, analyzable via Social Network Analysis (SNA) and NLP. Aggregated social dynamics can inform Discrete-Event Simulation (DES) for large-scale modeling, allowing calibration and performance evaluation using historical data layers and updating context to forecast new scenarios.

5.2. Using the DT to Test (and Refine) Deliberative Democracy Rules

Our goal is not to replace real deliberative communities with digital replicas but to use DTs as sandboxes for testing procedural rules. DTs allow experimentation that may be impractical or unethical in real settings (Helbing and Sánchez-Vaquerizo 2023, 65), especially under the constraints outlined in Section 3.1. We illustrate this through the three main procedural rule sets discussed in Section 2.

a) *Pre-deliberation rules.* DTs, using ABM, can test how different recruitment strategies—random, weighted, voluntary, or stakeholder-targeted—affect diversity, inclusivity, and discussion quality. 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). 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.

   Interestingly, deliberation quality can serve as a proxy for representational success. For instance, measuring “argument diversity” — using metrics such as entropy indices or the number of distinct argument types (Grimmer and Stewart 2013) — can help determine whether including minority voices meaningfully influences discussion content or if they remain overshadowed in practice.

   Informational preparation is another key rule. Here, ABM and SNA can model how information is absorbed and spreads, identifying influential voices. The goal isn't to suppress influence but to reduce its concentration, promoting balanced participation and minimizing uncritical deference. DES can help determine optimal preparation duration. Simulations might show expert presentations boost knowledge but risk excessive deference, while self-guided reading encourages independent 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.

Since perfect representativeness is often unattainable, DTs can simulate real-world constraints like budget, time, and availability to refine selection strategies. DES or system dynamics can estimate the required resources for each rule to maximize inclusion within limits.

b) *Discussion rules*. DT simulations can optimize discussion formats by testing various session structures. For example, simulations can assess whether deeper argumentation arises from single sessions or multiple rounds (e.g., alternating homogeneous and heterogeneous panels). Argument complexity can be measured by reasoning layers (e.g., premises and rebuttals) and degree of justification (e.g., evidence use and counterpoint engagement) (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) suggest that although multiple rounds outperform simple information provision, their incremental utility decreases over time, with limited effects on reducing polarization.

Simulations can also examine how participants express opinions and rebut arguments and how these affect fairness and engagement. DES can model the impact of time constraints on argument quality, while ABM with Natural Language Processing (NLP) can assess whether rebuttals deepen or stall discussion. SD can simulate how rebuttal styles (e.g., iterative vs. paired) shape opinion shifts and consensus.

Facilitation styles can also be tested: DTs can evaluate whether frequent moderator interventions support or hinder argument diversity and flow. SNA can detect power imbalances, such as facilitators over-engaging specific participants.

Decision-making strategies—key to discussion rules—can be explored using ABM and game-theoretic models, viewing participants as strategic (rational or boundedly rational) rather than purely sincere (Parsons, Gymtrasiewicz, and Wooldridge 2012). Simulations may model cooperative vs. aggressive behaviors (e.g., hawk-dove games) and test if rules like mediation or facilitation promote cooperation (Amadae and Watts 2023). All these tools help weigh trade-offs: consensus-building may deepen dialogue and participation but slow decisions, while majority voting speeds processes but risks marginalizing certain groups, reducing fairness and engagement.

c) *Post-deliberation rules*. Unlike pre- and during-deliberation rules, post-deliberation processes should rely on real-world data—such as surveys, voting records, and participatory policy histories—since these interactions are highly context-specific and challenging to simulate. As a result, the focus shifts from simulation to data analytics.

Analytics and predictive modeling can assess adequate documentation and reporting strategies. For example, they can compare whether open-access vs. restricted reports increase trust or risk misinterpretation. Historical feedback data can also reveal how feedback methods (anonymous vs. named, immediate vs. delayed, group reflection) affect engagement and reflection depth.

Sentiment analysis and NLP help identify participant reactions and feedback focus areas but should be supplemented with qualitative analysis. For instance, if open reports elicit more positive sentiment (e.g., trust, willingness to re-engage), they may boost public confidence. Conversely, adverse reactions (e.g., anger and skepticism) could signal data overload or miscommunication.

It is crucial to stress 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. Designers might focus on specific aspects—e.g., decision-making—if other elements like recruitment are already well-established.

6. Limitations and Future Research

DTs offer a powerful tool for empirical research on procedural rules, addressing the challenges of replicating, iterating, and scaling deliberative experiments in physical or lab settings. However, their effectiveness relies on accurate behavioral assumptions, high-quality input data, and their ability to generalize to real-world democracy. Oversimplified agent behaviors (e.g., disengagement patterns) or flawed decision rules can bias outcomes. Modeling complex, human-based social systems—especially network dynamics—is inherently difficult (Caldarelli et al. 2023).

These challenges underscore the trade-offs in building large-scale societal DTs. Full digital replicas may be infeasible, favoring issue-specific DTs instead. Yet, DTs still provide value through controlled, reproducible experimentation and support for adaptive experimental design to explore vast variables and interventions (Offer-Westort, Coppock, and Green 2021). Comparing AI and human responses to similar challenges can help improve DT's ecological validity.

In fact, DTs, long in development, can support a spectrum of data-driven, AI-managed societies—from digital democracies to more centralized models. As with urban DTs, they range from closely mirroring real environments to decoupled simulations for testing alternatives. In democratic deliberation, this allows DTs to serve as predictive tools or interactive systems, evolving alongside human processes and enabling hybrid intelligence involving humans and AI (e.g., Human-in-the-Loop collaboration).

Key research priorities should ensure Digital Twins (DTs) align with liberal and deliberative democratic values:

- *Avoid surveillance*. DTs should not replicate individuals exactly but use "noisy" hypothetical personas with representative traits to protect privacy and prevent tracking or targeting.

- *Ensure transparency and accountability*. Data, software, and procedures must be transparent to build public trust, prevent manipulation, and support accountability throughout design and implementation.

- *Open, plural, and fair discussion*. DT platforms must ensure a level playing field, enabling respectful dialogue and inclusive expression of diverse views and interests without fear.

- *Avoid mis- and disinformation*. Platforms should promote verified, evidence-based information, emphasizing critical thinking and context rather than raw data prone to misinterpretation.

- *Human-centered approach*. DTs should prioritize human needs over efficiency or control. Cognitive architectures should reflect human reasoning, not just data-driven outputs.

- *Respect multi-dimensional values*. DTs must consider legal, ethical, cultural, and emotional values such as dignity, trust, love, and creativity—not just utility or efficiency.

- *Enhance human agency*. Platforms should empower people to co-create solutions through participatory engagement.

In essence, DTs should support open, inclusive collaboration, not top-down control. Like a "peace room", they should offer a space for multiple stakeholders to explore, deliberate, and simulate solutions before real-world action.

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

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.

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.

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.

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.

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.

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. 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.
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.
Fishkin, James. 2009. *When the People Speak: Deliberative Democracy and Public Consultation*. OUP Oxford.
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.
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.
Gastil, John. 2000. *By Popular Demand: Revitalizing Representative Democracy Through Deliberative Elections*. First Edition. Berkeley, CA: University of California Press.
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.
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.
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-oa/book/9781800887374/book-part-9781800887374-13.xml.
Iyengar, Shanto, Robert Luskin, Austin James, and S. Fishkin. 2003. 'Facilitating Informed Public Opinion: Evidence from Face-to-Face and Online Deliberative Polls', January.
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.
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.
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.
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.

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.

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.

Offer-Westort, Molly, Alexander Coppock, and Donald P. Green. 2021. ‘Adaptive Experimental Design: Prospects and Applications in Political Science’. *American Journal of Political Science* 65 (4): 826–44. https://doi.org/10.1111/ajps.12597.

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.

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.

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.

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.

Ruijer, 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.

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.

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.

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.

---

[i] This commentary is a condensed version of a longer paper currently available on public repositories.